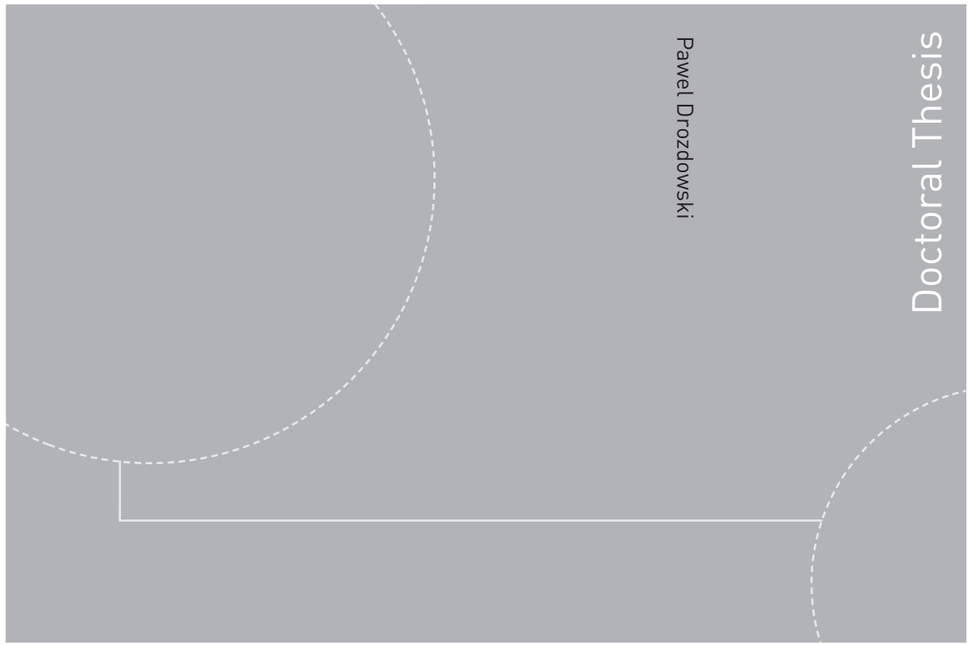
ISBN 978-82-326-4578-7 (printed version)
ISBN 978-82-326-4579-4 (electronic version)
ISSN 1503-8181



Doctoral theses at NTNU, 2020:115

Pawel Drozdowski

Efficient privacy-preserving biometric identification in large-scale multibiometric systems

NTNU
Norwegian University of
Science and Technology
Faculty of Information Technology
and Electrical Engineering
Department of Information Security
and Communication Technology

Doctoral theses at NTNU, 2020:115

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

Pawel Drozdowski

Efficient privacy-preserving biometric identification in large-scale multibiometric systems

Thesis for the degree of Philosophiae Doctor

Gjøvik, May 2020

Norwegian University of Science and Technology
Faculty of Information Technology
and Electrical Engineering
Department of Information Security and Communication Technology

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Information Technology
and Electrical Engineering
Department of Information Security and Communication
Technology

ISBN 978-82-326-4578-7 (printed version)
ISBN 978-82-326-4579-4 (electronic version)
ISSN 1503-8181

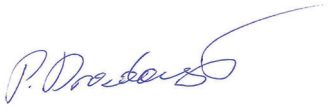
Doctoral theses at NTNU, 2020:115

Printed by Skipnes Kommunikasjon as

Declaration of Authorship

I, Pawel Drozdowski, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed:



(Pawel Drozdowski)

Date: 2019–11–11

Summary

In recent years, applications of biometric systems on national and international scale have appeared. Biometric identification is one of the important operational modes of such systems. It entails ascertaining the data subject identity corresponding to a given biometric sample, solely using the information from said biometric sample, *i.e.* effectively conducting a nearest-neighbour search. The naïve search method, *i.e.* an exhaustive (linear) search of the biometric enrolment database, suffers from two drawbacks, namely: high computational workload and increased probability of false positive occurrences.

Consequently, research into computationally efficient methods of biometric identification is necessary; it is the main topic covered in this thesis. Specifically, the key contributions of this thesis are:

- Formulation of a taxonomy for conceptual categorisation of methods of efficient biometric identification. A comprehensive survey of the relevant existing publications and organisation thereof in the context of the developed taxonomy.
- Development of methods which substantially decrease (by space search and/or template comparison cost reduction) the computational workload requirements of the biometric identification transactions, including:
  - Methods which take advantage of the intrinsic properties of certain types of biometric characteristics and/or biometric feature representations.
  - Methods which can be applied irrespective of the type of biometric characteristic and the biometric feature representation.
  - Methods which utilise biometric information fusion.
- Development of methods (both general purpose and biometric characteristic specific) of biometric template protection in the aforementioned context of computationally efficient biometric identification systems.
- Development of methods relevant to other (*e.g.* stress testing and usability) aspects of the operational biometric identification systems.

Acknowledgements

I would like to thank my supervisors: Prof. Dr. Busch and Dr. Rathgeb. I greatly appreciate their commitment and time spent on guiding and advising me throughout the period of my doctoral studies. They consistently provided ideas, critical insights, and feedback, all of which were immensely valuable for my research and writing.

I thank my co-authors from the research articles included in this thesis for the interesting, friendly, and fruitful collaborations.

I acknowledge the financial support of the German Federal Ministry of Education and Research (BMBF), the Hessian State Ministry for Higher Education, Research and the Arts (HMWK) within CRISP and ATHENE, and the Hessian State Offensive for the Development of Scientific and Economic Excellence (LOEWE-3).

Last but foremost, I express my utmost gratitude to my family and girlfriend, whose consistent support and encouragement made my doctoral studies and the writing of this thesis possible.

Contents

Summary i

Acknowledgements iii

Contents v

List of Figures ix

List of Tables xiii

List of Algorithms xvii

I Overview 1

1 Introduction 3
1.1 Motivation . . . 3
1.2 Background . . . 4
1.3 Overview of Research Projects . . . 10
1.4 Thesis Organisation . . . 11
1.5 Bibliography . . . 12

2 Thesis Scope 17
2.1 Research Questions . . . 17
2.2 Evaluation Environment . . . 20
2.3 Bibliography . . . 24

3 Contribution 29
3.1 Research Articles . . . 29
3.2 Bibliography . . . 47

II Related Work 49

4 Computational Workload in Biometric Identification Systems: An Overview 51
4.1 Introduction . . . 51

4.2 Background 55
4.3 Computational Workload Reduction Approaches 60
4.4 Discussion 74
4.5 Summary 83
4.6 Bibliography 84

III Research Articles **103**

5 Database Binning and Retrieval in Multi-Fingerprint Identification Systems **105**
5.1 Introduction 106
5.2 Fingerprint class statistics 108
5.3 Multi-fingerprint binning and retrieval 111
5.4 Performance evaluation 114
5.5 Conclusion 116
5.6 Bibliography 118

6 Multi-Iris Indexing and Retrieval: Fusion Strategies for Bloom Filter-based Search Structures **121**
6.1 Introduction 121
6.2 Related Work 122
6.3 Bloom Filter-based Search Structure 123
6.4 Multi-Iris Indexing and Retrieval 126
6.5 Experimental Setup 128
6.6 Results and Discussion 130
6.7 Conclusion 135
6.8 Bibliography 135

7 Privacy-Preserving Indexing of Iris-Codes with Cancelable Bloom Filter-based Search Structures **139**
7.1 Introduction 139
7.2 Privacy-preserving Indexing of Iris-Codes 142
7.3 Experiments 145
7.4 Summary 149
7.5 Bibliography 149

8 Benchmarking Binarisation Schemes for Deep Face Templates **153**
8.1 Introduction 153
8.2 Related Work 154
8.3 Binarisation of Deep Face Templates 155
8.4 Experiments 157
8.5 Summary 161
8.6 Bibliography 161

9 On the Application of Homomorphic Encryption to Face Identification 165
9.1 Introduction . . . 165
9.2 Proposed System . . . 167
9.3 Experiments . . . 170
9.4 Technical Considerations . . . 171
9.5 Summary . . . 173
9.6 Bibliography . . . 173

10 Multi-biometric Identification with Cascading Database Filtering 177
10.1 Introduction . . . 177
10.2 Background and Related Work . . . 179
10.3 Proposed System . . . 183
10.4 Experimental Setup . . . 188
10.5 Results . . . 195
10.6 Discussion . . . 198
10.7 Summary . . . 202
10.8 Bibliography . . . 203

11 Turning a Vulnerability into an Asset: Accelerating Facial Identification with Morphing 211
11.1 Introduction . . . 211
11.2 Background and Related Work . . . 212
11.3 Proposed System . . . 214
11.4 Experiments . . . 215
11.5 Summary . . . 217
11.6 Bibliography . . . 219

12 Towards Pre-alignment of Near-infrared Iris Images 223
12.1 Introduction . . . 223
12.2 Related Work . . . 225
12.3 Results . . . 228
12.4 Conclusion . . . 234
12.5 Bibliography . . . 235

13 Detection of Glasses in Near-infrared Ocular Images 239
13.1 Introduction . . . 239
13.2 Experimental Setup . . . 241
13.3 Impact of Glasses on Iris Recognition . . . 242
13.4 Automatic Detection Approaches . . . 243
13.5 Conclusion . . . 249
13.6 Bibliography . . . 250

14 SIC-Gen: A Synthetic Iris-Code Generator 253
14.1 Introduction . . . 253
14.2 Proposed Method . . . 255

14.3 Validation 258
14.4 Conclusion and Future Work 260
14.5 Bibliography 261

15 Score Fusion Strategies in Single-Iris Dual-Probe Recognition Systems 265
15.1 Introduction 265
15.2 Fusion Strategies 266
15.3 Performance Evaluation 268
15.4 Summary 272
15.5 Bibliography 273

IV Conclusions 275

16 Summary of Results 277
16.1 Research Question 1 277
16.2 Research Question 2 277
16.3 Research Question 3 278
16.4 Research Question 4 278
16.5 Research Question 5 279
16.6 Summary 279
16.7 Bibliography 280

17 Future Work 283
17.1 Scalability 283
17.2 Unconstrained Data 283
17.3 Deep Learning 284
17.4 Standardisation 284
17.5 Bibliography 285

V Appendix 287

A Bloom Filter-based Search Structures for Indexing and Retrieving Iris-Codes 289
A.1 Introduction 290
A.2 Workload Reduction in Iris Biometric Systems 291
A.3 Methodology 294
A.4 Experimental Setup 302
A.5 Results 305
A.6 Conclusion and Future Research 310
A.7 Bibliography 311

Nomenclature 315

List of Figures

1.1 Examples of biometric characteristics (images taken from from publicly available research databases). 4
1.2 A conceptual overview of the components and information flow in a biometric system (from ISO/IEC 19795-1 [14]). 5

4.1 Example images of some biometric characteristics commonly used in large-scale biometric identification systems (taken from the MCYT, FRGC, and IITD databases) 52
4.2 Taxonomy of methods used for the purpose of speeding-up biometric identification 56
4.3 Conceptual view of pre-filtering approaches 61
4.4 Conceptual view of binning approaches 63
4.5 Conceptual view of data-structures approaches 66

5.1 Example fingerprints for each of the five classes displaying minutiae, core and delta points (images generated using Synthetic Fingerprint Generator (SFinGe) [5]) 109
5.2 Sample images from the SD9 database 109
5.3 System overview 112

6.1 Indexing and retrieval in the Bloom filter-based system. In this case, the retrieval follows the bold arrow path down to a leaf, where the final decision is made. 125
6.2 Example images from the datasets 129
6.3 Filtered experimental results. The iris-code baseline is not visible, as it is located at $F = 1.0$ and $TP_{0.01} \approx 1.0$. 133

7.1 An overview of the proposed system. 140
7.2 Indexing and retrieval in the Bloom filter-based system. In this case, the retrieval follows the bold arrow path down to a leaf, where the final decision is made. 144
7.3 Example images from the chosen datasets. 145

7.4 DET curves for the proposed system. The faint colours around the curves represent the 95% confidence interval, while the black line represents the baseline (with EER of 0.66) – an Iris-Code system performing an exhaustive search and using ± 4 bit-shifts for sample alignment compensation. 146

8.1 Processing chain 154
8.2 Quantisation 156
8.3 Example images after pre-processing 158
8.4 DET curves 159

9.1 System overview sketch 168
9.2 DET curves for biometric identification with original and quantised features (with 95% CI) 172

10.1 Overview of the proposed system 184
10.2 Determining s_ϵ based on a training CMC curve and ϵ 186
10.3 Example images from the selected datasets 189
10.4 Baseline results 194
10.5 Estimation of the shortlist sizes 197
10.6 Proposed system's results 198
10.7 Summary of the results – best configuration for each of the tested fusion methods 201

11.1 Proposed system overview (here, $n = 2$) 212
11.2 Morphing example (from Scherhag *et al*. [23]) 214
11.3 Template comparisons per identification transaction 215
11.4 Results (with errorbars denoting the 95% confidence interval) . . 216

12.1 Iris image with eye corner landmarks (red), the rotation center (green), the horizon line and the frame for cropping and rotation (a) as well as the resulting image (b). 227
12.2 Insertion of an iris image to a high resolution frontal face image. . 228
12.3 The 9 landmarks automatically detected by the model on a sample image. The curves show locating the eye corners by fitting circles (green) and polynomials (blue) to the eyelid landmarks. . . 229
12.4 Example images from the BioSecure database. 229
12.5 Eye with muscles responsible for torsional movement in the eye socket highlighted. Images by Patrick J. Lynch, medical illustrator (CC BY 2.5). . . 231
12.6 Example images with landmarks detected by the proposed approaches: FaceLD - basic (black), EyeLD - Corners (red), EyeLD - Polynomial (green), EyeLD - Circle (blue). 232
12.7 Biometric performance comparison for the evaluated approaches (note the logarithmic scale of the y-axis). 232
12.8 Cumulative distributions of the distance from the optimal alignment achieved by the presented pre-alignment approaches. 233

12.9 Kernel density estimate of impostor scores from no (red) to $K = \pm 24$ bits (blue) rotation compensation. 234

13.1 Example images from the CASIA-Thousand dataset. Samples (a) and (b) are captured from the same eye instance. 241
13.2 Segmentation failures caused by glasses 242
13.3 BSIF-based approach 244
13.4 Reflection detection with a relative brightness measure. The two specular reflections caused by the glasses are clearly observed by this proposed metric. 246
13.5 Edge detection and measurement 248
13.6 A scatter plot of edge and reflection scores for all images from the CASIA-Thousand dataset, which shows significant separation between the two image classes 249
13.7 Examples of incorrectly classified images from all 3 methods. Figures (a)-(b) falsely classified as glasses, figures (c)-(f) falsely classified as non-glasses. 250

14.1 The process of generating an Iris-Code pair with SIC-Gen 256
14.2 Example Iris-Codes produced from real eye images and generated by the proposed method 258
14.3 Distributions of Hamming distances for a large number of comparisons between synthetic templates 259
14.4 Visualisation of lengths of sequences of consecutive bits in real data from BioSecure database, SIC-Gen synthetic templates and synthetic templates generated with Daugmann's HMM 260
14.5 Example error patterns for comparisons between the real Iris-Codes from the BioSecure dataset and between the synthetic Iris-Codes 260

15.1 Single-iris dual-sample iris recognition 267
15.2 Example images from the datasets 268
15.3 Iris recognition processing chain 269
15.4 ROC curves 270
15.5 Scatter plots for w-ARP scheme showing the dependence of biometric performance on the a parameter 270
15.6 Kernel density estimates for the score distributions 272

A.1 Indexing and retrieval in the Bloom filter-based system. In this case, the retrieval follows the bold path down to a leaf, where the final decision is made. 298
A.2 Lookup in the Bloom filter-based system 301
A.3 Example images from the datasets 303

A.4 Iris recognition processing chain: (a) iris detection in the raw image, (b) normalized pre-processed iris texture, and (c)-(d) iris-codes of applied feature extractor. Image taken from CASIA-v4-Interval iris database [3]. 304
A.5 Fit between the model and real data for all the relevant system configurations 305
A.6 The correlation between biometric performance and top tree level node filling 306
A.7 ROC curves comparison for different system versions 309

List of Tables

2.1 Used datasets 21
2.2 Summary of the data processing pipelines 22

3.1 Relations between the research articles and the research questions 31

4.1 Examples of currently operational and planned large-scale biometric identification systems around the world 53
4.2 Pre-filtering approaches 62
4.3 Binning approaches 64
4.4 Data-structures approaches 67
4.5 Feature transformation approaches 69
4.6 Other approaches 71

5.1 Most relevant fingerprint classification approaches proposed in the last five years 107
5.2 Fingerprint class distributions 108
5.3 Distributions of fingerprint class combinations for two contiguous fingers 110
5.4 Distributions of fingerprint class combinations for three contiguous fingers 110
5.5 Distributions of fingerprint class combinations for four contiguous fingers 111
5.6 Distributions of fingerprint class combinations for five contiguous fingers 112
5.7 CCR at a confidence interval of 95% for the classification of single fingerprints 115
5.8 Single-finger binning and retrieval results 116
5.9 Multi-finger binning and retrieval results 117

6.1 Evaluation dataset overview 128
6.2 Results of basic traversal approaches 131
6.3 Results of path fusion traversal approaches 132
6.4 Best operating point in terms of τ for each of the experiments . . 134

7.1 Results 148

8.1 Encoding schemes . . . 157
8.2 Overview of the data used for experiments . . . 157
8.3 Results (best one(s) for each dataset/extractor pair marked in bold) . . . 160
8.4 CPU instructions per template comparison . . . 161

10.1 Used datasets . . . 190
10.2 Data processing pipelines . . . 192
10.3 Configurations per experiment . . . 193
10.4 Baseline results (with 95% CI) . . . 195
10.5 Proposed system's results (with 95% CI) . . . 199
10.6 Summary of the results – best configuration for each of the tested fusion methods (with 95% CI) . . . 200

11.1 Pre-selection results . . . 217
11.2 Baseline results . . . 217
11.3 Two-stage system results . . . 218

12.1 Baseline and groundtruth results (in %). . . . 230
12.2 Algorithmic results (in %). . . . 232
12.3 Parameters of the impostor score distributions. . . . 233

13.1 Overview of the CASIA-Thousand dataset . . . 241
13.2 Impact of glasses on iris recognition . . . 243
13.3 Topology of the DNN-based approach . . . 245
13.4 Results of the evaluation (with 95% CI) . . . 249

15.1 Dataset overview . . . 268
15.2 Numbers of comparisons performed during experiments. ("Fusion" refers to all three fusion experiments, *i.e.* ARP, w-ARP and Min-or-ARP, since for each one of those the transactions numbers are identical) . . . 269
15.3 Results . . . 271
15.4 Distribution statistics . . . 271

A.1 Related works (results as reported by the authors, or if unavailable, extracted from the presented plots) . . . 292
A.2 Approximation of filling a Bloom filter resulting from a block of height H_B and width W_B with random data (lower values reflect higher data representation sparseness and fewer potential collisions) . . . 296
A.3 Evaluation dataset overview . . . 302
A.4 Dataset split (templates) for the experiments . . . 303
A.5 % of bits set to 1 at the top levels of the basic, single-tree system (level 0 is the tree root) . . . 306
A.6 The results of the 3 configurations with best performance in the single and multiple tree schemes . . . 307

A.7 The results of the Bloom filter scheme with selective tree traversal 308
A.8 A summary of the results for various system improvements . . . 308

List of Algorithms

10.1 Shortlist size estimation 186

Part I

Overview

Chapter 1

Introduction

Abstract

This chapter presents the motivation and general background for the research work conducted in this thesis. Furthermore, an outline of the thesis contents and organisation is provided.

1.1 Motivation

Biometric systems can be used as a replacement or supplement for the traditional knowledge (*e.g.* password) and token (*e.g.* RFID chip) based identity management systems. The current and future value of the biometrics market has been estimated in tens of billions of dollars by various market studies [1, 28, 36]. The number, scope, and scale of the personal, corporate, and governmental applications are quickly increasing. In recent years, biometric solutions have been applied extensively in various contexts and domains. Prominent examples include, but are not limited to:

- National citizen inventory.
- Identity documents and passports.
- Voter registration during elections.
- Automated border security and surveillance in general.
- Law enforcement forensics.
- Financial services.
- Personal and corporate access control systems.
- Signing of legal documents.

With the growing size (in terms of enrolled data subjects) of such systems (see *e.g.* [5, 9, 10, 39]), a need for research into computationally efficient biometric solutions has arisen. The research conducted in this thesis focuses on this topic; specifically, it concentrates on matters associated with biometric identification, information fusion, and data protection.

1.2 Background

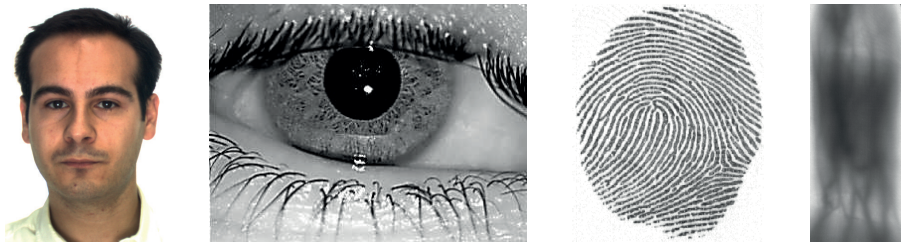
The following subsections provide an introduction and further reading references for the key research areas relevant in the context of this thesis.

1.2.1 Biometrics

Biometrics is a science which deals with the task of establishing or verifying the identity of individuals. The international standard ISO/IEC 2382-37 [16] defines biometrics as (quote):

> "automated recognition of individuals based on their biological and behavioural characteristics"

Certain characteristics which are (nearly) universally possessed by all humans (*e.g.* anatomical, such as iris) are highly distinctive and can be used to distinguish between different individuals with a very high degree of confidence. Figure 1.1 shows example images of several popular types of biometric characteristics. The four depicted characteristics were used in the research conducted in the scope of this thesis (see chapter 2 for more details).



(a) Face (b) Iris (c) Fingerprint (d) Vein

Figure 1.1: Examples of biometric characteristics (images taken from from publicly available research databases).

An automated biometric recognition system relies on algorithms which process biometric samples (often, but not necessarily, images) to extract distinguishing features, which are subsequently compared to establish the degree of similarity between two biometric samples. Regardless of the chosen type of biometric characteristic, the elements of such a system are generalisable into a modular framework. A conceptual overview of a generic biometric system is depicted in figure 1.2.

In the figure, the overall system is divided into five subsystems, while the arrows represent the data transmission paths between the subsystems. Below, the tasks handled by the subsystems are described briefly.

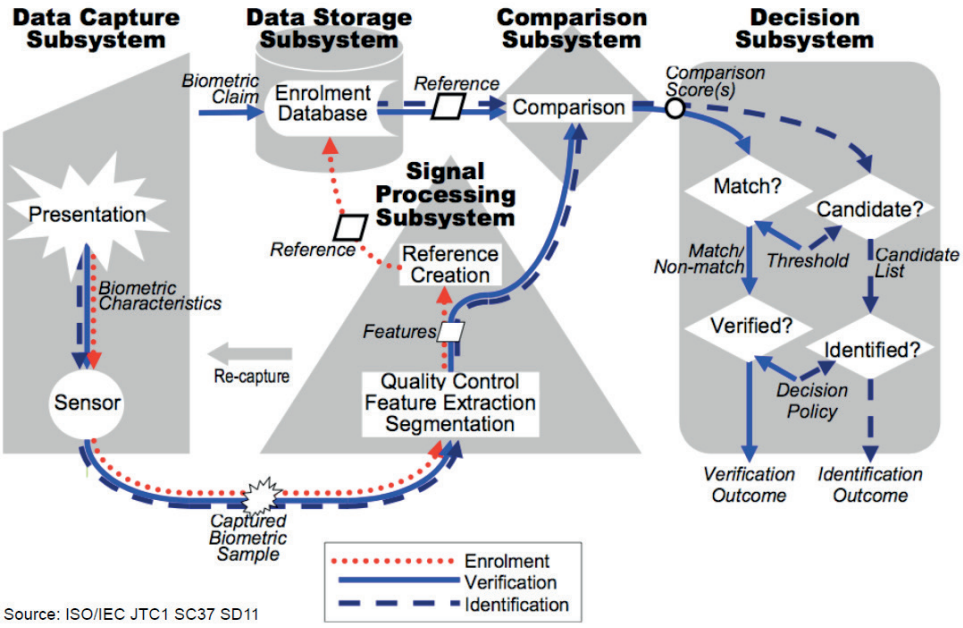


Figure 1.2: A conceptual overview of the components and information flow in a biometric system (from ISO/IEC 19795-1 [14]).

Data capture: Responsible for the acquisition of a biometric sample (*e.g.* a facial image) with a capture device containing one or multiple sensor(s) (*e.g.* a camera).

Signal processing: Responsible for processing the raw biometric sample. This includes *e.g.* steps such as: segmentation of the region of interest, extraction of distinguishing features, and quality control. The extracted features are used to create a so-called biometric template.

Data storage: This module is typically a database where the biometric templates and the personal details (*e.g.* user ID) associated therewith are stored.

Comparison: This module makes it possible to ascertain the similarity of two biometric templates by comparing them with each other. For example, templates with biometric features represented as vectors of floating point numbers could be compared using the Euclidean distance.

Decision: This module is used to reach a decision (subject verified or identified, see subsection 1.2.2) based on the comparison score(s), as well as the policies and thresholds set by the system operator.

The research in this thesis concerns especially the comparison subsystem (see chapter 2 for more details). For a much more comprehensive introduction to biometrics, the reader is referred to *e.g.* [26] and the handbook series [2, 18, 25, 27, 37].

1.2.2 Biometric Identification and Computational Workload

The main two of the possible operational modes of biometric systems are (definitions quoted directly from ISO/IEC 2382-37 [16]):

Biometric verification: Referring to the "process of confirming a biometric claim through biometric comparison".

Biometric identification: Referring to the "process of searching against a biometric enrolment database to find and return the biometric reference identifier(s) attributable to a single individual".

Within biometric identification, the closed-set and open-set scenarios can be distinguished. In a closed-set identification, it is assumed that all the potential system users are present in the enrolment database, whereas in an open-set identification, it is possible and tested for some users (impostors) not to be enrolled in the system. Arguably, the latter scenario is more realistic and challenging from the practical point of view; the research in this thesis focuses chiefly on the open-set identification systems.

Biometric identification systems need to ascertain the presence in the enrolment database and the identity of a data subject solely based on the information extracted from a biometric sample (*i.e.* without a biometric identity claim, as is the case in the biometric verification). Thus, in the worst case, reaching a decision requires an exhaustive database search (*i.e.* comparing the probe against all the references stored in the enrolment database). This naïve search approach encounters two non-trivial challenges:

Computational workload: With the enrolment database size increasing, the system response time becomes proportionally slower. From the operational point of view, this may end up requiring software optimisations and/or hardware investment in order to facilitate more data subjects.

False positives: The probability of making false positive errors is likewise increased with the growing size of the biometric enrolment database.

Those challenges necessitate research into methods of efficient biometric identification, which is the main topic of this thesis. Specifically, methods of computational workload reduction are of interest; more precisely concerning the following two key aspects:

Search space: Aiming to create algorithms and datastructures with sub-linear or logarithmic complexity in biometric identification transactions.

Individual template comparisons: Aiming to create compact biometric feature representations, whose similarity can be ascertained using computationally efficient biometric comparators.

The biometric data has certain properties (*i.e.* lack of inherent ordering, within-subject variability, and high dimensionality), due to which many traditional approaches (*e.g.* normal database indexing) become unsuitable or perform poorly [12]. Consequently, methods of computational workload reduction which are specifically tailored to the aforementioned properties need to be developed. Research in this area exists; however, this topic is by no means solved with many research avenues remaining relatively unexplored, especially nowadays with the rise of numerous large-scale deployments and the challenges associated therewith. Some of those areas are addressed by this thesis (see chapter 2 for details). A detailed overview of this research area is given in chapter 4, which contains a comprehensive survey [7] of the current state-of-the-art in this research area, a formulation of a taxonomy for categorising those approaches, as well as a discussion of the current trends and challenges. In addition to the scientific literature surveyed and systematised in the abovementioned chapter, two recent theses (by Li [24] and Schuch [35]) have been at least partially devoted to the topic of efficient biometric identification. Both of them focused exclusively on fingerprint-based systems. The existing works notwithstanding, many research avenues had been (and/or remain) open in this field; those include, but are by no means limited to: a general framework for the purposes of categorisation of the existing approaches, the creation of methods for other types of biometric characteristics, the incorporation of information fusion and data protection techniques, as well as the development of generic methods which can be applied irrespective of the chosen types of biometric characteristics and representations of their features. Several of those matters are addressed in this thesis (see chapter 2 where its scope is defined).

1.2.3 Multi-Biometrics

Due to the increasing operational and security demands, the focus of many biometric systems, especially large-scale ones, is shifting from single to multi-biometrics. Multi-biometric systems rely on information fusion, whereby information obtained from multiple sources is consolidated with the goal of improving the biometric performance, sample quality, or other quantifiable aspects w.r.t. a single information source system. Illustrating the concept with a finger-based biometric system, the different information sources could be:

- Sensors (*e.g.* capacitive and optical).
- Algorithms (*e.g.* fingerprint texture and minutiae-based).
- Samples (*e.g.* multiple acquisitions of the same fingerprint).
- Instances (*e.g.* multiple fingers).
- Types of characteristics (*e.g.* fingerprints and fingerveins).

The fusion of the acquired information can happen on multiple operational levels of a biometric system. Most generally speaking, two categories of information fusion can be distinguished in this context:

- Prior to the template comparisons, *i.e.* sensorial (see *e.g.* [17]) or feature (see *e.g.* [20]) level fusion.
- After the template comparisons, *i.e.* score (see *e.g.* [19]), rank (see *e.g.* [23]), or decision (see *e.g.* [30]) level fusion.

Information fusion is an active field of research within biometrics. An introductory overview of this topic is available in, for example [34], while [6] provides a comprehensive survey of the current state-of-the-art. Lastly, the ISO/IEC 24722 [15] is also of interest in this context. The large amount of the existing research notwithstanding, the matter of computational workload reduction coupled with (or by means of) a biometric information fusion has received relatively little attention and is one of the topics investigated in this thesis.

1.2.4 Biometric Data Protection

A number of data privacy and security concerns has arisen w.r.t. biometrics. If compromised or leaked, the biometric data can be misused in a variety of ways, including but not limited to:

- Identity theft.
- Tracking or profiling.
- Extraction of sensitive information (*e.g.* illness).

Therefore, a strong demand (*e.g.* by the general population, various non-governmental organisations and advocacy groups, as well as policymakers) for data protection exists. Recently, the General Data Protection Regulation (GDPR) [8] has been introduced by the European Union. Under this regulation, biometric data is categorised within "special categories" of personal data (formerly, "sensitive personal data") and hence allotted extensive legal

protections. Those societal and legal trends indicate the need of research into secure, privacy-preserving biometric systems (see *e.g.* [4] for more details). According to ISO/IEC 24745 [13], biometric systems must be designed and operated so that they satisfy a number of security and privacy related requirements. While the security ones are mostly related to the operational details and implementation of a biometric system (thus out of scope covered by this thesis), the following properties are of interest w.r.t. the biometric data itself:

Irreversibility: Recreating the original biometric template from the secured template must be computationally infeasible.

Unlinkability: Cross-correlating protected templates across different systems and databases must not be possible in order to avoid profiling (without consent) of the data subject.

Renewability: In case of being compromised, revoking and reissuing a new (different) protected biometric reference should be possible and straightforward.

Performance: The biometric performance of the protected system must not be (severely) impaired by the template protection scheme.

Collectively, methods which aim to satisfy those properties are referred to as "biometric template protection". There exists a body of work on this subject in the scientific literature, with the key categories of approaches being (see *e.g.* [3], [33], and [32] for more details):

Biometric cryptosystems: Originally aimed at securing or deriving cryptographic keys, such schemes can also be used for biometric template protection directly (see *e.g.* [38]). However, in most cases, the biometric comparators need to use error-correcting codes, thereby being computationally expensive, which constitutes a major limitation for large-scale biometric identification.

Cancelable biometrics: Methods relying on the application of a non-invertible transform (see *e.g.* [31]) or a salt (see *e.g.* [22]) to the biometric data. The aim is to create a protected template which maintains the fundamental statistical properties of the original data. In many cases, the protected template comparison can be performed using the same comparators as for the plain, unprotected templates. For the purpose of biometric identification, such methods will generally be superior (in terms of computational workload) to biometric cryptosystem schemes.

General purpose: Relying on methods not necessarily limited to biometrics, such as homomorphic encryption (see *e.g.* [11]).

Coupling and addressing the challenges of biometric template protection and computational workload reduction has not yet been sufficiently explored in the scientific literature and is therefore of interest for this thesis.

In addition to the technological challenges of privacy-preserving biometrics, many legal, societal and ethical issues are associated with this research area (see *e.g.* [21] and [29]). However, those (fascinating) non-technical matters are out of scope covered by this thesis.

1.3 Overview of Research Projects

The research for the articles included in this thesis was conducted in the context of two research projects. The projects and their respective funding agencies are briefly described in the following subsections.

1.3.1 BioIndex

This project was conducted in the context of CRISP, which is one of the national IT-security research centres in Germany. It was funded by the German Federal Ministry of Education and Research (BMBF) and the Hessen State Ministry for Higher Education, Research and the Arts (HMWK). The CRISP research centre focuses on close contacts to the industry stakeholders and conducts applied research into the application-oriented issues of cybersecurity and privacy.

The BioIndex project description states (quote[1]):

> "Nowadays, biometric recognition represents an integral component of identity management systems. The aim of the BioIndex subproject of CRISP is to accelerate biometric systems operating in identification mode without decreasing the recognition accuracy of the overall system. This represents a challenging issue since generic biometric recognition systems do not provide the scalability needed for large-scale applications. Within the BioIndex project diverse techniques will be investigated and developed in order to provide real-time identification on large-scale biometric databases."

1.3.2 BioBiDa

This project was sponsored by the Development of Scientific and Economic Excellence (LOEWE-3) initiative. It was a collaboration between academia (Hochschule Darmstadt), an industry partner (iCOGNIZE GmbH), and the German Federal Police (Bundeskriminalamt). The LOEWE-3 funds projects

[1] `https://dasec.h-da.de/projects/bioindex/`

which strengthen the cooperation between small and medium Hessian companies and universities, as well as non-university research institutions. The overarching goal of the projects is facilitating applied research with high relevance for end-users, industry, and public institutions.

The BioBiDa project description states (quote[2]):

> "With the recent rapid growth of biometric systems' sizes and popularity, technologies supporting efficient and accurate processing of large amounts of biometric data are sought for. The goal of this 2-year project is development of efficient algorithms and datastructures for biometric identification, which can perform search queries on large biometric datasets in real-time, while simultaneously facilitating biometric data protection. The project will focus on systems based on biometric characteristics from hands and faces. The application of the developed schemes will be twofold – a robust and quick search for use both in the identification scenario with cooperative subjects, as well as forensic investigations. Furthermore, by virtue of development of privacy enhancing concepts, the societal acceptance of biometric technologies is expected to be strengthened."

1.4 Thesis Organisation

The main contents of this thesis are presented as a collection of interrelated research articles. The thesis is divided into five parts:

- Part I consists of 3 overview chapters. A general topic introduction is given in chapter 1. In chapter 2, the scope of the thesis is defined along with the research questions and the experimental evaluation environment. Chapter 3 outlines the contributions of this thesis, specifically containing a list and summary of the research articles written within the scope of this thesis, as well as their relation to the research questions.
- Part II contains one of the larger articles written in the course of the doctoral studies. This article contains a comprehensive overview of the current state-of-the-art and related works. Furthermore, it discusses the pertinent challenges and issues from both the academic and industry perspective.
- Part III is the main technical body of this thesis and comprises all the other individual research articles, which collectively address the research questions.

[2] https://dasec.h-da.de/projects/biobida/

- Part IV concludes the thesis by summarising the findings and results, as well as by answering the research questions. Lastly, a discussion of the potential future research avenues is provided.
- Part V contains the appendices.

Note, that following the regulations, in this thesis the overview and research articles are reproduced verbatim (*i.e.* as written for the scientific conferences and journals). The only changes pertain to the layout and typesetting (*e.g.* migrating from double to single column format and using a consistent style for the references).

1.5 Bibliography

[1] BHUTANI, A., AND BHARDWAJ, P. Biometrics market size by application. Tech. Rep. GMI493, Global Market Insights, August 2017.

[2] BOWYER, K., AND BURGE, M. J. *Handbook of iris recognition*. Springer, 2016.

[3] BREEBAART, J., BUSCH, C., GRAVE, J., AND KINDT, E. A reference architecture for biometric template protection based on pseudo identities. In *BIOSIG: Biometrics and Electronic Signatures* (September 2008), Gesellschaft für Informatik e. V., pp. 25–37.

[4] CAMPISI, P. *Security and privacy in biometrics*, vol. 24. Springer, June 2013.

[5] CONSORTIUM FOR ELECTIONS AND POLITICAL PROCESS STRENGTHENING. Assessment of electoral preparations in the Democratic Republic of the Congo. Tech. rep., CEPPS, May 2018.

[6] DINCA, L. M., AND HANCKE, G. P. The fall of one, the rise of many: A survey on multi-biometric fusion methods. *IEEE Access 5* (April 2017), 6247–6289.

[7] DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Computational workload in biometric identification systems: An overview. *IET Biometrics 8*, 6 (November 2019), 351–368.

[8] EUROPEAN PARLIAMENT. Regulation (EU) 2016/679. *Official Journal of the European Union L119* (April 2016), 1–88.

[9] EUROPEAN UNION AGENCY FOR THE OPERATIONAL MANAGEMENT OF LARGE-SCALE IT SYSTEMS IN THE AREA OF FREEDOM, SECURITY AND JUSTICE. Eurodac storage capacity increased. `https://www.eulisa.europa.eu/Newsroom/News/Pages/Eurodac-`

storage-capacity-increased.aspx, April 2016. Last accessed: 2020–03–11.

[10] GEMALTO. DHS's automated biometric identification system IDENT - the heart of biometric visitor identification in the USA. https://www.gemalto.com/govt/customer-cases/ident-automated-biometric-identification-system, March 2019. Last accessed: 2020–03–11.

[11] GOMEZ-BARRERO, M., MAIORANA, E., GALBALLY, J., CAMPISI, P., AND FIERREZ, J. Multi-biometric template protection based on homomorphic encryption. *Pattern Recognition 67* (July 2017), 149–163.

[12] HAO, F., DAUGMAN, J., AND ZIELINSKI, P. A fast search algorithm for a large fuzzy database. *Transactions on Information Forensics and Security (TIFS) 3*, 2 (June 2008), 203–212.

[13] ISO/IEC JTC1 SC27 IT SECURITY TECHNIQUES. *ISO/IEC 24745:2011. Information technology – Security techniques – Biometric information protection*. International Organization for Standardization and International Electrotechnical Committee, June 2011.

[14] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 19795-1:2006. Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*. International Organization for Standardization and International Electrotechnical Committee, April 2006.

[15] ISO/IEC JTC1 SC37 BIOMETRICS. Iso/iec tr 24722:2015. information technology – biometrics – multimodal and other multibiometric fusion. Tech. rep., International Organization for Standardization, December 2015.

[16] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 2382-37:2017. Information technology – Vocabulary – Part 37: Biometrics*, 2 ed. International Organization for Standardization and International Electrotechnical Committee, February 2017.

[17] JAIN, A., AND ROSS, A. Fingerprint mosaicking. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (May 2002), vol. 4, IEEE, pp. IV–4064–IV–4067.

[18] JAIN, A. K., FLYNN, P., AND ROSS, A. *Handbook of biometrics*. Springer, 2007.

[19] JAIN, A. K., NANDAKUMAR, K., AND ROSS, A. Score normalization in multimodal biometric systems. *Pattern recognition 38*, 12 (December 2005), 2270–2285.

[20] KANHANGAD, V., KUMAR, A., AND ZHANG, D. Contactless and pose invariant biometric identification using hand surface. *Transactions on Image Processing (TIP) 20*, 5 (May 2011), 1415–1424.

[21] KINDT, E. J. *Privacy and data protection issues of biometric applications*, vol. 1. Springer, 2016.

[22] KONG, A., CHEUNG, K.-H., ZHANG, D., KAMEL, M., AND YOU, J. An analysis of BioHashing and its variants. *Pattern recognition 39*, 7 (July 2006), 1359–1368.

[23] KUMAR, A., AND SHEKHAR, S. Personal identification using multibiometrics rank-level fusion. *Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 41*, 5 (September 2011), 743–752.

[24] LI, Q. *Innovative methods for large-scale ngerprint identication systems - facilitating searching in a large-scale database*. Ph.D. thesis, Norwegian University of Science and Technology, October 2016.

[25] LI, S. Z., AND JAIN, A. K. *Handbook of face recognition*. Springer, 2004.

[26] LI, S. Z., AND JAIN, A. K. *Encyclopedia of biometrics*. Springer, 2015.

[27] MALTONI, D., MAIO, D., JAIN, A. K., AND PRABHAKAR, S. *Handbook of fingerprint recognition*. Springer, 2009.

[28] MARKETS AND MARKETS. Biometric system market by authentication type - global forecast to 2023. Tech. Rep. SE 3449, Markets and Markets, July 2018.

[29] MORDINI, E., AND TZOVARAS, D. *Second generation biometrics: The ethical, legal and social context*, vol. 11. Springer Science & Business Media, 2012.

[30] PRABHAKAR, S., AND JAIN, A. K. Decision-level fusion in fingerprint verification. *Pattern Recognition 35*, 4 (April 2002), 861–874.

[31] RATHGEB, C., BREITINGER, F., AND BUSCH, C. Alignment-free cancelable iris biometric templates based on adaptive Bloom filters. In *International Conference on Biometrics (ICB)* (June 2013), IEEE, pp. 1–8.

[32] RATHGEB, C., AND BUSCH, C. *Biometric template protection: State-of-the-art, issues and challenges*. Institution of Engineering and Technology, November 2017, ch. 8, pp. 173–191.

[33] RATHGEB, C., AND UHL, A. A survey on biometric cryptosystems and cancelable biometrics. *EURASIP Journal on Information Security 2011*, 1 (September 2011), 1–25.

[34] ROSS, A., NANDAKUMAR, K., AND JAIN, A. K. *Handbook of multibiometrics*. Springer, 2006.

[35] SCHUCH, P. *Deep Learning for Fingerprint Recognition Systems*. Ph.D. thesis, Norwegian University of Science and Technology, October 2019.

[36] THAKKAR, D. Global biometric market analysis: Trends and future prospects. https://www.bayometric.com/global-biometric-market-analysis/, August 2018. Last accessed: 2020–03–11.

[37] UHL, A., MARCEL, S., BUSCH, C., AND VELDHUIS, R. N. J. *Handbook of Vascular Biometrics*. Springer, 2020.

[38] ULUDAG, U., PANKANTI, S., PRABHAKAR, S., AND JAIN, A. K. Biometric cryptosystems: issues and challenges. *Proceedings of the IEEE 92*, 6 (June 2004), 948–960.

[39] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Aadhaar dashboard. https://www.uidai.gov.in/aadhaar_dashboard/. Last accessed: 2020–03–11.

Chapter 2

Thesis Scope

Abstract

This chapter presents the objectives of this thesis by defining and elaborating upon its main research questions. Furthermore, the environment for evaluating the developed methods is described.

2.1 Research Questions

Five research questions are addressed in this thesis. In the subsequent subsections, each of them is stated and elaborated upon. Following the main research body of the thesis, the research questions are discussed and answered in chapter 16 based on the obtained results.

2.1.1 Research Question 1

RQ1: Is it possible to vastly (*i.e.* by 90% or more) decrease the computational workload of a biometric identification system, while retaining high biometric performance of the naïve, exhaustive search approach?

The computational costs of a biometric identification transaction are typically dominated by the computations of the template comparisons (*i.e.* other computational costs, such as feature extraction, tend to be relatively trivial). In this context, two types of approaches are relevant: reducing the search space (*i.e.* the number of template comparisons necessary for a biometric identification transaction) and reducing the computational cost of the individual template comparisons. Furthermore, since machine learning has been increasingly successful in biometric recognition, computational workload in the context of features extracted by deep learning methods is also investigated.

This research question defines the overarching goal of this thesis – research into computational workload reduction in biometric identification systems. Two distinct research objectives are defined within the scope of this research question:

- Creation of algorithms and datastructures which facilitate search space reduction.
- Creation of algorithms and datastructures which facilitate more efficient individual template comparisons.

2.1.2 Research Question 2

RQ2: Is it possible to categorise different methods of efficient biometric identification across biometric modalities into a single, coherent taxonomy? Is it possible to create metrics suitable for a modality agnostic evaluation and reporting of computational workload reduction methods?

Although substantial amount of research has been published in the area of computational workload reduction, the field lacks a unified taxonomy of approaches, as well as evaluation metrics. Those matters, as well as other issues relevant for operational biometric identification systems are the focus of this research question.

In summary, three distinct research objectives are defined within the scope of this research question:

- Conducting a comprehensive study of biometric workload reduction literature across modalities and systematisation of the approaches into a generalised, biometric characteristic-agnostic taxonomy.
- Development of a framework and metrics for computational workload reduction and a proposal submission to the revision of the International Standard ISO/IEC 19795-1 Biometric performance testing - Part 1: Principles and framwork [12].
- Investigation of other factors relevant for practical large-scale biometric systems' deployments (*e.g.* usability and facilitating stress testing by creating large synthetic datasets of biometric data).

2.1.3 Research Question 3

RQ3: Is it possible to incorporate biometric information fusion for the benefit of the approaches aimed at computational workload reduction?

The underlying idea of multi-biometric systems is combining the biometric information obtained from multiple sources. This can be done at several different stages of the biometric systems' pipeline and using a multitude of methods. In any case, by doing so, higher discriminative power (and hence,

higher biometric performance) can be achieved. Although biometric information fusion has been applied extensively in the context of biometric recognition, coupling it with computational workload reduction in the biometric identification scenario has not yet been exhaustively explored. Thus, this research question couples the **RQ1** with the additional stipulation of utilising information fusion, either alongside other methods of computational workload reduction or for the explicit purpose of computational workload reduction.

In summary, three distinct research objectives are defined within the scope of this research question:

- Creation of methods which take advantage of biometric information fusion on various levels (*e.g.* signal, feature, score, decision) of the biometric processing pipeline.
- Creation of methods which take advantage of biometric information fusion utilising one (*e.g.* multi-sample and multi-instance) type of biometric characteristic.
- Creation of methods which take advantage of biometric information fusion utilising multiple types of biometric characteristics.

2.1.4 Research Question 4

RQ4: Is it possible to develop computational workload reduction methods which work prior to feature extraction or even irrespective of the chosen feature representation? Can those be applied within a system utilising biometric information fusion?

Most of the already published methods for computational workload reduction are tied to a certain feature representation (*i.e.* they somehow rely on its inherent properties). This research question couples the **RQ1** and **RQ3** with the additional stipulation that the developed method be independent of the chosen feature representation or happen prior to the feature extraction step.

In summary, two distinct research objectives are defined within the scope of this research question:

- Creation of methods which can be used irrespective of the feature representation.
- Creation of methods which can be used prior to the feature extraction step.

2.1.5 Research Question 5

RQ5: Is it possible to create a privacy-preserving (cancelable) biometric system, which maintains a high biometric performance and a low computational workload in the biometric identification scenario?

Data security and privacy is an important issue in the context of biometric recognition. While the research has yielded numerous promising biometric template protection methods, their use together with computational workload reduction methods has not yet been explored. This research question thus couples the **RQ1** with the additional stipulation of fulfilling the template protection properties demanded by ISO/IEC 24745 [11].

In summary, two distinct research objectives are defined within the scope of this research question:

- Creation of a hierarchical retrieval system, which exhibits template protection (cancelable) properties and reduces the computational workload associated with the biometric identification transactions.
- Investigation of general-purpose methods which can be incorporated into a biometric identification system to strengthen the security and privacy of its enrolled data subjects.

2.2 Evaluation Environment

This section contains the basic information regarding the experimental setups used in the research articles. In particular, the datasets (subsection 2.2.1) and processing pipelines (subsection 2.2.2), as well as the evaluation metrics (subsection 2.2.3) are described briefly. The individual research articles provide more details on their respective experimental setups.

2.2.1 Datasets

Initially, three main types of biometric characteristics were selected for the experimental work (face, fingerprint, and iris). The selected characteristics are well-established in the scientific community and widely used in the practical biometric systems around the world (*cf*. table 4.1). Later on, vascular data (specifically, fingervein) was also included due to its promising potential (*e.g*. good biometric performance and relative robustness against presentation attacks), as well as the growing interest in this characteristic both from the academic and industry side.

Table 2.1 shows the datasets used in the experiments. The key criteria for the dataset choices are the image quality (compliance with the quality

requirements set in ISO/IEC 29794 [13]) and a large (for a research dataset) size in terms of the number of data subjects. The thesis focuses on cooperative biometric recognition; therefore not considering poor quality, in-the-wild, or automatically scraped datasets. Hence, some of the images are excluded to meet those criteria; this especially the case for the facial datasets, which often deliberately contain *e.g.* images with accessories or imperfect lighting conditions. In some of the experiments, the datasets are merged to facilitate a larger or chimeric multi-modal experimental set-up.

Table 2.1: Used datasets

Characteristic	Dataset	Subjects	Instances	Samples
Face	FERET [28]	1,199	1,199	14,051
	FRGC [27]	569	569	40,084
	AR Face [21]	126	126	4,000
	FEI [33]	200	200	2,800
	BioSecure (subset) [24]	210	210	840
	CAS-PEAL [9]	1,040	1,040	30,863
	CASIA NIR-VIS [17]	725	725	17,580
Fingerprint	NIST SD 9 [23]	2,700	27,000	54,000
	MCYT [25]	330	3,330	39,600
Fingervein	UTFVP [34]	60	360	1,440
	IDIAP [36]	110	220	440
	PolyU [16]	156	312	3,132
	SCUT-FV [29]	100	600	3,600
	FV-USM [2]	123	492	5,904
	SDUMLA [38]	106	636	3,816
Iris	CASIA-V4-Interval [4]	249	395	2,639
	CASIA-V4-Thousand [4]	1,000	2,000	20,000
	IITDv1 [15]	224	448	1,120
	BioSecure [24]	210	420	1,680
	ND-Iris-Template-Aging [8]	322	644	22,156

2.2.2 Processing Pipelines

The capability to extract discriminative features from the biometric samples is a prerequisite for a successful biometric system. In the context of the image-based biometrics, various general purpose (see *e.g.* a survey [19]) and biometric characteristic specific (see *e.g.* the handbook series [3, 18, 20, 35]) feature extraction methods exist. Improving the existing and developing entirely new feature extraction methods is an active research area; however, it is out of scope for this thesis. Its focus lies elsewhere, namely developing methods of computational workload reduction which can be applied prior (*i.e.* on the samples) or after (*i.e.* on the feature vectors and/or their comparators) the signal processing steps. Additionally, one of the research questions stipulates development of computational workload reduction meth-

ods which work irrespective of the chosen type of biometric characteristics and feature representations.

Therefore, all the research articles contained in this thesis utilise existing frameworks for the purposes of biometric data pre-processing (*e.g.* region of interest segmentation, feature extraction, *etc.*). The used tools are listed below, while more detailed descriptions and images from the processing pipelines are included in the research articles themselves. All the used frameworks and pre-trained models are open-source and achieve state-of-the-art biometric performance rates in biometric systems' evaluations. To facilitate reproducible research, all of the used tools and frameworks are open-source.

Iris: OSIRIS [26] and USIT [30].

Fingerprint: FingerJetFX [6], FingerNet [32], and sourceAFIS [37].

Fingervein: PLUS OpenVein [14] and spectral minutiae [22].

Face: FaceNet [31], OpenFace [1], and ArcFace [5].

Table 2.2 summarises the information about the data processing pipelines used for the biometric recognition. More detailed information about the processing pipelines is given in the individual research articles contained in this thesis.

Table 2.2: Summary of the data processing pipelines

Characteristic	Features	Representation	Size	Comparison
Face	Embedding	1-D vector	512 floats	Euclidean distance
Fingerprint	Minutiae	Set of triplets	Variable	Minutiae pairing
Fingervein	Spectral minutiae	2-D matrix	256×128 floats	Correlation
Iris	Wavelet demodulation	2-D matrix	20×256 bits	Hamming distance

2.2.3 Visualisation and Metrics

Depending on the focus of the individual research article, one or multiple of the following aspects need to be considered and evaluated quantitatively:

- Biometric performance.
- Computational workload.
- Template protection.

Accordingly, the methods used for results' visualisation and reporting are briefly outlined in the subsections below. The individual research articles provide more detailed information in their respective experimental protocol sections.

2.2.3.1 Biometric Performance

Biometric performance evaluation methodology and some metrics are standardised trough ISO/IEC 19795-1 [12]. They are followed whenever possible in this work.

In the context of the biometric performance assessment, the most important visualisation tools used in the research articles are:

- Histogram of genuine and impostor comparison scores, which makes it possible to visually assess the ranges and overlap of their distributions.
- DET curve, which shows the trade-off between type I and type II error rates depending on the chosen decision threshold.

In the context of the biometric performance assessment, the most important metrics used in the research articles are:

- Equal-error rate, which is the point at which the type I and type II error rates are equal.
- Hit rate (for a pre-selection algorithm), which denotes the proportion of genuine attempts where the enrolment record corresponding to the probe is contained in the subset of templates pre-selected from the enrolment database.
- Descriptive statistics of the genuine and impostor scores, such as: mean and median, minimum and maximum, standard deviation, skewness, and excess kurtosis.
- Sensitivity/Decidability index, which measures the degree of separation between two distributions.

The above methods facilitate a quantitative benchmark of different systems (*e.g.* state-of-the-art baseline *vs.* a proposed method) or different configurations of the same system, thereby making it possible to ascertain their relative strengths and weaknesses in the context of the biometric recognition performance.

2.2.3.2 Computational Workload

As opposed to the aspects considered in the previous subsection, no standardised methodology for computational workload reduction in biometric systems exists at the time of this writing. ISO/IEC 19795-1:2006 [12] does define the penetration rate (see below); however, it is not sufficient for all the scenarios considered in the research articles.

In the context of computational workload assessment, the most important metrics used in the research articles are:

- Penetration rate (for a pre-selection algorithm), which denotes the remaining proportion of the enrolment database that has to be considered after the pre-selection step.
- Workload fraction, as defined in [7], a metric which in addition to penetration rate and enrolment database size also considers the cost of the individual template comparisons. The metric expresses the computational workload of a proposed method as a fraction of the computational workload of a baseline method.
- Operations counts, *i.e.* the number of intrinsic CPU instructions necessary for a certain computation.
- Execution time, measured on commodity hardware.

Lastly, combining the aspects of biometric performance and computational workload assessment, the Euclidean distance from the optimal operation point (*i.e.* no errors and almost no computational workload) is used where appropriate.

2.2.3.3 Template Protection

The ISO/IEC 24745 [11] defines objectives which need to be fulfilled by a biometric template protection system. However, the standard currently does not define specific metrics, hence the metrics currently used in the scientific literature are adopted.

In the context of template protection assessment, the most important metrics used in the research articles are:

- Unlinkability, which is measured using the methodology and metric proposed in [10], which provide an estimation of the degree of the global linkability of a system.
- Irreversibility, which refers to the probability of an attacker guessing an original biometric template given a protected template.
- Renewability, which is measured by computing the available key space.
- Performance preservation, for which the previously outlined methods (see subsection 2.2.3.1) are used.

2.3 Bibliography

[1] AMOS, B., LUDWICZUK, B., AND SATYANARAYANAN, M. OpenFace: A general-purpose face recognition library with mobile applications. Tech. Rep. CMU-CS-16-118, CMU School of Computer Science, 2016.

[2] ASAARI, M. S. M., SUANDI, S. A., AND ROSDI, B. A. Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics. *Expert Systems with Applications 41*, 7 (June 2014), 3367–3382.

[3] BOWYER, K., AND BURGE, M. J. *Handbook of iris recognition*. Springer, 2016.

[4] CHINESE ACADEMY OF SCIENCES' INSTITUTE OF AUTOMATION. CASIA iris image database. `http://biometrics.idealtest.org/`, December 2010. Last accessed: 2020–03–11.

[5] DENG, J., GUO, J., AND ZAFEIRIOU, S. ArcFace: Additive angular margin loss for deep face recognition. *Computing Research Repository (CoRR)* (January 2018), 1–11.

[6] DIGITALPERSONA, INC. FingerJetFX OSE – fingerprint feature extractor, open source edition. `https://github.com/FingerJetFXOSE/FingerJetFXOSE`, 2011. Last accessed: 2020–03–11.

[7] DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Bloom filter-based search structures for indexing and retrieving Iris-Codes. *IET Biometrics 7*, 3 (May 2018), 260–268.

[8] FENKER, S. P., AND BOWYER, K. W. Analysis of template aging in iris biometrics. In *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (June 2012), IEEE, pp. 45–51.

[9] GAO, W., CAO, B., SHAN, S., CHEN, X., ZHOU, D., ZHANG, X., AND ZHAO, D. The CAS-PEAL large-scale chinese face database and baseline evaluations. *Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 38*, 1 (January 2008), 149–161.

[10] GOMEZ-BARRERO, M., GALBALLY, J., RATHGEB, C., AND BUSCH, C. General framework to evaluate unlinkability in biometric template protection systems. *Transactions on Information Forensics and Security (TIFS) 13*, 6 (June 2018), 1406–1420.

[11] ISO/IEC JTC1 SC27 IT SECURITY TECHNIQUES. *ISO/IEC 24745:2011. Information technology – Security techniques – Biometric information protection*. International Organization for Standardization and International Electrotechnical Committee, June 2011.

[12] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 19795-1:2006. Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*. International Organization for Standardization and International Electrotechnical Committee, April 2006.

[13] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 29794-1:2016. Information technology – Biometric sample quality – Part 1: Framework*. International Organization for Standardization and International Electrotechnical Committee, September 2016.

[14] KAUBA, C. PLUS OpenVein Toolkit. `http://www.wavelab.at/sources/OpenVein-Toolkit/`. Last accessed: 2020–03–11.

[15] KUMAR, A., AND PASSI, A. Comparison and combination of iris matchers for reliable personal authentication. *Pattern Recognition 43*, 3 (March 2010), 1016–1026.

[16] KUMAR, A., AND ZHOU, Y. Human identification using finger images. *Transactions on Image Processing (TIP) 21*, 4 (April 2012), 2228–2244.

[17] LI, S., YI, D., LEI, Z., AND LIAO, S. The CASIA NIR-VIS 2.0 face database. In *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (June 2013), IEEE, pp. 348–353.

[18] LI, S. Z., AND JAIN, A. K. *Handbook of face recognition*. Springer, 2004.

[19] LIU, L., CHEN, J., FIEGUTH, P., ZHAO, G., CHELLAPPA, R., AND PIETIKÄINEN, M. From BoW to CNN: Two decades of texture representation for texture classification. *International Journal of Computer Vision 127*, 1 (January 2019), 74–109.

[20] MALTONI, D., MAIO, D., JAIN, A. K., AND PRABHAKAR, S. *Handbook of fingerprint recognition*. Springer, 2009.

[21] MARTÍNEZ, A. M., AND BENAVENTE, R. The AR face database. Tech. Rep. 24, CVC, June 1998.

[22] MOKROSS, B.-A. Efficient biometric identification in large-scale palm vein databases. MSc thesis, Hochschule Darmstadt, Germany, December 2017.

[23] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. Special Database 9. `https://www.nist.gov/srd/nist-special-database-9`, August 2010. Last accessed: 2020–03–11.

[24] ORTEGA-GARCIA, J., FIERREZ, J., ALONSO-FERNANDEZ, F., GALBALLY, J., FREIRE, M. R., ET AL. The multiscenario multienvironment BioSecure multimodal database (BMDB). *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 32*, 6 (June 2010), 1097–1111.

[25] ORTEGA-GARCIA, J., FIERREZ-AGUILAR, J., SIMON, D., GONZALEZ, J., FAUNDEZ-ZANUY, M., ET AL. MCYT baseline corpus: a bimodal biometric database. *IEE Proceedings – Vision, Image and Signal Processing 150*, 6 (December 2003), 395–401.

[26] OTHMAN, N., DORIZZI, B., AND GARCIA-SALICETTI, S. OSIRIS: An open source iris recognition software. *Pattern Recognition Letters 82*, 2 (September 2016), 124–131.

[27] PHILLIPS, P. J., FLYNN, P. J., SCRUGGS, T., BOWYER, K. W., CHANG, J., ET AL. Overview of the face recognition grand challenge. In *Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2005), vol. 1, IEEE, pp. 947–954.

[28] PHILLIPS, P. J., MOON, H., RIZVI, S. A., AND RAUSS, P. J. The FERET evaluation methodology for face-recognition algorithms. *Transactions on pattern analysis and machine intelligence (TPAMI) 22*, 10 (October 2000), 1090–1104.

[29] QIU, X., KANG, W., TIAN, S., JIA, W., AND HUANG, Z. Finger vein presentation attack detection using total variation decomposition. *Transactions on Information Forensics and Security (TIFS) 13*, 2 (February 2018), 465–477.

[30] RATHGEB, C., UHL, A., WILD, P., AND HOFBAUER, H. Design decisions for an iris recognition SDK. In *Handbook of Iris Recognition*, K. Bowyer and M. J. Burge, Eds., 2 ed., Advances in Computer Vision and Pattern Recognition. Springer, July 2016, pp. 359–396.

[31] SCHROFF, F., KALENICHENKO, D., AND PHILBIN, J. FaceNet: A unified embedding for face recognition and clustering. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2015), IEEE, pp. 815–823.

[32] TANG, Y., GAO, F., FENG, J., AND LIU, Y. FingerNet: An unified deep network for fingerprint minutiae extraction. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 108–116.

[33] THOMAZ, C. E. FEI face database. `https://fei.edu.br/~cet/facedatabase.html`. Last accessed: 2020–03–11.

[34] TON, B. T., AND VELDHUIS, R. N. J. A high quality finger vascular pattern dataset collected using a custom designed capturing device. In *International Conference on Biometrics (ICB)* (June 2013), IEEE, pp. 1–5.

[35] UHL, A., MARCEL, S., BUSCH, C., AND VELDHUIS, R. N. J. *Handbook of Vascular Biometrics*. Springer, 2020.

[36] VANONI, M., TOME, P., EL SHAFEY, L., AND MARCEL, S. Cross-database evaluation using an open finger vein sensor. In *Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS)* (October 2014), IEEE, pp. 30–35.

[37] VAŽAN, R. SourceAFIS – opensource fingerprint matcher. https://sourceafis.machinezoo.com/, 2019. Last accessed: 2020–03–11.

[38] YIN, Y., LIU, L., AND SUN, X. SDUMLA-HMT: A multimodal biometric database. In *Chinese Conference on Biometric Recognition (CCBR)* (December 2011), Springer, pp. 260–268.

Chapter 3

Contribution

Abstract

In this chapter, the research articles written during the course of the doctoral studies are listed with short descriptions of their respective backgrounds, goals, contributions, and relation to the research questions of this thesis.

3.1 Research Articles

This thesis consists of a compilation of interrelated research articles which collectively address the research questions and objectives (section 2.1).

Following research articles are included in the main body of this thesis:

- DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Computational workload in biometric identification systems: An overview. *IET Biometrics 8*, 6 (November 2019), 351–368.
- DROZDOWSKI, P., FISCHER, D., RATHGEB, C., SCHIEL, C., AND BUSCH, C. Database binning and retrieval in multi-fingerprint identification systems. In *International Workshop on Information Forensics and Security (WIFS)* (December 2018), IEEE, pp. 1–7.
- DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Multi-iris indexing and retrieval: Fusion strategies for Bloom filter-based search structures. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 46–53.
- DROZDOWSKI, P., GARG, S., RATHGEB, C., GOMEZ-BARRERO, M., CHANG, D., AND BUSCH, C. Privacy-preserving indexing of IrisCodes with cancelable Bloom filter-based search structures. In *European Signal Processing Conference (EUSIPCO)* (September 2018), IEEE, pp. 2360–2364.
- DROZDOWSKI, P., STRUCK, F., RATHGEB, C., AND BUSCH, C. Benchmarking binarisation schemes for deep face templates. In *International Conference on Image Processing (ICIP)* (October 2018), IEEE, pp. 191–195.

- DROZDOWSKI, P., BUCHMANN, N., RATHGEB, C., MARGRAF, M., AND BUSCH, C. On the application of homomorphic encryption to face identification. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2019), IEEE, pp. 173–180.
- DROZDOWSKI, P., RATHGEB, C., MOKROSS, B.-A., AND BUSCH, C. Multi-biometric identification with cascading database filtering. *Transactions on Biometrics, Behavior, and Identity Science (TBIOM)*, (March 2020), 1–14.
- DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Turning a vulnerability into an asset: Accelerating facial identification with morphing. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (May 2019), IEEE, pp. 2582–2586.
- DROZDOWSKI, P., RATHGEB, C., HOFBAUER, H., WAGNER, J., UHL, A., AND BUSCH, C. Towards pre-alignment of near-infrared iris images. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 359–366.
- DROZDOWSKI, P., STRUCK, F., RATHGEB, C., AND BUSCH, C. Detection of glasses in near-infrared ocular images. In *International Conference on Biometrics (ICB)* (February 2018), IEEE, pp. 202–208.
- DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. SIC-Gen: A synthetic Iris-Code generator. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2017), IEEE, pp. 61–69.
- DROZDOWSKI, P., WIEGAND, N., RATHGEB, C., AND BUSCH, C. Score fusion strategies in single-iris dual-probe recognition systems. In *International Conference on Biometric Engineering and Applications (ICBEA)* (May 2018), ACM, pp. 13–17.

Following research article is included in the appendix of this thesis (due to being partially based on work previous to the doctoral studies):

- DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Bloom filter-based search structures for indexing and retrieving Iris-Codes. *IET Biometrics 7*, 3 (May 2018), 260–268.

Following additional research articles were co-authored during the course of the doctoral studies, but are not included in this thesis (2nd authorship):

- OSORIO-ROIG, D., DROZDOWSKI, P., RATHGEB, C., MORALES-GONZÁLEZ, A., GAREA-LLANO, E., AND BUSCH, C. Iris recognition in visible wavelength: Impact and automated detection of glasses. In *International Conference on Signal-Image Technology Internet-Based Systems (SITIS)* (November 2018), IEEE, pp. 542–546.

- MOKROSS, B.-A., DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. *Efficient Identification in Large-Scale Vein Recognition Systems Using Spectral Minutiae Representations*. Springer, 2020, ch. 9.

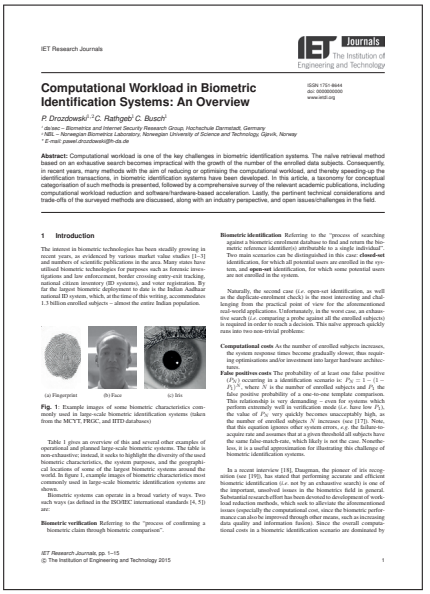
The research articles are listed and briefly summarised in the subsections below, while table 3.1 shows the relations between them and the research questions.

Table 3.1: Relations between the research articles and the research questions

Chapter	Reference	RQ1	RQ2	RQ3	RQ4	RQ5
4	[8]	●	●			
5	[3]	●		●	●	
6	[5]	●		●		
7	[4]	●				●
8	[12]	●				
9	[2]	●				●
10	[11]	●		●	●	
11	[9]	●		●	●	
12	[10]	●			●	
13	[13]		●			
14	[6]		●			
15	[14]			●		
Appendix A	[7]	●	●			
—	[19]		●			
—	[18]	●				

3.1.1 Main

The main research articles included in this thesis are listed below.



IET Research Journals

Journals
The Institution of Engineering and Technology

ISSN 1751-8644
doi: 0000000000
www.ietdl.org

Computational Workload in Biometric Identification Systems: An Overview

P. Drozdowski[1,2] C. Rathgeb[1] C. Busch[1]

[1] da/sec – Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
[2] NBL – Norwegian Biometrics Laboratory, Norwegian University of Science and Technology, Gjøvik, Norway
* E-mail: pawel.drozdowski@h-da.de

Abstract: Computational workload is one of the key challenges in biometric identification systems. The naive retrieval method based on an exhaustive search becomes impractical with the growth of the number of the enrolled data subjects. Consequently, in recent years, many methods with the aim of reducing or optimising the computational workload, and thereby speeding-up the identification transactions, in biometric identification systems have been developed. In this article, a taxonomy for conceptual categorisation of such methods is presented, followed by a comprehensive survey of the relevant academic publications, including computational workload reduction and software/hardware-based acceleration. Lastly, the pertinent technical considerations and trade-offs of the surveyed methods are discussed, along with an industry perspective, and open issues/challenges in the field.

1 Introduction

The interest in biometric technologies has been steadily growing in recent years, as evidenced by various market value studies [1–3] and numbers of scientific publications in the area. Many states have utilised biometric technologies for purposes such as forensic investigations and law enforcement, border crossing entry-exit tracking, national citizen inventory (ID systems), and voter registration. By far the largest biometric deployment to date is the Indian Aadhaar national ID system, which, at the time of this writing, accommodates 1.3 billion enrolled subjects – almost the entire Indian population.

(a) Fingerprint (b) Face (c) Iris

Fig. 1: Example images of some biometric characteristics commonly used in large-scale biometric identification systems (taken from the MCYT, FRGC, and ITD databases)

Table 1 gives an overview of this and several other examples of operational and planned large-scale biometric systems. The table is non-exhaustive; instead, it seeks to highlight the diversity of the used biometric characteristics, the system purposes, and the geographical locations of some of the largest biometric systems around the world. In figure 1, example images of biometric characteristics most commonly used in large-scale biometric identification systems are shown.

Biometric systems can operate in a broad variety of ways. Two such ways (as defined in the ISO/IEC international standards [4, 5]) are:

Biometric verification Referring to the "process of confirming a biometric claim through biometric comparison",

Biometric identification Referring to the "process of searching against a biometric enrolment database to find and return the biometric reference identifier(s) attributable to a single individual".

Two main scenarios can be distinguished in this case: **closed-set** identification, for which all potential users are enrolled in the system, and **open-set** identification, for which some potential users are not enrolled in the system.

Naturally, the second case (*i.e.* open-set identification, as well as the duplicate-enrolment check) is the most interesting and challenging from the practical point of view for the aforementioned real-world applications. Unfortunately, in the worst case, an exhaustive search (*i.e.* comparing a probe against all the enrolled subjects) is required in order to reach a decision. This naïve approach quickly runs into two non-trivial problems:

Computational costs As the number of enrolled subjects increases, the system response times become gradually slower, thus requiring optimisations and/or investment into larger hardware architectures.

False positives costs The probability of at least one false positive (P_N) occurring in a identification scenario is: $P_N \geq 1-(1-P_1)^N$, where N is the number of enrolled subjects and P_1 the false positive probability of a one-to-one template comparison. This relationship is very demanding – even for systems which perform extremely well in verification mode (*i.e.* have low P_1), the value of P_N very quickly becomes unacceptably high, as the number of enrolled subjects N increases (see [17]). Note, that this equation ignores other system errors, *e.g.* the failure-to-acquire rate and assumes that at a given threshold all subjects have the same false-match-rate, which likely is not the case. Nonetheless, it is a useful approximation for illustrating this challenge of biometric identification systems.

In a recent interview [18], Daugman, the pioneer of iris recognition (see [19]), has stated that performing accurate and efficient biometric identification (*i.e.* not by an exhaustive search) is one of the important, unsolved issues in the biometrics field in general. Substantial research effort has been devoted to development of workload reduction methods, which seek to alleviate the aforementioned issues (especially the computational cost, since the biometric performance can also be improved through other means, such as increasing data quality and information fusion). Since the overall computational costs in a biometric identification scenario are dominated by

Publication reference: DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Computational workload in biometric identification systems: An overview. *IET Biometrics 8*, 6 (November 2019), 351–368.
DOI: `https://doi.org/10.1049/iet-bmt.2019.0076`
Thesis chapter: 4
Addressed research question(s): RQ1, RQ2
Background: With the raise of the popularity and scale of biometric (identification) systems worldwide, substantial amount of research has been conducted in the field of computational workload reduction. However, no survey of this research area has been conducted previously. Likewise, a systematic categorisation of the existing approaches was lacking.
Contribution: The main contribution of this research article is a comprehensive overview of the research area this thesis focuses on. Furthermore, a taxonomy is proposed, which allows to conceptually categorise the methods for efficient biometric identification irrespective of the biometric characteristic. Furthermore, a broad survey of the existing methods is conducted and put in the context of the proposed taxonomy. Finally, the article discusses the relevant technical and practical considerations, as well as future research perspectives from both the industry and academic points of view.

Database Binning and Retrieval in Multi-Fingerprint Identification Systems

P. Drozdowski*†, D. Fischer*, C. Rathgeb*, C. Schiel‡, and C. Busch*
* da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
† Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway
‡ Bundeskriminalamt (BKA), Wiesbaden, Germany
{pawel.drozdowski,daniel.fischer,christian.rathgeb,christoph.busch}@h-da.de
Christopher.Schiel@bka.bund.de

Abstract

The increasingly large scale of deployed biometric systems necessitates approaches for computational workload reduction in order to perform identification queries efficiently. Simple database binning based on classification of features in biometric samples is amongst the most frequently used and researched methods for achieving said goal. However, multi-instance database binning appears to be a neglected topic in the scientific literature: best to the authors' knowledge, for fingerprints there exists only one, entirely theoretical, study on this subject. In this paper, we propose a retrieval algorithm based on multi-instance binning of fingerprint databases, along with usage of statistical information on fingerprint classes and their correlations.

The aforementioned statistics are obtained from NIST SD9 database and data obtained from the German Federal Criminal Police Office. Subsequently, the experimental evaluation of the proposed algorithm is performed on the NIST SD9 database. The proposed system is evaluated using a classifier based on the PCASYS tool and neuronal networks. The results show a significant workload reduction from a baseline exhaustive search scenario - down to 12.7% for this particular classifier and 5.8% for a theoretical perfect (completely accurate) classifier. The proposed method could be seamlessly integrated into operational systems, as it relies on well-established features and compatibility with the current acquisition methods.

1. Introduction

Nowadays, biometric technologies are already deployed in numerous nation-wide large-scale applications, such as the Indian Aadhaar project [21]. With the rapid growth of biometric systems' sizes and popularity, technologies supporting efficient and accurate processing of large amounts of biometric data are vital in order to guarantee practical response times. Conventional biometric systems require exhaustive one-to-many comparisons in order to identify biometric probes, *i.e.* comparison time frequently dominates the overall computational workload of an identification attempt. In past years, researchers have invested significant efforts to tackle the challenge of computational workload reduction in biometric identification systems. Basically, four different key concepts can be distinguished: *classification* or "binning", *indexing*, a *serial combination* of a computationally efficient and an accurate (but more complex) algorithm and *hardware-based acceleration*. Depending on the used biometric characteristic, the vast majority of classification approaches are designed to reliably extract human understandable attributes from a biometric sample, *e.g.* sex or ethnicity for face. While not necessarily unique to an individual, those attributes allow for a binning of biometric databases according to a predefined number of classes, *i.e.* the search space ($\in$ computational workload) for a given biometric probe can be reduced to one (or a few) bin(s). In contrast, biometric indexing approaches introduce hierarchical search structures (tolerating a certain amount of biometric variance), where the process of search space reduction might not be reproducible by human experts. Lastly, the latter two categories do not aim at reducing the complexity of an identification attempt but response times.

Focusing on fingerprint recognition systems, the classification model of Henry [12] has been widely used by researchers, as well as commercial vendors, for computational workload reduction in identification scenarios. The five fingerprint classes (or types), *i.e. arch*, *tented arch*, *right loop*, *left loop* and *whorl*, which are depicted in figure 1, are unevenly distributed in the population. Fingerprint classes are mainly determined based on the global (level-1) features, in particular ridge line flow (orientation map) and the singular points, *i.e.* core and delta, derived from it. Numerous approaches, which either directly employ or further process those features, have been proposed for the purpose of distinguishing between said classes. For more

1

Publication reference: DROZDOWSKI, P., FISCHER, D., RATHGEB, C., SCHIEL, C., AND BUSCH, C. Database binning and retrieval in multi-fingerprint identification systems. In *International Workshop on Information Forensics and Security (WIFS)* (December 2018), IEEE, pp. 1–7.
DOI: https://doi.org/10.1109/WIFS.2018.8630763
Thesis chapter: 5
Addressed research question(s): RQ1, RQ3, RQ4
Background: Biometric enrolment database binning is a relatively simple method of computational workload reduction. It relies on classifying (by somewhat discriminative features, *e.g.* fingerprint type) the biometric samples into a discrete number of categories (bins). During a biometric identification transaction, the search is conducted only within the bin(s) most likely corresponding to that of the probe sample. The existing binning approaches tend to concentrate on single instance data, whereas the use of multiple data instances has not yet been explored in detail.
Contribution: The contributions of this research article are twofold: firstly, the statistics of fingerprint types and their correlations across multiple fingers are computed and summarised. Secondly, a multi-instance binning method is presented and evaluated, with the results indicating a substantial reduction in the computational workload. The proposed method relies on an auxiliary feature and can therefore be used prior to the actual feature extraction step. In addition to using publicly available data, the experiments in this research article were validated using data provided by the German Federal Police (Bundeskriminalamt).

Multi-Iris Indexing and Retrieval:
Fusion Strategies for Bloom Filter-based Search Structures

P. Drozdowski*†, C. Rathgeb* and C. Busch*

* da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
† Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway

{pawel.drozdowski,christian.rathgeb,christoph.busch}@h-da.de

Abstract

We present a multi-iris indexing system for efficient and accurate large-scale identification. The system is based on Bloom filters and binary search trees. We describe and empirically evaluate several possible information fusion strategies for the system. Those experiments are performed using a combination of several publicly available datasets; the proposed system is tested in an open-set identification scenario consisting of 6,000 genuine and 100,000 impostor transactions. The system maintains the near-optimal biometric performance of an iris-code, score fusion based baseline system, while reducing the required lookup workload to less than 1% thereof.

1. Introduction

The increased popularity of biometrics worldwide is reflected in the appearance of several large-scale deployments. Of these, by far the largest is the Indian National ID project - at the time of this writing, more than 1 billion subjects have been enrolled [22] with biometric data from iris, face and fingerprints.

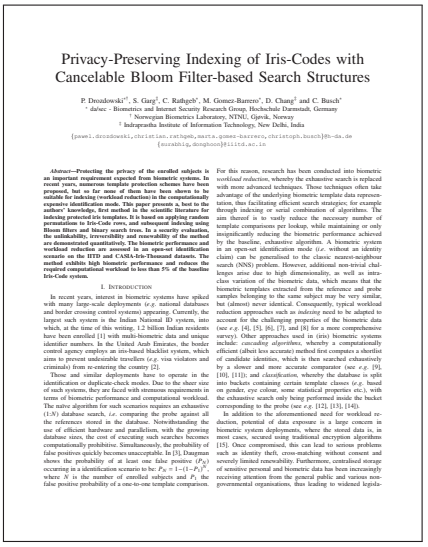
In (open-set) identification and de-duplication scenarios, one of the key challenges is the system accuracy, especially in terms of false positive occurrences. In a naïve, brute-force approach 1:N comparisons per retrieval of the enrolled reference are performed, i.e. the probe template is compared against every template in the biometric reference database. Hence, for large databases, the possibility of false positive occurrences quickly becomes unacceptable (see [1]). Fusing information from multiple sources can be used to increase the discriminative power of a biometric system [11]. In this paper, we utilise information from multiple instances of the same biometric characteristic - images of the left and right irides. Since the operational systems often already capture images of both irides (e.g. the aforementioned deployment in India), the proposed approach would not incur additional hardware costs or acquisition time during enrolment and could be easily integrated into existing systems.

The main contribution of this paper is a, best to the author's knowledge, first system for multi-iris indexing in the biometric literature and a large-scale, open-set identification evaluation of several information fusion strategies for said system. The key goal was to explore different strategies for multi-instance Bloom filter-based indexing and conduct a benchmark in terms of biometric performance and workload reduction. The paper is organised as follows: section 2 outlines related work; the basics of Bloom filter-based indexing are described in section 3, while section 4 shows how it can be extended to support multi-iris templates and which information fusion strategies can be applied. Section 5 describes the experimental setup; the results, discussion thereof and conclusions are presented in sections 6 and 7.

2. Related Work

The task solved by a biometric system in an open-set identification mode (i.e. where no identity claim is made) can be generalised to the classic nearest-neighbour search (NNS) problem. Additional non-trivial challenges are caused by high dimensionality, as well as fuzziness of the biometric data, meaning that the reference and probe sample from the same subject may be very similar, but never identical. In large systems, it is desirable to avoid the necessity of a naïve, brute-force lookup for every search query, as such retrieval method is computationally expensive and prone to false positive occurrences. Database indexing is a commonly used method for achieving this goal. In such systems, an index data-structure (e.g. a tree or hash table), which allows to quickly locate the approximate location of the data, is created and maintained. In other words, indexing systems utilise additional storage space in order to decrease response times. For biometric data, the indexing schemes must take into account the aforementioned issues

Publication reference: DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Multi-iris indexing and retrieval: Fusion strategies for Bloom filter-based search structures. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 46–53.
DOI: https://doi.org/10.1109/BTAS.2017.8272681
Thesis chapter: 6
Addressed research question(s): RQ1, RQ3
Background: Information fusion can be used to increase the discriminative power (and hence the biometric performance) of biometric recognition systems. However, those benefits are often counterbalanced by other costs, *e.g.* sensing and usability overhead. Coupling information fusion and computational workload reduction while simultaneously avoiding the above costs has not yet received enough attention in the scientific literature.
Contribution: The main contribution of this article is a method for indexing of multi-instance iris data. This work builds on the previous work in [7], by incorporating several information fusion strategies, whereby the biometric templates from the left and right iris can be consolidated. Additionally, several heuristics for traversing the constructed hierarchical data-structures are presented. The proposed method could be incorporated without additional sensing and usability overhead, as many iris recognition systems already capture both irides during the acquisition step. On the other hand, the proposed method offers substantial benefits in terms of computational workload and biometric performance.



Privacy-Preserving Indexing of Iris-Codes with Cancelable Bloom Filter-based Search Structures

P. Drozdowski[*†], S. Garg[‡], C. Rathgeb[*], M. Gomez-Barrero[*], D. Chang[‡] and C. Busch[*]
[*] da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
[†] Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway
[‡] Indraprastha Institute of Information Technology, New Delhi, India
{pawel.drozdowski,christian.rathgeb,marta.gomez-barrero,christoph.busch}@h-da.de
{saraahlg,dongkoo}@iiitd.ac.in

***Abstract*—Protecting the privacy of the enrolled subjects is an important requirement expected from biometric systems. In recent years, numerous template protection schemes have been proposed, but so far none of them have been shown to be suitable for indexing (workload reduction) in the computationally expensive identification mode. This paper presents a, best to the authors' knowledge, first method in the scientific literature for indexing protected iris templates. It is based on applying random permutations to Iris-Code rows, and subsequent indexing using Bloom filters and binary search trees. In a security evaluation, the unlinkability, irreversibility and renewability of the method are demonstrated quantitatively. The biometric performance and workload reduction are assessed in an open-set identification scenario on the IITD and CASIA-Iris-Thousand datasets. The method exhibits high biometric performance and reduces the required computational workload to less than 5% of the baseline Iris-Code system.**

I. INTRODUCTION

In recent years, interest in biometric systems have spiked with many large-scale deployments (*e.g.* national databases and border crossing control systems) appearing. Currently, the largest such system is the Indian National ID system, into which, at the time of this writing, 1.2 billion Indian residents have been enrolled [1] with multi-biometric data and unique identifier numbers. In the United Arab Emirates, the border control agency employs an iris-based blacklist system, which aims to prevent undesirable travellers (*e.g.* visa violators and criminals) from re-entering the country [2].

Those and similar deployments have to operate in the identification or duplicate-check modes. Due to the sheer size of such systems, they are faced with strenuous requirements in terms of biometric performance and computational workload. The naive algorithm for such scenarios requires an exhaustive (1:*N*) database search, *i.e.* comparing the probe against all the references stored in the database. Notwithstanding the use of efficient hardware and parallelism, with the growing database sizes, the cost of executing such searches becomes computationally prohibitive. Simultaneously, the probability of false positives quickly becomes unacceptable. In [3], Daugman shows the probability of at least one false positive (P_N) occurring in a identification scenario to be: $P_N = 1-(1-P_1)^N$, where N is the number of enrolled subjects and P_1 the false positive probability of a one-to-one template comparison. For this reason, research has been conducted into biometric *workload reduction*, whereby the exhaustive search is replaced with more advanced techniques. Those techniques often take advantage of the underlying biometric template data representation, thus facilitating efficient search strategies; for example through indexing or serial combination of algorithms. The aim thereof is to vastly reduce the necessary number of template comparisons per lookup, while maintaining or only insignificantly reducing the biometric performance achieved by the baseline, exhaustive algorithm. A biometric system in an open-set identification mode (*i.e.* without an identity claim) can be generalised to the classic nearest-neighbour search (NNS) problem. However, additional non-trivial challenges arise due to high dimensionality, as well as intra-class variation of the biometric data, which means that the biometric templates extracted from the reference and probe samples belonging to the same subject may be very similar, but (almost) never identical. Consequently, typical workload reduction approaches such as *indexing* need to be adapted to account for the challenging properties of the biometric data (see *e.g.* [4], [5], [6], [7], and [8] for a more comprehensive survey). Other approaches used in (iris) biometric systems include: *cascading algorithms*, whereby a computationally efficient (albeit less accurate) method first computes a shortlist of candidate identities, which is then searched exhaustively by a slower and more accurate comparator (see *e.g.* [9], [10], [11]); and *classification*, whereby the database is split into buckets containing certain template classes (*e.g.* based on gender, eye colour, some statistical properties etc.), with the exhaustive search only being performed inside the bucket corresponding to the probe (see *e.g.* [12], [13], [14]).

In addition to the aforementioned need for workload reduction, potential of data exposure in a large concern in biometric system deployments, where the stored data is, in most cases, secured using traditional encryption algorithms [15]. Once compromised, this can lead to serious problems such as identity theft, cross-matching without consent and severely limited renewability. Furthermore, centralised storage of sensitive personal and biometric data has been increasingly receiving attention from the general public and various non-governmental organisations, thus leading to widened legisla-

Publication reference: DROZDOWSKI, P., GARG, S., RATHGEB, C., GOMEZ-BARRERO, M., CHANG, D., AND BUSCH, C. Privacy-preserving indexing of Iris-Codes with cancelable Bloom filter-based search structures. In *European Signal Processing Conference (EUSIPCO)* (September 2018), IEEE, pp. 2360–2364.
DOI: https://doi.org/10.23919/EUSIPCO.2018.8553053
Thesis chapter: 7
Addressed research question(s): RQ1, RQ5
Background: Due to public scrutiny and new legislation (*e.g.* GDPR in the European Union [15]), biometric systems (among others) are faced with increasingly stringent data privacy and security requirements. The ISO/IEC 24745 [16] stipulates several properties (unlinkability, irreversibility, renewability, biometric performance preservation) to be fulfilled by the biometric template protection schemes.
Contribution: The contribution of this research article builds on the previous research articles included in this thesis ([7] and [5]). In particular, it extends the Bloom filter-based hierarchical indexing algorithm with data privacy and security properties. A row-wise permutation of the iris feature vectors is applied prior to the indexing step. The resulting system exhibits strong template protection properties, while simultaneously significantly reducing the computational workload of the biometric identification transactions.

BENCHMARKING BINARISATION SCHEMES FOR DEEP FACE TEMPLATES

P. Drozdowski*†, F. Struck*, C. Rathgeb* and C. Busch*

* da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
† Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway

{pawel.drozdowski,christian.rathgeb,christoph.busch}@h-da.de
florian.struck@stud.h-da.de

ABSTRACT

Feature vectors extracted from biometric characteristics are often represented using floating point values. It is, however, more appealing to store and compare feature vectors in a binary representation, since it generally requires less storage and facilitates efficient comparators which utilise intrinsic bit operations. Furthermore, the binary representations are very often necessary for some specific application scenarios, e.g. template protection and indexing.

In recent years, usage of deep neural networks for facial recognition has vastly improved the biometric performance of said systems. In this paper, various binarisation schemes are applied to such feature vectors and benchmarked for biometric performance. It is shown that with only a negligible drop in biometric performance, the storage space and computational requirements can be vastly decreased.

Index Terms— Biometrics, face recognition, binarisation, deep face templates

1. INTRODUCTION

Face is one of the most widely used biometric characteristics. Various methods have been proposed over the span of last three decades [1, 2]. In recent years, methods based on deep learning (e.g. [3, 4, 5, 6]) have been proposed, and significantly improved on the biometric performance of the heretofore existing methods. With this improved biometric performance, face has become an attractive characteristic for large-scale identification systems.

The deep face feature representations typically involve float-valued vectors, for which the template comparison is performed using metrics such as Euclidean distance (L^2 norm) or Chi-square distance (χ^2). Those metrics are computationally expensive – thus creating a potential efficiency bottleneck for large-scale biometric identification systems, where 1:N template comparisons are performed during lookup. Additionally, transmission of such feature vectors from low-cost mobile devices to central systems requires a compact encoding, specifically when the bandwith of mobile networks is limited. *Binarisation* of feature vectors offers an attractive alternative – such templates can be stored efficiently and be compared quickly in the Hamming domain utilising intrinsic CPU operations (i.e. xor and popcount) [7].

Over time, many methods of binarising data have been proposed, mostly with template protection as motivation (to transform the features to certain input forms required by the different cryptographic primitives) [8] and shown to work with, among others, classical facial recognition systems. However, it is unknown, whether or not those approaches are suitable for the vectors produced by deep learning based face feature extractors and their potential biometric performance degradation due to information loss has not been studied thoroughly. With this uncertainty as the motivation, the main contribution of this paper is such a benchmark, where various binarisation methods are evaluated in terms of biometric performance and computational workload incurred at the comparison stage.

The remainder of this paper is organised as follows: Section 2 introduces the related work. In section 3, binarisation schemes for deep facial templates are described. In section 4, the experimental setup and results are presented, while section 5 contains a summary of the paper and future work items.

2. RELATED WORK

In the recent decade, several data binarisation approaches have been proposed. Kevenaar *et al.* [9] extract the most reliable components of facial feature vectors and binarise them for use in a template protection scheme. Chen *et al.* [10] present a detection rate optimized bit allocation (DROBA) principle, which is biometric characteristic-agnostic. Based on the discriminative power of the features, it assigns more or fewer bits to them during binarisation, thus improving the biometric performance of the binarised feature representation. Bringer *et al.* [11] transform fingerprint minutiae set using a vicinity-based approach, which in addition to producing a compact feature representation, also exhibits self-alignment property. When presenting a novel fingerprint minutiae representation scheme, Cappelli *et al.* [12] note that it can also

Publication reference: DROZDOWSKI, P., STRUCK, F., RATHGEB, C., AND BUSCH, C. Benchmarking binarisation schemes for deep face templates. In *International Conference on Image Processing (ICIP)* (October 2018), IEEE, pp. 191–195.
DOI: https://doi.org/10.1109/ICIP.2018.8451291
Thesis chapter: 8
Addressed research question(s): RQ1
Background: Facial recognition systems based on neural networks have achieved breakthrough biometric performances in recent years. The feature vectors typically extracted by those systems are represented using floating point values and compared using metrics such as Euclidean or χ^2 distance. From the computational efficiency point of view, however, it would be more appealing to store and compare the feature vectors using a binary representation, as it would generally require less storage space and facilitate efficient comparators utilising intrinsic bit operations.
Contribution: The main contribution of this research article is a benchmark of biometric performance and computational workload achieved by two neural network-based facial recognition methods. The benchmark is conducted prior to and after the application of various existing floating point data quantisation and encoding methods. It is shown that there exists a trade-off, whereby significant computational efficiency gains can be achieved by sacrificing relatively little biometric performance.

On the Application of Homomorphic Encryption to Face Identification

P. Drozdowski*†‡, N. Buchmann‡, C. Rathgeb*, M. Margraf‡, and C. Busch*
* da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
† Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway
‡ Freie Universität Berlin, Germany
{pawel.drozdowski,christian.rathgeb,christoph.busch}@h-da.de
{nicolai.buchmann,marlin.margraf}@fu-berlin.de

***Abstract*—The data security and privacy of enrolled subjects is a critical requirement expected from biometric systems. This paper addresses said topic in facial biometric identification. In order to fulfil the properties of unlinkability, irreversibility, and renewability of the templates required for biometric template protection schemes, homomorphic encryption is leveraged. In addition to ensuring template confidentiality, subjects, the use of homomorphic encryption ensures that the biometric performance remains completely unaffected by the template protection scheme.**

The main contributions of this paper are: It proposes an architecture of a system capable of performing biometric identification in the encrypted domain, as well as provides and evaluates an implementation using two existing homomorphic encryption schemes. Furthermore, it discusses the pertinent technical considerations and challenges in this context.

***Index Terms*—Biometric Identification, Template Protection, Face, Homomorphic Encryption**

I. Introduction

Data exposure is a potential risk in biometric system deployments, which typically store their data secured using traditional encryption algorithms [1]. If this protection were to be compromised, serious problems would arise, including but not limited to, identity theft, cross-matching without consent, and severely limited renewability. Increasingly, the public and non-governmental organisations pay attention to those (and other) issues associated with centralised storage of sensitive personal and biometric data. In some areas, this contributes to the political process of widening legislation against privacy violations (*e.g.* GDPR in Europe [2]), which entail significant responsibilities for the data controllers.

Recently, biometric *template protection* (see *e.g.* [3] for a survey) has been an active research field attempting to address said security and privacy challenges. The ISO/IEC IS 24745 [4] mandates several properties, which must be guaranteed by such schemes:

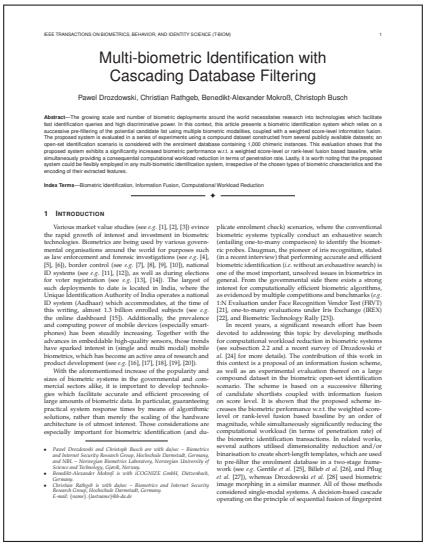
- **Unlinkability** referring to making it infeasible to determine if two or more protected templates were derived from the same instance. By fulfilling this property, cross-matching across different databases is prevented.
- **Irreversibility** referring to making it infeasible to reconstruct the original biometric data given a protected template and its corresponding secret. With this property fulfilled, the privacy of the users' data is increased, and additionally the security of the system is increased against presentation and replay attacks.
- **Renewability** referring to making it possible to revoke old protected templates and creating new ones from the same biometric instance and/or sample. With this property fulfilled, it is possible to revoke and reissue the templates in case of the database being compromised, thereby preventing misuse.
- **Performance preservation** referring to the requirement of the biometric performance not being significantly impaired by the protection scheme.

Three main biometric template protection approach classes can be distinguished: (1) biometric cryposystems, which use the biometric data to bind or extract a key [5], (2) cancelable biometrics, which utilise irreversible transformations to the biometric samples or templates [6], and (3) (homomorphic) encryption of biometric data [7].

Homomorphic encryption (henceforth referred to as "HE") makes it possible to compute operations in the encrypted domain, which render the same result as those in the plaintext domain. Thus, provided that it is possible to implement a given biometric comparator to feasibly operate within the homomorphic domain, such a template protection scheme would operate without any loss of biometric performance, whereas some impairment is often inevitable in biometric cryposystems and cancelable biometrics. In general, an encryption algorithm E has the homomorphic property for an operation $\odot$ if it holds $E(m_1) \odot E(m_2) = E(m_1 \odot m_2), m_1, m_2 \in M$ where M is the set of all possible messages. HE schemes are classified depending on the number and type of supported $\odot$ operations (see *e.g.* [8] for a detailed survey). The following three HE scheme types exist today:

- **Partially Homomorphic Encryption (PHE)** schemes are defined as allowing only a single operation type an unlimited number of times. PHE schemes have been around for over 30 years and are the oldest HE schemes like the classical RSA scheme [9] and the ElGamal scheme [10]

Publication reference: DROZDOWSKI, P., BUCHMANN, N., RATHGEB, C., MARGRAF, M., AND BUSCH, C. On the application of homomorphic encryption to face identification. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2019), IEEE, pp. 173–180.
DOI: N/A
Thesis chapter: 9
Addressed research question(s): RQ1, RQ5
Background: Homomorphic encryption makes it possible to carry out meaningful computations on encrypted data. This unique property is of interest in various contexts where the privacy and security of the data needs to be guaranteed. Recent public scrutiny and legislation (*e.g.* GDPR in the European Union [15]) means that increasingly stringent data privacy and security requirements are put onto biometric (and other) systems.
Contribution: While homomorphic encryption has previously been shown to be able to satisfy the biometric template protection goals of ISO/IEC 24745 [16], it has only been done for biometric verification. The main contribution of this research article is an architecture and protocol proposal for performing biometric identification in the homomorphically encrypted domain, as well as an implementation and evaluation thereof using open-source frameworks. As an additional contribution, the research article discusses the relevant challenges and technical considerations in the context of the efficient biometric identification and computational workload reduction in the homomorphically encrypted domain.



IEEE TRANSACTIONS ON BIOMETRICS, BEHAVIOR, AND IDENTITY SCIENCE (TBIOM) 1

Multi-biometric Identification with Cascading Database Filtering

Pawel Drozdowski, Christian Rathgeb, Benedikt-Alexander Mokroß, Christoph Busch

Abstract—The growing scale and number of biometric deployments around the world necessitates research into technologies which facilitate fast identification queries and high discriminative power. In this context, this article presents a biometric identification system which relies on a successive pre-filtering of the potential candidate list using multiple biometric modalities, coupled with a weighted score-level information fusion. The proposed system is evaluated in a series of experiments using a compound dataset constructed from several publicly available datasets; an open-set identification scenario is considered with the enrolment database containing 1 000 chimeric instances. This evaluation shows that the proposed system exhibits a significantly increased biometric performance w.r.t. a weighted score-level or rank-level fusion based baseline, while simultaneously providing a consequential computational workload reduction in terms of penetration rate. Lastly, it is worth noting that the proposed system could be flexibly employed in any multi-biometric identification system, irrespective of the chosen types of biometric characteristics and the encoding of their extracted features.

Index Terms—Biometric Identification, Information Fusion, Computational Workload Reduction

1 INTRODUCTION

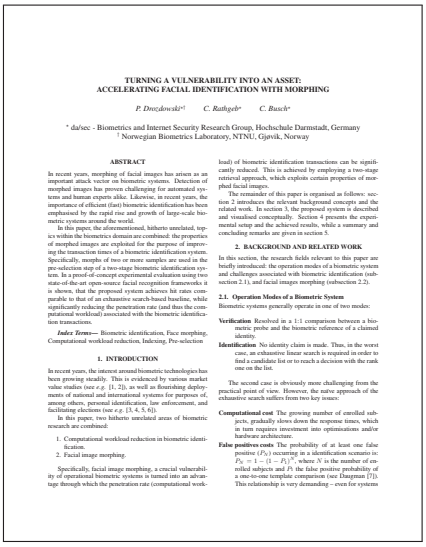
Various market value studies (see *e.g.* [1], [2], [3]) evince the rapid growth of interest and investment in biometric technologies. Biometrics are being used by various governmental organisations around the world for purposes such as law enforcement and forensic investigations (see *e.g.* [4], [5], [6]), border control (see *e.g.* [7], [8], [9], [10]), national ID systems (see *e.g.* [11], [12]), as well as during elections for voter registration (see *e.g.* [13], [14]). The largest of such deployments to date is located in India, where the Unique Identification Authority of India operates a national ID system (Aadhaar) which accommodates, at the time of this writing, almost 1.3 billion enrolled subjects (see *e.g.* the online dashboard [15]). Additionally, the prevalence and computing power of mobile devices (especially smartphones) has been steadily increasing. Together with the advances in embeddable high-quality sensors, those trends have sparked interest in (single and multi modal) mobile biometrics, which has become an active area of research and product development (see *e.g.* [16], [17], [18], [19], [20]).

With the aforementioned increase of the popularity and sizes of biometric systems in the governmental and commercial sectors alike, it is important to develop technologies which facilitate accurate and efficient processing of large amounts of biometric data. In particular, guaranteeing practical system response times by means of algorithmic solutions, rather than merely the scaling of the hardware architecture is of utmost interest. Those considerations are especially important for biometric identification (and duplicate enrolment check) scenarios, where the conventional biometric systems typically conduct an exhaustive search (entailing one-to-many comparisons) to identify the biometric probes. Daugman, the pioneer of iris recognition, stated (in a recent interview) that performing accurate and efficient biometric identification (*i.e.* without an exhaustive search) is one of the most important, unsolved issues in biometrics in general. From the governmental side there exists a strong interest for computationally efficient biometric algorithms, as evidenced by multiple competitions and benchmarks (*e.g.* 1:N Evaluation under Face Recognition Vendor Test (FRVT) [21], one-to-many evaluations under Iris Exchange (IREX) [22], and Biometric Technology Rally [23]).

In recent years, a significant research effort has been devoted to addressing this topic by developing methods for computational workload reduction in biometric systems (see subsection 2.2 and a recent survey of Drozdowski *et al.* [24] for more details). The contribution of this work in this context is a proposal of an information fusion scheme, as well as an experimental evaluation thereof on a large compound dataset in the biometric open-set identification scenario. The scheme is based on a successive filtering of candidate shortlists coupled with information fusion on score level. It is shown that the proposed scheme increases the biometric performance w.r.t. the weighted score-level or rank-level fusion based baseline by an order of magnitude, while simultaneously significantly reducing the computational workload (in terms of penetration rate) of the biometric identification transactions. In related works, several authors utilised dimensionality reduction and/or binarisation to create short-length templates, which are used to pre-filter the enrolment database in a two-stage framework (see *e.g.* Gentile *et al.* [25], Billeb *et al.* [26], and Pflug *et al.* [27]), whereas Drozdowski *et al.* [28] used biometric image morphing in a similar manner. All of those methods considered single-modal systems. A decision-based cascade operating on the principle of sequential fusion of fingerprint

- *Pawel Drozdowski and Christoph Busch are with dasec – Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany, and NBL – Norwegian Biometrics Laboratory, Norwegian University of Science and Technology, Gjøvik, Norway.*
- *Benedikt-Alexander Mokroß is with iCOGNIZE GmbH, Dietzenbach, Germany.*
- *Christian Rathgeb is with dasec – Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany.*
  E-mail: {surname}.{lastname}@h-da.de

Publication reference: DROZDOWSKI, P., RATHGEB, C., MOKROSS, B.-A., AND BUSCH, C. Multi-biometric identification with cascading database filtering. *Transactions on Biometrics, Behavior, and Identity Science (TBIOM)*, (March 2020), 1–14.
DOI: `https://doi.org/10.1109/TBIOM.2020.2977215`
Thesis chapter: 10
Addressed research question(s): RQ1, RQ3, RQ4
Background: By combining information from multiple sources (*e.g.* multiple biometric modalities), the discriminative power (and hence biometric performance) of a biometric recognition system can be increased. On the other hand, the use of multiple information sources tends to make the system more computationally demanding, typically in terms of additional template comparisons being necessary.
Contribution: The main contribution of this research article is a concept for a system which successively filters the potential candidate lists in biometric identification transactions using multiple types of biometric characteristics. The proposed concept is evaluated experimentally, where it is shown that it not only significantly reduces the computational workload, but also improves the biometric performance w.r.t. a weighted score-level fusion baseline. The proposed method works irrespective of the chosen type of biometric characteristic or their corresponding feature representations, since it only relies on the comparison scores and ranked candidate lists.



TURNING A VULNERABILITY INTO AN ASSET:
ACCELERATING FACIAL IDENTIFICATION WITH MORPHING

P. Drozdowski[*†] *C. Rathgeb*[*] *C. Busch*[*]

[*] da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
[†] Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway

ABSTRACT

In recent years, morphing of facial images has arisen as an important attack vector on biometric systems. Detection of morphed images has proven challenging for automated systems and human experts alike. Likewise, in recent years, the importance of efficient (fast) biometric identification has been emphasised by the rapid rise and growth of large-scale biometric systems around the world.

In this paper, the aforementioned, hitherto unrelated, topics within the biometrics domain are combined: the properties of morphed images are exploited for the purpose of improving the transaction times of a biometric identification system. Specifically, morphs of two or more samples are used in the pre-selection step of a two-stage biometric identification system. In a proof-of-concept experimental evaluation using two state-of-the-art open-source facial recognition frameworks it is shown, that the proposed system achieves hit rates comparable to that of an exhaustive search-based baseline, while significantly reducing the penetration rate (and thus the computational workload) associated with the biometric identification transactions.

Index Terms— Biometric identification, Face morphing, Computational workload reduction, Indexing, Pre-selection

1. INTRODUCTION

In recent years, the interest around biometric technologies has been growing steadily. This is evidenced by various market value studies (see *e.g.* [1, 2]), as well as flourishing deployments of national and international systems for purposes of, among others, personal identification, law enforcement, and facilitating elections (see *e.g.* [3, 4, 5, 6]).

In this paper, two hitherto unrelated areas of biometric research are combined:

1. Computational workload reduction in biometric identification.
2. Facial image morphing.

Specifically, facial image morphing, a crucial vulnerability of operational biometric systems is turned into an advantage through which the penetration rate (computational workload) of biometric identification transactions can be significantly reduced. This is achieved by employing a two-stage retrieval approach, which exploits certain properties of morphed facial images.

The remainder of this paper is organised as follows: section 2 introduces the relevant background concepts and the related work. In section 3, the proposed system is described and visualised conceptually. Section 4 presents the experimental setup and the achieved results, while a summary and concluding remarks are given in section 5.

2. BACKGROUND AND RELATED WORK

In this section, the research fields relevant to this paper are briefly introduced: the operation modes of a biometric system and challenges associated with biometric identification (subsection 2.1), and facial images morphing (subsection 2.2).

2.1. Operation Modes of a Biometric System

Biometric systems generally operate in one of two modes:

Verification Resolved in a 1:1 comparison between a biometric probe and the biometric reference of a claimed identity.

Identification No identity claim is made. Thus, in the worst case, an exhaustive linear search is required in order to find a candidate list or to reach a decision with the rank one on the list.

The second case is obviously more challenging from the practical point of view. However, the naïve approach of the exhaustive search suffers from two key issues:

Computational cost The growing number of enrolled subjects, gradually slows down the response times, which in turn requires investment into optimisations and/or hardware architecture.

False positives costs The probability of at least one false positive (P_N) occurring in a identification scenario is: $P_N = 1 - (1 - P_1)^N$, where N is the number of enrolled subjects and P_1 the false positive probability of a one-to-one template comparison (see Daugman [7]). This relationship is very demanding – even for systems

Publication reference: DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Turning a vulnerability into an asset: Accelerating facial identification with morphing. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (May 2019), IEEE, pp. 2582–2586.
DOI: `https://doi.org/10.1109/ICASSP.2019.8683326`
Thesis chapter: 11
Addressed research question(s): RQ1, RQ3, RQ4
Background: Facial morphing has recently been established as a potential attack vector against facial recognition systems. During the morphing process, images of multiple data subjects are combined together to produce a single image which exhibits a similarity (both in the context of human expert perception and automated recognition systems) to each of the contributing data subjects. Current research predominantly concentrates on methods for creation of better morphs and detection of the morphed images.
Contribution: The contribution of this research article is the insight that morphed images (a vulnerability) can be used in a positive way (an asset). Specifically, it is shown that by using morphing, the the biometric enrolment database can be organised into a two-step retrieval system. In a biometric identification transaction, the morphed images can be used to conduct a rough pre-selection of the candidate list (and hence computational workload reduction); subsequently, the individual images from the candidates contributing to the pre-selected morphs are checked using the normal procedure. Furthermore, the proposed scheme functions on the image-level, *i.e.* prior to feature extraction, which enables the use of any system susceptible to morphing attacks for the pre-selection step, whereas the feature representation for the second step can be chosen freely.

Towards Pre-alignment of Near-infrared Iris Images

P. Drozdowski[*†], C. Rathgeb[*], H. Hofbauer[‡], J. Wagner[*], A. Uhl[‡] and C. Busch[*]
[*] da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
[†] Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway
[‡] Multimedia Signal Processing and Security Lab, University of Salzburg, Austria
{pawel.drozdowski,christian.rathgeb,johannes.wagner,christoph.busch}@h-da.de
{hhofbaue,uhl}@cosy.sbg.ac.at

Abstract

The necessity of biometric template alignment imposes a significant computational load and increases the probability of false positive occurrences in biometric systems. While for some modalities, automatic pre-alignment of biometric samples is utilised, this topic has not yet been explored for systems based on the iris.

This paper presents a method for pre-alignment of iris images based on the positions of automatically detected eye corners. Existing work in the area of automatic eye corner detection has hitherto only involved visible wavelength images; for the near-infrared images, used in the vast majority of current iris recognition systems, this task is significantly more challenging and as of yet unexplored. A comparative study of two methods for solving this problem is presented in this paper. The eye corners detected by the two methods are then used for the pre-alignment and biometric performance evaluation experiments. The system utilising image pre-alignment is benchmarked against a baseline iris recognition system on the iris subset of the BioSecure database. In the benchmark, the workload associated with alignment compensation is significantly reduced, while the biometric performance remains unchanged or even improves slightly.

1. Introduction

The iris is one of the main biometric characteristics used in biometric systems around the world. At the time of this writing, the Indian Aadhaar system has enrolled over 1 billion subjects' multi-modal (including iris) biometric data [27]. The border control system of United Arab Emirates checks every traveller against a growing blacklist consisting of hundreds of thousands of subjects [2]. The deployments of this size and importance face strenuous requirements in terms of, among other matters, biometric performance and computational efficiency.

Following Daugman's approach [7], which is the core of most public operational systems, four major modules constitute an iris recognition system: (1) acquisition of the near-infrared image, where most current deployments require subjects to fully cooperate with the capture device in order to capture images of sufficient quality; (2) preprocessing, which involves a detection of inner and outer iris boundaries, a detection of eyelids, an exclusion of eyelashes as well as contact lens rings, a scrubbing of specular reflections and an estimation of quality factors [13]. Subsequently, the iris is mapped to dimensionless coordinates, i.e. a normalized rectangular texture, and an according noise mask is stored; (3) feature extraction, in which a two-dimensional binary feature vector, i.e. iris-code, is generated by applying adequate filters to the pre-processed iris texture. This binary data representation enables compact storage and rapid (4) comparison, which is based on the estimation of Hamming distance (HD) scores between pairs of iris-codes. In the comparison stage circular bit-shifts are applied to iris-codes and HD scores are estimated at $\pm K$ different shifting positions, i.e. relative tilt angles, in order to compensate the biometric sample misalignment. The minimal obtained HD, which corresponds to an optimal alignment, represents the final score.

Considering multiple shifting positions during a template comparison increases the computational workload of the system and the probability of a false match with K [8]. This is especially important for identification systems, where an exhaustive search of the reference database is performed during an authentication attempt. By pre-aligning the eye images, the aforementioned cost (in terms of computational workload and biometric performance degradation) could be significantly reduced, thus partially alleviating the issues created by the necessity of alignment compensation. For the biometric references, the pre-alignment could be performed at enrollment stage, while any additional computational cost of pre-alignment of the biometric

Publication reference: DROZDOWSKI, P., RATHGEB, C., HOFBAUER, H., WAGNER, J., UHL, A., AND BUSCH, C. Towards pre-alignment of near-infrared iris images. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 359–366.
DOI: https://doi.org/10.1109/BTAS.2017.8272718
Thesis chapter: 12
Addressed research question(s): RQ1, RQ4
Background: Most of the currently operational iris recognition systems utilise some version of the Daugman algorithm (Iris-Code). The algorithm includes sample misalignment (*i.e.* relative tilt angles) compensation at the template comparison stage: the Hamming distance between two Iris-Codes (binary matrices) is computed at multiple relative positions by circularly shifting the matrices, thereby incurring additional computational costs. Since the comparison score at the optimal alignment is chosen, the score distribution of (zero-effort) impostors tends to shift towards the genuine score distribution, which typically increases the probability of false positives.
Contribution: The main contribution of this research article is a method of iris pre-alignment, which works on the biometric sample level. By detecting the eye corners in the iris images, they can be rotated onto a common plane, so that the eye corners form a horizontal line. As a consequence, the relative misalignment of the samples is (on average) smaller overall. Thus, fewer relative shifting positions need to be considered during the template comparison, hence reducing the computational workload and improving the biometric performance. An additional contribution are the methods for eye corner detection, as previous research in this area only concerned visible wavelength and not near-infrared iris images.

Detection of Glasses in Near-infrared Ocular Images

P. Drozdowski*†, F. Struck*, C. Rathgeb* and C. Busch*
* da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
† Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway
{pawel.drozdowski,christian.rathgeb,christoph.busch}@h-da.de
florian.struck@stud.h-da.de

Abstract—**Eyeglasses change the appearance and visual perception of facial images. Moreover, under objective metrics, glasses generally deteriorate the sample quality of near-infrared ocular images and as a consequence can worsen the biometric performance of iris recognition systems. Automatic detection of glasses is therefore one of the prerequisites for a sufficient quality, interactive sample acquisition process in unsupervised iris recognition applications. Three different approaches (i.e. a statistical method, a deep learning based method and an algorithmic method based on detection of edges and reflections) for automatic detection of glasses in near-infrared iris images are presented. Those approaches are evaluated using cross-validation on the CASIA-IrisV4-Thousand dataset, which contains 20000 images from 1000 subjects. Individually, they are capable of correctly classifying 95-98% of images, while a majority vote based fusion of the three approaches achieves a correct classification rate (CCR) of 99.54%.**

***Keywords*—Biometrics; Iris Recognition; Glasses Detection;**

I. INTRODUCTION

In recent years, iris recognition has become a popular modality for biometric systems and is used in many large-scale deployments (e.g. the Indian National ID project [22]). The technology is also increasingly being used in automatic (without human operator supervision) systems, such as smart borders/airport gates and mobile devices [14]. Operational systems typically capture iris images in the near-infrared light spectrum, in which the iris patterns are much more pronounced than in the visible light spectrum, even for darkly pigmented irides [6]. According to recent reports [20], [21], over 50% of adult population in the developed world wear eyeglasses. The pervasiveness of short-sightedness (myopia) has been on an extreme rise in Eastern Asia and around the world in general; a recent report in Nature News [7] states:

> East Asia has been gripped by an unprecedented rise in myopia, also known as short-sightedness. Sixty years ago, 10-20% of the Chinese population was short-sighted. Today, up to 90% of teenagers and young adults are. In Seoul, a whopping 96.5% of 19-year-old men are short-sighted. Other parts of the world have also seen a dramatic increase in the condition, which now affects around half of young adults in the United States and Europe - double the prevalence of half a century ago. By some estimates, one-third of the world's population - 2.5 billion people - could be affected by short-sightedness by the end of this decade.

Due to specular reflections, blur, scratches and other factors, glasses tend to decrease the biometric sample quality and consequently often the biometric performance of the systems. While several researchers have investigated the impact of glasses on face recognition systems, the scientific literature on iris recognition contains very little related work on this subject, except for a paper in which a small-scale quantification of the effects of glasses on iris image pre-processing is presented [13] and glasses being mentioned as a significant noise factor (e.g. [3], [1], [9]). ISO/IEC 29794-6 biometric sample quality standard [10] specifically recommends to instruct data subjects to remove glasses during acquisition or to perform the acquisition with additional care.

Therefore, and due to the prevalence of glasses in the world population, automatic detection of glasses is an important matter in iris recognition (as will also be substantiated by the experiments described in section III). It is of particular interest for automatic sample acquisition systems, where such a detection module would enable an interactive sample acquisition and thus facilitate higher sample quality. While this is a well-researched topic in systems working with images of the facial region (e.g. [23], [2]), doing so in images of ocular region alone has not received enough attention. In this paper, three methods for accomplishing said task are presented and benchmarked.

This paper is organised as follows: in section II, the used dataset and experimental setup are described. Section III provides an overview of the impact of glasses on iris recognition. In section IV the three proposed automatic glasses detection approaches are presented and evaluated. Concluding remarks are given in section V.

II. EXPERIMENTAL SETUP

The Thousand subset of the CASIA-IrisV4 database [5] (henceforth referred to as "CASIA-Thousand dataset") was chosen for the experiments performed for this paper. Said dataset contains near-infrared iris images of size 640 × 480 pixels and, due to its size, is suitable for large-scale testing.

Publication reference: DROZDOWSKI, P., STRUCK, F., RATHGEB, C., AND BUSCH, C. Detection of glasses in near-infrared ocular images. In *International Conference on Biometrics (ICB)* (February 2018), IEEE, pp. 202–208.
DOI: `https://doi.org/10.1109/ICB2018.2018.00039`
Thesis chapter: 13
Addressed research question(s): RQ2
Background: The appearance and visual perception of facial and ocular images changes with the presence of eyeglasses. Furthermore, eyeglasses tend to deteriorate the objective sample quality of ocular images (*e.g.* due to blur, reflections, and scratches); as a consequence, the biometric performance of iris recognition systems can be negatively affected. With the quickly raising prevalence of myopia (nearsightedness), especially in the technologically developed world, eyeglasses can present an additional challenge for the operational iris recognition systems.
Contribution: The main contributions of this research article are twofold: firstly, an experimental evaluation of the impact of eyeglasses on iris recognition using a large near-infrared ocular dataset is conducted. Previously, only small studies have been reported in the literature. Secondly, three methods (and a fusion thereof) are shown to reliably detect eyeglasses in near-infrared ocular images. An additional contribution are the labels (with or without eyeglasses) for one of the largest publicly available near-infrared ocular datasets.

SIC-Gen: A Synthetic Iris-Code Generator

Paweł Drozdowski[1,2], Christian Rathgeb[2] and Christoph Busch[2]
[1]*Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway*
[2]*da/sec – Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany*
{pawel.drozdowski,christian.rathgeb,christoph.busch}@h-da.de

***Abstract*—Nowadays large-scale identity management systems enrol more than one billion data subjects. In order to limit transaction times, biometric indexing is a suitable method to reduce the search space in biometric identifications. Effective testing of such biometric identification systems and biometric indexing approaches requires large datasets of biometric data. Currently, the size of the publicly available iris datasets is insufficient, especially for system scalability assessments. Synthetic data generation offers a potential solution to this issue; however, it is challenging to generate data that is both statistically sound and visually realistic – for the iris, the currently available approaches prove unsatisfactory.**

In this paper, we present a method for generation of synthetic binary iris-based templates, i.e. Iris-Codes, which are the de facto standard in currently deployed major biometric deployments around the world. We validate the statistical properties of the synthetic templates and show that they closely resemble ones produced from real ocular images. With the proposed approach, large databases of synthetic Iris-Codes with flexibly adjustable properties can be generated.

***Index Terms*—Biometrics, Iris Recognition, Iris-Code, Synthetisation**

I. INTRODUCTION

The iris is one of the most widely applied biometric modalities. In recent years, several large-scale deployments have been created, most notably the Indian National ID program [1], which has, at the time of this writing, enrolled over one billion subjects with biometric data including the irides. Despite using efficient comparators (e.g. Hamming distance for the iris) and parallelism, the computational load faced by such deployments in the identification scenario is extremely high. With biometric workload reduction as a motivation, many approaches for indexing of iris data have been developed [2]. However, evaluation of such approaches and their scalability is often questionable due to lack of large test datasets. While various publicly available iris databases with near-infrared (NIR) data exist, they are relatively small. At the time of this writing, some of the largest publicly available datasets, CASIA-IrisV4-Thousand and ND-CrossSensor-Iris-2013, contain merely 20.000 images from 1000 subjects and 146.550 images from 676 subjects, respectively. This is several orders of magnitude smaller than some of the large-scale deployments nowadays.

Synthetic data generation is one possible way of dealing with the issue of testing efficient indexing methods. Most of the existing approaches for synthetic iris generation attempt to synthesise an entire iris image or texture [3]–[11]. The main issues with such approaches include the computational costs and the difficulty in guaranteeing the statistical properties of the real data. The vast majority of operational iris biometric systems are based on the Iris-Code [12], making it a *de facto* standard. Generating Iris-Codes (feature vectors) directly is therefore also viable and may offer better control over the statistical properties of the synthetic data. Recently, two such approaches have been proposed. Proença and Neves [11] provide a method of Iris-Code synthesis based on bit correlations; the method is shown to attain some of the desired statistical properties (the shapes of the genuine and impostor distributions). It is also somewhat flexible with adjustable parameters; however, it does not allow to generate a set of templates following a desired score distribution. Furthermore, the liker-response resulting from the typical feature extraction process is not modelled (in other words, the produced synthetic Iris-Codes scantily resemble the ones produced from real iris images through the commonly used iris processing pipeline). Lastly, typical error patterns between two mated templates are not modelled. Daugman [13] proposed to use a simple hidden Markov model to generate a stream of bits and showed that it can be adjusted, so that the produced templates mimic the impostor distribution of real iris templates. However, the produced streams are 1-dimensional (i.e. do not model the correlation between the Iris-Code rows); furthermore, the method does not offer a way to generate more than one template per subject (i.e. it is not possible to use it for simulating genuine comparisons). As such, it might only be useful for stress-testing of iris identification systems.

In this paper, we present a synthetic Iris-Code generator, which both reflects the statistical properties of the real Iris-Codes and resembles the real templates visually. An important feature of the proposed approach is its flexibility, in that it allows to generate Iris-Codes with an arbitrary resolution and an arbitrary score distribution of mated templates, unlike any of the approaches currently in the literature. To facilitate reproducible research, the software written in Python3 programming language, is released to the scientific community under a permissive license.

The remainder of this paper is organised as follows: section II describes the proposed method of synthetic Iris-Code generation. In section III the properties of the generated templates are validated, while section IV contains concluding remarks.

Publication reference: DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. SIC-Gen: A synthetic Iris-Code generator. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2017), IEEE, pp. 61–69.
DOI: `https://doi.org/10.23919/BIOSIG.2017.8053520`
Thesis chapter: 14
Addressed research question(s): RQ2
Background: Large datasets of biometric data are required in order to conduct effective tests of biometric identification systems and computational workload reduction methods. Currently, however, the publicly available datasets are often insufficient in size, especially for the assessment of system scalability. A potential solution to this issue is synthetic data generation, whereby arbitrarily large datasets can be created within short timespans. However, synthetically generating data which is both visually and statistically realistic is considered challenging.
Contribution: The main contribution of this research article is a conceptual framework for generation of synthetic binary iris templates (*i.e.* Iris-Codes). The implementation of the concept has been provided as an open-source code release. The templates generated using the proposed method closely resemble the ones extracted from the real ocular images and mimic their statistical properties. The proposed approach is highly flexible and parametrisable, thereby allowing the generation of large databases for a variety of stress testing scenarios.

Score Fusion Strategies in Single-Iris Dual-Probe Recognition Systems

Pawel Drozdowski
da/sec - Biometrics and Internet Security Research Group
Hochschule Darmstadt, Germany
Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway
pawel.drozdowski@h-da.de

Nikolai Wiegand
da/sec - Biometrics and Internet Security Research Group,
Hochschule Darmstadt, Germany
nikolai.wiegand@stud.h-da.de

Christian Rathgeb
da/sec - Biometrics and Internet Security Research Group,
Hochschule Darmstadt, Germany
christian.rathgeb@h-da.de

Christoph Busch
da/sec - Biometrics and Internet Security Research Group,
Hochschule Darmstadt, Germany
christoph.busch@h-da.de

ABSTRACT

Multiple samples can be utilised at the comparison stage of a biometric system in order to increase its biometric performance via information fusion or decision heuristics. It has been shown, that in a single-instance dual-probe setup, fusing the probe scores yields significant biometric performance increase over the single-probe baseline. Additionally, using the probe-probe comparison score was demonstrated to further improve the biometric performance of a fingerprint recognition system in a study by Cheng et al. In this paper, through a benchmark on the CASIA-IrisV4-Interval dataset and on the iris corpus of the BioSecure dataset, the aforementioned method is shown to be viable for an iris recognition system. However, since it requires an additional parameter, which must be estimated empirically, we propose a simpler method which exhibits similar biometric performance, while requiring no additional parameterisation.

CCS Concepts

• **Security and privacy** → **Biometrics**; • **Computing methodologies** → **Biometrics**;

Keywords

Biometrics, Biometric Information Fusion, Iris Recognition

1. INTRODUCTION

In past years, several multi-biometric iris recognition systems have been proposed [1, 2], some of which consolidate information from multiple samples of a single eye instance during enrolment. Some of these single-instance multi-sample fusion approaches have been found to significantly improve the recognition accuracy of iris recognition systems. The vast majority of proposed iris-based multi-sample fusion schemes process multiple extracted feature vectors, i.e. binary iris-codes, at the time of enrolment. The first conceptual scheme of this kind was presented in [3], in which a majority vote-based coding is applied for each bit position of an odd number of iris-codes, with the goal of reducing the intra-class variation between the resulting reference and probe iris-codes. In [4], a weighted majority voting was proposed to improve the accuracy of an iris recognition system. A weight map, which indicates the stability of iris-code bits, is obtained from several iris-codes at enrolment. Comparison scores are then estimated as a weighted sum of mis-matching bits. A similar approach based on personalised weight maps has been presented in [5]. In [6], so-called "fragile" bits, i.e. bits which exhibit a higher probability than others to flip their value during a genuine comparison, are detected by comparing several iris-codes obtained from a single eye. Incorporating those bits into noise-masks extracted in the iris segmentation stage was shown to improve the overall biometric performance of the iris recognition system. In contrast to the aforementioned approaches, a signal-level fusion of iris texture information extracted from multiple frames of a video was proposed in [7]. Based on a pixel-wise averaging, a single normalised iris texture is obtained. Such textures exhibit higher quality/reliability, and have been shown to improve the biometric performance of an iris recognition system. This scheme has been derived from a concept which was first introduced for face recognition [8]. Similar schemes have been proposed for fingerprint recognition systems [9, 10]. In [11], a score fusion of single-fingerprint dual-probe is proposed, where in addition to utilising the reference-probe comparison scores, the probe-probe comparison score is incorporated into a score fusion. In this paper, said score fusion method, along with proposed of further heuristics are applied in an iris-based system and benchmarked.

The remainder of this paper is organised as follows: in section 2, the employed fusion strategies for single-iris dual-probe iris recognition are described. In section 3, the experimental setup and results are presented, while section 4 contains a summary of the paper.

2. FUSION STRATEGIES

State-of-the-art iris recognition systems capture multiple samples during acquisition stage for the purpose of supporting compensation of pose or gaze variations or for providing some fundamental presentation attack detection (PAD) [12]. These additional samples can then be utilised at comparison stage. Specifically, in a system where two probe samples are present at comparison stage, three comparison scores can be computed as shown in figure 1: two (HD_1 and HD_2) between the reference and each probe and one (HD_3) between the two probes themselves. It is then possible to fuse the scores, for example, in following ways:

ICBEA '18, May 24–26, 2018, Amsterdam, Netherlands

ACM ISBN 978-1-4503-6358-5/18/05…$15.00
https://doi.org/10.1145/3230820.3230823

Publication reference: DROZDOWSKI, P., WIEGAND, N., RATHGEB, C., AND BUSCH, C. Score fusion strategies in single-iris dual-probe recognition systems. In *International Conference on Biometric Engineering and Applications (ICBEA)* (May 2018), ACM, pp. 13–17.
DOI: `https://doi.org/10.1145/3230820.3230823`
Thesis chapter: 15
Addressed research question(s): RQ3
Background: By acquiring and using multiple samples of a biometric probe, the biometric performance of a biometric recognition system can be improved, *e.g.* by applying quality assurance and/or information fusion methods.
Contribution: In this research article, an existing method of dual-sample fingerprint recognition is transferred to iris recognition and benchmarked. The main contribution is an extension of the original method by proposing a heuristic, which, as opposed to the original method, is parameter-free and hence requires no training step.

3.1.2 Additional

Several additional research articles have been published during the course of the doctoral studies research. They are listed below.

IET Biometrics

Research Article

IET Journals
The Institution of Engineering and Technology

Bloom filter-based search structures for indexing and retrieving iris-codes

ISSN 2047-4938
Received on 20th January 2017
Revised 18th July 2017
Accepted on 8th August 2017
doi: 10.1049/iet-bmt.2017.0007
www.ietdl.org

Paweł Drozdowski[1] ✉, Christian Rathgeb[1], Christoph Busch[1]
[1]da/sec Biometrics and Internet Security Research Group, Hochschule Darmstadt, Darmstadt, Germany
✉ E-mail: pawel.drozdowski@h-da.de

Abstract: Large-scale biometric deployments are becoming ubiquitous. The computational workload of the conventional retrieval method, requiring $1:N$ comparisons in the identification mode, is impractical for such systems. In recent years, many approaches for efficient biometric identification were proposed, but their scalability is often questionable. Furthermore, the lack of a unified methodology for biometric workload reduction reporting often makes a direct benchmark or a thorough evaluation of the proposed schemes cumbersome. We present an iris indexing scheme based on Bloom filters and binary search trees. With a statistical model, the system is shown to be theoretically scalable for arbitrarily many enrollees. We evaluate this system on a combined database from several publicly available datasets, containing a total of 11,936 iris images from 1477 instances. In an open-set identification scenario, the system maintains the biometric performance of an iris-code $1:N$ baseline – a true positive identification rate of approximately 98% measured at 0.1% false positive identification rate, at only 10% of the baseline workload. In a proof-of-concept multi-iris indexing experiment, the true positive identification rate is further increased to over 99%, without additional workload costs. Lastly, we define several prerequisites necessary for a transparent and comprehensive methodology of biometric workload reduction results dissemination.

1 Introduction

In recent years, several large-scale biometric systems have been introduced worldwide. By far the largest of these is the Indian National ID project, which at the time of this writing has successfully enrolled over one billion subjects [1] with biometric data from iris, face and fingerprints. Two main challenges associated with large-scale biometric identification are the computational workload and the response time. The conventional (*brute force*) approach is to perform template comparisons between the probe and all enrolled reference templates (i.e. $1:N$ comparisons). Even with excellent hardware and reliance on parallelism, the computational cost quickly becomes prohibitive. Similarly, the possibility of false positive occurrences quickly becomes unacceptable. In [2], Daugman shows the probability of at least one false positive (P_N) occurring in an identification scenario to be calculated using (1), where N is the number of enrolled subjects and P_1 the false positive probability of a one-to-one template comparison

$$P_N = 1 - (1 - P_1)^N \quad (1)$$

A biometric system which performs well in the verification mode (i.e. has a low P_1) is not necessarily suitable for the much more demanding identification mode. Observe that for values of P_1 which are acceptable for biometric verification, the value of P_N might very quickly becomes unacceptably high as the number of enrolled subjects N increases. The equation ignores other system errors (e.g. failure-to-acquire rate (FTA)).

Following Daugman's approach [3], which is the core of most public operational systems, four major modules constitute an iris recognition system: (i) image acquisition, where most current deployments require subjects to fully cooperate with the capture device in order to capture images of sufficient quality; (ii) pre-processing, which involves a detection of inner and outer iris boundaries, a detection of eyelids, an exclusion of eyelashes as well as contact lens rings, a scrubbing of specular reflections and an estimation of quality factors. Subsequently, the iris is mapped to dimensionless coordinates, i.e. a normalised rectangular texture, and an according noise mask is stored; (iii) feature extraction, in which a two-dimensional (2D) binary feature vector consisting of $W_{\mathrm{IC}} \times H_{\mathrm{IC}}$ bits, i.e. iris-code, is generated by applying adequate filters to the pre-processed iris texture. This binary data representation enables compact storage and rapid (iv) comparison, which is based on the estimation of Hamming distance (HD) scores between pairs of iris-codes. In the comparison stage circular bit shifts are applied to iris-codes and HD scores are estimated at K different shifting positions, i.e. relative tilt angles, resulting in $W_{\mathrm{IC}} \times H_{\mathrm{IC}} \times K$ bit comparisons. The minimal obtained HD, which corresponds to an optimal alignment, represents the final score.

Despite the rapid comparison and high resilience against false matches offered by the iris-code representation, the sheer scale of the major biometric deployments makes biometric workload reduction a very relevant and important research topic. Recently, a promising workload reduction approach employing a Bloom filter-based representation of iris-codes and binary search trees has been proposed in a proof-of-concept study [4]. In this paper, we further analyse and expand upon said idea.

1.1 Contribution of work and paper organisation

The remainder of this paper is organised as follows:

i. The related works in the area of biometric workload reduction for iris recognition are outlined (Section 2).
ii. A proposal for a standardised way of biometric workload reduction reporting is made (Section 2).
iii. A basic Bloom filter-based system for biometric indexing is described in detail. This system forms the basis of the work performed for this paper (Section 3).
iv. A general model for Bloom filter-based indexing is introduced along with a number of improvements for the basic system; those include scalability to an arbitrary number of enrollees and a first of its kind attempt of multi-iris indexing (Section 3).
v. The experiment methodology and results are presented and discussed, along with future work items and concluding remarks (Sections 4–6).

Publication reference: DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Bloom filter-based search structures for indexing and retrieving Iris-Codes. *IET Biometrics 7*, 3 (May 2018), 260–268.
DOI: `https://doi.org/10.1049/iet-bmt.2017.0007`
Thesis chapter: Appendix A
Addressed research question(s): RQ1, RQ2
Background: Currently, there exists no unified methodology for benchmarking the computational workload (and reduction thereof) in biometric identification systems. This makes it difficult to directly compare the results achieved by various computational workload reduction methods which are presented in the different scientific publications.
Contribution: This research article is included in the appendix, as a substantial part of its contents is based on previous work. Specifically, the theory and methods therein have been developed and described in the context of the M.Sc. thesis [1] of this author. On the other hand, the preparation of a large-scale experimental setup, biometric performance and computational workload reduction evaluation, as well as the process of article writing/revisioning have been conducted after the hand-in of the aforementioned thesis (*i.e.* during the course of the doctoral studies). In addition to the research component, where an efficient hierarchical retrieval method for iris recognition systems has been presented and shown to be arbitrarily scalable using a statistical model, this article has laid some of the groundwork for the submissions of comments to the ISO/IEC 19795-1 [17] standardisation project.

Iris Recognition in Visible Wavelength:
Impact and automated Detection of Glasses

D. Osorio-Roig*, P. Drozdowski†‡, C. Rathgeb†, A. Morales-González*, E. Garea-Llano* and C. Busch†
*Advanced Technologies Application Center (CENATAV), La Habana, Cuba
†da/sec – Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
‡Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway
{dosorio,egarea,amorales}@cenatav.co.cu
{pawel.drozdowski,christian.rathgeb,christoph.busch}@h-da.de

Abstract—The prevalence of visual impairment around the world is rapidly increasing, causing large numbers of people to wear glasses. Glasses are generally considered an important noise source in iris recognition; under objective metrics, they have recently been shown to deteriorate the sample quality of near-infrared (NIR) ocular images (consequently impairing the segmentation accuracy and biometric performance). Automatically and robustly detecting glasses in ocular images is therefore one of the important challenges for operational iris recognition systems. While this issue has recently been addressed for NIR ocular images, it remains an open issue in the visible wavelength (VW) spectrum. As the popularity of VW iris recognition increases (due to e.g. deployment of iris recognition in consumer grade mobile devices and general improvements in VW recognition algorithms), it becomes a matter of interest to quantitatively evaluate the impact of glasses on such systems, as well as develop methods for automatic detection of glasses in VW ocular images.

In this paper, the impact of glasses on VW iris segmentation performance is investigated using the UBIRIS.v2 and MobBIO iris databases. It is shown that the presence of glasses significantly degrades the accuracy of iris segmentation. In addition, a state-of-the-art iris segmentation method which can perform a semantic segmentation of ocular images (including the segmentation of eyeglasses) is proposed. On the evaluation of glasses in the UBIRIS.v2 used database, correct classification rates (CCRs) of 98.57% and 83.62% are obtained, respectively.

Index Terms—**biometrics, iris recognition, iris segmentation, glasses detection**

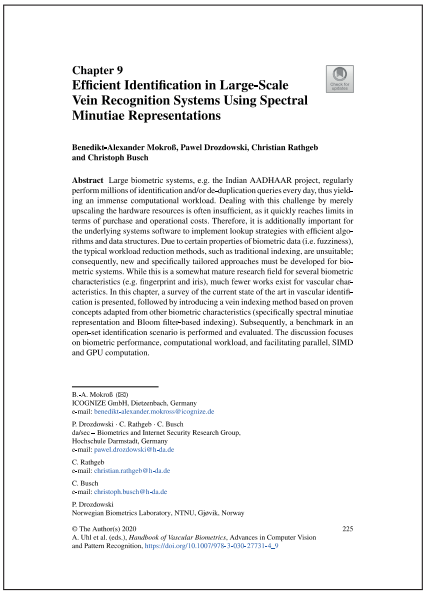
I. Introduction

In the past years, biometric recognition has become ubiquitous in various applications ranging from automated border control to forensic investigations. While some technologies, e.g. face or fingerprint recognition, are already commercially deployed in numerous application scenarios, the potential of others still needs to be explored. In particular, non-cooperative iris recognition based on images captured at VW represents a challenging task [1]. In contrast to iris images captured under NIR light, the iris tends to exhibit less textural information when acquired at VW, depending on the eye colour of a data subject. Furthermore, within VW iris images possible artefacts, such as specular reflections or shadows, are more pronounced. Those issues generally lead to an increased intra-class variation, which can cause a severe drop in the biometric performance. Accurate segmentation of the iris region in VW images represents one of the most critical tasks [2] in the processing pipeline and has also been the topic of several competitions (e.g. MICHE [3], [4] and NICE [5], which concentrated on mobile devices and noisy images, respectively) aimed at improving the accuracy of the contemporary algorithms. The segmentation of the iris involves a detection of inner and outer iris boundaries, a detection of eyelids, an exclusion of eyelashes and contact lense rings, as well as scrubbing of specular reflections [6]. More recently, methods based on deep learning, e.g. [7]–[10], revealed encouraging results for the task of iris segmentation.

Visual impairment is becoming an increasingly common affliction around the world. By some recent estimates (e.g. [11], [12]), over 50% of adults in the developed world are glass-wearers. In Eastern Asia, the prevalence of short-sightedness (myopia) has been rapidly increasing to unprecedented levels [13]. Several researchers mention glasses as a significant noise factor for iris recognition systems (e.g. [14]–[16]). However, very little related work on this subject is available in the contemporary scientific literature. In [17], the impact of glasses on the pre-processing pipeline of NIR iris images was evaluated in a small-scale study. Recently, a more thorough investigation on the impact of glasses for NIR iris images was done in [18]. In both works the drastic impact of glasses on NIR iris image pre-processing and recognition is demonstrated. Furthermore, in ISO/IEC 29794-6 biometric sample quality standard [19], it is recommended to exercise increased care during image acquisition from data subjects wearing glasses, or to outright instruct them to remove their glasses. Due to the non-trivial negative impact of glasses on the biometric performance of iris recognition systems, as well as the aforementioned pervasiveness of vision impairment (and, consequently, of glasses in the world population), automatic detection of glasses is an important matter in iris recognition. This is particularly the case for automatic sample acquisition systems, where higher sample quality could be facilitated through interactive sample acquisition with a glasses detection module.

In [18] methods based on texture descriptors, deep learning, edge/reflection detection, and a fusion thereof have been shown to achieve near-optimal results for glasses detection in

Publication reference: OSORIO-ROIG, D., DROZDOWSKI, P., RATHGEB, C., MORALES-GONZÁLEZ, A., GAREA-LLANO, E., AND BUSCH, C. Iris recognition in visible wavelength: Impact and automated detection of glasses. In *International Conference on Signal-Image Technology Internet-Based Systems (SITIS)* (November 2018), IEEE, pp. 542–546.
DOI: `https://doi.org/10.1109/SITIS.2018.00088`
Thesis chapter: This research article is not included in this thesis, as it has been decided to only include first-authorship research articles.
Addressed research question(s): RQ2
Background: The appearance and visual perception of facial and ocular images changes with the presence of eyeglasses. Furthermore, eyeglasses tend to deteriorate the objective sample quality of ocular images (*e.g.* due to blur, reflections, and scratches); as a consequence, the biometric performance of iris recognition systems can be negatively affected. With the quickly raising prevalence of myopia (nearsightedness), especially in the technologically developed world, eyeglasses can present an additional challenge for the operational iris recognition systems.
Contribution: This is a complementary research article to one of the main research articles ([13]). This research article considers the visible wavelength iris recognition, which in recent years is increasingly receiving attention as an alternative to the near-infrared-based iris recognition. The contribution is an experimental assessment of the impact of eyeglasses on visible wavelength iris recognition, as well as a method for detection of eyeglasses in visible wavelength iris images using a semantic ocular image segmentation neural network.



Chapter 9
Efficient Identification in Large-Scale Vein Recognition Systems Using Spectral Minutiae Representations

Benedikt-Alexander Mokroß, Paweł Drozdowski, Christian Rathgeb and Christoph Busch

Abstract Large biometric systems, e.g. the Indian AADHAAR project, regularly perform millions of identification and/or de-duplication queries every day, thus yielding an immense computational workload. Dealing with this challenge by merely upscaling the hardware resources is often insufficient, as it quickly reaches limits in terms of purchase and operational costs. Therefore, it is additionally important for the underlying systems software to implement lookup strategies with efficient algorithms and data structures. Due to certain properties of biometric data (i.e. fuzziness), the typical workload reduction methods, such as traditional indexing, are unsuitable; consequently, new and specifically tailored approaches must be developed for biometric systems. While this is a somewhat mature research field for several biometric characteristics (e.g. fingerprint and iris), much fewer works exist for vascular characteristics. In this chapter, a survey of the current state of the art in vascular identification is presented, followed by introducing a vein indexing method based on proven concepts adapted from other biometric characteristics (specifically spectral minutiae representation and Bloom filter-based indexing). Subsequently, a benchmark in an open-set identification scenario is performed and evaluated. The discussion focuses on biometric performance, computational workload, and facilitating parallel, SIMD and GPU computation.

B.-A. Mokroß (✉)
IDNOMIZE GmbH, Dietzenbach, Germany
e-mail: benedikt-alexander.mokross@idnomize.de

P. Drozdowski · C. Rathgeb · C. Busch
da/sec – Biometrics and Internet Security Research Group,
Hochschule Darmstadt, Germany
e-mail: pawel.drozdowski@h-da.de

C. Rathgeb
e-mail: christian.rathgeb@h-da.de

C. Busch
e-mail: christoph.busch@h-da.de

P. Drozdowski
Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway

A. Uhl et al. (eds.), *Handbook of Vascular Biometrics*, Advances in Computer Vision and Pattern Recognition, https://doi.org/10.1007/978-3-030-27731-4_9

Publication reference: MOKROSS, B.-A., DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. *Efficient Identification in Large-Scale Vein Recognition Systems Using Spectral Minutiae Representations*. Springer, 2020, ch. 9.
DOI: `https://doi.org/10.1007/978-3-030-27731-4_9`
Thesis chapter: This research article is not included in this thesis, as it has been decided to only include first-authorship research articles.
Addressed research question(s): RQ1
Background: Vascular (*e.g.* palm and finger vein) biometric characteristics are of interest for practical biometric applications due to their high discriminative power and the relative difficulty of conducting successful presentation attacks against them. However, many of the current state-of-the-art vascular recognition algorithms (*e.g.* vein skeleton and minutiae based) are computationally intensive, thus necessitating research into methods of computational workload reduction.
Contribution: This research article is a chapter in the handbook of vascular biometrics [20]. The contribution is a proposal and evaluation of several methods for indexing of vascular data, some of which are based on the concepts from other research articles contained in this thesis ([7] and [5]).

3.2 Bibliography

[1] DROZDOWSKI, P. Efficient biometric identification in large-scale iris databases. MSc thesis, Technical University of Denmark, June 2016.

[2] DROZDOWSKI, P., BUCHMANN, N., RATHGEB, C., MARGRAF, M., AND BUSCH, C. On the application of homomorphic encryption to face identification. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2019), IEEE, pp. 173–180.

[3] DROZDOWSKI, P., FISCHER, D., RATHGEB, C., SCHIEL, C., AND BUSCH, C. Database binning and retrieval in multi-fingerprint identification systems. In *International Workshop on Information Forensics and Security (WIFS)* (December 2018), IEEE, pp. 1–7.

[4] DROZDOWSKI, P., GARG, S., RATHGEB, C., GOMEZ-BARRERO, M., CHANG, D., AND BUSCH, C. Privacy-preserving indexing of IrisCodes with cancelable Bloom filter-based search structures. In *European Signal Processing Conference (EUSIPCO)* (September 2018), IEEE, pp. 2360–2364.

[5] DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Multi-iris indexing and retrieval: Fusion strategies for Bloom filter-based search structures. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 46–53.

[6] DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. SIC-Gen: A synthetic Iris-Code generator. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2017), IEEE, pp. 61–69.

[7] DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Bloom filter-based search structures for indexing and retrieving Iris-Codes. *IET Biometrics 7*, 3 (May 2018), 260–268.

[8] DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Computational workload in biometric identification systems: An overview. *IET Biometrics 8*, 6 (November 2019), 351–368.

[9] DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Turning a vulnerability into an asset: Accelerating facial identification with morphing. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (May 2019), IEEE, pp. 2582–2586.

[10] DROZDOWSKI, P., RATHGEB, C., HOFBAUER, H., WAGNER, J., UHL, A., AND BUSCH, C. Towards pre-alignment of near-infrared iris images. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 359–366.

[11] DROZDOWSKI, P., RATHGEB, C., MOKROSS, B.-A., AND BUSCH, C. Multi-biometric identification with cascading database filtering. *Transactions on Biometrics, Behavior, and Identity Science (TBIOM)* (March 2020), 1–14.

[12] DROZDOWSKI, P., STRUCK, F., RATHGEB, C., AND BUSCH, C. Benchmarking binarisation schemes for deep face templates. In *International Conference on Image Processing (ICIP)* (October 2018), IEEE, pp. 191–195.

[13] DROZDOWSKI, P., STRUCK, F., RATHGEB, C., AND BUSCH, C. Detection of glasses in near-infrared ocular images. In *International Conference on Biometrics (ICB)* (February 2018), IEEE, pp. 202–208.

[14] DROZDOWSKI, P., WIEGAND, N., RATHGEB, C., AND BUSCH, C. Score fusion strategies in single-iris dual-probe recognition systems. In *International Conference on Biometric Engineering and Applications (ICBEA)* (May 2018), ACM, pp. 13–17.

[15] EUROPEAN PARLIAMENT. Regulation (EU) 2016/679. *Official Journal of the European Union L119* (April 2016), 1–88.

[16] ISO/IEC JTC1 SC27 IT SECURITY TECHNIQUES. *ISO/IEC 24745:2011. Information technology – Security techniques – Biometric information protection*. International Organization for Standardization and International Electrotechnical Committee, June 2011.

[17] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 19795-1:2006. Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*. International Organization for Standardization and International Electrotechnical Committee, April 2006.

[18] MOKROSS, B.-A., DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. *Efficient Identification in Large-Scale Vein Recognition Systems Using Spectral Minutiae Representations*. Springer, 2020, ch. 9.

[19] OSORIO-ROIG, D., DROZDOWSKI, P., RATHGEB, C., MORALES-GONZÁLEZ, A., GAREA-LLANO, E., AND BUSCH, C. Iris recognition in visible wavelength: Impact and automated detection of glasses. In *International Conference on Signal-Image Technology Internet-Based Systems (SITIS)* (November 2018), IEEE, pp. 542–546.

[20] UHL, A., MARCEL, S., BUSCH, C., AND VELDHUIS, R. N. J. *Handbook of Vascular Biometrics*. Springer, 2020.

Part II

Related Work

Chapter 4

Computational Workload in Biometric Identification Systems: An Overview

Abstract

Computational workload is one of the key challenges in biometric identification systems. The naïve retrieval method based on an exhaustive search becomes impractical with the growth of the number of the enrolled data subjects. Consequently, in recent years, many methods with the aim of reducing or optimising the computational workload, and thereby speeding-up the identification transactions, in biometric identification systems have been developed. In this article, a taxonomy for conceptual categorisation of such methods is presented, followed by a comprehensive survey of the relevant academic publications, including computational workload reduction and software/hardware-based acceleration. Lastly, the pertinent technical considerations and trade-offs of the surveyed methods are discussed, along with an industry perspective, and open issues/challenges in the field.

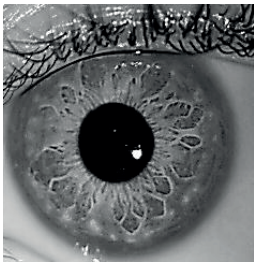
Addressed research question(s): RQ1, RQ2
Reference: DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Computational workload in biometric identification systems: An overview. *IET Biometrics 8*, 6 (November 2019), 351–368.

4.1 Introduction

The interest in biometric technologies has been steadily growing in recent years, as evidenced by various market value studies [10, 122, 182] and numbers of scientific publications in the area. Many states have utilised biometric technologies for purposes such as forensic investigations and law enforcement, border crossing entry-exit tracking, national citizen inventory (ID systems), and voter registration. By far the largest biometric deployment to date is the Indian Aadhaar national ID system, which, at the time of this writing, accommodates 1.3 billion enrolled subjects – almost the entire Indian population.

(a) Fingerprint

(b) Face



(c) Iris

Figure 4.1: Example images of some biometric characteristics commonly used in large-scale biometric identification systems (taken from the MCYT, FRGC, and IITD databases)

Table 4.1 gives an overview of this and several other examples of operational and planned large-scale biometric systems. The table is not exhaustive; instead, it seeks to highlight the diversity of the used biometric characteristics, the system purposes, and the geographical locations of some of the largest biometric systems around the world. In figure 4.1, example images of biometric characteristics most commonly used in large-scale biometric identification systems are shown.

Biometric systems can operate in a broad variety of ways. Two such ways (as defined in the ISO/IEC international standards [86, 87]) are:

Biometric verification Referring to the "process of confirming a biometric claim through biometric comparison".

Biometric identification Referring to the "process of searching against a biometric enrolment database to find and return the biometric reference identifier(s) attributable to a single individual". Two main scenarios can be distinguished in this case: **closed-set** identification, for which all potential users are enrolled in the system, and **open-set** identification, for which some potential users are not enrolled in the system.

Naturally, the second case (*i.e.* open-set identification, as well as the duplicate enrolment check) is the most interesting and challenging from the practical point of view for the aforementioned real-world applications. Unfortunately, in the worst case, an exhaustive search (*i.e.* comparing a probe against all the enrolled subjects) is required in order to reach a decision. This naïve approach quickly runs into two non-trivial problems:

Table 4.1: Examples of currently operational and planned large-scale biometric identification systems around the world

Status	System	Location	Characteristic(s)	Subjects	Purpose
Operational	Aadhaar [184, 185]	India	Fingerprint and iris (operational), face (potential future use)	1.3 billion	National ID
	EURODAC [49, 51]	EU	Fingerprint	7 million	Border-control
	IDENT/US-VISIT [63]	USA	Fingerprint (operational), face and iris (pilots ongoing)	200 million	Entry-exit
	CODIS [53]	USA	DNA	17.5 million	Law enforcement
	CENI [25]	DR Congo	Fingerprint	46 million	Voter registration
Planned	HART [37, 140]	USA	Fingerprint, Face, Iris	(expected) up to 500 million	Entry-exit
	EES [47, 50]	EU	Fingerprint, Face	(expected) up to 200 million	Entry-exit

Computational costs As the number of enrolled subjects increases, the system response times become gradually slower, thus requiring optimisations and/or investment into larger hardware architectures.

False positives costs The probability of at least one false positive (P_N) occurring in a identification scenario is: $P_N = 1 - (1 - P_1)^N$, where N is the number of enrolled subjects and P_1 the false positive probability of a one-to-one template comparison. This relationship is very demanding – even for systems which perform extremely well in verification mode (*i.e.* have low P_1), the value of P_N very quickly becomes unacceptably high, as the number of enrolled subjects N increases (see [32]). Note, that this equation ignores other system errors, *e.g.* the failure-to-acquire rate and assumes that at a given threshold all subjects have the same false-match-rate, which likely is not the case. Nonetheless, it is a useful approximation for illustrating this challenge of biometric identification systems.

In a recent interview [83], Daugman, the pioneer of iris recognition (see [34]), has stated that performing accurate and efficient biometric identification (*i.e.* not by an exhaustive search) is one of the important, unsolved issues in the biometrics field in general. Substantial research effort has been devoted to development of workload reduction methods, which seek to alleviate the aforementioned issues (especially the computational cost, since the biometric performance can also be improved through other means, such as increasing data quality and information fusion). Since the overall computational costs in a biometric identification scenario are dominated by performing the biometric comparisons, most approaches are aimed at that step in the system pipeline. Specialised data representations and search algorithms are utilised to reduce the computational effort required for a single template comparison, and/or to reduce the overall number of required template comparisons. However, biometric data exhibits certain properties, which present challenges or outright invalidate many traditional approaches aimed at retrieval speed improvement:

Ordering Biometric data has no inherent logical ordering (as opposed to, for example, text data, which can be indexed *e.g.* alphabetically).

Within-subject variability The samples acquired from the same subject (even within short time intervals) are almost never exactly identical (*i.e.* they are fuzzy). Some variations are nearly inevitable due to numerous noise sources in the acquisition process (*e.g.* distance and angle from the sensor, environmental conditions, occlusions *etc.*).

Dimensionality The biometric feature vectors are typically high dimensional; many search and indexing methods perform poorly in such spaces [68].

Consequently, computational workload reduction methods tailored specifically to the particular properties of biometric data have been developed in recent years. Such methods will be surveyed in the following sections. For a general overview of search structures and algorithms used for fast similarity searches across various disciplines, the reader is referred to *e.g.* [1, 21, 79, 148, 149, 150, 189, 205]. The reader is expected to possess certain background knowledge on biometric recognition systems in general and the typical algorithms used in their signal processing pipelines. For quick primers, the reader is referred to the encyclopedia of biometrics [111], as well as the renowned handbook series: [90] for biometrics in general and [14, 110, 119, 183] specifically for fingerprint, face, iris, and vascular characteristics, respectively.

While previously there have been surveys on biometric workload reduction methods (*e.g.* [145] for iris and [168] for fingerprint), they tend to concentrate on particular methods and/or biometric characteristics, rather than the overall spectrum of available research. Although the emphasis of this article is on the *academic* research, a discussion from the industry perspective and the interplay between academia and industry are included. The main contributions of this article are thus as follows:

Taxonomy which conceptually categorises the computational workload reduction methods in biometric identification.

Comprehensive survey of the methods reported in the scientific literature. It is organised by the relevant concepts, rather than by biometric characteristics. Instead of concentrating on one biometric characteristic only, the (arguably) most popular ones (in terms of actual use in industry and scientific research interest) are surveyed.

Discussion of relevant technical considerations and trade-offs, along with an industry perspective, and open issues/challenges pertaining to this research field.

The remainder of this article is organised as follows: section 4.2 gives an overview of relevant background information; in particular it introduces and defines key concepts used throughout the article, as well as outlines the current methodologies for results reporting and issues associated therewith. Section 4.3 contains a comprehensive survey of the existing computational workload reduction approaches reported in the scientific literature, conceptually organised within the framework of the proposed taxonomy. Section 4.4 discusses the topic from the purely academic, as well as industrial perspective, and outlines open issues/challenges. A summary and concluding remarks are given in section 4.5.

4.2 Background

This section gives an overview of relevant background information. Subsection 4.2.1 contains a list and short descriptions of the pertinent concepts and nomenclature, whereas in subsection 4.2.2, the dilemma associated with biometric result reporting and benchmarking is outlined.

4.2.1 Concepts and Nomenclature

Throughout this article, the nomenclature from the biometric vocabulary [87] and biometric performance testing and reporting [86] ISO/IEC international standards are used whenever applicable. However, as of this writing, many concepts relating to computational workload in biometric systems have not yet been put into standards by ISO/IEC (although efforts in this direction are ongoing, especially as some of the key standards are now/soon up for a revision). In this context, the present standards only defines the terms (quoted directly from the standards):

Pre-selection algorithm Referring to the "algorithm to reduce the number of templates that need to be matched in an identification search of the enrolment database".

Pre-selection error Referring to "the error that occurs when the corresponding enrolment template is not in the pre-selected subset of candidates when a sample from the same biometric characteristic on the same user is given".

Baseline performance Referring to the "performance of a biometric system in a reference evaluation environment".

Those terms are insufficient to capture the whole spectrum of issues and methods relevant in the aforementioned context. Therefore, several key concepts listed below are defined based on their actual use in the surveyed scientific literature:

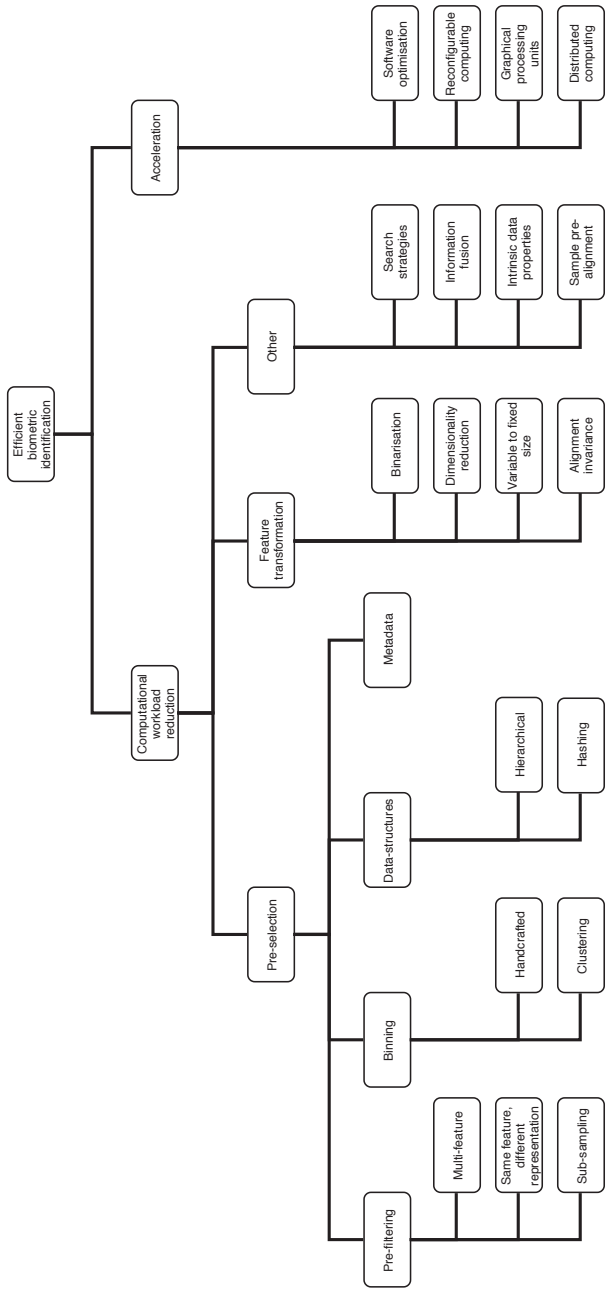


Figure 4.2: Taxonomy of methods used for the purpose of speeding-up biometric identification

Baseline system A state-of-the-art algorithm performing an exhaustive database search during a biometric identification transaction.

Computational workload The total computational effort of a single transaction (or a set of transactions) in a biometric (identification) system, including: number of intrinsic operations, execution time, memory and storage requirements.

Computational workload reduction The extent to which a method reduces the computational requirements (workload) of a biometric transaction (in a biometric identification system). See also subsection 4.2.2.

Pre-filtering (also "pre-selection", "cascading algorithms", "serial combination of algorithms", "guided search", "continuous classification") Computationally efficient, but somewhat inaccurate, comparator(s) are used to compare the biometric probe against the enrolled templates to produce successively smaller short-lists of candidate identities. In the end, the actual accurate, but computationally expensive, comparator is applied only to a fraction of the entries from the candidate short-list.

Binning (also "(discrete/exclusive) classification", "clustering") Splitting of the enrolment database into a number of subset (*i.e.* bins) based on coarse-level features. Those features can be tangible sample meta-data (*e.g.* sex, ethnicity, age) or based on intrinsic statistical properties of a template representation. During retrieval, the search space is reduced by only searching within the bins(s) most likely corresponding to the biometric probe.

Data-structures Organising the enrolment database to take advantage of efficient ordering principles (*e.g.* based on trees or fuzzy hashing), thus enabling searching in sub-linear/logarithmic time.

Indexing An often used umbrella term (in the biometric literature – *e.g.* a recent survey [168] and many individual publications) for all pre-selection methods (*i.e.* pre-filtering, binning, and data-structures). Simultaneously, it also has specific meaning outside the biometrics community. In order to avoid ambiguities, the term is not used in this article. Instead, the publications which present "indexing" methods, are assigned conceptually to one of the aforementioned categories.

Feature transformation The act of deriving additional features from a biometric template with the goal of attaining some desirable properties (*e.g.* smaller template size, ability to use a faster comparator, biometric sample alignment invariance).

Acceleration (hardware and/or software based) Utilisation of specialised hardware, hardware-software co-design, parallelism, distributed computing, and other methods in order to increase the efficiency/speed of a system compared to execution on standard CPUs.

In section 4.3, a taxonomy, which encompasses the abovementioned concepts and terms is presented.

4.2.2 Results Reporting

In subsequent subsections, tables which summarise the surveyed publications are presented. They include, among other matters, biometric performance and computational workload details. The metrics used for measuring biometric performance are well-defined and standardised [86]. The most relevant, in the context of this article, is the pre-selection error rate (complement of the hit rate; incidentally the hit rate is preferred in the vast majority of the works referenced later on in this article). While, in theory, this should make it possible to directly compare different methods, the reality is rather disappointing. First of all, some of the listed publications pre-date or ignore this standard, *i.e.* use a wide range of other metrics. Secondly, there inevitably exist other confounding issues and discrepancies in the experimental protocol, such as *e.g.* mode of operation (closed or open set), choice of dataset (hence, crucially, data quality), as well as size and partitioning thereof (*i.e.* training/testing partitions, number of biometric mated and non-mated comparison trials). Furthermore, at the time of this writing, metrics for measuring computational workload and its reduction are not standardised in any way whatsoever; many different metrics do appear in the scientific literature, for example:

- Penetration rate, which measures what fraction of the database is searched during a biometric identification transaction.
- Biometric template and/or model size, which determines how computationally expensive a single biometric comparison is.
- The fraction or percentage between the computational workload of a proposed system and a baseline system.
- Computational time, which measures the average execution time on some specific hardware configuration.

Additionally, it is often the case, that the publications present various parameter configurations with different trade-off spectra *etc.* for the proposed systems. It is therefore not always clear, which result to choose to present in a survey table, and how to select the single operational point which best

encompasses all the aspects of the proposed systems. As such, the choices in this survey were made as follows:

1. If the authors have provided a single representative result (operational point) in the publication text (*e.g.* in the abstract or summary) for the biometric performance and/or computational workload, those values are taken directly.

2. Otherwise, a single operational point is chosen in good faith from the presented plots and tables. If possible, this is done based on what is commonly reported elsewhere in the literature, *e.g.* equal-error-rate or other recognised metric. For the sake of consistency, if results for multiple ranks (*e.g.* CMC curve) are available, rank-1 results are preferred.

3. Computational time results are not reproduced, since they depend on a specific hardware configuration (which is most likely obsolete anyway). Where possible, the relative speed-up between the baseline and the proposed method is (calculated and) reported.

Due to the aforementioned issues, directly comparing the results from the surveyed publications is problematic, if not impossible. Furthermore, different systems require different considerations and trade-offs w.r.t. the biometric performance and the computational workload, as well as additional matters such as user convenience, software and hardware infrastructure, financial costs, and others. Consequently, the readers interested in benchmarking and/or utilising the surveyed methods are strongly recommended to investigate the relevant publications by themselves in order to obtain full-picture information of the proposed methods along with the biometric performance and computational workload trade-offs associated therewith.

4.2.3 Feature Extraction

Extracting sufficiently discriminative features is a critical prerequisite for any biometric system. This is especially a concern in biometric identification systems, due to the significantly increased risk of false positive errors (see section 4.1). Over time, various general purpose and biometric characteristic specific feature extraction methods have been proposed and used in this context. However, comprehensively surveying and comparing those would tremendously extend the already significant scope of this article. Therefore, the reader interested in a detailed treatment of this subject is referred to a recently published comprehensive survey of general purpose texture based feature extraction methods [116], as well as the handbook series: [90] for biometrics in general and [14, 110, 119, 183] specifically for fingerprint, face, iris, and vascular characteristics, respectively.

4.3 Computational Workload Reduction Approaches

In this section, the current state-of-the-art is presented. Firstly, the proposed taxonomy around which this section is structured is introduced and described below. Thereafter, a comprehensive survey of existing methods is given and put in the context of the taxonomy.

Figure 4.2 shows the proposed taxonomy under which the existing approaches to speeding-up the biometric identification can be categorised. Note, that in many cases the approaches can be combined into multi-level frameworks, *e.g*. a binning followed by tree-based hierarchical retrieval, implemented utilising hardware acceleration or pre-selection based on multiple levels of complementary features. Two main approaches to improving the computational efficiency of biometric identification can be distinguished: workload reduction (subsections 4.3.1 to 4.3.5) and acceleration (subsection 4.3.6). The latter does not reduce the computational workload *per se* – instead, it seeks to perform the same amount of computations in a more efficient manner (*e.g*. by utilising specialised hardware or optimising the software implementation). The goal of the former is to reduce the amount of computations necessary to perform a biometric identification transaction. For those approaches, two main categories can be distinguished: concentrating on reducing the penetration rate, the aim of the pre-selection approaches (subsections 4.3.1 to 4.3.3) is to narrow down the search space by taking advantage of auxiliary features, metadata, or search structures, which can be extracted or created from the samples. On the other hand, the goal of feature transformation approaches (subsection 4.3.4) is to reduce the computational cost of individual template comparisons, *e.g*. by reducing their dimensionality or utilising more computationally efficient comparators. The vast majority of the approaches can be assigned to one of those categories. The remaining few ones (subsection 4.3.5) are based *e.g*. on augmenting the search strategy of the retrieval algorithm or rely on certain intrinsic properties of specific biometric data.

This section is organised to facilitate selective reading: firstly, a very broad overview of the efficient biometric identification research areas has been given above by introducing and describing the proposed taxonomy. The following subsections' text outlines the relevant high-level concepts and ideas, while the tables contain more detailed information w.r.t. specific tools, algorithms, and datasets used, as well as the achieved results. Finally, the considerations and trade-offs associated with the different approach categories are discussed in subsection 4.4.1.

4.3.1 Pre-filtering

Figure 4.3 shows a conceptual overview of pre-filtering approaches, while table 4.2 summarises the surveyed methods.

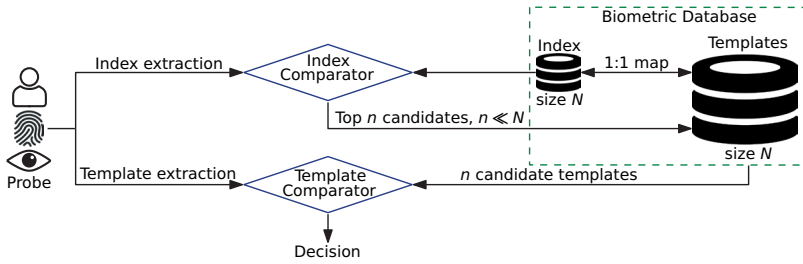


Figure 4.3: Conceptual view of pre-filtering approaches

4.3.1.1 Multi-Feature

The key idea behind the multi-feature approaches is the extraction of one or several auxiliary features, which in themselves do not have sufficient discriminative power for unique identification, but can nonetheless significantly reduce the search space (*i.e.* by acting as an index, which allows to determine a candidate short-list).

Auxiliary features such as orientation field, ridge density, local (minutiae-based) and global (*e.g.* fingerprint types, which have been in use for decades for the purposes of manual indexing of analog ten-fingerprint records with the Henry Classification System, see *e.g.* [77, 78], and subsection 4.3.2) can be extracted from fingerprint images; some of them also pertain to other characteristics, such as vascular and palmprint patterns. Several authors (*e.g.* [9, 15, 36, 54, 97, 109, 113, 191]) utilise such coarse features as an index in a pre-filtering step. In other cases, the methods proposed in the scientific literature do not rely on specific, biometric characteristic-dependent features as above; instead, to create an index, they utilise general-purpose algorithms, such as texture extractors (*e.g.* [23, 39]), principal component analysis (*e.g.* [128]), and, more recently, deep learning (*e.g.* [187, 188]). It should be noted, that the pre-filtering can happen in a cascading manner, over two (*e.g.* [72, 73, 102]) or multiple (*e.g.* [55, 153, 202]) levels, which successively produce smaller candidate lists, or through direct application of information fusion strategies to the extracted features (*e.g.* [139]). However, an in-depth analysis and evaluation concerning which of the methods (cascades or fusion) performs better has not yet been reported in the scientific literature.

4.3.1.2 Same feature, different representation

The key idea behind this category of approaches is transformation of the original feature representation into a more compact one, whereby the computational costs of comparisons are decreased (often at the cost of losing some discriminative power). The compact templates can then be used to re-

Table 4.2: Pre-filtering approaches

Taxonomy	Characteristic	Publication	Method	Database	Biometric Performance	Computational Workload
Multi-Feature	Fingerprint	Ratha *et al.* [153]	Metadata (only conceptual), fingerprint type, ridge density	NIST-9 subset	80% accuracy, 10% reject rate	12.5% search space
		De Boer *et al.* [36]	Directional field, FingerCode, and minutiae triplets	FVC2000	100% hit rate	18% penetration rate
		Bhanu [9]	Minutiae triplets, geometric features	NIST SD4	85% hit rate	10% penetration rate
		Feng *et al.* [54]	Minutiae points and types, local ridge structures	FVC2002	100% hit rate	22% penetration rate
		Li *et al.* [109]	Ridge structure, symmetrical filters	NIST DB4	98% hit rate	32.7% penetration rate
		Liang *et al.* [113]	Minutiae neighbourhoods, Delaunay triangulation	FVC2002, FVC2004	100% hit rate	18.1%, 20.9% penetration rate
		Wang *et al.* [191]	2D Fourier expansion coefficients	NIST SD 14	100% hit rate	10% penetration rate
		Feng *et al.* [55]	Fingerprint type, singular points, orientation field	NIST SD27 (search attempts), NIST SD4, SD14 and SD27 (background)	97.3% accuracy	39% penetration rate
		Cappelli [15]	Ridge-line orientations and frequencies	NIST SD4, SD14, FVC2000 (DB2, DB3), FVC2002 (DB1)	96.5%, 96.5%, 99%, 93.5%, 99% hit rate	10% penetration rate
		Paulino *et al.* [139]	Orientation field, ridge period, singular points, minutiae triplets, simplified MCC	NIST SD27 (search attempts), in-house (background)	90.3% hit rate	20% penetration rate
	Fingerprint, face	Gyaourova *et al.* [72, 73]	Index-codes from non-mated comparison trials, fusion	FERET, FRGC, WVU	100% hit rate	84% reduction
	Face	Mohanty *et al.* [128]	Affine approximation, PCA	FERET	—	20-fold reduction
		Chen *et al.* [23]	LBP, semantic codewords from metadata	LFW, Pubfig	18.6%, 21.0% MAP	—
		Wang *et al.* [187, 188]	Deep features and COTS	LFW, IJB-A	0.25 MAP at 1% FAR, 0.175 MAP at 1% FAR	30-fold time reduction
	Iris	Konrad *et al.* [102]	Rotationally invariant representation	CASIA-V1, CASIA-V3 Interval, MMU	92% IR, 0% FAR; 89% IR, 0.85% FAR; 79% IR, 0.85% FAR	70-80% time reduction
		Gadde *et al.* [57]	BWT	CASIA-V3-Interval	99.83% hit rate	17.23 % penetration rate
		Dey *et al.* [39]	Gabor energy features, multi-sample enrolment	Bath, CASIA-V3-Interval, CASIA-V4-Thousand, MMU2, WVU	98.2%, 91.1%, 90.7%, 85.2%, 96% hit rate	11.3%, 14.5%, 16.3%, 13.5%, 10.3% penetration rate
	Fingervein	Kavati *et al.* [97]	Delaunay triangulation	NTU NIR, NTU FIR	100% hit rate	17.99%, 11.75% penetration rate
	Palmprint	You *et al.* [202]	Global geometry, global texture energy, fuzzy "interest" line, local texture	In-house	6.13% FRR at 11.77% FAR	2-fold speed-up
Same feature, different representations	Face	Wu *et al.* [197]	Binary template pre-screening	In-house	Better than baseline	∼10-fold reduction
	Iris	Gentile *et al.* [64]	Short-length IrisCodes	MMU	7% pre-selection error	12-fold reduction
	Fingervein	Tang *et al.* [179]	Binary vein encoding	PKU	98.4% hit rate	250-fold time reduction
	Voice	Billeb *et al.* [11]	Binary template pre-screening	Unknown, text-independent	same or better than baseline	95% speed-up
	Ear	Pflug *et al.* [143]	Binary template pre-screening	PolyU, UND-J2	100% hit rate	30% penetration rate
Sub-sampling	Fingerprint	Iqbal *et al.* [84]	Incremental matching	FVC2002	99% hit rate	26% penetration rate
	Fingerprint, palmprint	Chen *et al.* [24]	Incremental matching	THU, NIST SD 4, in-house	90.4% IR, 85.3% IR, 75.3% IR	∼50% reduction
	Face	Yi *et al.* [200]	Incremental matching	FERET, in-house	Same as baseline	7.5-fold speed-up
	Iris	Hao *et al.* [75]	BGS, incremental matching	UAE	0% FAR, 0.64% FRR	0.006% penetration rate
		Ross *et al.* [164]	Partial matching	UPOL	0.62% EER	10% of baseline
		Hämmerle-Uhl *et al.* [74]	Partial matching	CASIA-V3 Interval	Same as baseline	1 order of magnitude reduction
		Rathgeb *et al.* [160]	Incremental matching	CASIA-V3 Interval	Same as baseline	95% fewer bit comparisons
	Fingervein	Surbiryala *et al.* [178]	Partial matching	Combined 7 fingervein DBs	8.05% pre-selection error	∼3-fold reduction

duce the search space (*i.e.* by acting as an index, which allows to determine a candidate short-list).

Conceptually similar approaches, where binarised (see also subsection 4.3.4.1) and/or shortened feature vectors are used as an index in the pre-filtering step, have been proposed *e.g.* in [11, 64, 143, 179, 197] for iris, face, fingervein, voice, and ear, respectively.

The difference between the key idea in this and previous subsection is subtle – here, the same feature is used to create the index template (*e.g.* through binarisation), whereas in the multi-feature concept, additional features are extracted from the sample (*e.g.* through texture or keypoint descriptors or high level geometric features).

4.3.1.3 Sub-sampling

The key idea behind sub-sampling is to utilise partial information from the original feature vectors once or in an incremental manner to facilitate search space reduction via accurate early rejection of unlikely candidates. In other words, parts of the original feature vector itself act as an index in this case. This can be done trivially by deterministically or randomly selecting the partial information or, in more advanced approaches, by reorganising the feature vectors based on reliability and discriminative power (see *e.g.* [80]), as well as utilising other heuristics. In the literature, numerous conceptually similar approaches have been presented *e.g.* in [24, 74, 84, 160, 164, 178, 200] for various biometric characteristics, including fingerprints, face, iris, and fingervein. In all the aforementioned publications, the computational workload is shown to be substantially reduced without causing degradation of the biometric performance. In [75] a more sophisticated approach, which relies on creating an auxiliary search guiding structure and an early search termination strategy, was presented with impressive results, albeit on proprietary data.

4.3.2 Binning

Figure 4.4 shows a conceptual overview of binning approaches, while table 4.3 summarises the surveyed methods.

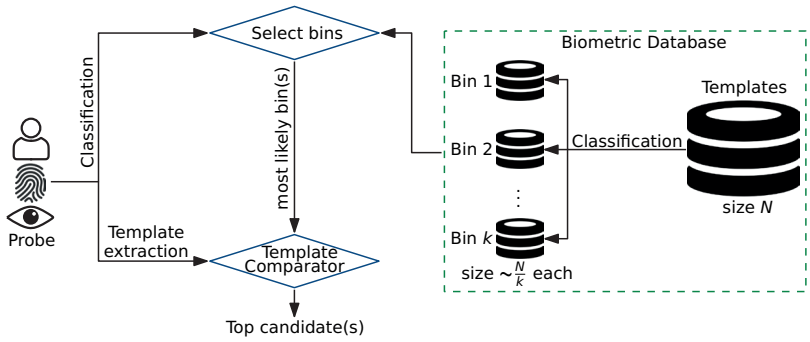


Figure 4.4: Conceptual view of binning approaches

Table 4.3: Binning approaches

Taxonomy	Characteristic	Publication	Method	Database	Biometric Performance	Computational Workload
Handcrafted	Fingerprint	Zheng *et al.* [207]	Classification, coarse-level matching, class-jumping, SURF	NIST DB 4	100% hit rate	15% penetration rate
		Drozdowski *et al.* [40]	Fingerprint types, multi-instance, variable search order	NIST DB 9, in-house Bundeskriminalamt (BKA) DB	Same as an exhaustive search	5-15% of an exhaustive search
	Face	Park *et al.* [138]	Facial marks, scars, and tattoos	PCSO (police mugshots)	7.1%, 0.5% rank-1 accuracy loss	7%,20% speed-up
	Iris	Yu *et al.* [203]	Box-counting, fractal dimensions	In-house	1.72% pre-selection error	Less than 40% time
		Puhan *et al.* [146]	Colour information in YCbCr space, set intersection	UBIRIS	97% hit rate	25% penetration rate
		Zhao [206]	Average RGB colour components, set union	UBIRIS	92.35% hit rate	28.28% penetration rate
	Palmprint	Palla *et al.* [136]	Geometric features, codebook vectors, Voronoi regions	In-house	100% hit rate rate	30% penetration rate
	Palmvein	Zhou *et al.* [208]	Principal orientation features	PolyU, CASIA, in-house	96.67%, 96.00%, 97.71% retrieval accuracy	14.29%, 14.50%, 14.28% penetration rate
Clustering	Fingerprint	Germain *et al.* [66]	Minutiae triplets, ridge skeleton, Flash algorithm	In-house	3.5% FNMR at 0.01% FMR	—
		Ross *et al.* [163]	Delaunay triangulation, geometric and ridge features, k-means clustering	FVC2002, FVC2004	100% hit rate	∼50% av. penetration rate
		Liu *et al.* [117]	Orientation field, average ridge distance, k-means clustering	NIST-DB 4	95.8% hit rate	20% penetration rate
		Biswas *et al.* [12]	Curvature, minutiae geometry, k-means clustering	IBM proprietary	90% rank-1 accuracy	5-fold decrease
		Iloanusi *et al.* [81, 82]	Minutiae quadruplets, k-means clustering	FVC2002, FVC2004	100% av. hit rate	∼12% av. penetration rate
	Face	Perronnin *et al.* [141]	Expectation maximisation clustering, anchor modelling	FERET	∼95% IR	6-7-fold reduction
		Chaari *et al.* [20]	Eigenfaces and Fisherfaces, k-means clustering	XM2VTS	87.5% IR at rank-1	40% penetration rate
		Klare *et al.* [99]	Spectral clustering, k-means and k-medoids clustering	LFW, PCSO	85% IR	50% reduction
	Iris	Mukherjee *et al.* [130]	Iris-Code, PCA, k-means clustering	CASIA-V3-Interval	80% hit rate	8% penetration rate
		Ross *et al.* [164]	Statistical texture features, Principal Direction Divisive Partitioning	UPOL	100% CCR	3-5-fold reduction
		Sun *et al.* [177]	Ordinal measures, hierarchical visual codebook, k-means clustering, SVM	CASIA Thousand	∼2% EER	less than 30%
		Nalla *et al.* [132]	Online dictionary learning, k-means clustering	UPOL	100% CCR	3-4-fold reduction
	Fingervein	Surbiryala *et al.* [178]	Maximum curvature, k-means clustering	Combined 7 fingervein DBs	97.47% hit rate	86.43% penetration rate
		Raghavendra *et al.* [151]	Self Organizing Map neural network, k-means or k-medoids clustering, multi-cluster search	Combined 7 fingervein DBs	92.42%; 99.02% hit rate	42.48%; 52.88% penetration rate
	Palmprint and signature	Mhatre *et al.* [126]	K-means clustering	Unknown	0% FRR, — FAR	5% penetration rate
	Ear	Pflug *et al.* [142]	K-means clustering, texture descriptors	UND-J2, AMI, IITK	3.11% pre-selection error rate	31.7% penetration rate

4.3.2.1 Handcrafted

Depending on the observed biometric characteristic, there exist classification approaches designed to reliably extract human understandable attributes from a biometric sample, *e.g.* sex or ethnicity for face, or fingerprint types. Such attributes are called "soft biometrics" (see *e.g.* [30] for a comprehensive survey).

Based on the global pattern formed by the ridge lines, fingerprints can

be classified into a number of classes/types initially proposed by Galton [60] and Henry [78] (currently typically 4 or 5, *i.e.* whorls, right and left loops, and (tented) arches, sometimes extended with additional sub-types). Over time, numerous approaches to automated fingerprint type classification have been proposed (see *e.g.* [58, 59] for a comprehensive survey). The classification accuracy on data of reasonable quality is near-optimal; however, it tends to vary somewhat across the different fingerprint types. Binning based on fingerprint classes has been evaluated for single fingerprints by *e.g.* [153, 207] and for multi-instance data in [40, 194]. Attributes extracted from iris data can also be used in this manner. Conceptually similar systems are presented in [203], [136], and [208], where binning based on biometric characteristic-specific geometric/texture features is proposed for irs, palmprint, and palmvein data, respectively. In [147, 172, 180], it has been demonstrated, that ethnicity and gender information can be extracted from iris images. When reliably extracted, such features could be used for simple database binning. Binning based on iris colour has been performed *e.g.* in [146, 206]. Although the vast majority of the human population has brown eyes, for certain population groups, the eye colour can be used as a somewhat distinguishing soft biometric trait. Currently, all practical iris recognition systems operate within the near-infrared (NIR) light spectrum. In recent years, significant advances in the visible-wavelength (VW) iris recognition have been made, hence potentially making it an emerging technology. See *e.g.* [31] for an investigation of the reliability of the iris colour as a soft biometric trait. Facial region is a rich source of potential soft biometric attributes. In addition to simple approaches based on sex, age, or ethnicity classification, binning based on marks, scars, and tattoos has been proposed [138].

While the aforementioned attributes are not discriminative enough to be directly used in biometric identification, they allow for a relatively straightforward binning of biometric databases according to a predefined number of classes. In other words, the potential search space for a given biometric probe can be narrowed down to one (or a few) bin(s), thereby reducing the penetration rate, and hence the computational workload.

4.3.2.2 Clustering

Cluster analysis or clustering refers to the unsupervised or semi-supervised classification of patterns (*i.e.* feature vectors, data items, or observations) into groups (referred to as clusters), wherein the items are, in some sense, similar to each other. With applications across many different disciplines, k-means clustering is currently one of the most popular and effective algorithms used in data mining [196].

Likewise, in the surveyed literature, k-means clustering (and its various extensions/derivatives) is by far the most popular method, used in *e.g.*

[12, 20, 81, 82, 99, 117, 126, 130, 132, 142, 151, 163, 177, 178]. Other methods include *e.g.* multimap clustering [66], expectation maximisation clustering [141], and principal direction divisive partitioning [164]. Comparing the various clustering methods is out of scope for this article. For more details regarding this field of research, the reader is referred to surveys, *e.g.* [2, 91]. Generally, the approaches referenced in this subsection extract certain biometric characteristic-specific features (*e.g.* orientation field or Delaunay triangles for fingerprint, or general-purpose texture descriptors for iris) to facilitate the clustering or apply it directly with the feature vectors (*e.g.* minutiae points) themselves. As a result, the search space is separated into distinct bins, whereby during biometric identification, candidates only from the most likely one(s) are retrieved. Hence, the penetration rate (and thereby the computational workload) is significantly reduced.

4.3.3 Data-Structures

Figure 4.5 shows a conceptual overview of hierarchical retrieval approaches, while table 4.4 summarises the surveyed methods. A multitude of methods, algorithms, and data-structures (whose detailed descriptions are out of scope for this article) has been used in the surveyed approaches. For a general introduction to on approximate searching, relevant concepts, and most commonly used data-structures, the reader is referred to existing surveys, *e.g.* [1, 21, 79] for theoretical, practical, and easily digestible perspectives, respectively.

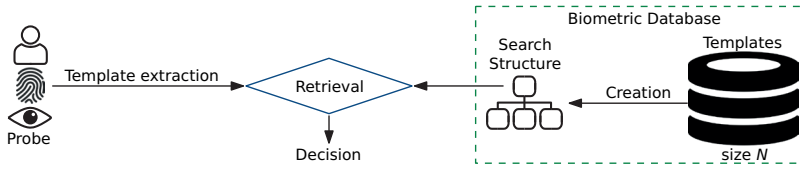


Figure 4.5: Conceptual view of data-structures approaches

4.3.3.1 Hierarchical

Approaches in this category are most often tree-based, most prominently utilising k-d trees (*e.g.* [6, 38, 94, 95]), b or b+ trees (*e.g.* [71, 98, 124, 125]), other tree-like search structures (*e.g.* [120, 130, 144, 145, 154, 190]), and forests thereof (*e.g.* [28, 29, 42, 43]). The differences between the various types of used trees (some of which are each other's generalisations) are out of scope for this article; instead, the reader is referred to *e.g.* [26, 100]. The key idea is to create a search structure, which repeatedly partitions the data (*i.e.* the search space – the enrolment database) into successively smaller subsets.

Table 4.4: Data-structures approaches

Taxonomy	Characteristic	Publication	Method	Database	Biometric Performance	Computational Workload
Hierarchical	Fingerprint	Mansukhani *et al.* [120]	Local minutiae neighbourhoods, unbalanced tree	FVC2002, FVC2004	81% accuracy	Almost constant w.r.t. enrolment DB size
	Face	Dewangan *et al.* [38]	SURF, kd-tree	FERET, FRGC, CalTech	95.57%, 97.00%, 92.31% hit rate	7.90%, 12.55% and 23.72% penetration rate
	Iris	Mukherjee *et al.* [130]	Blockwise texture SPLDH, tree-like structure	CASIA-V3-Interval	84% hit rate	30% penetration rate
		Mehrotra *et al.* [124]	DCT, subband coding, energy histogram, b-tree	CASIA Interval, BATH, IITK	95% hit rate	25% penetration rate
		Khalaf *et al.* [98]	DCT, DWT, SVD, subband coding, energy histogram, b-tree	CASIA Interval, BATH, IITK	~97.5%,~97.5%, 95% hit-rate	20% penetration rate
		Jayaraman *et al.* [95]	Iris colour, SURF, kd-tree	UBIRISv2, UPOL	98.7%, 98.5% av. hit rate	7.5%, 1.5% av. penetration rate
		Barbu *et al.* [6]	HOG, kd-tree	UPOL	85% precision and recall	—
		Rathgeb *et al.* [154]	Bloom filters, binary search trees	IITD	same or better than baseline	6% penetration rate
		Drozdowski *et al.* [42]	Bloom filters, binary search trees, multi-instance fusion	Combination of 4 iris datasets	99.41% TPIR at 0.01% FPIR	<1% of baseline
		Drozdowski *et al.* [43]	Bloom filters, binary search trees	Combination of 4 iris datasets	98% TPIR at 0.1% FPIR	10% of baseline
		Damer *et al.* [29]	LSH-forest	ISYN1	99.85% single instance, 99.99% multi instance hit rate	0.4% penetration rate
		Damer *et al.* [28]	General Borda count, LSH-forest, multi-instance	ISYN1	>99.5% hit rate	0.1% penetration rate
		Proença *et al.* [144, 145]	Multi-resolution decomposition, n-ary trees	CASIA-V4-Thousand, UBIRISv2	95% hit rate	20%, 80% penetration rate
	Iris, Signature, Face, Ear	Jayaraman *et al.* [94]	Dimensionality reduction, feature-level fusion, kd-tree	IITK	97.33% hit-rate at 0.66% FRR	—
	Fingervein	Wang *et al.* [190]	Local textons, vocabulary tree	PolyU, SDUMLA, MMCBNU, FV-USM	~99% hit rate at rank-5	Up to 5-fold speedup
	Palmprint	Mhatre *et al.* [125]	Geometric features, spatial hashing, b-tree	unknown	0% FRR, — FAR	8.86% penetration rate
	Ear	Gupta *et al.* [71]	Division into quadrants, wavelet decomposition, b-tree	IITK	95.8% accuracy	34% penetration rate
Hashing	Fingerprint	Shuai *et al.* [169]	SIFT, LSH	FVC2000, FVC2002	98%, 96% hit rate	10% penetration rate
		He *et al.* [76]	SIFT, SURF, DAISY, LSH	FVC2000, FVC2002	99%, 90% hit rate	10% penetration rate
		Capelli *et al.* [18]	MCC, LSH, voting	NIST SD4, 14, FVC2000, 2002	95% hit rate	<10% penetration rate
		Yuan *et al.* [204]	Minutiae triplets, two-level hashtable	FVC2000, 2002, 2004	100% hit rate	22%, 9.9%, 11.7% av. penetration rate
		Wang *et al.* [193]	MCC, Markov random field theory, geometric dictionary	FVC2002 DB1	100% hit rate	10% penetration rate
		Li *et al.* [107]	MCC, binarisation, LSH	FVC2002, FVC2004, FVC2006	7.5%, 22.5%, 4% preselection error rate	10%, 10%, 5% penetration rate
	Face	Kaushik *et al.* [96]	SURF, geometric hashing, voting	FERET	100% hit rate	4% penetration rate
	Iris	Mehrotra *et al.* [123]	SIFT, geometric hashing, voting	BATH, CASIA-V3-Interval, IITK, UBIRIS	98.29%, 98.55%, 99.61%, 97.57% EER	Order of magnitude faster than baseline
		Rathgeb *et al.* [158]	Iris texture hashes, Karnaugh map	CASIA-V3 Interval	90% accuracy	3% of baseline
		Jayaraman *et al.* [93]	Iris-Code, LSH, voting	CASIA-V3-Interval	94.07% hit rate	10.63% penetration rate
		Panda *et al.* [137]	SIFT, geometric hashing	CASIA-V3-Interval, UBIRISv1	98.25%, 97.62% accuracy	~75% of baseline time
	Palmprint	Badrinath *et al.* [5]	SURF, geometric hashing	IITK, CASIA, PolyU	100% hit rate	22.5%, 22.8%, 31.9% penetration rate

For this partitioning, the highly discriminative (and high-dimensional) feature vectors themselves and/or the more coarse auxiliary features can be used. By doing so, sub-linear or even logarithmic lookup complexity can be achieved, thereby substantially reducing the computational workload of biometric identification.

4.3.3.2 Hashing

Hashing makes it possible to map the highly-dimensional biometric feature vectors into compact hashtables or similar data-structures, which facilitate

efficient retrieval. Since biometric data is inherently fuzzy (recall section 4.1), many traditional hashing approaches are not suitable. Nevertheless, there exist methods, which can deal with fuzzy data. One of such method is locality-sensitive hashing (LSH) [68], which refers to a family of functions, which can be used to map data points into buckets in such a way, that it is highly probable for data points which are close to each to be located in the same buckets; conversely, data points which are distant from each other, are likely located in different buckets. Several authors utilised LSH and variations/extensions thereof to facilitate efficient retrieval of (in most cases) fingerprint data [18, 76, 93, 107, 169, 204]. Geometric hashing [105], which was originally developed for object recognition (matching similar geometric shapes irrespective of translation, rotation, and scaling), has also been applied in the context of biometrics by coupling it with general-purpose keypoint detectors [5, 96, 123, 137].

Deeper descriptions of the various hashing algorithms and their extensions are out of scope for this article – the reader is referred to *e.g.* [189, 195]. Generally, by significantly reducing the dimensionality of the data and facilitating retrieval of a subset of candidate identities, general purpose fuzzy hashing methods adapted to the biometric data can be used to greatly reduce the computational workload. Aside from potential biometric performance degradation due to hashtable/bucket misses, the storage requirements of the system (especially in the case of geometric hashing) are typically increased.

4.3.4 Feature Transformation

This subsection surveys methods based on creating efficient representations of biometric templates, which reduce the computational cost of a single template comparison. This can typically be achieved through *e.g.* reducing the template dimensionality, creating fully or partially alignment invariant representations, or utilising more efficient template comparators (for instance, based on bit instead of floating-point operations). In other words, the goal is often to transform the original template (or create an unrelated alternative representation), so that it obtains certain desirable properties, while predominantly maintaining the discriminative power. Templates utilising such alternative or transformed representations can then be used on their own in an exhaustive search, or in more advanced approaches, *e.g.* act as a pre-selector (see subsection 4.3.1) in a multi-stage retrieval system. Table 4.5 summarises the surveyed methods.

4.3.4.1 Binarisation

Comparison of float-based feature vectors is relatively expensive computationally, due to use of comparators based on *e.g.* Euclidean or χ^2 distances.

Table 4.5: Feature transformation approaches

Taxonomy	Characteristic	Publication	Method	Database	Biometric Performance	Computational Workload
Binarisation	Fingerprint	Capelli *et al.* [17]	Binarised minutia cylinder-code	FVC2006	<1% average EER	At least an order of magnitude faster
	Face	Schlett *et al.* [166]	Multi-scale block LBP, binarisation	FERET, Extended-Yale-B	15% FNMR at 10% FMR	20-fold speed-up
		Drozdowski *et al.* [46]	Benchmark of various quantisation and encoding methods	FERET, FRGC	0.3% EER, 2.3% EER	An order of magnitude fewer CPU operations required
Dimensionality reduction	Iris	Gentile *et al.* [65]	Short-length Iris-Codes	MMU	79.4% FNR at 1% FPR	12-fold size reduction
		Rathgeb *et al.* [161]	Most discriminative bits, selective algorithm fusion	CASIA-V3-Interval	1.15% EER	∼50% fewer bits
Variable to fixed size	Fingerprint	Jain *et al.* [92]	FingerCode	NIST SD9, MSU_DBI	∼15% FRR at 1% FAR; ∼8% FRR at 1% FAR	—
		Xu *et al.* [198]	Spectral minutiae	MCYT	3.13% EER	—
		Yang *et al.* [199]	Tessellated invariant moment features	FVC2002	3.57% average EER	3-fold reduction
Alignment invariance	Iris	Rathgeb *et al.* [155]	Bloom filters	CASIA-V3 interval	1.5% EER	20% of baseline
		Damer *et al.* [27]	Translation-invariant transformation	SYN1	0.646%EER, 1.213% EER	6.56% of baseline, 2.45% of baseline

In many cases, such feature vectors can be quantised and encoded into binary strings, whereby utilisation of comparators based on *e.g.* Hamming distance is possible. Such comparators can take advantage of the more efficient bitwise operators, thereby reducing the computational workload. An illustrative example can be seen in [46] (and a simpler one in [166]), where various bit allocation schemes for float-based feature vectors generated by neural network-based systems are benchmarked. In [17], a new representation is extracted from minutiae points, which can be further binarised to accelerate the biometric template comparisons. Although some information is lost through the binarisation process, both publications show only negligible biometric performance loss in relation to their respective baselines, while achieving a significant speed-up. Finally, binarised feature vectors are an essential component in the context of many template protection schemes (see *e.g.* [114] for more details on this subject).

4.3.4.2 Dimensionality reduction

Templates produced through dimensionality reduction can be used directly as a replacement for the full-sized templates (*e.g.* through PCA). Additionally, they can serve as a first pre-filtering step in a two-stage system (see subsection 4.3.1 for examples). An illustrative example is[65] (and a similar approach in [161]), where the so-called "short-length Iris-Codes', which comprise the most discriminative parts of the normal Iris-Codes, are presented. The transformed templates are an order of magnitude smaller than the original ones, and exhibit somewhat impaired biometric performance when benchmarked against the original templates, thereby making them good candidates for a pre-filtering step.

4.3.4.3 Variable to fixed size

Comparisons of variable-size feature vectors are computationally demanding and often suffer from other domain-specific drawbacks. In biometrics, most prominently used variable-sized feature representation is that of fingerprint minutiae. The number of minutiae points can be inherently different between different data subjects and can further be augmented depending on the sample acquisition conditions (*i.e.* the so-called missing and spurious minutiae). In the literature, a number of alternative approaches to the traditional minutiae-based fingerprint comparison algorithm has been proposed by several authors [92, 198, 199]. All of those methods achieve biometric performance and computational workload results competitive with those of the traditional variable-size, minutiae-based algorithm.

4.3.4.4 Alignment Invariance

An important issue in biometrics, and especially fingerprint and iris recognition is the necessity of compensating for the relative sample misalignment caused by roll pose variations. This is typically done by considering multiple relative shifting positions of the Iris-Codes matrix and choosing the one with best comparison score, thereby increasing the computational cost of a single template comparison. In [155] and [27] feature transformations are presented, which ensure that sample misalignment (to a certain degree, reasonable from practical point of view) is intrinsically compensated for by the resulting feature vectors. Both approaches achieve substantial speed-up in an exhaustive search without significantly impairing the baseline biometric performance. Several other (not feature transformation based) approaches tackling the issue of iris alignment are also listed in subsection 4.3.5.

4.3.5 Other

This subsection presents computational workload reduction approaches which do not fit into the previous categories. Table 4.6 summarises the surveyed methods. A simple method of reducing the computational workload in an exhaustive search is performing an early exit strategy, *i.e.* finishing the search once first (not necessarily best) suitable candidate is found. This is sometimes referred to as "one-to-first" search. In [104] this search strategy is analysed extensively for iris recognition in order to assess potential degradation of biometric performance. It is discovered, that the biometric performance degradation is strongly dependent on the decision thresholds (accuracy target) and size of the enrolment database. For some parameters, the biometric performance of an exhaustive search can be maintained, while the computational workload is significantly reduced. In [16] several strategies were proposed, which reduce candidate lists (produced by other meth-

ods) through analysis of comparison scores. In [157] an approach to reduce the number of relative shifting positions of the Iris-Codes which need to be considered in a template comparison was presented. The method is based on an analysis of the intrinsic properties of the iris data and achieves a considerable speed-up without impairing the biometric performance. In [45] a pre-alignment of raw iris images is performed. The method is based on automatic detection of eye corners and several other points in raw iris images, and subsequently aligning the eye corners onto a horizontal line. Thus, at a later point, once features are extracted, fewer relative shifting positions need to be considered during template comparisons. The approach of [44] relies on morphing (signal-level fusion). The facial images from the enrolment database are morphed (in 2s, 4s, or 8s), whereby biometric information from multiple subjects is fused into one image. The morphed images are then utilised for pre-filtering (see subsection 4.3.1). In addition to being explicitly used in some computational workload reduction schemes surveyed in this article, information fusion is an important aspect in ensuring the scalability of biometric systems in terms of biometric performance. For a comprehensive survey of this topic, the reader is referred to *e.g.* [173].

Table 4.6: Other approaches

Taxonomy	Characteristic	Publication	Method	Database	Biometric Performance	Computational Workload
Search strategies	Iris	Kuehlkamp *et al.* [104]	1-to-first search	Notre-Dame	see paper	50-70% of baseline
	Fingerprint	Cappelli *et al.* [16]	Analysis of comparison scores, ruleset/criteria	FVC	1% average error rate	27% penetration rate (from indexing) reduced to 3.9%
Intrinsic data properties	Iris	Rathgeb *et al.* [157]	Iris-Code analysis, fewer relative shifting positions at comparison	CASIA-V4 interval	<1% EER	4-fold reduction
Sample pre-alignment	Iris	Drozdowski *et al.* [45]	Pre-alignment of raw samples based on eye corner and pupil center locations	BioSecure	~2.5% EER	2-fold reduction
Information fusion	Face	Drozdowski *et al.* [44]	Morphing	FERET	98.82% RR-1	52.5% penetration rate

4.3.6 Acceleration

Hardware acceleration can facilitate massive execution speed gains for certain types of computations. In the following subsections, the use of reconfigurable computing (subsection 4.3.6.1) and graphical processing units (subsection 4.3.6.2) in biometric systems is surveyed. The references in those two subsections are by no means exhaustive, due to the focus of this article being elsewhere. Instead, they outline the relevant concepts and highlight a few systems created for the different biometric characteristics. Lastly, they focus on the more recent publications due to the fast pace of developments within hardware components. For a quick general comparison of the capabilities, along with the advantages and disadvantages of those two types of hardware, the reader is referred to *e.g.* an industry white paper in [8], or a general survey of various Big Data analysis platforms and methods [171].

Although hardware acceleration cannot be strictly considered a method of workload reduction (since the *amount* of computations is not reduced – it is merely parallelised, distributed, or executed more efficiently), it is also mentioned here as an important aspect of speeding-up transactions in large-scale biometric identification systems. There appears to be a substantial research interest in the area of hardware-based acceleration utilising FPGAs and GPUs. Some of the existing publications present convincing and well-substantiated results, whereby massive speed gains (up to two orders of magnitude) are achieved in the benchmarks. It should be noted, however, that in some cases the experimental protocols of the benchmarks are questionable; in particular, it is not always clear if the external latency factors (unrelated to the algorithms themselves) have been accounted for in the evaluation. Furthermore, the degree of the CPU-based baseline algorithm optimisation is often not clearly outlined. The results must therefore be closely scrutinised, as it could be that the speed gains result merely from a poor baseline implementation. This caveat notwithstanding, using reconfigurable computing and/or graphical processing units could be a promising avenue for speeding up the execution of various components (or even entire pipelines) in many different biometric modalities. On the other hand, factors such as difficulty of implementation, as well as purchase and maintenance cost have to be taken into consideration for real-world systems.

Lastly, software acceleration and optimisation are also worth mentioning in this context; although there does not seem to be many scientific publications on the topic. In [156], an extensive analysis of possible speed-ups in CPU-based Iris-Code comparisons is presented. The authors consider possible improvements through low-level implementations, manual loop unrolling, caching and pre-computing certain parts of data, analysis of memory access bottlenecks, multi-threading, as well as statistical optimisation of micro-operations. In [118], a hardware-software co-design of iris recognition pipeline is proposed. The authors benchmark highly optimised software code, coupled with a hardware-based implementation of several of the pipeline components. Both publications show that substantial speed-ups (but not computational workload reduction) can be achieved through code optimisations, which do not in themselves change the underlying algorithms or biometric feature representations.

4.3.6.1 Reconfigurable Computing

Field Programmable Gate Arrays (FPGAs) are integrated circuits containing an array/matrix of programmable logic blocks (of different types, *e.g.* general logic, memory, arithmetic), which can be programmably interconnected with each other and with input/output blocks. The programming/configuring is generally done using a hardware description language (*e.g.* VHDL or Verilog) or (nowadays rarely) circuit diagrams, and takes place *after* the chip

has been manufactured. In other words, the FPGAs can be configured and re-configured to execute arbitrary digital circuits, and thus are capable of solving any computable problem. FPGAs can utilise hardware parallelism and deep pipelining extensively, thereby completing many more computations per clock cycle as opposed to a normal sequential execution. Additionally, they rely on much fewer layers of abstraction than the general purpose CPUs, thus facilitating lower-level programming, as well as custom memory and I/O interfaces. Those properties can be exploited to yield potentially massive speed-ups for certain applications (see *e.g.* [165, 174]). For a more detailed view of the current FPGA state-of-the-art, advantages and disatvantages, as well as future outlook and challenges, the reader is referred to fundamentals, *e.g.* [4]. Due to the abovementioned advantages, reconfigurable computing has been extensively applied to solve a variety of problems in many fields (see *e.g.* [181] for a survey), including computer vision, signal processing and pattern matching (see *e.g.* [61]), and neural networks (see *e.g.* [22]). Algorithms from those domains are cornerstones of various biometric systems; hence, substantial research effort has also been devoted to development of FPGA-based processing of biometric data.

FPGA based implementations of biometric systems' components or complete data processing pipelines were published *e.g.* for iris [152], fingerprint [56], face [175], (finger)vein [88], retina [103], and voice [13]. There, speed-ups over traditional CPU-based algorithms of up to two orders of magnitude were reported.

4.3.6.2 Graphical Processing Units

As the name suggests, traditional Graphical Processing Units (GPUs) were designed for very efficient processing of two and three dimensional graphics and have a rigid set of functions and programmable features. Over time, the ease of use/programmability and the range of applications for GPUs have steadily increased, especially with the introduction of general purpose frameworks for GPU programming such as CUDA [135] and OpenCL [176]. Taking advantage of the single program multiple-data (SPMD) programming model, the data can be processed in highly parallel ways. Thus, adapting code to run on GPUs can yield massive execution speed gains for many applications, *e.g.* linear algebra, sorting and searching, differential equations, or more generally floating-point operations on vectorisable data. For a general introduction to GPU computing, the reader is referred to *e.g.* [127]. Some tasks at which GPUs excel are important in typical biometric processing pipelines. Hence, there has been interest in the scientific community to leverage the power of GPUs in this domain as well. A good general introduction to usage of GPUs in biometrics, along with a brief survey of applications for fingerprint-based systems can be found in [106].

GPU based implementations of biometric systems' components or complete data processing pipelines were reported. It should also be noted, that GPUs (and more recently, specialised tensor processing units (TPUs) [7]), have also been utilised extensively in problems involving machine learning and deep neural networks, see *e.g.* [112]. In recent years, those technologies have also been applied to biometrics (*e.g.* facial recognition deep neural networks [167]), highlighting possibilities of hardware-acceleration use beyond efficient biometric identification, more specifically in the algorithm training phase. *E.g.* for iris [186], fingerprint [67], face [201], and sclera-vein [115], similarly to FPGAs (see subsection 4.3.6.1), speed-ups over traditional CPU-based algorithms of up to two orders of magnitude were reported.

4.4 Discussion

In this section, several matters relevant to the topic of this article are discussed, namely: the considerations and trade-offs of computational workload reduction approaches (subsection 4.4.1), a brief digression into data security (subsection 4.4.2), a perspective on how real large-scale biometric systems deal with large-scale biometric identification (subsection 4.4.3), and finally an outline of open issues and challenges in this research field (subsection 4.4.4).

4.4.1 Considerations and Trade-offs

As evidenced by previous sections, there exists a plethora of approaches which seek to reduce the computational workload requirements in biometric identification systems. Below, a systematic (qualitative, due to the infeasibility of directly comparing the results – recall subsection 4.2.2) discussion of noteworthy matters w.r.t. the different approach categories is given, concentrating on their general impact on: 1) computational workload, 2) biometric performance, and 3) disk/memory storage.

Pre-filtering

Computational workload The potential speed-up depends on the discriminative power and size of the index templates. Given strongly discriminative index templates, a much smaller short-list of candidates can be produced, thereby minimising the number of the necessary template comparisons with the expensive (and accurate) comparator. On the other hand, the size of the index templates determines the computational cost of the pre-filtering step, as the probe index is compared exhaustively against the index templates. Naturally, those two parameters typically counterbal-

ance each other – smaller size of the index templates typically entails lower discriminative power.

Biometric performance Since the features used for pre-filtering typically have limited discriminative power, errors may occur, so that the sought identity is not among the returned candidate short-list, thereby increasing the false-negative rates. CMC curves are useful in assessing the efficacy of such features and can help decide on a reasonable size of the candidate short-list.

Storage Since additional information (index) is stored in order to facilitate the pre-filtering step, the storage requirements are increased.

Binning

Computational workload The potential speed-up benefits are limited by the number of bins. It tends to be rather small, especially for the handcrafted classes/types. Additionally, the handcrafted classes/types are very often inherently unevenly distributed (due to genetics and environmental influences). Consequently, computational workload reduction obtained through binning varies accordingly with the relative frequencies of the bins. A good example is binning based on fingerprint types. Whorls and loops generally exhibit the highest prevalence and there are variations across different ethnic groups (see Rife [162]). Nevertheless, it is still a feasible approach, and it has been used in operational systems, *e.g.* in AFIS' (see *e.g.* [52, 129]) – initially using the explicit classes, more recently utilising machine learning to develop non-exclusive classes that lead to more balanced bin sizes. In some cases, however, a severely non-uniform distribution across the bins can invalidate the binning approach entirely. For instance, people of many ethnic groups (or entire countries) have predominantly brown eyes, thus little to no speed-up can be achieved by binning using eye colour in such systems.

Biometric performance In order for the system to be viable, the classification accuracy must be near-optimal. Otherwise, the probability of false-negative errors increases due to misclassification and consequently searching in the wrong bin(s) (*i.e.* pre-selection errors).

Storage Typically not significantly increased, since only the metadata (*e.g.* the fingerprint types) need to be stored.

Data-structures

Computational workload By often relying on divide-and-conquer approaches, the complexity of the retrieval algorithm can often be reduced from the linear complexity down to the (near-)logarithmic complexity.

Biometric performance Due to wrong paths being taken during the search structure traversal, the potential for making false-negative error increases. For many approaches this is especially relevant for the higher levels of the structure, where the information stored by the nodes is denser than near the leaves. On the other hand, the potential for false positive errors is typically reduced due to the lower penetration rate.

Storage The storage requirements are typically increased, since additional hashtables and/or tree-like data-structures have to be maintained.

Feature transformation

Computational workload Although the individual template comparisons are computed much more efficiently, the identification is still carried out over the entire search space (exhaustive search), thereby severely limiting the potential computational workload reduction.

Biometric performance The more compact template representations and/or more efficient comparators may suffer from a decrease in discriminative power and hence a lower biometric performance.

Storage Typically decreased, due to more compact template representations.

Acceleration

Computational workload Not reduced *per se*, merely computed more efficiently (*e.g.* paralellised, distributed, or otherwise optimised).

Biometric performance Typically unaffected, as functionally equivalent algorithms are somehow implemented or optimised to achieve faster computation speeds.

Storage Possibly increased, as it may be necessary to port the biometric data to the specifics of the utilised system (*e.g.* CUDA) or distribute them across a network.

Due to varying system requirements and policies, it is important to enable the biometric systems' operators to make well-informed decisions w.r.t. the used algorithms. Therefore, for any proposed computational workload reduction methods, it is crucial to include the above information, as well

as benchmarks against a current state-of-the-art algorithm performing an exhaustive search (baseline). By doing so, the trade-offs (biometric performance, computational workload, storage) of the proposed methods can be evaluated, thereby facilitating informed decisions on the systems' design and policies. In some cases, it may even be possible to probabilistically model the impact of the proposed methods on the biometric performance, which could potentially be very useful in establishing the pertinent trade-offs even prior to the experimental evaluations on real data. However, such models rarely appear in the surveyed literature. Examples include, *e.g.* [32], which discusses the binning approaches in general and [43], where a statistical model for the proposed Bloom filter-based hierarchical retrieval method is included. Hence, the development of such models could be an interesting avenue of future research.

4.4.2 Data Security

In addition to the need for computational workload reduction, which was the core topic of this article, the potential of data exposure is a large concern in biometric system deployments, where the stored data is, in most cases, secured using traditional encryption algorithms (see *e.g.* [133]). From the technical point of view, this means that should the data be compromised, serious problems such as identity theft, cross-matching without consent arise, furthermore the renewability of such biometric templates is severely limited. Additionally, the centralised storage of sensitive personal and biometric data has increasingly been under scrutiny, both by the general public and various non-governmental organisations, which recently has led to widened legislation against privacy violations (*e.g.* GDPR in Europe [48]). The ISO/IEC international standard on biometric information protection [85] stipulates several properties required for biometric template protection schemes. While many approaches have been proposed for normal biometric systems (see *e.g.* [159] for a survey), template protection coupled together with computational workload reduction has received relatively little attention in the scientific literature. Some early proof-of-concept works and trade-off analyses have been carried out *e.g.* in [41, 108, 192].

4.4.3 Real-World Systems

Due to confidentiality constraints (*i.e.* company or state secrets), the availability of details for real-world systems is not nearly as abundant as that of scientific publications. Nevertheless, this subsection will give several examples, based on the existing literature and personal communications.

Aadhaar As of this writing, the Indian National ID Programme (see *e.g.* [170, 185]) encompasses acquisition and usage of the largest biomet-

ric system in the world in terms of number of the enrolled data subjects. During enrolment, a de-duplication check must be carried out, in order to avoid issuing multiple unique identification numbers to the same individual. The de-duplication proceeds in several steps: first, obvious duplicates are pre-filtered out based on metadata (demographic information), either through exact matches or fuzzy matches followed by a preemptive biometric check. Subsequently, an exhaustive search of the biometric database is performed by one of the COTS systems provided by three different vendors. Each system uses a different implementation and information fusion (data from two irides and ten fingerprints is used) strategies. Any potential duplicate is verified by another system, and if the need arises, it is also adjudicated manually by a trained biometric system operator. The whole system design is distributed and massively parallelisable, and claimed to be scalable through use of commodity hardware and enterprise Big Data solutions.

UAE The border control system in United Arab Emirates (see *e.g.* [3, 35]) takes advantage of the intrinsically efficient iris representation (IrisCodes, see [33]) and distributed architecture of COTS components for quick biometric identification queries. It is reported, that an exhaustive search against the database of close to 1 million subjects can be executed within a couple seconds. I/O latency issues are avoided by pre-loading the entire enrolment database into random access memory.

NEXUS In Canada, automated self-service kiosks for selected airports and (frequent) travellers are offered in order to expedite the border control process. The system uses iris as the biometric characteristic, and performs 1-to-first searches on the database of over 0.5 million enrolled data subjects for each biometric identification (see *e.g.* [69]).

EES Due to the specifics of the legal mandate [50], the biometric systems for the EU visa and entry-exit system will be forced to perform exhaustive database searches. Due to the operational scenarios (such as border control), stringent requirements for quick (real-time) query responses have been imposed on the potential biometric systems vendors. It can therefore be expected that efficient data representations, information fusion schemes, as well as parallel/distributed design will be essential components of the forthcoming infrastructure solutions.

AFIS Deployments of the Automated Fingerprint Identification Systems (see *e.g.* [62, 101, 129]) are ubiquitous around the world and used for instance in the context of criminal investigations. Such systems are known to utilise computational workload reduction methods based on

demographics, coarse fingerprint data (such as fingerprint type), along with highly optimised software algorithms to facilitate fast response times. Prominent examples include the Integrated AFIS (IAFIS) ran by the Federal Bureau of Investigation (FBI) in the USA and the database of the German Bundeskriminalamt (BKA).

Industry In order to provide competitive search speeds for large biometric identification systems, commercial vendors of biometric recognition technologies, for example the German company Dermalog (information acquired through personal communication), are known to utilise methods of workload reduction in their products. National Institute of Standards and Technology (NIST) carried out an evaluation of 1:N face recognition vendors [70]. Among the details, it has been stated that several submitted algorithms take the expense of constructing fast search data-structures at enrolment, in order to achieve sub-linear search duration growth (with respect to the size of the biometric reference database).

From the above examples, it is clear that computational workload is a critical consideration in operational systems and certain methods used to expedite the high number of queries handled by the existing large-scale systems around the world. To summarise in the context of the proposed taxonomy (recall figure 4.2), the following methods are represented in the list above:

- Pre-filtering by metadata (*e.g.* demographic and geographic).
- Search strategies (*e.g.* 1-to-first search).
- Binning with coarse features (*e.g.* fingerprint types).
- Intrinsically efficient feature representations (*e.g.* IrisCodes).
- Software optimisation.
- Massive parallelisation and distribution of computations.

Notably absent are the methods of hardware acceleration (*i.e.* reconfigurable computing and graphical processing units) – it could be, that the practical matters (*e.g.* monetary implementation and maintenance costs) outweigh the benefits of potential speed improvements. Another potential important issue is vendor lock-in and the necessity of tailoring for specific software/hardware combination, which was deliberately avoided in *e.g.* the UAE and Aadhaar systems by making the design decision to use commodity CPU-based hardware (see *e.g.* [3, 185]). The pre-selection methods, which are heavily researched, seem to be seldom used in those deployments. In

other words, in this case there does appear to be (or perhaps is perpetuated by the information scarcity in this area), at least to a certain degree, a mismatch between what is researched in academia and what is actually used by the industry. This and other open issues are discussed in more detail in the next subsection.

4.4.4 Open Issues

Several open issues/challenges remain in research related to computational workload reduction in biometric identification systems:

Standardisation As described in subsection 4.2.2, as of yet there exists no standardised way of reporting results for biometric computational workload and its reduction. This leads to a multitude of methodologies and metrics in the scientific literature, thereby making direct benchmarks and comparative assessment of the proposed methods extremely cumbersome to carry out. Moreover, in many publications, even the baseline results (*i.e.* exhaustive search with a state-of-the-art algorithm) are not reported, which further exacerbates this issue. In [43], experimental prerequisites and metrics for such evaluations are proposed; as of this writing, there is an ongoing effort to include metrics to measure computational workload and its reduction in the current revision of the ISO/IEC IS 19795-1. This effort notwithstanding, the standardisation in this research area remains an open debate subject within the standardisation committee and in general.

Scalability Due to limited availability of large-scale biometric data for academic research, many, if not most, of the surveyed approaches were tested on relatively small databases (mostly up to hundreds or thousands of subjects; tens of thousands of samples). Hence, the scalability of many of the proposed methods remains questionable or unproven in practice (*cf.* table 4.1 w.r.t. the sizes of large real-world deployments of biometric systems). Some authors are fortunate enough to receive access to large-scale, sequestered databases (*e.g.* law enforcement) in order to evaluate and validate their approaches (*e.g.* [75, 188]), but such cases are rare in the surveyed literature, likely as a consequence of the significant practical and legal hurdles associated with accessing the sensitive biometric/personal data.

Biometric performance trade-off It appears that many approaches are capable of delivering significant (*e.g.* one or two orders of magnitude) decrease in computational workload requirements. However, further reductions prove elusive for most methods due to rapidly degenerating biometric performance. Some of the surveyed methods incorporate in their designs information fusion (from multiple biometric char-

acteristics or multiple instances of the same characteristics) to mitigate this issue to a certain degree.

Dissonance between academia and industry There seem to exist, to a certain degree, discrepancies between methods and goals of the research conducted within the academia, and the actual practical use cases in the industry. Several examples (also partially related to the aforementioned issues of standardisation and scalability) relevant to the topic of this article are:

Evaluation protocol Whereas a substantial number of the surveyed publications perform their experiments using the closed-set identification, the real-world systems are essentially universally required to perform the much more challenging open-set identification.

Decision thresholds Almost all of the surveyed publications do not report decision thresholds at which the given biometric performance was achieved. Furthermore, evaluations on different datasets with fixed decision thresholds are rarely performed. This is in stark contrast with the industry/law enforcement practices, where decision thresholds and fixed operational points are used extensively in the operational systems.

Results reporting The surveyed publications often report results using metrics, which are of limited value in the industry practice. For instance, operational systems would not typically operate at EER (even if it ever was an operational point at all), but rather at fixed false-positive identification-error rates acceptable within their respective system policies. Many publications use rank-based reporting and CMC curves, which imply the less interesting (from the industry point of view) closed-set evaluation protocol, as mentioned above.

Acceleration As evidenced by subsection 4.3.6, significant research efforts have been devoted to develop biometric algorithms suitable for FPGA or GPU computations. However, as mentioned in subsection 4.4.3, the existing deployments of large-scale biometric systems lean towards distributed architectures of commodity CPU hardware due to other practical considerations; additionally, software optimisations, which play an important role in the commercial systems, are only superficially treated in the scientific literature.

Thus, aside from the technical challenges (and potential limits) of improving on the trade-offs between biometric performance and computational workload reduction, there are two key areas that should be consid-

ered by the large-scale biometric identification practitioners from academia and industry alike:

Academia and industry cooperation Much tighter integration between the academic research and industry requirements is needed. The academics should seek out industry partners to validate their proposed systems and solutions outside of a lab setting; conversely, the industry should engage in outreach initiatives to academia in order to promote the actual prerequisites, requirements, and challenges of the commercial systems. More frequent and deeper joint (research or otherwise) projects and partnerships between academia and industry could potentially help to ameliorate the aforementioned dissonances. Jain *et al*. [89] recently published the short "Guidelines for Best Practices in Biometric Research". This document can serve as a good starting point outlining the absolute essentials for legitimate and practical reporting of results within biometrics research. Furthermore, the representatives from all the stakeholders should begin or continue to actively engage in the international standardisation efforts, for instance the ones by the ISO/IEC JTC 1/SC 37 (who are responsible for, among others, the biometric performance evaluation and harmonized biometric vocabulary standards [86, 87]) and/or of other national agencies, such as NIST in the USA and BSI in Germany. This way, meaningful consensus w.r.t. evaluation protocols, metrics, and benchmarks that reflect the real use cases can be attained.

Practical evaluations Entities (*e.g*. governmental agencies, universities, companies) in possession of large amounts of biometric data should support large-scale evaluations, thus facilitating scalability assessment and fair benchmarks between the systems developed by the academic researchers and the commercial vendors. Such benchmarks make available an API against which algorithms can be coded, then submitted to a central server, and finally ran and evaluated there using the same experimental protocol and metrics. Such tests offer an additional advantage, in that the vast majority of the image data used for evaluation remains unseen by the algorithm authors, which in turn facilitates higher generalisably of the proposed algorithms. Lastly, synthetically generated data could be used to some extent in the context of biometric scalability testing (see *e.g*. [131]). Interesting existing initiatives in this area are, for example: the BEAT platform [121] with the aim of developing a general standard framework biometric technologies' evaluation, along with several more constrained initiatives such as the indexing competition under Fingerprint Verification Competition (FVC-onGoing) [19], the 1:N Evaluation under Face Recognition Vendor Test

(FRVT) [134], one-to-many evaluations under Iris Exchange (IREX) – for fingerprint, face, and iris characteristics, respectively.

4.5 Summary

Large-scale biometric identification systems are confronted with high computational workload. This is especially the case for an exhaustive search, where the computational effort required during retrieval grows linearly with the number of enrollees. Methods that seek to alleviate this issue aim at reducing the number of template comparisons necessary per retrieval (penetration rate), the computational costs associated with individual template comparisons, or at optimising the software/hardware system implementations. In this article, a taxonomy for conceptual categorisation of such methods is presented, followed by a comprehensive survey of publications pertaining therewith. The article is concluded with a discussion of matters to take note of with the various categories of approaches, a digression on usage of such methods in real-world systems, as well as an outline of remaining relevant challenges.

As the number and scale of the biometric systems deployments worldwide steadily increases, computational workload reduction in biometric identification systems can be expected to remain an active field of research, especially since a number of open issues/challenges remains unresolved regardless of the significant advances made through the academic and commercial research. To solve said challenges, standardisation and much tighter cooperation between the academia, industry, governmental agencies, and other concerned parties is necessary. There exists an urgent need of a unified methodology for reporting of computational workload and its reduction. Furthermore, another important matter is the development of experimental protocols, benchmarks, and metrics which closely correspond with the actual prerequisites and use cases of the real-world deployments. To accomplish this, the international standardisation efforts are a promising avenue, albeit continuous engagement from all the concerned stakeholders is necessary to establish a suitable and broad consensus.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF), by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP), and the LOEWE-3 BioBiDa Project (594/18-17).

4.6 Bibliography

[1] ABBASIFARD, M. R., GHAHREMANI, B., AND NADERI, H. A survey on nearest neighbor search methods. *International Journal of Computer Applications 95*, 25 (2014).

[2] AGGARWAL, C. C., AND REDDY, C. K. *Data Clustering: Algorithms and Applications*. Chapman & Hall/CRC, 2013.

[3] AL-RAISI, A. N., AND AL-KHOURI, A. M. Iris recognition and the challenge of homeland and border control security in UAE. *Telematics and Informatics 25*, 2 (May 2008), 117–132.

[4] ANDINA, J. J. R., DE LA TORRE ARNANZ, E., AND VALDES, M. D. *FPGAs: Fundamentals, Advanced Features, and Applications in Industrial Electronics*. CRC Press, January 2017.

[5] BADRINATH, G. S., GUPTA, P., AND MEHROTRA, H. Score level fusion of voting strategy of geometric hashing and SURF for an efficient palmprint-based identification. *Journal of real-time image processing 8*, 3 (September 2013), 265–284.

[6] BARBU, T., AND LUCA, M. Content-based iris indexing and retrieval model using spatial acces methods. In *International Symposium on Signals, Circuits and Systems (ISSCS)* (July 2015), IEEE, pp. 1–4.

[7] BARRUS, J. Cloud TPU machine learning accelerators now available in beta. `https://cloud.google.com/blog/products/gcp/cloud-tpu-machine-learning-accelerators-now-available-in-beta`, February 2018. Last accessed: 2020–03–11.

[8] BERTEN DIGITAL SIGNAL PROCESSING. GPU vs FPGA performance comparison. Tech. Rep. BWP001 v1.0, Berten DSP, May 2016.

[9] BHANU, B., AND TAN, X. Fingerprint indexing based on novel features of minutiae triplets. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 25*, 5 (May 2003), 616–622.

[10] BHUTANI, A., AND BHARDWAJ, P. Biometrics market size by application. Tech. Rep. GMI493, Global Market Insights, August 2017.

[11] BILLEB, S., RATHGEB, C., BUSCHBECK, M., REININGER, H., AND KASPER, K. Efficient two-stage speaker identification based on universal background models. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2014), IEEE, pp. 1–6.

[12] BISWAS, S., RATHA, N. K., AGGARWAL, G., AND CONNELL, J. Exploring ridge curvature for fingerprint indexing. In *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2008), IEEE, pp. 1–6.

[13] BOURAOUI, H., JERAD, C., CHATTOPADHYAY, A., AND HADJ-ALOUANE, N. B. Hardware architectures for embedded speaker recognition applications: A survey. *Transactions on Embedded Computing Systems (TECS) 16*, 3 (2017), 78.

[14] BOWYER, K., AND BURGE, M. J. *Handbook of iris recognition*. Springer, 2016.

[15] CAPPELLI, R. Fast and accurate fingerprint indexing based on ridge orientation and frequency. *Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 41*, 6 (December 2011), 1511–1521.

[16] CAPPELLI, R., FERRARA, M., AND MAIO, D. Candidate list reduction based on the analysis of fingerprint indexing scores. *Transactions on Information Forensics and Security (TIFS) 6*, 3 (September 2011), 1160–1164.

[17] CAPPELLI, R., FERRARA, M., AND MALTONI, D. Minutia Cylinder-Code: A new representation and matching technique for fingerprint recognition. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 32*, 12 (December 2010), 2128–2141.

[18] CAPPELLI, R., FERRARA, M., AND MALTONI, D. Fingerprint indexing based on minutia cylinder-code. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 33*, 5 (May 2011), 1051–1057.

[19] CAPPELLI, R., FERRARA, M., AND MALTONI, D. FIDXICB-2013. `https://biolab.csr.unibo.it/fvcongoing/UI/Form/ICB2013FIDX.aspx`, 2013. Last accessed: 2020–03–11.

[20] CHAARI, A., LELANDAIS, S., AND AHMED, M. B. A pruning approach improving face identification systems. In *International Conference on Advanced Video and Signal Based Surveillance* (September 2009), IEEE, pp. 85–90.

[21] CHÁVEZ, E., NAVARRO, G., BAEZA-YATES, R., AND MARROQUÍN, J. L. Searching in metric spaces. *Compututing Surveys (CSUR) 33*, 3 (September 2001), 273–321.

[22] CHEN, A.-J., BIGLARI-ABHARI, M., WANG, K.-K., , BOUZERDOUM, A., AND TIVIVE, F. Convolutional neural network acceleration with hardware/software co-design. *Applied Intelligence 48*, 5 (2018), 1288–1301.

[23] CHEN, B. C., CHEN, Y. Y., KUO, Y. H., AND HSU, W. H. Scalable face image retrieval using attribute-enhanced sparse codewords. *Transactions on Multimedia 15*, 5 (August 2013), 1163–1173.

[24] CHEN, F., HUANG, X., AND ZHOU, J. Hierarchical minutiae matching for fingerprint and palmprint identification. *Transactions on Image Processing 22*, 12 (December 2013), 4964–4971.

[25] CONSORTIUM FOR ELECTIONS AND POLITICAL PROCESS STRENGTHENING. Assessment of electoral preparations in the Democratic Republic of the Congo. Tech. rep., CEPPS, May 2018.

[26] CORMEN, T. H., LEISERSON, C. E., RIVEST, R. L., AND STEIN, C. *Introduction to algorithms*. MIT press, 2009.

[27] DAMER, N., TERHÖRST, P., BRAUN, A., AND KUIJPER, A. Efficient, accurate, and rotation-invariant iris code. *Signal Processing Letters 24*, 8 (August 2017), 1233–1237.

[28] DAMER, N., TERHÖRST, P., BRAUN, A., AND KUIJPER, A. General Borda count for multi-biometric retrieval. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 420–428.

[29] DAMER, N., TERHÖRST, P., BRAUN, A., AND KUIJPER, A. Indexing of single and multi-instance iris data based on LSH-forest and rotation invariant representation. In *International Conference on Computer Analysis of Images and Patterns (CAIP)* (2017), Springer, IEEE, pp. 190–201.

[30] DANTCHEVA, A., ELIA, P., AND ROSS, A. What else does your biometric data reveal? A survey on soft biometrics. *Transactions on Information Forensics and Security (TIFS) 11*, 3 (March 2016), 441–467.

[31] DANTCHEVA, A., ERDOGMUS, N., AND DUGELAY, J.-L. On the reliability of eye color as a soft biometric trait. In *Workshop on Applications of Computer Vision (WACV)* (January 2011), IEEE, pp. 227–231.

[32] DAUGMAN, J. Biometric decision landscapes. Tech. Rep. UCAM-CL-TR-482, University of Cambridge - Computer Laboratory, January 2000.

[33] DAUGMAN, J. How iris recognition works. *Transactions on Circuits and Systems for Video Technology (TCSVT) 14*, 1 (January 2004), 21–30.

[34] DAUGMAN, J. History of iris recognition. https://www.cl.cam.ac.uk/ jgd1000/history.html, 2019. Last accessed: 2020–03–11.

[35] DAUGMAN, J., AND MALHAS, I. Iris recognition border-crossing system in the UAE. *International Airport Review 8*, 2 (2004), 1–5.

[36] DE BOER, J., BAZEN, A. M., AND GEREZ, S. H. Indexing fingerprint databases based on multiple features. In *Annual Workshop on Circuits, Systems and Signal Processing* (November 2001), Technology Foundation (STW), pp. 300–306.

[37] DEPARTMENT OF HOMELAND SECURITY. DHS/ALL-041 external biometric records (EBR) system of records. `https://www.regulations.gov/docket?D=DHS-2017-0039`, April 2018. Last accessed: 2020–03–11.

[38] DEWANGAN, J., DEY, S., AND SAMANTA, D. Face images database indexing for person identification problem. *International Journal of Biometrics and Bioinformatics 7*, 2 (2013), 93–122.

[39] DEY, S., AND SAMANTA, D. Iris data indexing method using Gabor energy features. *Transactions on Information Forensics and Security (TIFS) 7*, 4 (August 2012), 1192–1203.

[40] DROZDOWSKI, P., FISCHER, D., RATHGEB, C., SCHIEL, C., AND BUSCH, C. Database binning and retrieval in multi-fingerprint identification systems. In *International Workshop on Information Forensics and Security (WIFS)* (December 2018), IEEE, pp. 1–7.

[41] DROZDOWSKI, P., GARG, S., RATHGEB, C., GOMEZ-BARRERO, M., CHANG, D., AND BUSCH, C. Privacy-preserving indexing of IrisCodes with cancelable Bloom filter-based search structures. In *European Signal Processing Conference (EUSIPCO)* (September 2018), IEEE, pp. 1–5.

[42] DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Multi-iris indexing and retrieval: Fusion strategies for Bloom filter-based search structures. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 1–8.

[43] DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Bloom filter-based search structures for indexing and retrieving Iris-Codes. *IET Biometrics 7*, 3 (May 2018), 260–268.

[44] DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Turning a vulnerability into an asset: Accelerating facial identification with morphing. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (May 2019), IEEE, pp. 1–5.

[45] Drozdowski, P., Rathgeb, C., Hofbauer, H., Wagner, J., Uhl, A., and Busch, C. Towards pre-alignment of near-infrared iris images. In *International Joint Conference on Biometrics (IJCB)* (2017), IEEE, pp. 359–366.

[46] Drozdowski, P., Struck, F., Rathgeb, C., and Busch, C. Benchmarking binarisation schemes for deep face templates. In *International Conference on Image Processing (ICIP)* (October 2018), IEEE, pp. 1–5.

[47] European Commission. Smart borders. https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/smart-borders_en, 2018. Last accessed: 2020–03–11.

[48] European Parliament. Regulation (EU) 2016/679. *Official Journal of the European Union L119* (April 2016), 1–88.

[49] European Union. Regulation (EU) no 603/2013 of the European Parliament and of the Council. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32013R0603, June 2013. Last accessed: 2020–03–11.

[50] European Union. Regulation (EU) 2017/2226 of the European Parliament and of the Council. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R2226, November 2017. Last accessed: 2020–03–11.

[51] European Union Agency for the Operational Management of Large-Scale IT Systems in the Area of Freedom, Security and Justice. Eurodac storage capacity increased. https://www.eulisa.europa.eu/Newsroom/News/Pages/Eurodac-storage-capacity-increased.aspx, April 2016. Last accessed: 2020–03–11.

[52] Federal Bureau of Investigation. *The Science of Fingerprints: Classification and Uses*. General Press, August 2013.

[53] Federal Bureau of Investigation. CODIS - NDIS statistics. https://www.fbi.gov/services/laboratory/biometric-analysis/codis/ndis-statistics, June 2018. Last accessed: 2020–03–11.

[54] Feng, J., and Cai, A. Fingerprint indexing using ridge invariants. In *International Conference on Pattern Recognition (ICPR)* (2006), vol. 4, IEEE, pp. 433–436.

[55] Feng, J., and Jain, A. K. Filtering large fingerprint database for latent matching. In *International Conference on Pattern Recognition (ICPR)* (December 2008), IEEE, pp. 1–4.

[56] FONS, M., FONS, F., AND CANTÓ, E. Fingerprint image processing acceleration through run-time reconfigurable hardware. *Transactions on Circuits and Systems II: Express Briefs 57*, 12 (December 2010), 991–995.

[57] GADDE, R. B., ADJEROH, D., AND ROSS, A. Indexing iris images using the Burrows-Wheeler transform. In *International Workshop on Information Forensics and Security (WIFS)* (December 2010), IEEE, pp. 1–6.

[58] GALAR, M., DERRAC, J., PERALTA, D., TRIGUERO, I., PATERNAIN, D., LOPEZ-MOLINA, C., GARCÍA, S., BENÍTEZ, J. M., PAGOLA, M., BARRENECHEA, E., ET AL. A survey of fingerprint classification part I: Taxonomies on feature extraction methods and learning models. *Knowledge-based systems 81* (June 2015), 76–97.

[59] GALAR, M., DERRAC, J., PERALTA, D., TRIGUERO, I., PATERNAIN, D., LOPEZ-MOLINA, C., GARCÍA, S., BENÍTEZ, J. M., PAGOLA, M., BARRENECHEA, E., ET AL. A survey of fingerprint classification part II: experimental analysis and ensemble proposal. *Knowledge-Based Systems 81* (June 2015), 98–116.

[60] GALTON, F. *Fingerprint directories*. Macmillan and Company, 1895.

[61] GARCÍA, G. J., JARA, C. A., POMARES, J., ALABDO, A., POGGI, L. M., AND TORRES, F. A survey on FPGA-based sensor systems: towards intelligent and reconfigurable low-power sensors for computer vision, control and signal processing. *Sensors 14*, 4 (2014), 6247–6278.

[62] GEMALTO. Automated Fingerprint Identification System (AFIS) - a short history. `https://www.gemalto.com/govt/biometrics/afis-history`, April 2019. Last accessed: 2020–03–11.

[63] GEMALTO. DHS's automated biometric identification system IDENT - the heart of biometric visitor identification in the USA. `https://www.gemalto.com/govt/customer-cases/ident-automated-biometric-identification-system`, March 2019. Last accessed: 2020–03–11.

[64] GENTILE, J. E., RATHA, N., AND CONNELL, J. An efficient, two-stage iris recognition system. In *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2009), IEEE, pp. 211–215.

[65] GENTILE, J. E., RATHA, N., AND CONNELL, J. SLIC: short-length iris codes. In *International Conference on Biometrics: Theory, Applications, and Systems (BTAS)* (2009), IEEE, pp. 1–5.

[66] GERMAIN, R. S., CALIFANO, A., AND COLVILLE, S. Fingerprint matching using transformation parameter clustering. *Computational Science and Engineering 4*, 4 (October 1997), 42–49.

[67] GHAFOOR, M., IQBAL, S., TARIQ, S. A., TAJ, I. A., AND JAFRI, N. M. Efficient fingerprint matching using GPU. *IET Image Processing 12*, 2 (2018), 274–284.

[68] GIONIS, A., INDYK, P., AND MOTWANI, R. Similarity search in high dimensions via hashing. In *International Conference on Very Large Data Bases (VLDB)* (September 1999), Morgan Kaufmann, pp. 518–529.

[69] GORODNICHY, D. O., AND CHUMAKOV, M. P. Analysis of the effect of ageing, age, and other factors on iris recognition performance using NEXUS scores dataset. *IET Biometrics* (July 2018).

[70] GROTHER, P., NGAN, M., AND HANAOKA, K. Ongoing face recognition vendor test (FRVT) part 2: Identification. Tech. Rep. NISTIR 8238, National Institute of Standards and Technology, November 2018.

[71] GUPTA, P., SANA, A., MEHROTRA, H., AND HWANG, C. J. An efficient indexing scheme for binary feature based biometric database. In *Biometric technology for human identification IV* (April 2007), vol. 6539, International Society for Optics and Photonics, pp. 1–10.

[72] GYAOUROVA, A., AND ROSS, A. A coding scheme for indexing multimodal biometric databases. In *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (June 2009), IEEE, pp. 93–98.

[73] GYAOUROVA, A., AND ROSS, A. Index codes for multibiometric pattern retrieval. *Transactions on Information Forensics and Security (TIFS) 7*, 2 (April 2012), 518–529.

[74] HÄMMERLE-UHL, J., PENN, G., PÖTZELSBERGER, G., AND UHL, A. Size-reduction strategies for iris codes. *International Journal of Computer, Electrical, Automation, Control and Information Engineering 9*, 1 (2015), 290–293.

[75] HAO, F., DAUGMAN, J., AND ZIELINSKI, P. A fast search algorithm for a large fuzzy database. *Transactions on Information Forensics and Security (TIFS) 3*, 2 (June 2008), 203–212.

[76] HE, S., ZHANG, C., AND HAO, P. Comparative study of features for fingerprint indexing. In *International Conference on Image Processing (ICIP)* (November 2009), IEEE, pp. 2749–2752.

[77] HEINDL, R. *Daktyloskopie*. W. de Gruyter & Company, 1927.

[78] HENRY, E. R. *Classification and uses of finger prints*. HM Stationery Office, 1900.

[79] HJALTASON, G. R., AND SAMET, H. Index-driven similarity search in metric spaces (survey article). *Transactions on Database Systems (TODS) 28*, 4 (December 2003), 517–580.

[80] HOLLINGSWORTH, K. P., BOWYER, K. W., AND FLYNN, P. J. The best bits in an iris code. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 31*, 6 (June 2009), 964–973.

[81] ILOANUSI, O. N. Fusion of finger types for fingerprint indexing using minutiae quadruplets. *Pattern Recognition Letters 38* (2014), 8–14.

[82] ILOANUSI, O. N., GYAOUROVA, A., AND ROSS, A. Indexing fingerprints using minutiae quadruplets. In *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (June 2011), IEEE, pp. 127–133.

[83] INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. Biometric council newsletter. `http://ieee-biometrics.org/images/pdf/Newsletter_Nov_2015_corrected.pdf`, November 2015. Last accessed: 2020–03–11.

[84] IQBAL, A., AND NAMBOODIRI, A. Cascaded filtering for fingerprint identification using random projections. In *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (June 2012), IEEE, pp. 77–82.

[85] ISO/IEC JTC1 SC27 IT SECURITY TECHNIQUES. *ISO/IEC 24745:2011. Information technology – Security techniques – Biometric information protection*. International Organization for Standardization and International Electrotechnical Committee, June 2011.

[86] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 19795-1:2006. Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*. International Organization for Standardization and International Electrotechnical Committee, April 2006.

[87] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 2382-37:2017. Information technology – Vocabulary – Part 37: Biometrics*, 2 ed. International Organization for Standardization and International Electrotechnical Committee, February 2017.

[88] JADHAV, M., AND NERKAR, P. M. Implementation of an embedded hardware of FVRS on FPGA. In *International Conference on Information Processing (ICIP)* (December 2015), IEEE, pp. 48–53.

[89] JAIN, A., KLARE, B., AND ROSS, A. Guidelines for best practices in biometrics research. In *International Conference on Biometrics (ICB)* (May 2015), IEEE, pp. 541–545.

[90] JAIN, A. K., FLYNN, P., AND ROSS, A. *Handbook of biometrics*. Springer, 2007.

[91] JAIN, A. K., MURTY, M. N., AND FLYNN, P. J. Data clustering: A review. *Computing Surveys (CSUR) 31*, 3 (September 1999), 264–323.

[92] JAIN, A. K., PRABHAKAR, S., HONG, L., AND PANKANTI, S. Filterbank-based fingerprint matching. *Transactions on Image Processing 9*, 5 (May 2000), 846–859.

[93] JAYARAMAN, U., AND GUPTA, P. Iris code hashing. In *International Conference on Communications (ICC)* (June 2013), IEEE, pp. 2123–2127.

[94] JAYARAMAN, U., PRAKASH, S., AND GUPTA, P. Indexing multimodal biometric databases using kd-tree with feature level fusion. In *International Conference on Information Systems Security* (2008), Springer, pp. 221–234.

[95] JAYARAMAN, U., PRAKASH, S., AND GUPTA, P. An efficient color and texture based iris image retrieval technique. *Expert Systems with Applications 39*, 5 (April 2012), 4915–4926.

[96] KAUSHIK, V. D., UMARANI, J., GUPTA, A. K., AND GUPTA, P. An efficient indexing scheme for face database using modified geometric hashing. *Neurocomputing 116* (2013), 208–221.

[97] KAVATI, I., PRASAD, M. V. N. K., AND BHAGVATI, C. Vein pattern indexing using texture and hierarchical decomposition of Delaunay triangulation. In *International Symposium on Security in Computing and Communication* (2013), Springer, pp. 213–222.

[98] KHALAF, E. T., MOHAMMED, M., AND MOORTHY, K. Robust partitioning and indexing for iris biometric database based on local features. *IET Biometrics* (February 2018).

[99] KLARE, B. F., BLANTON, A., AND KLEIN, B. Efficient face retrieval using synecdoches. In *International Joint Conference on Biometrics (IJCB)* (September 2014), IEEE, pp. 1–7.

[100] KNUTH, D. Sorting and searching. *The art of computer programming 3* (1998).

[101] KOMARINSKI, P. *Automated Fingerprint Identification Systems (AFIS)*. Elsevier, 2005.

[102] KONRAD, M., STÖGNER, H., UHL, A., AND WILD, P. Computationally efficient serial combination of rotation-invariant and rotation compensating iris recognition algorithms. In *International Conference on Computer Vision Theory and Applications (VISAPP)* (May 2010), SciTePress, pp. 85–90.

[103] KOUKOUNIS, D., TTOFIS, C., PAPADOPOULOS, A., AND THEOCHARIDES, T. A high performance hardware architecture for portable, low-power retinal vessel segmentation. *Integration, the VLSI journal 47*, 3 (2014), 377–386.

[104] KUEHLKAMP, A., AND BOWYER, K. Found a good match: Should i keep searching? - accuracy and performance in iris matching using 1-to-first search. *Image and Vision Computing 73* (2018), 17–27.

[105] LAMDAN, Y., AND WOLFSON, H. J. Geometric hashing: A general and efficient model-based recognition scheme. In *International Conference on Computer Vision (ICCV)* (December 1988), IEEE, pp. 238–249.

[106] LASTRA, M., GUTIÉRREZ, P. D., BENÍTEZ, J. M., AND HERRERA, F. GPU processing for biometric big data based identification. why and what for? *Biostatistics and Biometrics Open Access Journal 2* (May 2017), 1–4.

[107] LI, G., YANG, B., AND BUSCH, C. A fingerprint indexing scheme with robustness against sample translation and rotation. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2015), IEEE, pp. 1–8.

[108] LI, G., YANG, B., AND BUSCH, C. A fingerprint indexing algorithm on encrypted domain. In *Trustcom/BigDataSE/ISPA* (August 2016), IEEE, pp. 1030–1037.

[109] LI, J., YAU, W.-Y., AND WANG, H. Fingerprint indexing based on symmetrical measurement. In *International Conference on Pattern Recognition (ICPR)* (2006), vol. 1, IEEE, pp. 1038–1041.

[110] LI, S. Z., AND JAIN, A. K. *Handbook of face recognition*. Springer, 2004.

[111] LI, S. Z., AND JAIN, A. K. *Encyclopedia of biometrics*. Springer, 2015.

[112] LI, X., ZHANG, G., HUANG, H. H., WANG, Z., AND ZHENG, W. Performance analysis of GPU-based convolutional neural networks. In *International Conference on Parallel Processing (ICPP)* (August 2016), IEEE, pp. 67–76.

[113] LIANG, X., BISHNU, A., AND ASANO, T. A robust fingerprint indexing scheme using minutia neighborhood structure and low-order delaunay triangles. *Transactions on Information Forensics and Security (TIFS) 2*, 4 (December 2007), 721–733.

[114] LIM, M.-H., TEOH, A. B. J., AND KIM, J. Biometric feature-type transformation: Making templates compatible for secret protection. *Signal Processing Magazine 32*, 5 (September 2015), 77–87.

[115] LIN, Y., DU, E. Y., ZHOU, Z., AND THOMAS, N. L. An efficient parallel approach for sclera vein recognition. *Transactions on Information Forensics and Security (TIFS) 9*, 2 (February 2014), 147–157.

[116] LIU, L., CHEN, J., FIEGUTH, P., ZHAO, G., CHELLAPPA, R., AND PIETIKÄINEN, M. From BoW to CNN: Two decades of texture representation for texture classification. *International Journal of Computer Vision 127*, 1 (January 2019), 74–109.

[117] LIU, M., JIANG, X., AND KOT, A. C. Efficient fingerprint search based on database clustering. *Pattern Recognition 40*, 6 (2007), 1793–1803.

[118] LÓPEZ, M., J.DAUGMAN, AND CANTÓ, E. Hardware-software co-design of an iris recognition algorithm. *IET Information Security 5*, 1 (2011), 60–68.

[119] MALTONI, D., MAIO, D., JAIN, A. K., AND PRABHAKAR, S. *Handbook of fingerprint recognition*. Springer, 2009.

[120] MANSUKHANI, P., TULYAKOV, S., AND GOVINDARAJU, V. A framework for efficient fingerprint identification using a minutiae tree. *Systems Journal 4*, 2 (June 2010), 126–137.

[121] MARCEL, S. BEAT – biometrics evaluation and testing. *Biometric technology today 2013*, 1 (2013), 5–7.

[122] MARKETS AND MARKETS. Biometric system market by authentication type - global forecast to 2023. Tech. Rep. SE 3449, Markets and Markets, July 2018.

[123] MEHROTRA, H., MAJHI, B., AND GUPTA, P. Robust iris indexing scheme using geometric hashing of SIFT keypoints. *Journal of Network and Computer Applications 33*, 3 (2010), 300–313.

[124] MEHROTRA, H., SRINIVAS, B. G., AND MAJHI, B. Indexing iris biometric database using energy histogram of DCT subbands. In *International Conference on Contemporary Computing (IC3)* (August 2009), vol. 40, Springer, pp. 194–204.

[125] MHATRE, A., CHIKKERUR, S., AND GOVINDARAJU, V. Indexing biometric databases using pyramid technique. In *International Conference on Audio-and Video-Based Biometric Person Authentication* (2005), Springer, pp. 841–849.

[126] MHATRE, A. J., PALLA, S., CHIKKERUR, S., AND GOVINDARAJU, V. Efficient search and retrieval in biometric databases. In *Biometric technology for human identification II* (2005), vol. 5779, International Society for Optics and Photonics, pp. 265–274.

[127] MITTAL, S., AND VETTER, J. S. A survey of CPU-GPU heterogeneous computing techniques. *Computing Surveys (CSUR) 47*, 4 (July 2015), 1–35.

[128] MOHANTY, P., SARKAR, S., KASTURI, R., AND PHILLIPS, P. J. Subspace approximation of face recognition algorithms: An empirical study. *Transactions on Information Forensics and Security (TIFS) 3*, 4 (December 2008), 734–748.

[129] MOSES, K. R., HIGGINS, P., MCCABE, M., PROBHAKAR, S., AND SWANN, S. *Fingerprint Sourcebook*. US Department of Justice, 2010, ch. Automated Fingerprint Identification System (AFIS), pp. 1–33.

[130] MUKHERJEE, R., AND ROSS, A. Indexing iris images. In *International Conference on Pattern Recognition (ICPR)* (December 2008), IEEE, pp. 1–3.

[131] MURPHY, T. M., BROUSSARD, R., RAKVIC, R., NGO, H., IVES, R. W., SCHULTZ, R., AND AGUAYO, J. T. Use of synthetic data to test biometric algorithms. *Journal of Electronic Imaging 25*, 4 (August 2016), 043023.

[132] NALLA, P. R., AND CHALAVADI, K. M. Iris classification based on sparse representations using on-line dictionary learning for large-scale de-duplication applications. *SpringerPlus 4*, 1 (2015).

[133] NANDAKUMAR, K., AND JAIN, A. K. Biometric template protection: Bridging the performance gap between theory and practice. *Signal Processing Magazine 32*, 5 (September 2015), 88–100.

[134] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. Face Recognition Vendor Test (FRVT) 1:N Evaluation. `https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-1n-2018-evaluation`, October 2017. Last accessed: 2020–03–11.

[135] NICKOLLS, J., BUCK, I., GARLAND, M., AND SKADRON, K. Scalable parallel programming with CUDA. In *SIGGRAPH classes* (April 2008), ACM, pp. 42–53.

[136] PALLA, S., CHIKKERUR, S., GOVINDARAJU, V., AND RUDRAVARAM, P. Classification and indexing in large biometric databases. In *Biometrics Consortium Conference* (September 2004), NIST, pp. 1–3.

[137] PANDA, A. K., MEHROTRA, H., AND MAJHI, B. Parallel geometric hashing for robust iris indexing. *Journal of real-time image processing 8*, 3 (September 2013), 341–349.

[138] PARK, U., LIAO, S., KLARE, B., VOSS, J., AND JAIN, A. K. Face finder: Filtering a large face database using scars, marks and tattoos. Tech. Rep. MSU-CSE-11-15, Michigan State University, 2011.

[139] PAULINO, A. A., LIU, E., CAO, K., AND JAIN, A. K. Latent fingerprint indexing: Fusion of level 1 and level 2 features. In *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2013), IEEE, pp. 1–8.

[140] PAYNTER, T. Northrop grumman wins $95 million award from department of homeland security to develop next-generation biometric identification services system. `https://news.northropgrumman.com/news/releases/northrop-grumman-wins-95-million-award-from-department-of-homeland-security-to-develop-next-generation-biometric-identification-services-system`, February 2018. Last accessed: 2020–03–11.

[141] PERRONNIN, F., AND DUGELAY, J.-L. Clustering face images with application to image retrieval in large databases. In *Biometric Technology for Human Identification II* (2005), vol. 5779, International Society for Optics and Photonics, pp. 256–265.

[142] PFLUG, A., BUSCH, C., AND ROSS, A. 2D ear classification based on unsupervised clustering. In *International Joint Conference on Biometrics (IJCB)* (September 2014), IEEE, pp. 1–8.

[143] PFLUG, A., RATHGEB, C., SCHERHAG, U., AND BUSCH, C. Binarization of spectral histogram models: An application to efficient biometric identification. In *International Conference on Cybernetics (CYBCONF)* (June 2015), IEEE, pp. 501–506.

[144] PROENÇA, H. Iris biometrics: Indexing and retrieving heavily degraded data. *Transactions on Information Forensics and Security (TIFS) 8*, 12 (December 2013), 1975–1985.

[145] PROENÇA, H., AND NEVES, J. Iris biometric indexing. In *Iris and Periocular Biometric Recognition*. Institution of Engineering and Technology, July 2017, pp. 101–124.

[146] PUHAN, N. B., AND SUDHA, N. A novel iris database indexing method using the iris color. In *Conference on Industrial Electronics and Applications* (June 2008), IEEE, pp. 1886–1891.

[147] QIU, X., SUN, Z., AND TAN, T. Global texture analysis of iris images for ethnic classification. *International Conference on Biometrics (ICB) 3832* (January 2006), 411–418.

[148] RACHKOVSKIJ, D. A. Index structures for fast similarity search for binary vectors. *Cybernetics and Systems Analysis 53*, 5 (2017), 799–820.

[149] RACHKOVSKIJ, D. A. Index structures for fast similarity search for real-valued vectors I. *Cybernetics and Systems Analysis 54*, 1 (2018), 152–164.

[150] RACHKOVSKIJ, D. A. Index structures for fast similarity search for real vectors II. *Cybernetics and Systems Analysis 54*, 2 (2018), 320–335.

[151] RAGHAVENDRA, R., SURBIRYALA, J., AND BUSCH, C. An efficient finger vein indexing scheme based on unsupervised clustering. In *International Conference on Identity, Security and Behavior Analysis (ISBA)* (March 2015), IEEE, pp. 1–8.

[152] RAKVIC, R. N., ULIS, B. J., BROUSSARD, R. P., IVES, R. W., AND STEINER, N. Parallelizing iris recognition. *Transactions on Information Forensics and Security (TIFS) 4*, 4 (December 2009), 812–823.

[153] RATHA, N. K., KARU, K., CHEN, S., AND JAIN, A. K. A real-time matching system for large fingerprint databases. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 18*, 8 (August 1996), 799–813.

[154] RATHGEB, C., BREITINGER, F., BAIER, H., AND BUSCH, C. Towards bloom filter-based indexing of iris biometric data. In *International Conference on Biometrics (ICB)* (May 2015), IEEE, pp. 422–429.

[155] RATHGEB, C., BREITINGER, F., BUSCH, C., AND BAIER, H. On application of bloom filters to iris biometrics. *IET Biometrics 3*, 4 (January 2014), 207–218.

[156] RATHGEB, C., BUCHMANN, N., HOFBAUER, H., BAIER, H., UHL, A., AND BUSCH, C. Methods for accuracy-preserving acceleration of large-scale comparisons in CPU-based iris ecognition systems. *IET Biometrics 7*, 4 (July 2018), 356–364.

[157] RATHGEB, C., HOFBAUER, H., UHL, A., AND BUSCH, C. TripleA: Accelerated accuracy-preserving alignment for iris-codes. In *International Conference on Biometrics (ICB)* (June 2016), IEEE, pp. 1–8.

[158] RATHGEB, C., AND UHL, A. Iris-biometric hash generation for biometric database indexing. In *International Conference on Pattern Recognition (ICPR)* (August 2010), IEEE, pp. 2848–2851.

[159] RATHGEB, C., AND UHL, A. A survey on biometric cryptosystems and cancelable biometrics. *EURASIP Journal on Information Security 2011*, 1 (September 2011), 1–25.

[160] RATHGEB, C., UHL, A., AND WILD, P. Incremental iris recognition: A single-algorithm serial fusion strategy to optimize time complexity. *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2010), 1–6.

[161] RATHGEB, C., UHL, A., AND WILD, P. On combining selective best bits of iris-codes. In *European Workshop on Biometrics and Identity Management (BioID)* (March 2011), Springer, pp. 227–237.

[162] RIFE, D. C. Finger prints as criteria of ethnic relationship. *American journal of human genetics 5*, 4 (1953), 389.

[163] ROSS, A., AND MUKHERJEE, R. Augmenting ridge curves with minutiae triplets for fingerprint indexing. In *Biometric Technology for Human Identification IV* (April 2007), vol. 6539, International Society for Optics and Photonics, SPIE, pp. 1–12.

[164] ROSS, A., AND SUNDER, M. S. Block based texture analysis for iris classification and matching. In *Conference on Computer Vision and Pattern Recognition - Workshops (CVPRW)* (June 2010), IEEE, pp. 30–37.

[165] SAEGUSA, T., MARUYAMA, T., AND YAMAGUCHI, Y. How fast is an FPGA in image processing? In *International Conference on Field Programmable Logic and Applications* (September 2008), IEEE, pp. 77–82.

[166] SCHLETT, T., RATHGEB, C., AND BUSCH, C. A binarization scheme for face recognition based on multi-scale block local binary patterns. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2016), IEEE, pp. 1–4.

[167] SCHROFF, F., KALENICHENKO, D., AND PHILBIN, J. FaceNet: A unified embedding for face recognition and clustering. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2015), IEEE, pp. 815–823.

[168] SCHUCH, P. Survey on features for fingerprint indexing. *IET Biometrics 8*, 1 (January 2019), 1–13.

[169] SHUAI, X., ZHANG, C., AND HAO, P. Fingerprint indexing based on composite set of reduced SIFT features. In *International Conference on Pattern Recognition (ICPR)* (December 2008), IEEE, pp. 1–4.

[170] SIMMHAN, Y., SHUKLA, A., AND VERMA, A. Benchmarking fast-data platforms for the Aadhaar biometric database. In *Big Data Benchmarking*. Springer, 2015, pp. 21–39.

[171] SINGH, D., AND REDDY, C. K. A survey on platforms for big data analytics. *Journal of big data 2*, 1 (October 2015), 1–20.

[172] SINGH, M., NAGPAL, S., VATSA, M., SINGH, R., NOORE, A., AND MAJUMDAR, A. Gender and ethnicity classification of iris images using deep class-encoder. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 666–673.

[173] SINGH, M., SINGH, R., AND ROSS, A. A comprehensive overview of biometric fusion. *Information Fusion 52* (December 2019), 187–205.

[174] SIROWY, S., AND FORIN, A. Where's the beef? why FPGAs are so fast. Tech. Rep. MSR-TR-2008-130, Microsoft Research, September 2008.

[175] STEKAS, N., AND V. D. HEUVEL, D. Face recognition using local binary patterns histograms (LBPH) on an FPGA-based system on chip (SoC). In *International Parallel and Distributed Processing Symposium Workshops (IPDPSW)* (May 2016), IEEE, pp. 300–304.

[176] STONE, J. E., GOHARA, D., AND SHI, G. OpenCL: A parallel programming standard for heterogeneous computing systems. *Computing in Science Engineering 12*, 3 (May 2010), 66–73.

[177] SUN, Z., ZHANG, H., TAN, T., AND WANG, J. Iris image classification based on hierarchical visual codebook. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 36*, 6 (June 2014), 1120–1133.

[178] SURBIRYALA, J., RAGHAVENDRA, R., AND BUSCH, C. Finger vein indexing based on binary features. In *Colour and Visual Computing Symposium (CVCS)* (August 2015), IEEE, pp. 1–6.

[179] TANG, D., HUANG, B., LI, R., AND LI, W. A person retrieval solution using finger vein patterns. In *International Conference on Pattern Recognition (ICPR)* (August 2010), IEEE, pp. 1306–1309.

[180] TAPIA, J. E., PEREZ, C. A., AND BOWYER, K. W. Gender classification from iris images using fusion of uniform local binary patterns. In *European Conference on Computer Vision (ECCV)* (March 2014), Springer, Springer, pp. 751–763.

[181] TESSIER, R., POCEK, K., AND DEHON, A. Reconfigurable computing architectures. *Proc. of the IEEE 103*, 3 (2015), 332–354.

[182] THAKKAR, D. Global biometric market analysis: Trends and future prospects. https://www.bayometric.com/global-biometric-market-analysis/, August 2018. Last accessed: 2020–03–11.

[183] UHL, A., MARCEL, S., BUSCH, C., AND VELDHUIS, R. N. J. *Handbook of Vascular Biometrics*. Springer, 2019.

[184] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Aadhaar dashboard. https://www.uidai.gov.in/aadhaar_dashboard/. Last accessed: 2020–03–11.

[185] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Role of biometric technology in Aadhaar enrollment. Tech. rep., UIDAI, January 2012.

[186] VANDAL, N. A., AND SAVVIDES, M. CUDA accelerated iris template matching on graphics processing units (GPUs). In *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2010), IEEE, pp. 1–7.

[187] WANG, D., AND JAIN, A. K. Face retriever: Pre-filtering the gallery via deep neural net. In *International Conference on Biometrics (ICB)* (May 2015), IEEE, pp. 473–480.

[188] WANG, D., OTTO, C., AND JAIN, A. K. Face search at scale. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 39*, 6 (June 2017), 1122–1136.

[189] WANG, J., SHEN, H. T., SONG, J., AND JI, J. Hashing for similarity search: A survey. *arXiv preprint arXiv:1408.2927* (2014), 1–29.

[190] WANG, K., YANG, L., SU, K., YANG, G., AND YIN, Y. Binary search path of vocabulary tree based finger vein image retrieval. In *International Conference on Biometrics (ICB)* (2016), IEEE, IEEE, pp. 1–8.

[191] WANG, Y., HU, J., AND PHILLIPS, D. A fingerprint orientation model based on 2D fourier expansion (FOMFE) and its application to singular-point detection and fingerprint indexing. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 29*, 4 (April 2007), 573–585.

[192] WANG, Y., WAN, J., GUO, J., CHEUNG, Y., AND YUEN, P. C. Inference-based similarity search in randomized Montgomery domains for privacy-preserving biometric identification. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 40*, 7 (July 2018), 1611–1624.

[193] WANG, Y., WANG, L., CHEUNG, Y. M., AND YUEN, P. C. Learning compact binary codes for hash-based fingerprint indexing. *Transactions on Information Forensics and Security (TIFS) 10*, 8 (August 2015), 1603–1616.

[194] WAYMAN, J. L. Multifinger penetration rate and ROC variability for automatic fingerprint identification systems. In *Automatic Fingerprint Recognition Systems*. Springer, 2004, pp. 305–316.

[195] WOLFSON, H. J., AND RIGOUTSOS, I. Geometric hashing: An overview. *Computational science and engineering 4*, 4 (1997), 10–21.

[196] WU, X., KUMAR, V., QUINLAN, J. R., GHOSH, J., YANG, Q., MOTODA, H., MCLACHLAN, G. J., NG, A., LIU, B., PHILIP, S. Y., ET AL. Top 10 algorithms in data mining. *Knowledge and information systems 14*, 1 (2008), 1–37.

[197] WU, Z., KE, Q., SUN, J., AND SHUM, H.-Y. Scalable face image retrieval with identity-based quantization and multi-reference reranking. In *Conference on. Computer Vision and Pattern Recognition (CVPR)* (2010), IEEE, pp. 3469–3476.

[198] XU, H., VELDHUIS, R. N. J., KEVENAAR, T. A. M., AKKERMANS, A. H. M., AND BAZEN, A. M. Spectral minutiae: A fixed-length representation of a minutiae set. In *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (June 2008), IEEE, pp. 1–6.

[199] YANG, J. C., AND PARK, D. S. A fingerprint verification algorithm using tessellated invariant moment features. *Neurocomputing 71*, 10-12 (2008), 1939–1946.

[200] YI, D., LEI, Z., HU, Y., AND LI, S. Z. Fast matching by 2 lines of code for large scale face recognition systems. *arXiv preprint arXiv:1302.7180* (2013).

[201] YI, S., YOON, I., OH, C., AND YI, Y. Real-time integrated face detection and recognition on embedded GPGPUs. In *Symposium on Embedded Systems for Real-time Multimedia (ESTIMedia)* (October 2014), IEEE, pp. 98–107.

[202] YOU, J., KONG, W.-K., ZHANG, D., AND CHEUNG, K. H. On hierarchical palmprint coding with multiple features for personal identification in large databases. *Transactions on Circuits and Systems for Video Technology (TCSVT) 14*, 2 (February 2004), 234–243.

[203] YU, L., ZHANG, D., WANG, K., AND YANG, W. Coarse iris classification using box-counting to estimate fractal dimensions. *Pattern Recognition 38*, 11 (2005), 1791–1798.

[204] YUAN, B., SU, F., AND CAI, A. Fingerprint retrieval approach based on novel minutiae triplet features. In *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2012), IEEE, pp. 170–175.

[205] ZEZULA, P., AMATO, G., DOHNAL, V., AND BATKO, M. *Similarity search: the metric space approach*. Springer, 2006.

[206] ZHAO, Q. A new approach for noisy iris database indexing based on color information. In *International Conference on Computer Science Education (ICCSE)* (August 2011), IEEE, pp. 28–31.

[207] ZHENG, R., ZHANG, C., HE, S., AND HAO, P. A novel composite framework for large-scale fingerprint database indexing and fast retrieval. In *International Conference on Hand-Based Biometrics (ICHB)* (November 2011), IEEE, pp. 1–6.

[208] ZHOU, Y., LIU, Y., FENG, Q., YANG, F., HUANG, J., AND NIE, X. Palm-vein classification based on principal orientation features. *PLOS ONE 9*, 11 (November 2014), 1–12.

Part III

Research Articles

Chapter 5

Database Binning and Retrieval in Multi-Fingerprint Identification Systems

Abstract

The increasingly large scale of deployed biometric systems necessitates approaches for computational workload reduction in order to perform identification queries efficiently. Simple database binning based on classification of features in biometric samples is amongst the most frequently used and researched methods for achieving said goal. However, multi-instance database binning appears to be a neglected topic in the scientific literature: best to the authors' knowledge, for fingerprints there exists only one, entirely theoretical, study on this subject. In this paper, we propose a retrieval algorithm based on multi-instance binning of fingerprint databases, along with usage of statistical information on fingerprint classes and their correlations.

The aforementioned statistics are obtained from NIST SD9 database and data obtained from the German Federal Criminal Police Office. Subsequently, the experimental evaluation of the proposed algorithm is performed on the NIST SD9 database. The proposed system is evaluated using a classifier based on the PCASYS tool and neuronal networks. The results show a significant workload reduction from a baseline exhaustive search scenario – down to 12.7% for this particular classifier and 5.8% for a theoretical perfect (completely accurate) classifier. The proposed method could be seamlessly integrated into operational systems, as it relies on well-established features and compatibility with the current acquisition methods.

Addressed research question(s): RQ1, RQ3, RQ4

Reference: DROZDOWSKI, P., FISCHER, D., RATHGEB, C., SCHIEL, C., AND BUSCH, C. Database binning and retrieval in multi-fingerprint identification systems. In *International Workshop on Information Forensics and Security (WIFS)* (December 2018), IEEE, pp. 1–7.

5.1 Introduction

Nowadays, biometric technologies are already deployed in numerous nationwide large-scale applications, such as the Indian Aadhaar project [21]. With the rapid growth of biometric systems' sizes and popularity, technologies supporting efficient and accurate processing of large amounts of biometric data are vital in order to guarantee practical response times. Conventional biometric systems require exhaustive one-to-many comparisons in order to identify biometric probes, *i.e.* comparison time frequently dominates the overall computational workload of an identification attempt. In past years, researchers have invested significant efforts to tackle the challenge of computational workload reduction in biometric identification systems. Basically, four different key concepts can be distinguished: *classification* or "binning", *indexing*, a *serial combination* of a computationally efficient and an accurate (but more complex) algorithm and *hardware-based acceleration*. Depending on the used biometric characteristic, the vast majority of classification approaches are designed to reliably extract human understandable attributes from a biometric sample, *e.g.* sex or ethnicity for face. While not necessarily unique to an individual, those attributes allow for a binning of biometric databases according to a predefined number of classes, *i.e.* the search space ($\widehat{=}$ computational workload) for a given biometric probe can be reduced to one (or a few) bin(s). In contrast, biometric indexing approaches introduce hierarchical search structures (tolerating a certain amount of biometric variance), where the process of search space reduction might not be reproducible by human experts. Lastly, the latter two categories do not aim at reducing the complexity of an identification attempt but response times.

Focusing on fingerprint recognition systems, the classification model of Henry [12] has been widely used by researchers, as well as commercial vendors, for computational workload reduction in identification scenarios. The five fingerprint classes (or types), *i.e.* *arch*, *tented arch*, *right loop*, *left loop* and *whorl*, which are depicted in figure 5.1, are unevenly distributed in the population. Fingerprint classes are mainly determined based on the global (level-1) features, in particular ridge line flow (orientation map) and the singular points, *i.e.* core and delta, derived from it. Numerous approaches, which either directly employ or further process those features, have been proposed for the purpose of distinguishing between said classes. For more details on the topic of fingerprint classification and a comprehensive survey of proposed approaches, the reader is referred to [9, 10]. State-of-the-art fingerprint classification schemes obtain near-optimal classification accuracy. Table 5.1 summarises most notable approaches and reported results in terms of Correct Classification Rate (CCR) of the last five years. Note, that all of these classification approaches aim to determine the class of a *single* fingerprint.

As opposed to the existing literature, this paper investigates fingerprint classification in multi-finger identification systems. This is motivated by the

Table 5.1: Most relevant fingerprint classification approaches proposed in the last five years

Ref.	Year	Method	Database(s)	Classes	CCR	Reject
[4]	2013	MCC	SD4	4 / 5	97.2% / 95.9%	—
[11]	2014	FCA	FVC00/02/04	4	92.74%	—
[15]	2014	KNN	SD4	4 / 5	96.8% / 94.6%	—
[24]	2014	FCA	FVC02-1	5	91.1%	—
			FVC04-1	5	91.8%	—
[8]	2015	FCA	SD4	5	80.51%	12%
			FVC02-1	5	90.11%	—
			FVC04-1	5	88.98%	—
[14]	2015	RDM	FVC00	4	91.1%	—
			FVC02	4	97.8%	—
			FVC04	4	97.3%	—
[7]	2016	FCA	SD4	4 / 5	88.3% / 92.13%	—
[22]	2016	ANN	SD4	4	91.4% / 93.1%	—
[2]	2017	ANN	FVC2000	-	97.56%	—
[19]	2017	MCC	SD4	5	92.97%	—
			SD14	5	93.76%	—
			SFinGe	5	94.38%	—

MCC . . . multiple classifier combination
FCA . . . fixed classifier approach
KNN . . . k-nearest neighbour
RDM . . . ridge distribution models
ANN . . . artificial neuronal networks

facts that large-scale identification systems leverage the information of multiple fingerprints of data subjects, *e.g.* [21], and modern fingerprint capture devices can acquire multiple fingerprints of data subject's hand simultaneously, *e.g.* [13]. It is well-known that the classes of fingerprints obtained from one hand are highly correlated. Nevertheless, to the best of the authors' knowledge, the potential of multi-fingerprint database binning has only been theoretically analysed by Wayman [23]. It was confirmed that bins formed by combinations of fingerprint classes highly vary in probability. Moreover, theoretical estimations about expected penetration rates are reported. However, so far the potential of improving the overall fingerprint retrieval accuracy by consolidating information obtained from single fingerprint classification scores has been neglected. In this work, we obtain universally valid statistics of fingerprint class distributions and correlations from two datasets, namely the NIST SD9 [17] and an in-house database of the German Federal Criminal Police Office (BKA). Those statistics are used to effectively retrieve bins representing combinations of fingerprint classes according to their likelihood. In experiments on the SD9 database the well-established, publicly available Pattern-level Classification Automation SYStem (PCASYS) tool [3] in conjunction with a neuronal network-based classifier are employed for the purpose of fingerprint classification. The proposed approach is shown to substantially reduce the computational workload by

combining the classifier scores obtained from up to five fingers of a data subject's hand.

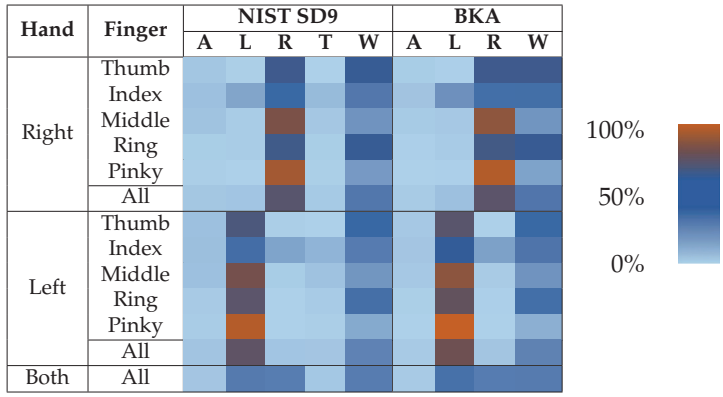
The remainder of this paper is organised as follows: fingerprint class distributions and correlations are analysed in section 5.2. Section 5.3 describes the proposed multi-finger binning and retrieval approach. Experimental results are reported in section 5.4. Conclusions are drawn in section 5.5.

Table 5.2: Fingerprint class distributions

(a) Percentages

Hand	Finger	NIST SD9					BKA			
		A	L	R	T	W	A	L	R	W
Right	Thumb	3.49%	0.71%	48.94%	0.22%	46.64%	1.49%	0.57%	49.02%	48.92%
	Index	5.61%	14.72%	39.43%	7.06%	33.18%	4.23%	22.41%	36.62%	36.74%
	Middle	4.76%	1.30%	69.48%	2.94%	21.52%	2.17%	2.66%	74.24%	20.93%
	Ring	1.19%	1.41%	49.61%	1.19%	46.60%	0.65%	1.56%	50.84%	46.95%
	Pinky	0.93%	0.19%	79.41%	0.82%	18.65%	0.38%	0.56%	83.47%	15.59%
	All	3.20%	3.66%	57.37%	2.45%	33.32%	1.79%	5.55%	58.84%	33.82%
Left	Thumb	5.50%	53.31%	0.93%	0.48%	39.78%	2.60%	57.75%	0.42%	39.23%
	Index	5.84%	37.72%	15.17%	9.63%	31.64%	3.60%	45.48%	16.35%	34.57%
	Middle	5.61%	67.36%	1.49%	5.06%	20.48%	2.66%	74.01%	1.70%	21.63%
	Ring	1.90%	58.66%	0.48%	1.67%	37.29%	0.81%	62.00%	0.62%	36.57%
	Pinky	1.26%	83.95%	0.22%	1.08%	13.49%	0.47%	88.07%	0.19%	11.27%
	All	4.02%	60.20%	3.66%	3.58%	28.54%	2.03%	65.46%	3.86%	28.65%
Both	All	3.61%	31.93%	30.52%	3.01%	30.93%	1.91%	35.76%	31.11%	31.22%

(b) Heatmap

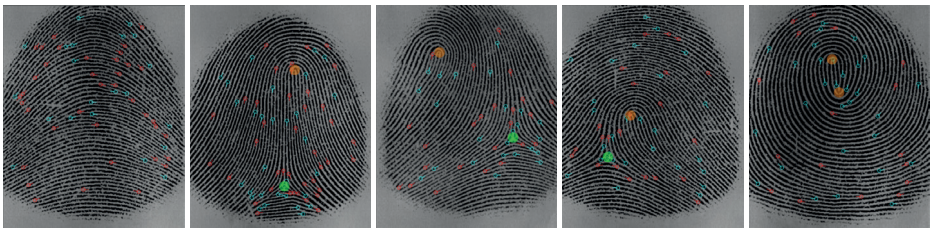


5.2 Fingerprint class statistics

In the following subsections, the used databases are presented, along with statistical distributions of fingerprint classes and their correlations.

5.2.1 Databases

Two databases were used for experiments in this paper:



(a) Arch (A) (b) Tented arch (T) (c) Left loop (L) (d) Right loop (R) (e) Whorl (W)

Figure 5.1: Example fingerprints for each of the five classes displaying minutiae, core and delta points (images generated using Synthetic Fingerprint Generator (SFinGe) [5])

Figure 5.2: Sample images from the SD9 database

SD9 NIST Special Database 9 [17], containing fingerprint images from scanned rolled-ink ten-print cards. 2 samples per finger are available for each of the 2,700 subjects, hence the total number of images is 54,000. Fingerprint class annotations made by professional forensic examiners are included. Example images from the database are shown in figure 5.2.

BKA A subset of the Automated Fingerprint Identification Systems (AFIS) data of the BKA consisting of fingerprint type statistical data from around 26,000 randomly selected subjects. Due to lack of actual images (data protection restrictions), this dataset was only used to validate the statistical results obtained on SD9 and not the computational workload reduction experiments. The data does not distinguish between arches and tented arches; instead classifying them together into one class.

The subjects in both databases were selected from their respective AFIS' randomly, hence ensuring a natural distribution of the fingerprint classes.

5.2.2 Distributions and Correlations

For the statistical analysis, only a single (first) sample from each finger is considered in order to avoid using redundant information. The class distri-

Table 5.3: Distributions of fingerprint class combinations for two contiguous fingers

Hand	Thumb, index			Index, middle			Middle, ring			Ring, pinky		
	Cls.	SD9	BKA	Cls.	SD9	BKA	Cls.	SD9	BKA	Cls.	SD9	BKA
Right	WW	26.53%	25.35%	RR	35.12%	36.66%	RR	41.66%	46.58%	RR	46.93%	49.70%
	RR	25.94%	26.31%	WW	16.69%	17.04%	RW	25.75%	26.00%	WR	29.73%	30.93%
	WR	11.82%	12.33%	WR	15.98%	19.14%	WW	19.47%	18.74%	WW	16.76%	14.69%
	RL	8.21%	11.57%	LR	10.33%	14.89%	AR	2.90%	1.76%	RW	1.67%	1.47%
	RW	6.43%	11.11%	TR	5.65%	—	TR	2.30%	—	TR	1.11%	—
	WL	6.02%	7.14%	LW	2.45%	2.07%	WR	2.04%	1.81%	LR	1.04%	1.13 %
	RT	5.13%	—	AA	2.45%	1.28%	RL	1.15%	0.94%	AR	0.59%	0.54%
	RA	3.23%	2.64%	AR	2.42%	3.01%	AA	0.93%	0.54%	RT	0.52%	—
	WT	1.64%	—	RW	2.16%	1.53%	RT	0.74%	—	AA	0.41%	0.11%
	AA	1.45%	0.71%	RT	1.00%	—	LR	0.71%	1.79%	RA	0.37%	0.37%
	Other	3.60%	2.84%	Other	5.75%	4.38%	Other	2.35%	1.84%	Other	0.87%	1.06%
Left	LL	24.05%	32.50%	LL	32.42%	40.77%	LL	46.51%	52.45%	LL	55.58%	58.05%
	WW	21.56%	20.83%	WW	16.17%	17.17%	LW	19.41%	18.74%	WL	25.69%	27.18%
	WL	11.34%	11.99%	WL	14.91%	16.60%	WW	17.10%	18.34%	WW	11.38%	10.41%
	LR	9.59%	9.44%	RL	10.71%	12.53%	TL	4.13%	—	LW	2.04%	1.60%
	LW	9.48%	13.13%	TL	7.21%	—	AL	3.42%	2.69%	TL	1.49%	—
	LT	7.10%	—	AA	2.94%	1.83%	WL	3.38%	4.01%	AL	0.93%	0.63%
	WR	4.91%	5.69%	LW	2.64%	3.32%	AA	1.38%	0.69%	AA	0.71%	0.31%
	LA	3.09%	2.58%	AL	2.12%	2.32%	RL	1.23%	1.09%	LT	0.48%	—
	AA	2.04%	0.86%	RW	1.45%	1.95%	LT	0.71%	—	LA	0.45%	0.43%
	AL	1.93%	1.32%	RT	1.45%	—	AT	0.56%	—	AT	0.26%	—
	Other	4.91%	1.66%	Other	7.98%	3.51%	Other	2.17%	1.99%	Other	0.99%	1.39%

Table 5.4: Distributions of fingerprint class combinations for three contiguous fingers

Hand	Thumb, index, middle			Index, middle, ring			Middle, ring, pinky		
	Cls.	SD9	BKA	Cls.	SD9	BKA	Cls.	SD9	BKA
Right	RRR	23.78%	24.47%	RRR	24.19%	25.43%	RRR	39.69%	44.77%
	WWW	15.01%	14.09%	WWW	15.46%	15.76%	RWR	18.06%	19.05%
	WWR	11.11%	11.00%	WRW	10.89%	11.14%	WWR	10.78%	10.94%
	WRR	10.00%	11.34%	RRW	9.92%	10.63%	WWW	8.70%	7.77%
	RLR	6.61%	9.84%	LRR	6.54%	10.86%	RWW	7.69%	6.72%
	RWR	4.72%	8.00%	WRR	5.02%	7.91%	ARR	2.42%	1.62%
	RTR	4.27%	—	TRR	4.46%	—	TRR	2.19%	—
	WLR	3.42%	4.82%	LRW	3.49%	3.84%	WRR	1.97%	1.67%
	WLW	2.04%	1.36%	LWW	2.04%	1.70%	RRW	1.52%	1.30%
	RWW	1.60%	2.89%	RWW	1.82%	1.28%	RLR	0.82%	0.77%
	Other	17.44%	12.19%	Other	16.17%	11.45%	Other	6.16%	5.39%
Left	LLL	21.26%	29.64%	LLL	23.35%	31.50%	LLL	44.24%	50.82%
	WWW	12.08%	12.36%	WWW	14.13%	14.54%	LWL	14.28%	14.85%
	WLL	9.37%	9.99%	RLL	8.85%	10.27%	WWL	10.97%	11.87%
	WWL	9.11%	8.27%	LLW	8.62%	8.67%	WWW	6.13%	6.47%
	LRL	7.10%	8.07%	WLW	8.40%	7.87%	LWW	4.94%	3.83%
	LWL	5.43%	8.13%	WLL	6.51%	8.64%	TLL	3.98%	—
	LTL	5.32%	—	TLL	5.99%	—	WLL	3.20%	3.75%
	LWW	3.87%	4.81%	LWW	2.12%	2.40%	ALL	3.01%	2.46%
	WRL	3.20%	4.18%	WWL	2.04%	2.55%	LLW	1.71%	1.34%
	ALL	1.52%	2.46%	ALL	1.82%	2.03%	RLL	1.15%	1.03%
	Other	21.74%	12.09%	Other	18.17%	11.53%	Other	6.39%	3.58%

butions for the SD9 and BKA datasets can be seen in table 5.2a, while table 5.2b presents the same information graphically in a heatmap format. It can be observed, that the loop classes are the most prevalent; overwhelmingly, their direction corresponds to the hand of the given finger (left loops on left hand and analogously for the right hand). Together with whorls, they account for around 95% of the total samples. The fingerprint class distri-

butions (both for overall percentages and individual fingers) obtained from both datasets tend to coincide. The largest (relative) discrepancies can be seen for the arches and tented arches. Overall, however, the findings from both datasets are very similar, which suggests the generality of the results.

As previously mentioned, the classes of fingers of one hand are also known to be correlated. Tables 5.3 to 5.6 show the occurrence frequency for the most prevalent (top 10 from SD9) class combinations between contiguous sequences of two to five adjacent fingers of each hand. For example, RL corresponds to a right and left loop fingerprint class (recall figure 5.1), for the pair of fingers noted in the table header. It can be observed, that combinations of loops and whorls again account for the vast majority of cases; furthermore, it is very often the case that same fingerprint classes are seen across multiple or even all fingers of a given hand. As was the case for the class distributions described earlier, the results for the class combinations within both datasets tend to largely coincide.

Table 5.5: Distributions of fingerprint class combinations for four contiguous fingers

Hand	Thumb, index, middle, ring			Index, middle, ring, pinky		
	Cls.	SD9	BKA	Cls.	SD9	BKA
Right	RRRR	18.99%	18.34%	RRRR	23.19%	24.35%
	WWWW	14.12%	13.21%	WWWR	8.18%	8.70%
	WWRW	8.10%	6.72%	WWWW	7.28%	7.03%
	WRRW	5.46%	4.82%	RRWR	6.91%	7.97%
	RLRR	5.02%	7.80%	WRWR	6.76%	7.57%
	WRRR	4.31%	6.44%	LRRR	6.21%	10.69%
	RRRW	4.12%	5.61%	WRRR	4.50%	7.54%
	RTRR	3.53%	—	TRRR	4.38%	—
	WWRR	3.01%	4.05%	WRWW	4.12%	3.52%
	RWRW	2.71%	4.14%	LRWR	3.08%	3.09%
	Other	30.63%	28.87%	Other	25.39%	19.54%
Left	LLLL	16.51%	24.14%	LLLL	22.08%	30.70%
	WWWW	10.86%	10.79%	WWWL	8.85%	9.36%
	LRLL	6.32%	6.87%	RLLL	8.70%	9.99%
	WWLW	5.87%	4.29%	LLWL	6.39%	7.18%
	WLLL	5.61%	6.44%	WLLL	6.10%	8.21%
	LTLL	4.65%	—	WLWL	5.91%	5.89%
	LLLW	4.54%	5.04%	TLLL	5.76%	—
	WLLW	3.64%	3.49%	WWWW	5.28%	5.18%
	WWLL	3.23%	3.95%	WLWW	2.45%	1.95%
	LWWW	3.12%	3.75%	LLWW	2.16%	1.46%
	Other	35.65%	31.24%	Other	26.32%	20.08%

5.3 Multi-fingerprint binning and retrieval

Figure 5.3 shows the overview of the proposed multi-fingerprint binning and retrieval algorithm. The binning step consists of enumerating all possible bins based on the combinations of fingerprint classes and accordingly assigning the data from the enrolled subjects to the bins. Subsections 5.3.1

Table 5.6: Distributions of fingerprint class combinations for five contiguous fingers

Hand	Thumb, index, middle, ring, pinky		
	Cls.	SD9	BKA
Right	RRRRR	18.43%	17.75%
	WWWWR	7.28%	6.97%
	WWWWW	6.84%	6.21%
	WWRWR	4.76%	4.28%
	RLRRR	4.76%	7.68%
	WRRRR	3.86%	6.10%
	RTRRR	3.53%	—
	RRRWR	3.42%	4.20%
	WWRWW	3.34%	2.61%
	WRRWR	3.23%	3.57%
	Other	40.55%	40.63%
Left	LLLLL	16.06%	23.58 %
	WWWWL	6.51%	6.61%
	LRLLL	6.25%	6.72%
	WLLLL	4.87%	6.21%
	LTLLL	4.42%	—
	WWWWW	4.35%	4.18%
	WWLWL	3.90%	3.29%
	LLLWL	3.61%	4.29%
	WWLLL	2.94%	3.69%
	LWLLL	2.94%	4.41%
	Other	44.15%	37.02%

and 5.3.2 describe the combination and adjustment of classifier outputs, as well as the utilised retrieval strategy.

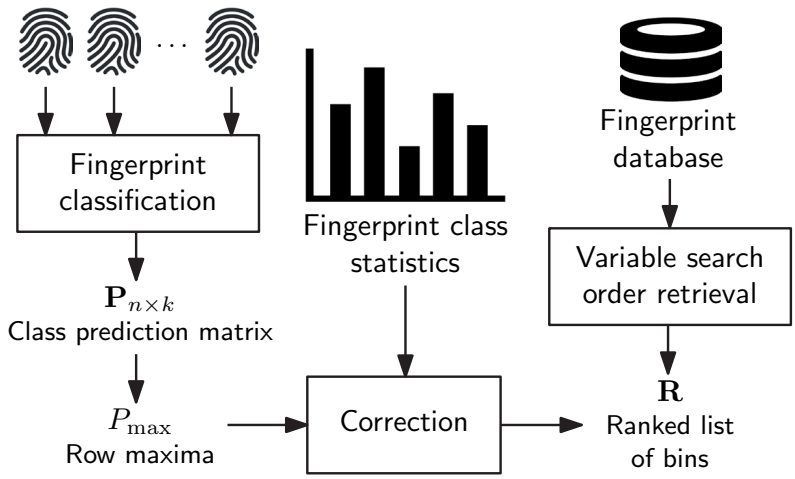


Figure 5.3: System overview

5.3.1 Combining classifier outputs

1. For each of the $\{n|2 \leq n \leq 5\}$ considered fingers of a hand, the classifier produces a list of k classification probabilities (p), where k is the number of possible fingerprint classes and $\sum_{i=0}^{k} p_i = 1$. In other words, a $\mathbf{P}_{n \times k}$ matrix of class predictions is obtained for the given hand, which may, for example, look as follows for $n = 5$ and $k = 4$:

$$\mathbf{P} = \begin{array}{c} \\ \text{Thumb} \\ \text{Index} \\ \text{Middle} \\ \text{Ring} \\ \text{Pinky} \end{array} \begin{array}{c} \begin{array}{cccc} \text{A} & \text{L} & \text{R} & \text{W} \end{array} \\ \begin{bmatrix} 4\% & 92\% & 1\% & 3\% \\ 1\% & 2\% & 1\% & 96\% \\ 5\% & 76\% & 2\% & 17\% \\ 1\% & 0\% & 0\% & 99\% \\ 1\% & 12\% & 2\% & 85\% \end{bmatrix} \end{array} \tag{5.1}$$

2. For each finger, the probability of the most probable class is determined (*i.e.* the row-wise maximum values). In this case:

$$P_{max} = \begin{array}{c} \begin{array}{ccccc} \text{Thumb} & \text{Index} & \text{Middle} & \text{Ring} & \text{Pinky} \end{array} \\ \begin{bmatrix} 92\% & 96\% & 76\% & 99\% & 85\% \end{bmatrix} \end{array} \tag{5.2}$$

3. All possible combinations (Cartesian product) of class labels for all fingers and corresponding probabilities taken from $\mathbf{P}$ are recorded in matrix $\mathbf{B}_{k^n \times n}$:

$$\mathbf{B} = \begin{array}{c} \text{Bin} \\ AAAAA \\ AAAAL \\ \ldots \\ WWWWW \end{array} \begin{array}{c} \begin{array}{ccccc} \text{Thumb} & \text{Index} & \text{Middle} & \text{Ring} & \text{Pinky} \end{array} \\ \begin{bmatrix} 4\% & 1\% & 5\% & 1\% & 1\% \\ 4\% & 1\% & 5\% & 1\% & 12\% \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 3\% & 96\% & 17\% & 99\% & 85\% \end{bmatrix} \end{array} \tag{5.3}$$

   For example, the second bin ($AAAAL$) consists of the A class probabilities for the thumb, index, middle, and ring fingers, along with the L class probability for the pinky finger taken from $\mathbf{P}$.

4. In order to fix unreliable classifier outputs, for each row (*i.e.* possible bin) in the above matrix, a classification correction algorithm is ran. It works based on the previously described correlation statistics (subsection 5.2.2). For every classification probability in the given bin, if the value is below a threshold (previously estimated on a disjoint training set), the statistical data is used to adjust it. In the current case ($n = 5$), data from table 5.6 is retrieved. For instance, if the algorithm is correcting the second (index finger) probability in the bin $AAAAA$:

a) The statistics for itself, along with $ALAAA$, $ARAAA$, and $AWAAA$ would be retrieved.

b) Subsequently, the sum of the statistical probabilities (p) for each of those bins is computed, *i.e.* $s = \text{sum}\,(p_{AAAAA}, p_{ALAAA}, p_{ARAAA}, p_{AWAAA})$.

c) Finally, the probability for the bin is divided by said sum and normalised by the maximum probability for the finger currently under processing (derived in step 2), *i.e.* $\frac{p_{AAAAA}}{s} * P_{max}(index)$.

In other words, the probabilities below the threshold are considered to contain no useful/significant information regarding the classification output, so the global (statistical) information is incorporated in a normalised manner to complement the classifier.

5. The probabilities are summed row-wise, and normalised to the interval $[0, 1] \in \mathbb{R}$. Thus, for each bin in the final list, the overall probability that it matches the fingerprint classes of the probe is recorded.

$$\mathbf{O} = \begin{array}{c} \text{Bin} \\ AAAAA \\ AAAAL \\ \cdots \\ WWWWW \end{array} \begin{array}{c} \text{Probability} \\ \left[\begin{array}{c} 2\% \\ 0.5\% \\ \cdots \\ 50\% \end{array}\right] \end{array} \tag{5.4}$$

5.3.2 Retrieval strategy

The *variable search order* strategy [16] is employed in the retrieval step. The previously acquired list of bin probabilities ($\mathbf{O}$) for the probe is first sorted in descending order of bin occurrence probability, thus producing a ranked list of bins ($\mathbf{R}$). Subsequently, the corresponding bins in the enrolment database are successively searched using the one-to-first strategy, *i.e.* until a match is found, whereupon the retrieval is concluded immediately.

5.4 Performance evaluation

The following subsections describe the experimental setup, the used fingerprint classification method, and the obtained results.

5.4.1 Experimental setup

Performance evaluations are conducted on the previously described SD9 database. A ten-fold cross-validation with randomly chosen disjoint training (20%) and test (80%) sets is performed using scikit-learn [18]. Classification accuracy is measured in terms of CCR, while the computational workload reduction is estimated in terms of the number of visited database bins and corresponding subjects for the identification transactions.

5.4.2 Feature extractor and classifiers

To facilitate the reproducibility of presented results, the publicly available PCASYS tool is employed for fingerprint classification. Extracted feature vectors comprise 128 elements, which are further processed using the Keras Framework [6] with Tensorflow 1.7 [1]. In order to obtain a suitable classifier input feature vector, the elements are normalised and scaled to the range $[-2, 2] \in \mathbb{R}$ using the Keras MinMaxScaler function. For the classification task, a neuronal network-based classifier[1], *i.e.* a Multi-Layer Perceptron, is trained. The network consists of three (hidden) dense layers (192/64/32 nodes), each with a ReLU activation kernel, which are initialised with the RandomNormal initialiser. The output layer comprises four nodes and is initialised with zeroes. Four fingerprint classes are used, *i.e.* arch and tented arch are represented as one class due to their rare occurrences. The output of the classifier is determined by the Sigmoid activation function. In the training (learning) step, a stochastic gradient descent is used with a learning rate of 0.005, a beta1 of 0.95 and a beta2 of 0.999. Training feature vectors are shuffled once and subsequently 150 epochs are performed with a batch size of 64.

Table 5.7: CCR at a confidence interval of 95% for the classification of single fingerprints

Class	Mean	Lower bound	Upper bound
A	63.50%	61.02%	65.98%
L	90.95%	90.11%	91.78%
R	90.19%	89.50%	90.88%
W	86.73%	85.69%	87.78%

5.4.3 Results

The performance of the employed method for the single fingerprint classification task is summarised in table 5.7. Compared to the current state-of-the-art (*cf.* table 5.1), the applied fingerprint classification scheme achieves a moderate accuracy. Particularly, a significantly lower CCR can be observed for the arch class, which results from natural (unbalanced) fingerprint class distribution in the training data. The resulting workload reduction obtained in a single-finger binning and retrieval strategy is listed in table 5.8 (average values are given for "visited" bins and "visited subjects").

With respect to the estimation of the maximum computational workload reduction (denoted "best possible"), a perfect fingerprint recognition sys-

[1]Parameters of the DNN-based classifier are summarised according to the guidelines provided by the IEEE Signal Processing Society.

Table 5.8: Single-finger binning and retrieval results

Hand	Finger	Visited bins	Visited subjects	% of naïve	Best possible
Right	Thumb	1.13	725.3	53.7%	45.9%
	Index	1.19	524.6	38.8%	30.7%
	Middle	1.14	821.7	60.8%	53.6%
	Ring	1.12	731.1	54.1%	46.5%
	Pinky	1.11	995.3	73.7%	66.6%
Left	Thumb	1.12	728.0	53.9%	44.7%
	Index	1.22	544.6	40.3%	29.1%
	Middle	1.15	806.7	59.7%	50.9%
	Ring	1.16	792.7	58.7%	48.6%
	Pinky	1.14	1,102.1	81.6%	72.4%

tem is assumed. This means, the retrieval is considered successful, when the fingerprints of the correct identity are reached in a closed set identification. This assumption is reasonable considering the accuracy reported for multi-fingerprint recognition systems [20]. Obtained workload reduction (denoted "% of naïve") for the proposed multi-finger binning and retrieval strategy for different number of contiguous fingers used (as described in subsection 5.2.2), and combinations thereof is summarised in table 5.9. The computational workload reduction is estimated by comparing the proposed scheme to a naïve system performing an exhaustive one-to-many search. The results for multi-fingerprint binning represent a significant improvement over a conventional single-finger binning, *cf.* table 5.8. Additionally, the computational workload could be further reduced, by employing a more accurate classifier

5.5 Conclusion

With the consistently growing size of deployed biometric databases, the need to reduce the computational requirements of the biometric identification scenario is clear. One of the popularly employed methods is arranging the enrolled database into bins based on the samples' tangible features (such as fingerprint classes). By doing so, during an identification transaction, only a small fraction of bins (and thereby biometric references) needs to be visited during the retrieval step. In this paper, the idea of multi-instance fingerprint binning is explored. Best to the author's knowledge, this is a neglected topic in the scientific literature, with only a single theoretical analysis done in the past.

In the proposed system, the classifier outputs for multiple fingers of one hand are combined and adjusted with statistical information about the oc-

Table 5.9: Multi-finger binning and retrieval results

Nr. fingers	Hand	First finger	Visited bins	Visited subjects	% of naïve	Best possible
2	Right	Thumb	1.47	349.4	25.9%	17.9%
		Index	1.49	384.9	28.5%	20.2%
		Middle	1.38	499.5	37.0%	28.3%
		Ring	1.33	584.3	43.3%	33.8%
	Left	Thumb	1.50	345.0	25.5%	15.4%
		Index	1.56	392.4	29.1%	18.3%
		Middle	1.46	545.2	40.4%	29.4%
		Ring	1.43	696.7	51.6%	39.1%
3	Right	Thumb	2.12	261.9	19.4%	12.1%
		Index	2.14	266.0	19.7%	12.4%
		Middle	1.84	419.7	31.1%	22.0%
	Left	Thumb	2.17	254.2	18.8%	10.0%
		Index	2.35	289.6	21.4%	11.6%
		Middle	2.09	493.4	36.5%	24.4%
4	Right	Thumb	3.81	199.1	14.7%	8.5%
		Index	3.69	227.1	16.8%	9.5%
	Left	Thumb	4.08	197.7	14.6%	7.0%
		Index	4.21	261.5	19.4%	9.3%
5	Right	Thumb	8.35	171.6	12.7%	6.6%
	Left	Thumb	9.41	182.8	13.5%	5.8%

currences of fingerprint classes and correlations among them obtained from two large databases (NIST SD9 and an in-house dataset of the German Federal Police). Subsequently, a variable search order strategy is applied to conduct a one-to-first search of the enrolled database. The experiments were conducted using the publicly available and well-established PCASYS tool and a neuronal network classifier. The results convince by significant reduction of the computational workload: for instance, when using all five fingers of a hand, it is reduced to less than 15% of the naïve exhaustive search. Additionally, the theoretical limits of the approach are established for a perfect (always accurate) classifier, with which the computational workload could be brought down to approximately 5% of the naïve exhaustive search.

Furthermore, several interesting observations can be made regarding the choice of the binning parameters:

- One or few fingers: Few bins (each containing many subjects) have to be visited, but more subjects have to be considered, *i.e.* the overall search space is larger.

- Many fingers: Many bins (each containing few subjects) have to be visited, but fewer subjects have to be considered, *i.e.* the overall search space is smaller.'

- Choice of fingers: the thumb appears to exhibit less correlation to other fingers, which makes its inclusion in the binning scheme beneficial, *cf*. table 5.9, where the lowest workload is always achieved by including the thumb.

By using data from both hands (*i.e*. ten fingers instead of five), the computational workload could presumably be further reduced, but this scenario would be less practical for operational deployments, which typically perform acquisition for a single hand or individual fingers only. Since the proposed approach utilises well-known and understood features, and is readily compatible with the current fingerprint sample acquisition methods, it could be seamlessly integrated into operational biometric deployments.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF), by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP), and the LOEWE-3 BioBiDa Project (594/18-17).

5.6 Bibliography

[1] ABADI, M., ET AL. TensorFlow: Large-scale machine learning on heterogeneous systems. `https://www.tensorflow.org/`, November 2015. Last accessed: 2020–03–11.

[2] BORRA, S., REDDY, G., AND REDDY, E. Classification of fingerprint images with the aid of morphological operation and AGNN classifier. *Applied Computing and Informatics 14*, 2 (July 2018), 166–176.

[3] CANDELA, G., GROTHER, P., WATSON, C., WILKINSON, R., AND WILSON, C. PCASYS – a pattern-level classification automation system for fingerprints. Tech. Rep. NISTIR5647, National Institute of Standards and Technology, August 1995.

[4] CAO, K., PANG, L., LIANG, J., AND TIAN, J. Fingerprint classification by a hierarchical classifier. *Pattern Recognition 46*, 12 (December 2013), 3186–3197.

[5] CAPPELLI, R., MAIO, D., AND MALTONI, D. SFinGe: an approach to synthetic fingerprint generation. In *International Workshop on Biometric Technologies* (June 2004), Kluwer, pp. 1–8.

[6] CHOLLET, F., ET AL. Keras: The Python deep learning library. `https://keras.io`, 2015. Last accessed: 2020–03–11.

[7] CHUA, S., WONG, E., AND TAN, A. A fuzzy rule-based fingerprint image classification. *International Journal of Applied Engineering Research 11*, 13 (January 2016), 7920–7925.

[8] DORASAMY, K., WEBB, L., TAPAMO, J., AND KHANYILE, N. Fingerprint classification using a simplified rule-set based on directional patterns and singularity features. In *International Conference on Biometrics (ICB)* (May 2015), IEEE, pp. 400–407.

[9] GALAR, M., DERRAC, J., PERALTA, D., TRIGUERO, I., PATERNAIN, D., LOPEZ-MOLINA, C., GARCÍA, S., BENÍTEZ, J. M., PAGOLA, M., BARRENECHEA, E., ET AL. A survey of fingerprint classification part I: Taxonomies on feature extraction methods and learning models. *Knowledge-based systems 81* (June 2015), 76–97.

[10] GALAR, M., DERRAC, J., PERALTA, D., TRIGUERO, I., PATERNAIN, D., LOPEZ-MOLINA, C., GARCÍA, S., BENÍTEZ, J. M., PAGOLA, M., BARRENECHEA, E., ET AL. A survey of fingerprint classification part II: experimental analysis and ensemble proposal. *Knowledge-Based Systems 81* (June 2015), 98–116.

[11] GUO, J.-M., LIU, Y.-F., CHANG, J.-Y., AND LEE, J.-D. Fingerprint classification based on decision tree from singular points and orientation field. *Expert Systems with Applications 41*, 2 (February 2014), 752–764.

[12] HENRY, E. R. *Classification and uses of finger prints*. HM Stationery Office, 1900.

[13] IDEMIA. MorphoWave Desktop: Capturing four fingerprints in less than a second. `https://www.idemia.com/morphowave-desktop`. Last accessed: 2020–03–11.

[14] JUNG, H.-W., AND LEE, J.-H. Noisy and incomplete fingerprint classification using local ridge distribution models. *Pattern Recognition 48*, 2 (February 2015), 473–484.

[15] LUO, J., SONG, D., XIU, C., GENG, S., AND DONG, T. Fingerprint classification combining curvelet transform and gray-level cooccurrence matrix. *Mathematical Problems in Engineering* (August 2014), 1–15.

[16] MALTONI, D., MAIO, D., JAIN, A. K., AND PRABHAKAR, S. *Handbook of fingerprint recognition*. Springer, 2009.

[17] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. Special Database 9. `https://www.nist.gov/srd/nist-special-database-9`, August 2010. Last accessed: 2020–03–11.

[18] PEDREGOSA, F., ET AL. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research 12* (October 2011), 2825–2830.

[19] PERALTA, D., TRIGUERO, I., GARCÍA, S., SAEYS, Y., ET AL. Distributed incremental fingerprint identification with reduced database penetration rate using a hierarchical classification based on feature fusion and selection. *Knowledge-Based Systems 126*, Supplement C (June 2017), 91–103.

[20] ROSS, A., NANDAKUMAR, K., AND JAIN, A. K. *Handbook of multibiometrics*. Springer, 2006.

[21] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Aadhaar dashboard. `https://www.uidai.gov.in/aadhaar_dashboard/`. Last accessed: 2020–03–11.

[22] WANG, R., HAN, C., AND GUO, T. A novel fingerprint classification method based on deep learning. In *International Conference on Pattern Recognition (ICPR)* (December 2016), IEEE, pp. 931–936.

[23] WAYMAN, J. L. Multifinger penetration rate and ROC variability for automatic fingerprint identification systems. In *Automatic Fingerprint Recognition Systems*. Springer, 2004, pp. 305–316.

[24] WEBB, L., AND MATHEKGA, M. Towards a complete rule-based classification approach for flat fingerprints. In *International Symposium on Computing and Networking* (December 2014), IEEE, pp. 549–555.

Chapter 6

Multi-Iris Indexing and Retrieval: Fusion Strategies for Bloom Filter-based Search Structures

Abstract

We present a multi-iris indexing system for efficient and accurate large-scale identification. The system is based on Bloom filters and binary search trees. We describe and empirically evaluate several possible information fusion strategies for the system. Those experiments are performed using a combination of several publicly available datasets; the proposed system is tested in an open-set identification scenario consisting of 6,000 genuine and 100,000 impostor transactions. The system maintains the near-optimal biometric performance of an iris-code, score fusion based baseline system, while reducing the required lookup workload to less than 1% thereof.

Addressed research question(s): RQ1, RQ3
Reference: DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Multi-iris indexing and retrieval: Fusion strategies for Bloom filter-based search structures. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 46–53.

6.1 Introduction

The increased popularity of biometrics worldwide is reflected in the appearance of several large-scale deployments. Of these, by far the largest is the Indian National ID project – at the time of this writing, more than 1 billion subjects have been enrolled [22] with biometric data from iris, face and fingerprints.

In (open-set) identification and de-duplication scenarios, one of the key challenges is the system accuracy, especially in terms of false positive occurrences. In a naïve, brute-force approach 1:N comparisons per retrieval of the enrolled reference are performed, *i.e.* the probe template is compared against every template in the biometric reference database. Hence, for large

databases, the possibility of false positive occurrenes quickly becomes unacceptable (see [3]). Fusing information from multiple sources can be used to increase the discriminative power of a biometric system [11]. In this paper, we utilise information from multiple instances of the same biometric characteristic – images of the left and right irides. Since the operational systems often already capture images of both irides (*e.g.* the aforementioned deployment in India), the proposed approach would not incur additional hardware costs or acquisition time during enrolment and could be easily integrated into existing systems.

The main contribution of this paper is a, best to the author's knowledge, first system for multi-iris indexing in the biometric literature and a large-scale, open-set identification evaluation of several information fusion strategies for said system. The key goal was to explore different strategies for multi-instance Bloom filter-based indexing and conduct a benchmark in terms of biometric performance and workload reduction. The paper is organised as follows: section 6.2 outlines related work; the basics of Bloom filter-based indexing are described in section 6.3, while section 6.4 shows how it can be extended to support multi-iris templates and which information fusion strategies can be applied. Section 6.5 describes the experimental setup; the results, discussion thereof and conclusions are presented in sections 6.6 and 6.7.

6.2 Related Work

The task solved by a biometric system in an open-set identification mode (*i.e.* where no identity claim is made) can be generalised to the classic nearest-neighbour search (NNS) problem. Additional non-trivial challenges are caused by high dimensionality, as well as fuzziness of the biometric data, meaning that the reference and probe sample from the same subject may be very similar, but never identical. In large systems, it is desirable to avoid the necessity of a naïve, brute-force lookup for every search query, as such retrieval method is computationally expensive and prone to false positive occurrences. Database indexing is a commonly used method for achieving this goal. In such systems, an index data-structure (*e.g.* a tree or hash table), which allows to quickly locate the approximate location of the data, is created and maintained. In other words, indexing systems utilise additional storage space in order to decrease response times. For biometric data, the indexing schemes must take into account the aforementioned issues of data fuzziness and high dimensionality.

Iris indexing is a relatively new research topic. Albeit the first work in the area by Mukherjee and Ross [15] has yielded only meagre results, it has demonstrated the feasibility of indexing iris data. Hao *et al.* [8] have developed a general fast search algorithm for fuzzy databases containing iris

codes or similar data, with the key idea of placing "beacons" in the search space, thereby shrinking the search range. Mehrotra *et al*. [14] developed an efficient scheme based on multi-resolution decomposition and B-trees, while Gadde *et al*. [7], utilised the Burrows-Wheeler Transform and binary pattern matching to obtain good results in terms of penetration and hit rates. Proença [17] presented methods for indexing low-quality iris data, as well as a more comprehensive survey of iris indexing approaches. Recently, in a proof-of-concept study of Rathgeb *et al*. [19], the concept of Bloom-filter based indexing and retrieval has been introduced with very promising results.

Information fusion in biometrics is used to consolidate data from multiple sources in order to improve the discriminative power of a system. In an identification system, several key categories of this data consolidation can be distinguished:

Image level Consolidating the raw data of same subject (*e.g*. two or more iris images).

Feature level Consolidating the feature vectors of various biometric instances from the same subject (*e.g*. two or more iris-codes).

Score level Combining the scores yielded by multiple comparators (*e.g*. two or more Hamming distance scores).

Rank level Combining the lists of candidate identities produced by multiple comparators.

Decision level Combining the decisions yielded by multiple comparators.

For each of the methods listed above, multiple heuristics and schemes have been proposed. Information fusion schemes have repeatedly been shown to improve the systems' biometric performance. The reader is referred to [21] and [18], as comprehensive references on the subject of multi-biometric systems and information fusion basics.

In summary, while many single-iris indexing schemes and many information fusion schemes exist, the field of *multi-biometric indexing* appears to be unexplored as of yet. In the subsequent sections, the basics of Bloom filter-based iris indexing and its extension to multi-iris indexing are described.

6.3 Bloom Filter-based Search Structure

The idea of using Bloom filters [1] and binary search trees for iris indexing has been presented in a proof-of-concept study [19]. In this section, the basics of Bloom-filter based indexing and retrieval are briefly reiterated.

The process of transforming the iris-code templates to the Bloom filter-based representation is as follows: First, the two-dimensional iris-code is divided into k blocks of equal size, $W_B \times H_B$. All these blocks are subsequently transformed to sets of integers of variable cardinality. The transformation function takes the decimal value of each of the binary columns $(\text{int}(c_1), \ldots, \text{int}(c_{W_B}))$ in a block and inserts them into the Bloom filter corresponding to that block (*i.e.* $\mathbf{b}[\text{int}(c_i)] = 1$). The column values are obviously always in range $0 \leq \text{int}(c_i) < 2^{H_B}$. Thereby, the resulting biometric template (denoted $\mathbf{B}$) is a fixed-length (k) sequence of Bloom filters, $[\mathbf{b}_1, \ldots, \mathbf{b}_k]$. A dissimilarity score of two such biometric templates is calculated as normalised Hamming distance between corresponding Bloom filters in the sequences, which was proposed in the proof-of-concept study [19], as shown in equation (6.1), where $\mathbf{B}$ and $\mathbf{B}'$ denote the reference and probe template, respectively.

$$\begin{aligned} DS(\mathbf{B}, \mathbf{B}') &= \frac{1}{k} \sum_{i=1}^{k} ds(\mathbf{b}_i, \mathbf{b}'_i) \\ ds(\mathbf{b}, \mathbf{b}') &= \frac{|\mathbf{b} \oplus \mathbf{b}'|}{(|\mathbf{b}| + |\mathbf{b}'|)} \end{aligned} \tag{6.1}$$

Where $|\cdot|$ represents the population count, *i.e.* Hamming weight. Note, that like in the case of the iris-code, the comparator utilises efficient bitwise instructions and can be trivially parallelised. It is also worth mentioning, that this data representation is, to a certain degree, rotation invariant [20], at the cost of loss of local information and loss of information about the number of identical columns in a block. In other words, the template alignment during comparison is not performed as in the case of the standard iris-code based approach. The block size determines the sparseness of the representation; the level to which a filter is filled is defined by the number of values present in the filter as a fraction of the number of possible values in it (*i.e.* $|\mathbf{b}| / 2^{H_B}$). As an example, suppose a system with block size 16 $\times$ 8. This means, that in each individual filter ($\mathbf{b}$) in the template ($\mathbf{B}$), the number of 1's will be *at most* 16 (in practice fewer, since the neighbouring iris-code columns are strongly correlated [5] and identical columns map to same value in the Bloom filter representation), out of possible 2^8 values, *i.e.* it will be *at most* only approximately 6.25% filled. This data representation sparseness is a crucial property in ensuring the system accuracy.

A binary tree data-structure is constructed from the templates of N enrolled subjects. First, the tree root is created from all enrolled templates (*i.e.* $\bigcup_{i=1}^{N} \mathbf{B}_i$), while the children of the root node each contain half of the enrolled templates (*i.e.* $\bigcup_{i=1}^{\frac{N}{2}} \mathbf{B}_i$ and $\bigcup_{i=\frac{N}{2}+1}^{N} \mathbf{B}_i$). The union of templates corresponds to OR'ing the individual binary filters stored in the children nodes. This process is repeated for node creation at subsequent tree levels,

until at the end, the individual templates $(\mathbf{B}_1, \ldots, \mathbf{B}_N)$ are inserted as tree leaves.

The complexity of a single lookup in such a tree is $O(2 * \log N)$. The tree construction and item retrieval processes are shown in figure 6.1. The lookup in an identification scenario begins at the tree root. The tree is traversed by calculating the dissimilarity scores (equation (6.1)) between the probe template and two nodes at the next tree level; subsequently choosing the one with lower score until a leaf is reached. The key idea is to take advantage of a sparse data representation in the nodes – in that case, for the probe comparisons against the tree, the relationship $DS_{genuine} \ll DS_{impostor}$ generally holds true. In other words, the genuine probes will be able to traverse the tree using the correct path to reach a matching leaf template. At the leaf, a final decision is made based on an acceptance threshold.

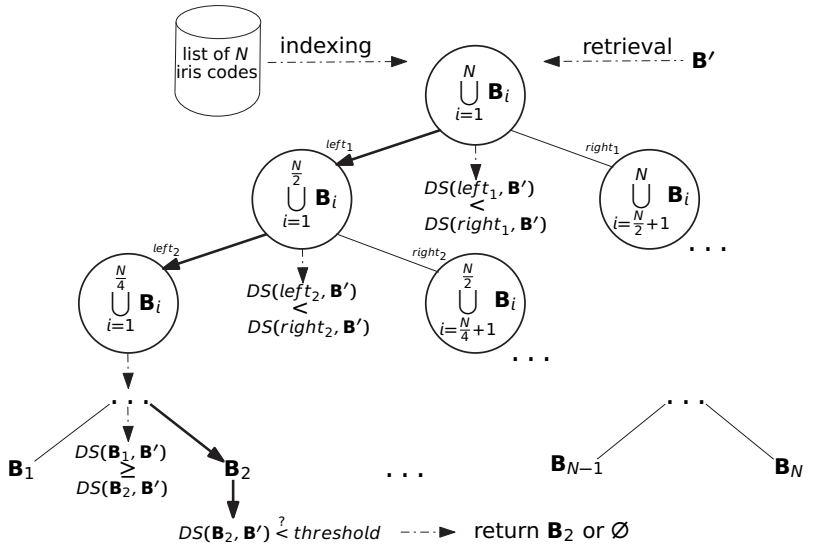


Figure 6.1: Indexing and retrieval in the Bloom filter-based system. In this case, the retrieval follows the bold arrow path down to a leaf, where the final decision is made.

In the aforementioned small-scale proof-of-concept study [19], this system has shown promising results both in terms of biometric performance and workload reduction. For large-scale experiments, the system can be extended by constructing multiple search trees (T) and, during a lookup, preselecting a subset consisting of the most promising (t) trees based on dissimilarity scores between the probe and tree roots. The trees are successively traversed until the first candidate identity is found or all the pre-selected trees have been checked. Note, that for the genuine transactions, the preselection step ensures that trees most likely to contain the sought identity

are traversed first. Additionally, as it is sufficient to pre-select only a small fraction of constructed trees, *i.e.* $t \ll T$, the lookup workload remains low, while arbitrarily many enrollees can be accommodated by the system.

6.4 Multi-Iris Indexing and Retrieval

This section describes how the Bloom Filter-based indexing scheme can be extended to accommodate multi-iris templates and how various information fusion techniques can be utilised during retrieval.

6.4.1 Bloom Filter Feature Fusion

The first matter that needs to be addressed is construction of trees with multi-iris templates. Let $\mathbf{B}_L$ and $\mathbf{B}_R$ represent Bloom filter-based iris templates from a subject's left and right eye, respectively. Prior to the construction of lookup trees, those two templates for each of the enrolled subjects are combined. Two viable strategies for doing so are described in subsections 6.4.1.1 and 6.4.1.2; one poised for better biometric performance, another for more compact storage. The tree construction with the fused templates works exactly as described in previous section and shown in figure 6.1.

6.4.1.1 Template Concatenation (TC)

One of the feature-level fusion methods is concatenation of feature vectors. The biometric templates extracted from the left and right irides of a subject ($\mathbf{B}_L$ and $\mathbf{B}_R$) can be combined in this way, thereby producing a single Bloom filter sequence of length $2k$: $\overbrace{[\mathbf{b}_{L_1}, \ldots, \mathbf{b}_{L_k}]}^{\mathbf{B}_L} \| \overbrace{[\mathbf{b}_{R_1}, \ldots, \mathbf{b}_{R_k}]}^{\mathbf{B}_R} = [\mathbf{b}_{L_1}, \ldots, \mathbf{b}_{L_k}, \mathbf{b}_{R_1}, \ldots, \mathbf{b}_{R_k}]$. This is equivalent to first concatenating the iriscodes and producing the sequence of Bloom filters as described in section 6.3.

6.4.1.2 Template Merging (TM)

Due to data representation sparseness, the Bloom filter-based representation lends itself to effortless combination of templates (which is the basis of building the lookup tree, see section 6.3). Since the left and right irides of a subject are mutually independent, the biometric templates $\mathbf{B}_R$ and $\mathbf{B}_L$ can be merged (rather than just appended unto each other) by taking a pairwise union of the two Bloom filter sequences: $\overbrace{[\mathbf{b}_{L_1}, \ldots, \mathbf{b}_{L_k}]}^{\mathbf{B}_L} \cup \overbrace{[\mathbf{b}_{R_1}, \ldots, \mathbf{b}_{R_k}]}^{\mathbf{B}_R} =$

$[\mathbf{b}_{L_1} \cup \mathbf{b}_{R_1}, \ldots, \mathbf{b}_{L_k} \cup \mathbf{b}_{R_k}]$. Observe that, contrary to the template concatenation, merging two templates does not increase the overall length (it remains k) of the Bloom filter sequence (and hence the cost of a single template comparison). On the other hand, some information loss may occur due to collisions.

6.4.2 Traversal Strategies

Let $\mathbf{B}'_L$ and $\mathbf{B}'_R$ represent the probe templates from left and right irides of a subject during a lookup. The presence of two templates makes it possible to design several tree traversal techniques, described below. They can be utilised with trees constructed using both the template concatenation and merging methods.

6.4.2.1 Basic (B)

The simplest strategy is to combine the probe templates using the same method as has been used during tree construction and traverse normally, as shown in figure 6.1. Observe that, for the concatenated templates, this method corresponds to a score-level fusion where comparison scores between tree nodes and $\mathbf{B}'_L$ and $\mathbf{B}'_R$ are averaged at each tree traversal step.

6.4.2.2 Iris-code at Leaf (ICL)

The basic lookup procedure is followed, however the template comparisons at the tree leaves (*i.e.* the last step of the tree traversal depicted in figure 6.1) are performed using iris-code based templates. The idea is that some of the information loss incurred by the Bloom filter-based representation may be mitigated, while still utilising the indexing scheme.

6.4.2.3 Path Fusion (PF)

Every individual tree is traversed using $\mathbf{B}'_L$ and $\mathbf{B}'_R$ simultaneously, whereby at each step (see figure 6.1), 4 dissimilarity scores are produced:

$$\overbrace{DS(\mathit{left}, \mathbf{B}'_L)}^{s_{L_l}}, \overbrace{DS(\mathit{right}, \mathbf{B}'_L)}^{s_{L_r}}, \overbrace{DS(\mathit{left}, \mathbf{B}'_R)}^{s_{R_l}}, \text{and } \overbrace{DS(\mathit{right}, \mathbf{B}'_R)}^{s_{R_r}}$$

These scores should satisfy the following relation: $s_{L_l} < s_{L_r} \wedge s_{R_l} < s_{R_r} \vee s_{L_l} \geq s_{L_r} \wedge s_{R_l} \geq s_{R_r}$, *i.e.* a step of the path is valid if and only if the same traversal direction has been chosen by both $\mathbf{B}'_L$ and $\mathbf{B}'_R$. Should the paths diverge (*i.e.* the above step scores relation is not satisfied), an early exit is performed and, subsequently, the next tree is checked in the same fashion. If none of the traversed trees yield an identical traversal path, the template is deemed to be an impostor. Observe, that an enrolled template for any

subject is stored as a leaf in one and only one tree – it is naturally expected, that both genuine $\mathbf{B}'_L$ and $\mathbf{B}'_R$ ought to be able to reach it when traversing the correct tree. If the traversal paths diverge, it can, with high confidence, be assumed that the sought template is absent from the traversed tree and the search can quickly proceed to the next one.

6.4.2.4 Pre-selection Fusion (PSF)

An additional heuristic for the path fusion: from the T constructed trees, subsets of most promising trees ($\mathbf{t}$) are found for the probe $\mathbf{B}'_L$ and $\mathbf{B}'_R$ based on their dissimilarity scores with the tree roots. Only trees that have been pre-selected for both the left and right template are traversed, *i.e.* $\mathbf{t} = \mathbf{t}_L \cap \mathbf{t}_R$. This heuristic is aimed especially at reducing the workload associated with the impostor transactions by means of the quicker rejection.

6.5 Experimental Setup

Table 6.1 lists the datasets used in the experiments, with example images from each one shown in figure 6.2. The images in the first two datasets are cropped, while the latter two are uncropped – as defined in the ISO/IEC standard on biometric data interchange formats [9].

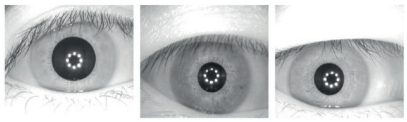
The CASIA-IrisV4-Thousand dataset was also considered, but was not used, since only just over half of the segmented images had more than 70% usable iris area (metric defined in [10]). Bearing in mind that in the real biometric deployments high-quality data acquisition is nowadays feasible and that specialised approaches for indexing of degraded data exist (*e.g.* [17]), it was deemed more interesting to work with data of higher quality.

Table 6.1: Evaluation dataset overview

Dataset	Instances	Images	Resolution
CASIA-IrisV4-Interval [2]	395	2639	320 × 280 px
IIT Delhi Iris Database v1 [13]	448	2240	320 × 240 px
BioSecure Iris Corpus [16]	420	1680	640 × 480 px
ND-Iris-Template-Aging[1] [6]	214	5377	640 × 480 px

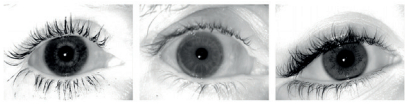
The raw near-infrared images were processed with the commonly used methods: After segmentation, where the iris and pupil boundaries are located, the iris textures were normalised according to the rubbersheet model [4] and subsequently enhanced by applying Contrast Limited Adaptive Histogram Equalization (CLAHE). Feature extraction was performed with the Daugman-like 1D-LogGabor algorithm (LG), generating iris-codes of size

[1] Only a single point in time from the dataset was used.

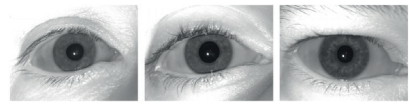


(a) CASIAv4-Interval

(b) IIT Delhi Iris Database v1



(c) BioSecure Iris Corpus



(d) ND-Iris-Template-Aging

Figure 6.2: Example images from the datasets

512 $\times$ 20 = 10,240 bits. Observe, that although a correlation between the imaginary and real LG filter response exists, the Bloom filter-based indexing does not, in any way, depend upon this fact. In other words, the method could also be utilised with *e.g.* only the real filter response, as is the case in some systems.

256 subjects from the first two datasets listed in the above table have been enrolled. *All* possible left-right iris combinations for each subject were considered when creating the multi-iris templates. In other words, *all* possible genuine identification transactions are performed. The remaining templates from all 4 datasets were cross-paired to produce a total of 100.000 multi-iris impostor transactions. For every fusion and traversal strategy combination (see section 6.4), there exists a large number of system configurations which may differ in following parameters:

- Number of trees constructed (T): 8 or 16
- Block size (H, W): between (8, 8) and (12, 32)
- Number of trees traversed (t): 1 to T
- Feature extractor: the results presented in next section stem from templates produced using the LG feature extractor; however, repeated experiments with other common feature extraction methods (*e.g.* Quadratic Spline Wavelet) yielded comparable outcomes.

In total, over 1500 experiments were performed and the following three metrics were used for the system performance and efficiency evaluation:

- $TP_{0.01}$ – the true positive identification rate measured at a false positive identification rate of 0.01%
- F – the fraction of the required (iris-code) baseline workload (bit comparisons) per lookup

- τ – inspired by the metric proposed in [17], the Euclidean distance from the optimal operating point ($TP_{0.01} = 1$ and $F \approx 0$), calculated as follows: $\tau = \sqrt{(TP_{0.01} - 1)^2 + F^2}$

6.6 Results and Discussion

The baseline results are established using an iris-code based system, which performs an exhaustive database search and averages the Hamming distance scores for the left and right iris. Expectedly, the results in terms of biometric performance are near-optimal ($TP_{0.01} > 0.999$), albeit at a high workload entailed by the naïve, brute-force search and the necessity of alignment correction (± 8 bit shifts, corresponding to approximately $\pm 5.625°$). As an intermediate step to show soundness of the template representation change, a Bloom filter-based baseline is likewise established (*i.e.* the same experiment as the iris-code baseline, only with the representation changed to Bloom filter-based and no tree construction). In tables 6.2 and 6.3, the experimental results of several generally best performing system configurations are listed, with the bold typeface showing the best result (in terms of τ) for each experiment.

In order to better visualise the results of more than 1500 experiments, a filtering has been applied. The idea is to consider only relevant system configurations for each experiment, as shown in equation (6.2). Essentially, for every experiment type, a system configuration (C) is considered relevant if and only if there exists no other configuration (C') with higher (or same) biometric performance and lower workload.

$$C \text{ is relevant} \Leftrightarrow \left\{C' | C'_{TP_{0.01}} \geq C_{TP_{0.01}} \wedge C'_F < C_F\right\} = \emptyset \qquad (6.2)$$

Figure 6.3 visualises the trade-off between workload (F) and performance ($TP_{0.01}$) after the results have been filtered to consider only the relevant configurations. Using the τ metric, the operating point closest to optimum is selected for each of the experiments and listed in table 6.4 as a summary.

As table 6.4 demonstrates, the Bloom filter baseline maintains the biometric performance of the iris-code baseline at a small fraction of the workload. It is evident, that the Bloom filter-based representation of an iris-code is very efficient (due to compact template size and rotational invariance) and does not suffer from too much information loss. Using the tree indexing scheme with some of the information fusion techniques allows to further improve on these results in terms of biometric performance and workload.

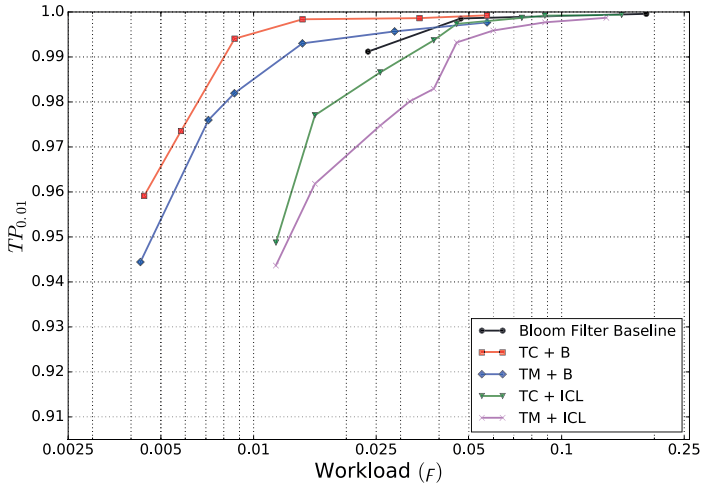
The most significant improvements are achieved with the template feature fusion techniques (subsections 6.4.1.1 and 6.4.1.2). The workload of a single template comparison between concatenated templates is naturally

Table 6.2: Results of basic traversal approaches

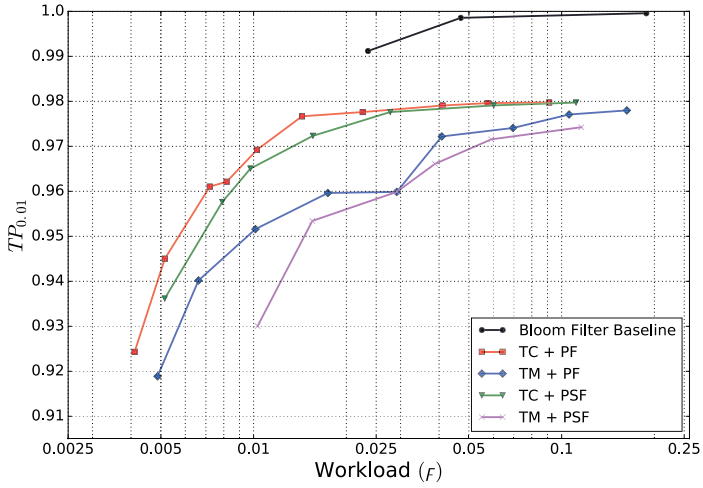
Type	Configuration				$TP_{0.01}$	F	τ
	T	H	W	t			
TC + B	16	8	16	2	0.9869	0.0116	0.0175
				4	0.9967	0.0173	0.0176
				8	0.9978	0.0287	0.0288
			32	2	0.9735	0.0058	0.0271
				4	**0.9941**	**0.0087**	**0.0105**
				8	0.9984	0.0144	0.0145
		10	16	2	0.9963	0.0459	0.0460
				4	0.9986	0.0684	0.0684
				8	0.9987	0.1133	0.1133
			32	2	0.9897	0.0232	0.0254
				4	0.9986	0.0346	0.0346
				8	0.9993	0.0572	0.0572
TM + B	16	8	16	2	0.9680	0.0058	0.0325
				4	0.9819	0.0087	0.0201
				8	**0.9930**	**0.0144**	**0.0160**
			32	2	0.8677	0.0029	0.1323
				4	0.9444	0.0043	0.0558
				8	0.9760	0.0071	0.0250
		10	16	2	0.9818	0.0231	0.0294
				4	0.9953	0.0345	0.0348
				8	0.9977	0.0573	0.0573
			32	2	0.9686	0.0116	0.0335
				4	0.9927	0.0173	0.0188
				8	0.9957	0.0287	0.0290
TC + ICL	16	8	16	2	**0.9866**	**0.0258**	**0.0291**
				4	0.9973	0.0456	0.0457
				8	0.9988	0.0860	0.0860
			32	2	0.9736	0.0207	0.0335
				4	0.9937	0.0385	0.0390
				8	0.9987	0.0743	0.0743
		10	16	2	0.9964	0.0558	0.0559
				4	0.9992	0.0882	0.0882
				8	0.9994	0.1566	0.1566
			32	2	0.9893	0.0359	0.0375
				4	0.9982	0.0600	0.0600
				8	0.9990	0.1096	0.1096
TM + ICL	16	8	16	2	**0.9686**	**0.0207**	**0.0376**
				4	0.9829	0.0384	0.0420
				8	0.9391	0.0743	0.0961
			32	2	0.8679	0.0181	0.1333
				4	0.9473	0.0348	0.0632
				8	0.8769	0.0684	0.1408
		10	16	2	0.9817	0.0359	0.0403
				4	0.9959	0.0600	0.0601
				8	0.9683	0.1096	0.1141
			32	2	0.9692	0.0258	0.0402
				4	0.9932	0.0456	0.0461
				8	0.9531	0.0860	0.0980

Table 6.3: Results of path fusion traversal approaches

Type	Configuration				$TP_{0.01}$	F	τ
	T	H	W	t			
TC + PF	16	8	16	2	0.9692	0.0103	0.0325
				4	**0.9767**	**0.0144**	**0.0274**
				8	0.9776	0.0226	0.0318
			32	2	0.9450	0.0051	0.0552
				4	0.9610	0.0072	0.0397
				8	0.9635	0.0113	0.0382
		10	16	2	0.9791	0.0410	0.0460
				4	0.9796	0.0575	0.0610
				8	0.9797	0.0911	0.0933
			32	2	0.9673	0.0206	0.0386
				4	0.9742	0.0288	0.0387
				8	0.9742	0.0455	0.0523
TM + PF	16	8	16	2	0.9402	0.0066	0.0602
				4	0.9516	0.0101	0.0494
				8	**0.9596**	**0.0174**	**0.0440**
			32	2	0.8259	0.0033	0.1741
				4	0.8947	0.0051	0.1054
				8	0.9214	0.0088	0.0791
		10	16	2	0.9597	0.0266	0.0483
				4	0.9722	0.0408	0.0494
				8	0.9741	0.0695	0.0742
			32	2	0.9362	0.0134	0.0652
				4	0.9573	0.0206	0.0474
				8	0.9590	0.0354	0.0542
TC + PSF	16	8	16	2	0.9556	0.0075	0.0450
				4	0.9651	0.0098	0.0362
				8	**0.9723**	**0.0156**	**0.0318**
			32	2	0.8882	0.0039	0.1119
				4	0.9362	0.0051	0.0640
				8	0.9576	0.0079	0.0431
		10	16	2	0.9703	0.0281	0.0409
				4	0.9766	0.0371	0.0439
				8	0.9791	0.0602	0.0637
			32	2	0.9503	0.0151	0.0519
				4	0.9715	0.0201	0.0349
				8	0.9735	0.0312	0.0409
TM + PSF	16	8	16	2	0.8840	0.0078	0.1163
				4	0.9301	0.0103	0.0707
				8	**0.9534**	**0.0155**	**0.0491**
			32	2	0.6190	0.0040	0.3810
				4	0.8138	0.0056	0.1863
				8	0.8989	0.0080	0.1014
		10	16	2	0.9389	0.0298	0.0680
				4	0.9661	0.0388	0.0515
				8	0.9716	0.0591	0.0656
			32	2	0.8872	0.0158	0.1139
				4	0.9379	0.0212	0.0656
				8	0.9580	0.0309	0.0521



(a) Basic traversal approaches



(b) Path fusion traversal approaches

Figure 6.3: Filtered experimental results. The iris-code baseline is not visible, as it is located at $F = 1.0$ and $TP_{0.01} \approx 1.0$.

twice as large as that between two merged templates. On the other hand, concatenating the Bloom filter-based templates avoids the information loss associated with merging (OR'ing) them. Thus larger block sizes and a smaller number of traversed trees can be successfully used in the concatenation-

Table 6.4: Best operating point in terms of τ for each of the experiments

Type	Configuration				$TP_{0.01}$	F	τ
	T	H	W	t			
Iris-code baseline	—				0.9993	1.0000	1.0000
Bloom filter baseline	—	8	32	—	0.9912	0.0235	0.0251
TC + B	16	8	32	4	0.9941	0.0087	0.0105
TM + B	16	8	16	8	0.9930	0.0144	0.0160
TC + ICL	16	8	16	2	0.9866	0.0258	0.0291
TM + ICL	16	8	16	2	0.9686	0.0207	0.0376
TC + PF	16	8	16	4	0.9767	0.0144	0.0274
TM + PF	16	8	16	8	0.9596	0.0174	0.0440
TC + PSF	16	8	16	8	0.9723	0.0156	0.0318
TM + PSF	16	8	16	8	0.9534	0.0155	0.0491

based feature fusion, thereby compensating for the increased cost of a single template comparison. Ultimately, for the feature-level fusion, a near-optimal biometric performance is attainable at around or less than 1% of the iris-code based baseline workload.

The results achieved by other fusion types are worse than those of the feature-level fusion. Any potential biometric performance gain due to usage of the iris-code based comparator at tree leaves (subsection 6.4.2.2) appears to have been unable to compensate for the increased computational cost of the method (alignment compensation for the iris-code templates at the leaf). While the early rejection heuristics (subsections 6.4.2.3 and 6.4.2.4) are efficient in quickly rejecting impostors, they unfortunately appear to have a non-negligible detrimental effect on the performance of the genuine transactions. Consequently, the basic feature-level fusion appears to be the preferred method in case of the Bloom filter-based indexing. Concatenating the templates yields best results in terms of biometric performance, while merged templates achieve comparable results with an added benefit of lower storage requirements.

In their *"Guidelines for best practices in biometrics research"* [12], in a point pertaining to biometric fusion, Jain *et al.*, state that:

> The improvement in recognition accuracy as a result of biometric fusion should be weighed against the associated overhead involved, such as additional sensing cost, enrolment and recognition times, computing resources, usability, etc.

Since, as has been mentioned in section 6.1, the current biometric deployments already often capture both irides during acquisition phase, the proposed information fusion scheme would generally not entail any additional overhead associated with sensing, enrolment and usability. Finally,

the workload (and thereby computing resources) associated with the proposed scheme is comparable with that of a single-iris Bloom filter-based indexing, while conferring significant biometric performance improvement.

6.7 Conclusion

In this paper, a system for indexing and retrieval of multi-iris templates has been presented. Best to the author's knowledge, this is first such attempt in the biometric literature. Numerous strategies of information fusion for using said system in the open-set identification scenario have been presented and evaluated using a large dataset. It has been concluded, that two types (concatenation and merging) of feature-level fusion with a standard search-tree traversal heuristic offer the best results. The proposed scheme maintains the near-optimal biometric performance of an iris-code score fusion based baseline, while reducing the necessary lookup workload to below 1% of said baseline. In future work, it is intended to investigate how the proposed scheme can be extended to perform indexing of multi-modal biometric data.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF) as well as by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP).

6.8 Bibliography

[1] BLOOM, B. H. Space/time trade-offs in hash coding with allowable errors. *Communications of the ACM 13*, 7 (July 1970), 422–426.

[2] CHINESE ACADEMY OF SCIENCES' INSTITUTE OF AUTOMATION. CASIA iris image database. `http://biometrics.idealtest.org/`, December 2010. Last accessed: 2020–03–11.

[3] DAUGMAN, J. Biometric decision landscapes. Tech. Rep. UCAM-CL-TR-482, University of Cambridge - Computer Laboratory, January 2000.

[4] DAUGMAN, J. How iris recognition works. *Transactions on Circuits and Systems for Video Technology (TCSVT) 14*, 1 (January 2004), 21–30.

[5] DAUGMAN, J. Information theory and the IrisCode. *Transactions on Information Forensics and Security 11*, 2 (February 2016), 400–409.

[6] FENKER, S. P., AND BOWYER, K. W. Analysis of template aging in iris biometrics. In *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (June 2012), IEEE, pp. 45–51.

[7] GADDE, R. B., ADJEROH, D., AND ROSS, A. Indexing iris images using the Burrows-Wheeler transform. In *International Workshop on Information Forensics and Security (WIFS)* (December 2010), IEEE, pp. 1–6.

[8] HAO, F., DAUGMAN, J., AND ZIELINSKI, P. A fast search algorithm for a large fuzzy database. *Transactions on Information Forensics and Security (TIFS) 3*, 2 (June 2008), 203–212.

[9] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 19794-6:2011. Information technology – Biometric data interchange formats – Part 6: Iris image data*. International Organization for Standardization and International Electrotechnical Committee, October 2011.

[10] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 29794-6:2015. Information technology – Biometric sample quality – Part 6: Iris image data*. International Organization for Standardization and International Electrotechnical Committee, July 2015.

[11] ISO/IEC JTC1 SC37 BIOMETRICS. Iso/iec tr 24722:2015. information technology – biometrics – multimodal and other multibiometric fusion. Tech. rep., International Organization for Standardization, December 2015.

[12] JAIN, A., KLARE, B., AND ROSS, A. Guidelines for best practices in biometrics research. In *International Conference on Biometrics (ICB)* (May 2015), IEEE, pp. 541–545.

[13] KUMAR, A., AND PASSI, A. Comparison and combination of iris matchers for reliable personal authentication. *Pattern Recognition 43*, 3 (March 2010), 1016–1026.

[14] MEHROTRA, H., SRINIVAS, B. G., AND MAJHI, B. Indexing iris biometric database using energy histogram of DCT subbands. In *International Conference on Contemporary Computing (IC3)* (August 2009), vol. 40, Springer, pp. 194–204.

[15] MUKHERJEE, R., AND ROSS, A. Indexing iris images. In *International Conference on Pattern Recognition (ICPR)* (December 2008), IEEE, pp. 1–3.

[16] ORTEGA-GARCIA, J., FIERREZ, J., ALONSO-FERNANDEZ, F., GALBALLY, J., FREIRE, M. R., ET AL. The multiscenario multienvironment BioSecure multimodal database (BMDB). *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 32*, 6 (June 2010), 1097–1111.

[17] PROENÇA, H. Iris biometrics: Indexing and retrieving heavily degraded data. *Transactions on Information Forensics and Security (TIFS) 8*, 12 (December 2013), 1975–1985.

[18] RADU, P., SIRLANTZIS, K., HOWELLS, G., DERAVI, F., AND HOQUE, S. A review of information fusion techniques employed in iris recognition systems. *International Journal of Advanced Intelligence Paradigms 4*, 3/4 (February 2012), 211–240.

[19] RATHGEB, C., BREITINGER, F., BAIER, H., AND BUSCH, C. Towards bloom filter-based indexing of iris biometric data. In *International Conference on Biometrics (ICB)* (May 2015), IEEE, pp. 422–429.

[20] RATHGEB, C., BREITINGER, F., BUSCH, C., AND BAIER, H. On application of Bloom filters to iris biometrics. *IET Biometrics 3*, 4 (December 2014), 207–218.

[21] ROSS, A., NANDAKUMAR, K., AND JAIN, A. K. *Handbook of multibiometrics*. Springer, 2006.

[22] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Aadhaar dashboard. `https://www.uidai.gov.in/aadhaar_dashboard/`. Last accessed: 2020–03–11.

Chapter 7

Privacy-Preserving Indexing of Iris-Codes with Cancelable Bloom Filter-based Search Structures

Abstract

Protecting the privacy of the enrolled subjects is an important requirement expected from biometric systems. In recent years, numerous template protection schemes have been proposed, but so far none of them have been shown to be suitable for indexing (workload reduction) in the computationally expensive identification mode. This paper presents a, best to the authors' knowledge, first method in the scientific literature for indexing protected iris templates. It is based on applying random permutations to Iris-Code rows, and subsequent indexing using Bloom filters and binary search trees. In a security evaluation, the unlinkability, irreversibility and renewability of the method are demonstrated quantitatively. The biometric performance and workload reduction are assessed in an open-set identification scenario on the IITD and CASIA-Iris-Thousand datasets. The method exhibits high biometric performance and reduces the required computational workload to less than 5% of the baseline Iris-Code system.

Addressed research question(s): RQ1, RQ5
Reference: DROZDOWSKI, P., GARG, S., RATHGEB, C., GOMEZ-BARRERO, M., CHANG, D., AND BUSCH, C. Privacy-preserving indexing of Iris-Codes with cancelable Bloom filter-based search structures. In *European Signal Processing Conference (EUSIPCO)* (September 2018), IEEE, pp. 2360–2364.

7.1 Introduction

In recent years, interest in biometric systems have spiked with many large-scale deployments (*e.g.* national databases and border crossing control systems) appearing. Currently, the largest such system is the Indian National ID system, into which, at the time of this writing, 1.2 billion Indian residents have been enrolled [30] with multi-biometric data and unique identi-

fier numbers. In the United Arab Emirates, the border control agency employs an iris-based blacklist system, which aims to prevent undesirable travellers (*e.g.* visa violators and criminals) from re-entering the country [1].

Those and similar deployments have to operate in the identification or duplicate-check modes. Due to the sheer size of such systems, they are faced with strenuous requirements in terms of biometric performance and computational workload. The naïve algorithm for such scenarios requires an exhaustive (1:N) database search, *i.e.* comparing the probe against all the references stored in the database. Notwithstanding the use of efficient hardware and parallelism, with the growing database sizes, the cost of executing such searches becomes computationally prohibitive. Simultaneously, the probability of false positives quickly becomes unacceptable. In [3], Daugman shows the probability of at least one false positive (P_N) occurring in a identification scenario to be: $P_N = 1 - (1 - P_1)^N$, where N is the number of enrolled subjects and P_1 the false positive probability of a one-to-one template comparison. For this reason, research has been conducted into biometric *workload reduction*, whereby the exhaustive search is replaced with more advanced techniques. Those techniques often take advantage of the underlying biometric template data representation, thus facilitating efficient search strategies; for example through indexing or serial combination of algorithms. The aim thereof is to vastly reduce the necessary number of template comparisons per lookup, while maintaining or only insignificantly reducing the biometric performance achieved by the baseline, exhaustive algorithm.

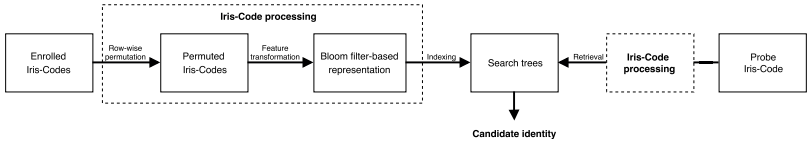


Figure 7.1: An overview of the proposed system.

A biometric system in an open-set identification mode (*i.e.* without an identity claim) can be generalised to the classic nearest-neighbour search (NNS) problem. However, additional non-trivial challenges arise due to high dimensionality, as well as intra-class variation of the biometric data, which means that the biometric templates extracted from the reference and probe samples belonging to the same subject may be very similar, but (almost) never identical. Consequently, typical workload reduction approaches such as *indexing* need to be adapted to account for the challenging properties of the biometric data (see *e.g.* [7, 11, 16, 17], and [21] for a more comprehensive survey). Other approaches used in (iris) biometric systems include: *cascading algorithms*, whereby a computationally efficient (albeit less accurate) method first computes a shortlist of candidate identities, which is

then searched exhaustively by a slower and more accurate comparator (see *e.g.* [8, 14, 25]); and *classification*, whereby the database is split into buckets containing certain template classes (*e.g.* based on gender, eye colour, some statistical properties etc.), with the exhaustive search only being performed inside the bucket corresponding to the probe (see *e.g.* [22, 28, 29]).

In addition to the aforementioned need for workload reduction, potential of data exposure is a large concern in biometric system deployments, where the stored data is, in most cases, secured using traditional encryption algorithms [18]. Once compromised, this can lead to serious problems such as identity theft, cross-matching without consent and severely limited renewability. Furthermore, centralised storage of sensitive personal and biometric data has been increasingly receiving attention from the general public and various non-governmental organisations, thus leading to widened legislation against privacy violations (*e.g.* GDPR in Europe [6]). Those matters have led to research into biometric *template protection* (see *e.g.* [24] and [27] for comprehensive surveys), with the aim of developing protection schemes especially dedicated for biometric data. Such systems must guarantee the properties stipulated by ISO/IEC IS 24745:2011 [12]:

Unlinkability It should be infeasible to determine whether or not two or more protected templates were derived from the same instance. This property prevents cross-matching across different databases.

Irreversibility Given a protected template and its corresponding secret, it should be infeasible to reconstruct the original biometric data. This property increases the security of the system against presentation and replay attacks.

Renewability It should be possible to issue new and revoke old protected templates from the same biometric instance and/or sample. This property ensures that in case of the biometric database being compromised, the data can be revoked and reissued, thereby preventing misuse.

Performance preservation The biometric performance is not significantly degraded by the template protection scheme.

With the aforementioned issues as motivation, this paper presents a, best to the authors' knowledge, first method in the scientific literature for indexing of protected iris templates. The method is based on Bloom filters and search trees (see [23] and [5]), which were previously shown to exhibit high workload reduction at an insignificant degradation to biometric performance, as well as scalability for an arbitrary number of enrollees. In this paper, said approach is extended by adapting ideas from [10] to accommodate cancelable iris templates which fulfil the aforestated properties and are suitable for indexing.

The remainder of this paper is organised as follows: in section 7.2, a method for privacy-preserving indexing of iris data is proposed. Section 7.3 presents the experiments and results, while section 7.4 contains a summary and concluding remarks.

7.2 Privacy-preserving Indexing of Iris-Codes

In this section, the key components of the proposed system are presented. Subsection 7.2.1 describes a row-based permutation of Iris-Codes, while their transformation to a Bloom filter-based representation, as well as indexing and retrieval are outlined in subsection 7.2.2. Figure 7.1 shows a schematic overview of the proposed system.

7.2.1 Row-based Permutation

To dissipate the statistical composition of the Iris-Code, a two-step feature rearrangement adapted from [10] is applied:

1. The Iris-Code is split into a small number of parts (IC_{parts}). The aim is to minimise the potential negative impact of the template protection on the biometric performance by preserving more spatial information. Several alternatives have been explored, namely: a) 2 parts – the real and imaginary response of the feature extractor; b) 4 or 8 parts – a further subdivision of each response into 2 or 4 parts, respectively.

2. A different row-based permutation is applied to each of the parts, which, as will be shown later (section 7.3), makes inversion attacks infeasible (even under the *full-disclosure attacker model*, where the attacker is in possession of the permutation key). Potential loss of discriminative power due to the permutation is (mostly) avoided, since the horizontal neighbourhoods within rows persist. Note, that a column-wise permutation would not have had the desirable effect, due to the nature of Bloom filter-based Iris-Code representation explained in subsection 7.2.2.

7.2.2 Indexing and Retrieval

Following the permutation of the Iris-Codes, the enrolled templates are organised into tree-based search structures following the methods of [23] and [5] described below.

1. The Iris-Codes are evenly split into j equally sized blocks of adjustable height and width ($H \times W$). Subsequently, a simple transformation function is applied to the blocks column-wise, whereby each column (a binary string), is mapped to its corresponding decimal value.

2. For each block, an empty (*i.e.* all bits set to 0) Bloom filter ($\mathbf{b}$) of length 2^H is created and the indices corresponding to the decimal column values are set to 1.

3. Hence, the resulting template ($\mathbf{B}$) is a sequence of j such Bloom filters - $[\mathbf{b}_1, \ldots, \mathbf{b}_j]$.

4. The dissimilarity (DS) between two Bloom filter-based templates (denoted $\mathbf{B}$ and $\mathbf{B}'$) can be efficiently computed (utilising intrinsic CPU operations and trivially parallelisable), as shown in the equation below, where $|\cdot|$ represents the population count, *i.e.* Hamming weight.

$$DS(\mathbf{B}, \mathbf{B}') = \frac{1}{j} \sum_{i=1}^{j} \frac{|\mathbf{b}_i \oplus \mathbf{b}'_i|}{|\mathbf{b}_i| + |\mathbf{b}'_i|}$$

The Bloom filter-based templates are, to a certain degree, rotation-invariant, which means that contrary to the Iris-Codes, no alignment compensation is needed during the template comparison stage. Furthermore, the data representation is sparse, which is a crucial property for the indexing step described below. The representation sparseness is guaranteed, since for each Bloom filter of length 2^H, at most W (in practice fewer – due to the bit correlations in the Iris-Codes) indices are activated, and for the considered system configurations $W \ll 2^H$.

1. The list of N enrolled templates is (approximately evenly) split and assigned to T trees. This step is needed (for any sizeable N values) to maintain the sparseness of the data representation.

2. Each node of a tree (containing $M = \frac{N}{T}$ templates) is constructed through a union of templates, which corresponds to the binary OR applied to the individual Bloom filters in the sequence. The tree root is constructed from all templates assigned to the respective trees (*i.e.* $\bigcup_{m=1}^{M} \mathbf{B}_m$), while the children at subsequent levels are created each from half of the templates from their parent node (*e.g.* at first level – the children of the root node – $\bigcup_{m=1}^{\frac{M}{2}} \mathbf{B}_m$ and $\bigcup_{m=\frac{M}{2}+1}^{M} \mathbf{B}_m$).

3. The templates ($\mathbf{B}_1, \ldots, \mathbf{B}_M$) are inserted as tree leaves.

After constructing the trees, the retrieval can be performed as shown below.

1. A small number of the most promising trees (t) out of T constructed trees can be pre-selected (denoted Υ_T) based on comparison scores between the probe and root nodes.

2. The chosen trees are successively checked until the first candidate identity is found or all the pre-selected trees have been visited. Note, that for the genuine transactions, thanks to the pre-selection step, the trees most likely to contain the sought identity are visited first.

   A tree is traversed by, at each level, computing the comparison score between its nodes and the probe, and choosing the path with the best score. Once a leaf is reached, a final comparison takes place. The idea is based on the representation sparseness: as long as, at each level, the relation $DS_{genuine} \ll DS_{impostor}$ generally holds true, the genuine probes will be able to traverse the tree using the correct path to reach a matching leaf template. Note, that the row-based permutation (subsection 7.2.1) does not, in any way, impair the representation sparseness, since the average number of activated indices remains identical for the Bloom filters produced from permuted and unpermuted Iris-Codes.

The complexity of a single lookup is $O(T + t * (2 * \log M))$. As it is sufficient to pre-select only a small fraction of the constructed trees, *i.e.* $t \ll T$, the lookup workload remains low, while arbitrarily many enrollees can be accommodated. For reference, figure 7.2 shows the indexing and retrieval in a single tree. If multiple trees are constructed, the search is trivially parallelisable by simultaneously traversing many trees at once.

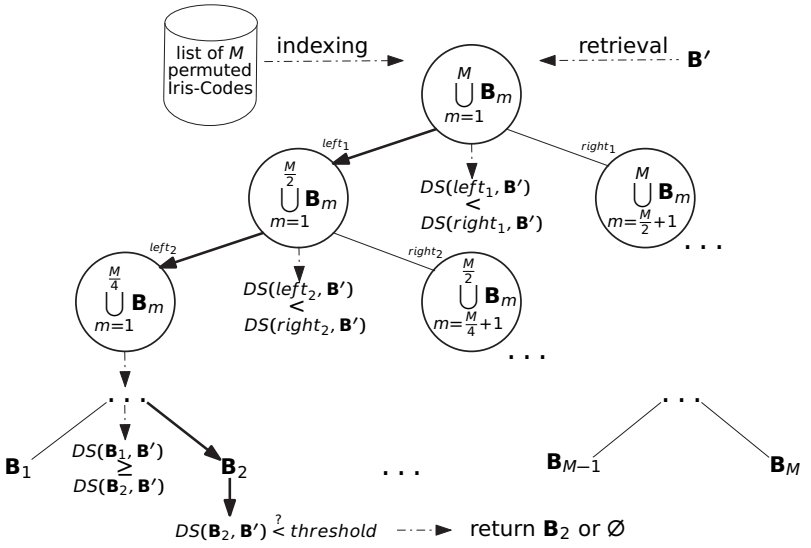


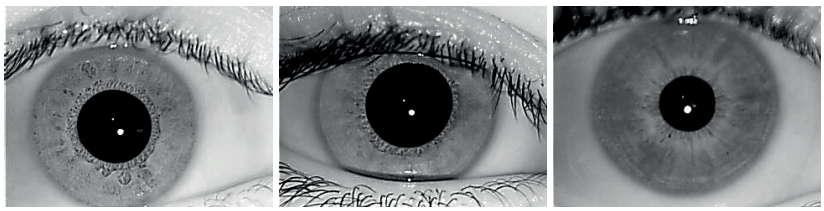
Figure 7.2: Indexing and retrieval in the Bloom filter-based system. In this case, the retrieval follows the bold arrow path down to a leaf, where the final decision is made.

7.3 Experiments

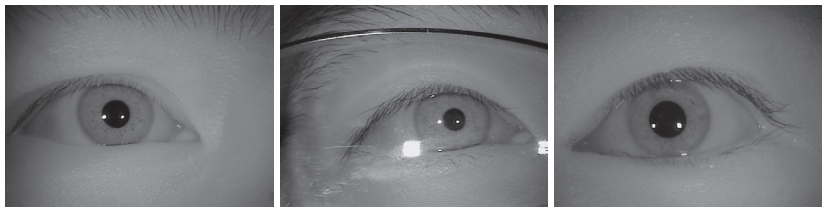
This section presents experiments performed to assess the proposed system. The experimental setup is outlined in subsection 7.3.1, while the performance and privacy evaluations are presented in subsection 7.3.2.

7.3.1 Experimental Setup

Two publicly available datasets of near-infrared iris images were chosen for the experiments: IITDv1 [15] and CASIA-IrisV4-Thousand [2] (henceforth referred to as "IITD" and "CASIA", respectively). They contain 1120 and 20000 images from 224 and 1000 subjects, respectively. Example images from the datasets are shown in figure 7.3.



(a) IITD

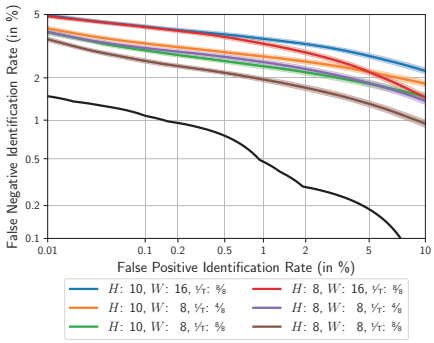


(b) CASIA

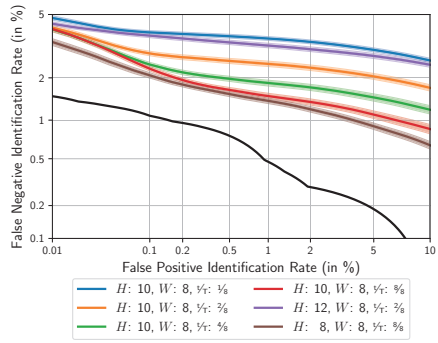
Figure 7.3: Example images from the chosen datasets.

The raw images were processed with the commonly used methods using open-source libraries: OSIRIS [20] and USIT [26]. After segmentation, where the iris and pupil boundaries are located, the iris textures were normalised according to the rubbersheet model [4] and subsequently enhanced by applying Contrast Limited Adaptive Histogram Equalization (CLAHE). Features were extracted with the Daugman-like 1D-LogGabor algorithm (LG), generating 512×20 bits Iris-Codes.

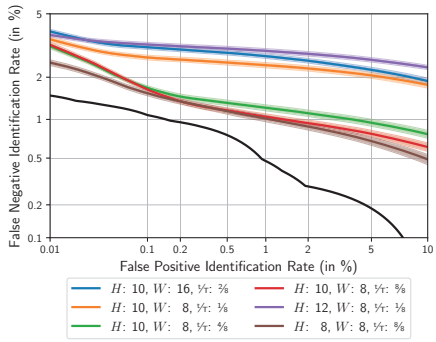
For the experiments, 256 references (from the IITD dataset – left and right eye instances are mutually independent and thus treated as separate subjects) were enrolled. The entire CASIA dataset together with the remainder of the IITD data are used to supply an ample number of impostor comparison trials. To make the evaluation more robust, 50 random permutations



(a) $IC_{parts} = 2$



(b) $IC_{parts} = 4$



(c) $IC_{parts} = 8$

Figure 7.4: DET curves for the proposed system. The faint colours around the curves represent the 95% confidence interval, while the black line represents the baseline (with EER of 0.66) – an Iris-Code system performing an exhaustive search and using ± 4 bit-shifts for sample alignment compensation.

are generated and used throughout the experiments. In other words, the performance evaluation for each system configuration is repeated 50 times with the different permutations of the Iris-Code templates.

Following metrics were used for evaluation of the various aspects of the proposed system:

Biometric performance: ISO/IEC IS 19795-1:2006 [13] metrics are used.

- The false positive and false negative identification rates plotted as detection error trade-off (DET) curves.
- The equal-error-rate (EER).

Workload: metrics from ISO/IEC IS 19795-1:2006 [13], and proposed in [5] is used.

- The penetration rate (p).
- The required number of bit comparisons per identification transaction expressed as a fraction (F) of the number of required Iris-Code baseline bit comparisons.

Template protection: metrics introduced in [9] are adopted.

- Unlinkability: the overall measure of the linkability of a given biometric template protection system ($\mathrm{D}_{\leftrightarrow}^{sys}$). It is computed in terms of the probabilities of having a mated or non-mated comparison for each possible linkage score between templates enrolled in different applications. It yields values in the closed range [0, 1], and reports a decreasing degree of unlinkability (*i.e.* increasing degree of linkability).
- Irreversibility: the success probability (P_{guess}) of guessing an original biometric template given a protected template under *full-disclosure attacker model* (*i.e.* the used permutation sequence is known to the attacker).
- Renewability: the number of possible permutation sequences, $|K|$ (*i.e.* the size of the key space).

7.3.2 Performance Evaluation and Protection Analysis

Figure 7.4 shows DET curves (with axes using a standard deviate scale [19]) for some of the best performing system configurations. Plots for each $IC_{parts} \in \{2, 4, 8\}$ show the three best configurations in terms of biometric performance and three best configurations in terms of workload reduction. The proposed protected indexing system exhibits high biometric performance, albeit naturally suffering a relatively small loss from the baseline Iris-Code based system. It can also be observed, that splitting the Iris-Code into more groups than just the real and imaginary parts prior to applying the permutation, is beneficial for the biometric performance. This is due to the fact that by splitting the Iris-Code into more parts, the potential for information loss due to permutation is decreased by preserving more spatial information. The plotted confidence intervals show that the biometric performance of the proposed system is stable across different permutations (in other words, changing the applied random permutation does not adversely affect the biometric performance of the system).

In table 7.1, the workload and security parameters of the proposed system (for the configurations plotted in figure 7.4) are listed. A significant

workload reduction is noticeable – the proposed system only requires between 1% and 10% of the workload incurred by the baseline system. This is achieved partly by decreasing the penetration rate as can be seen in the table, and partly by reducing the size of the biometric templates in terms of number of bits. Table 7.1 also shows the unlinkability, irreversibility and renewability of the proposed system[1]. It can be readily seen, that the keyspace ($|K|$) for the proposed system is huge, thereby ensuring renewability and contributing to the infeasibility of reversing the protected templates (P_{guess}), which is further enhanced by the nature of the Bloom filter based representation (some loss of local information). Lastly, the measure of global unlinkability ($\mathrm{D}_{\leftrightarrow}^{sys}$) for the tree leaves puts the proposed system (depending on the configuration) in close to fully unlinkable and semi-unlinkable region (as defined in [9]). Thus, for appropriate configuration selection, the security goals of a cancelable template protection scheme are accomplished.

Table 7.1: Results

IC_{parts}	H	W	$\frac{1}{T}$	EER	p	F	$\mathrm{D}_{\leftrightarrow}^{sys}$	P_{guess}	$\lvert K\rvert$
2	8	8	$\frac{8}{8}$	1.96	0.31	0.063	0.32	2^{-960}	2^{10097}
			$\frac{4}{8}$	2.15	0.19	0.038	0.29		
		16	$\frac{8}{8}$	3.01	0.31	0.063	0.45	2^{-1472}	2^{4414}
	10	8	$\frac{8}{8}$	2.11	0.31	0.063	0.09	2^{-960}	2^{10097}
			$\frac{4}{8}$	2.71	0.19	0.038	0.10		
		16	$\frac{8}{8}$	2.84	0.31	0.063	0.19	2^{-1536}	2^{4414}
4	8	8	$\frac{8}{8}$	1.46	0.31	0.063	0.31	2^{-960}	2^{20195}
	10	8	$\frac{8}{8}$	1.55	0.31	0.063	0.10		
			$\frac{4}{8}$	1.92	0.19	0.038	0.10		
			$\frac{2}{8}$	2.04	0.11	0.022	0.09		
			$\frac{1}{8}$	2.97	0.07	0.014	0.09		
	12	8	$\frac{2}{8}$	2.87	0.11	0.022	0.07	2^{-1080}	
8	8	8	$\frac{8}{8}$	1.12	0.31	0.063	0.31	2^{-960}	2^{40390}
	10	8	$\frac{8}{8}$	0.92	0.31	0.063	0.11		
			$\frac{4}{8}$	1.15	0.19	0.038	0.09		
			$\frac{1}{8}$	2.51	0.07	0.014	0.09		
		16	$\frac{2}{8}$	2.36	0.11	0.022	0.16	2^{-1536}	2^{17655}
	12	8	$\frac{2}{8}$	1.99	0.11	0.022	0.06	2^{-1080}	2^{40390}

[1] In calculations, the average number of activated bits in the Bloom filters must be rounded to the nearest integer, thus in some cases the resulting P_{guess} may be equal for different H values (particularly when $H = 8$ or $H = 10$). Furthermore, since the *full-disclosure attacker model* is used, the further effort of reversing the row-wise permutation (which would have been differ depending on H values) is not included in P_{guess}, since the attacker is assumed to be in possession of the used permutation sequences.

7.4 Summary

In this paper, an approach for indexing cancelable iris templates has been proposed. The approach is based on a row-wise permutation of the IrisCode rows and indexing them in Bloom filter-based tree structures. The experiments show that the proposed system fulfils the pre-requisites stipulated by ISO/IEC IS 24745:2011 for biometric template protection schemes (unlinkability, irreversibility, renewability and biometric performance), and additionally vastly reduces the workload associated with identification scenario – to less than 5% of the baseline system.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF) as well as by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP). One of the authors is supported by the TCS PhD Fellowship at IIIT-Delhi.

7.5 Bibliography

[1] AL-RAISI, A. N., AND AL-KHOURI, A. M. Iris recognition and the challenge of homeland and border control security in UAE. *Telematics and Informatics 25*, 2 (May 2008), 117–132.

[2] CHINESE ACADEMY OF SCIENCES' INSTITUTE OF AUTOMATION. CASIA iris image database. `http://biometrics.idealtest.org/`, December 2010. Last accessed: 2020–03–11.

[3] DAUGMAN, J. Biometric decision landscapes. Tech. Rep. UCAM-CL-TR-482, University of Cambridge - Computer Laboratory, January 2000.

[4] DAUGMAN, J. How iris recognition works. *Transactions on Circuits and Systems for Video Technology (TCSVT) 14*, 1 (January 2004), 21–30.

[5] DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Bloom filter-based search structures for indexing and retrieving Iris-Codes. *IET Biometrics 7*, 3 (May 2018), 260–268.

[6] EUROPEAN PARLIAMENT. Regulation (EU) 2016/679. *Official Journal of the European Union L119* (April 2016), 1–88.

[7] GADDE, R. B., ADJEROH, D., AND ROSS, A. Indexing iris images using the Burrows-Wheeler transform. In *International Workshop on Information Forensics and Security (WIFS)* (December 2010), IEEE, pp. 1–6.

[8] Gentile, J. E., Ratha, N., and Connell, J. An efficient, two-stage iris recognition system. In *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2009), IEEE, pp. 211–215.

[9] Gomez-Barrero, M., Galbally, J., Rathgeb, C., and Busch, C. General framework to evaluate unlinkability in biometric template protection systems. *Transactions on Information Forensics and Security (TIFS) 3*, 6 (June 2018), 1406–1420.

[10] Gomez-Barrero, M., Rathgeb, C., Li, G., Ramachandra, R., Galbally, J., and Busch, C. Multi-biometric template protection based on Bloom filters. *Information Fusion 42* (July 2018), 37–50.

[11] Hao, F., Daugman, J., and Zielinski, P. A fast search algorithm for a large fuzzy database. *Transactions on Information Forensics and Security (TIFS) 3*, 2 (June 2008), 203–212.

[12] ISO/IEC JTC1 SC27 IT Security techniques. *ISO/IEC 24745:2011. Information technology – Security techniques – Biometric information protection*. International Organization for Standardization and International Electrotechnical Committee, June 2011.

[13] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC 19795-1:2006. Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*. International Organization for Standardization and International Electrotechnical Committee, April 2006.

[14] Konrad, M., Stögner, H., Uhl, A., and Wild, P. Computationally efficient serial combination of rotation-invariant and rotation compensating iris recognition algorithms. In *International Conference on Computer Vision Theory and Applications (VISAPP)* (May 2010), SciTePress, pp. 85–90.

[15] Kumar, A., and Passi, A. Comparison and combination of iris matchers for reliable personal authentication. *Pattern Recognition 43*, 3 (March 2010), 1016–1026.

[16] Mehrotra, H., Srinivas, B. G., and Majhi, B. Indexing iris biometric database using energy histogram of DCT subbands. In *International Conference on Contemporary Computing (IC3)* (August 2009), vol. 40, Springer, pp. 194–204.

[17] Mukherjee, R., and Ross, A. Indexing iris images. In *International Conference on Pattern Recognition (ICPR)* (December 2008), IEEE, pp. 1–3.

[18] Nandakumar, K., and Jain, A. K. Biometric template protection: Bridging the performance gap between theory and practice. *Signal Processing Magazine 32*, 5 (September 2015), 88–100.

[19] NAUTSCH, A., MEUWLY, D., RAMOS, D., LINDH, J., AND BUSCH, C. Making likelihood ratios digestible for cross-application performance assessment. *Signal Processing Letters 24*, 10 (October 2017), 1552–1556.

[20] OTHMAN, N., DORIZZI, B., AND GARCIA-SALICETTI, S. OSIRIS: An open source iris recognition software. *Pattern Recognition Letters 82*, 2 (September 2016), 124–131.

[21] PROENÇA, H. Iris biometrics: Indexing and retrieving heavily degraded data. *Transactions on Information Forensics and Security (TIFS) 8*, 12 (December 2013), 1975–1985.

[22] QIU, X., SUN, Z., AND TAN, T. Global texture analysis of iris images for ethnic classification. *International Conference on Biometrics (ICB) 3832* (January 2006), 411–418.

[23] RATHGEB, C., BREITINGER, F., BAIER, H., AND BUSCH, C. Towards bloom filter-based indexing of iris biometric data. In *International Conference on Biometrics (ICB)* (May 2015), IEEE, pp. 422–429.

[24] RATHGEB, C., AND UHL, A. A survey on biometric cryptosystems and cancelable biometrics. *EURASIP Journal on Information Security 2011*, 1 (September 2011), 1–25.

[25] RATHGEB, C., UHL, A., AND WILD, P. Incremental iris recognition: A single-algorithm serial fusion strategy to optimize time complexity. *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2010), 1–6.

[26] RATHGEB, C., UHL, A., WILD, P., AND HOFBAUER, H. Design decisions for an iris recognition SDK. In *Handbook of Iris Recognition*, K. Bowyer and M. J. Burge, Eds., 2 ed., Advances in Computer Vision and Pattern Recognition. Springer, July 2016, pp. 359–396.

[27] RATHGEB, C., WAGNER, J., AND BUSCH, C. Iris biometric template protection. In *Iris and Periocular Biometric Recognition*. Institution of Engineering and Technology, August 2017, pp. 317–340.

[28] ROSS, A., AND SUNDER, M. S. Block based texture analysis for iris classification and matching. In *Conference on Computer Vision and Pattern Recognition - Workshops (CVPRW)* (June 2010), IEEE, pp. 30–37.

[29] SUN, Z., ZHANG, H., TAN, T., AND WANG, J. Iris image classification based on hierarchical visual codebook. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 36*, 6 (June 2014), 1120–1133.

[30] Unique Identification Authority of India. Aadhaar dashboard. https://www.uidai.gov.in/aadhaar_dashboard/. Last accessed: 2020–03–11.

Chapter 8

Benchmarking Binarisation Schemes for Deep Face Templates

Abstract

Feature vectors extracted from biometric characteristics are often represented using floating point values. It is, however, more appealing to store and compare feature vectors in a binary representation, since it generally requires less storage and facilitates efficient comparators which utilise intrinsic bit operations. Furthermore, the binary representations are very often necessary for some specific application scenarios, e.g. template protection and indexing.

In recent years, usage of deep neural networks for facial recognition has vastly improved the biometric performance of said systems. In this paper, various binarisation schemes are applied to such feature vectors and benchmarked for biometric performance. It is shown that with only a negligible drop in biometric performance, the storage space and computational requirements can be vastly decreased.

Addressed research question(s): RQ1
Reference: DROZDOWSKI, P., STRUCK, F., RATHGEB, C., AND BUSCH, C. Benchmarking binarisation schemes for deep face templates. In *International Conference on Image Processing (ICIP)* (October 2018), IEEE, pp. 191–195.

8.1 Introduction

Face is one of the most widely used biometric characteristics. Various methods have been proposed over the span of last three decades [7, 23]. In recent years, methods based on deep learning (*e.g.* [15, 20, 21, 22]) have been proposed, and significantly improved on the biometric performance of the heretofore existing methods. With this improved biometric performance, face has become an attractive characteristic for large-scale identification systems.

The deep face feature representations typically involve float-valued vectors, for which the template comparison is performed using metrics such as Euclidean distance (L^2 norm) or Chi-square distance (χ^2). Those metrics are

computationally expensive – thus creating a potential efficiency bottleneck for large-scale biometric identification systems, where 1:N template comparisons are performed during lookup. Additionally, transmission of such feature vectors from low-cost mobile devices to central systems requires a compact encoding, specifically when the bandwith of mobile networks is limited. *Binarisation* of feature vectors offers an attractive alternative – such templates can be stored efficiently and be compared quickly in the Hamming domain utilising intrinsic CPU operations (*i.e.* xor and popcount) [18].

Over time, many methods of binarising data have been proposed, mostly with template protection as motivation (to transform the features to certain input forms required by the different cryptographic primitives) [12] and shown to work with, among others, classical facial recognition systems. However, it is unknown, whether or not those approaches are suitable for the vectors produced by deep learning based face feature extractors and their potential biometric performance degradation due to information loss has not been studied thoroughly. With this uncertainty as the motivation, the main contribution of this paper is such a benchmark, where various binarisation methods are evaluated in terms of biometric performance and computational workload incurred at the comparison stage.

The remainder of this paper is organised as follows: Section 8.2 introduces the related work. In section 8.3, binarisation schemes for deep facial templates are described. In section 8.4, the experimental setup and results are presented, while section 8.5 contains a summary of the paper and future work items.

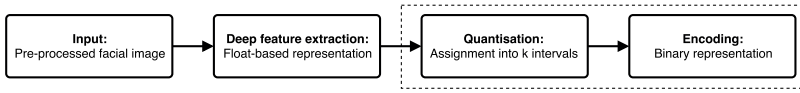


Figure 8.1: Processing chain

8.2 Related Work

In the recent decade, several data binarisation approaches have been proposed. Kevenaar *et al*. [8] extract the most reliable components of facial feature vectors and binarise them for use in a template protection scheme. Chen *et al*. [5] present a detection rate optimized bit allocation (DROBA) principle, which is biometric characteristic-agnostic. Based on the discriminative power of the features, it assigns more or fewer bits to them during binarisation, thus improving the biometric performance of the binarised feature representation. Bringer *et al*. [2] transform fingerprint minutiae set using a vicinity-based approach, which in addition to producing a compact feature representation, also exhibits self-alignment property. When present-

ing a novel fingerprint minutiae representation scheme, Cappelli *et al*. [3] note that it can also operate in binarised mode, without significantly decreasing the biometric performance of the scheme. Lee *et al*. [10] binarise facial PCA/ERE-based templates using a generalised Linnartz and Tuyl's quantisation index modulation (QIM) scheme for the purpose of template protection. Chen *et al*. [4] present a generic (for arbitrary characteristics with float-valued feature vectors) binarisation scheme using pairwise adaptive phase quantization and long-short pairing strategy. Lim *et al*. [13] describe a DROBA-based approach, in which bit statistics (reliability and discriminability) to improve the biometric performance of the binarised representation of facial features. In Lim *et al*. [11], the authors propose two new encoding schemes (LSSC and PLSSC – (Partially) Linearly Separable Subcode) which exhibit full-ideal and near-ideal separability capabilities, respectively. Schlett *et al*. [19] describe a simple, yet effective, scheme for binarising multi-scale LBP histograms.

In general, the results presented in the summarised state-of-the-art show, that various float-value based feature representations can be effectively transformed into compact binary strings, without a significant drop in biometric performance, when benchmarked against the original data representation.

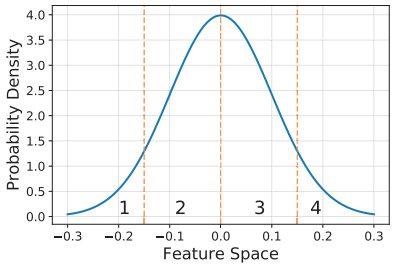
8.3 Binarisation of Deep Face Templates

Figure 8.1 shows a high-level view of the facial image processing chain used in this paper with the key steps (for this paper) highlighted. First, common pre-processing steps including region of interest detection, alignment and normalisation are applied. The current state-of-the-art deep facial recognition frameworks then extract feature vectors consisting of a predefined number of floating point values. Specific details regarding the preprocessing and feature extraction steps are given in subsection 8.4.1. To avoid computational overhead during comparison stage (computing Euclidean distance with floating-point numbers, as mentioned in section 8.1), the feature vector can be binarised. Normally, this process consists of two steps [12]: 1. Quantisation (subsection 8.3.1) and 2. Encoding (subsection 8.3.2).

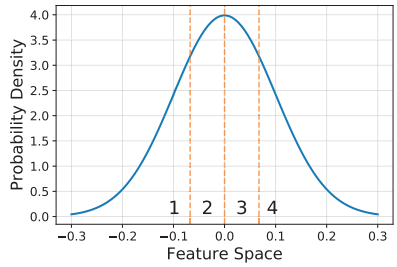
8.3.1 Quantisation

During quantisation, the values from the feature vector are mapped to a number of integer-labelled intervals over the feature space probability density (feature extraction algorithm dependent, obtained via a training set). In this paper, two quantisation schemes listed below are utilised and visualised in figure 8.2.

- Equal-width quantile: the feature space is divided into segments of equal size (figure 8.2a)
- Equal-probable quantile: the feature space is divided into segments containing equal population probability mass (figure 8.2b)



(a) Equal-width



(b) Equal-probable

Figure 8.2: Quantisation

8.3.2 Encoding

After quantisation, in the encoding step, the aforementioned quantised intervals (represented as integers) are mapped to short binary strings, which are subsequently concatenated to produce the final feature representation. The dissimilarity of two such templates can be then computed using Hamming distance. The encodings used in this paper are listed below.

- **Boolean** The simplest scheme, where the feature space is quantised into 2 sub-spaces (*i.e.* the resulting encoding is a single 0 or 1).
- **DBR** (Direct Binary Representation) In which the decimal numbers from quantisation are converted directly into their binary representations.
- **BRGC** (Binary Reflected Gray Code [6]) In which the encoding is done so that the distance in the Hamming domain between codewords resulting from successive decimal values is always 1.
- **LSSC** (Linearly Separable Subcode [11]) A more recent approach, which offers ideal separability, *i.e.* the distances between two values are the same in the decimal and Hamming domains.
- **Sparse** A simple scheme, in which the number of encoded bits is equal to that of quantised sub-spaces (k) and only one bit is set to 1 – that

corresponding to the sub-space index resulting from the quantisation step. When quantising into larger number of sub-spaces, this can result in a sparse binary feature vector.

Table 8.1 shows an example with 4 quantisation intervals and the encoding methods described above. Intuitively, there exists a trade-off between the ability to obtain better separation, representation sparsity and the required length of the encoding. In the next section, the schemes are put to test by assessing their biometric performance with deep facial feature vectors.

Table 8.1: Encoding schemes

Quantisation Interval	**Encoding**				
	Boolean	DBR	BRGC	LSSC	Sparse
1	0	00	00	000	0001
2	1	01	01	001	0010
3	—	10	11	011	0100
4	—	11	10	111	1000

8.4 Experiments

This section contains the evaluation of the binarisation schemes described earlier. In subsection 8.4.1, the used datasets and experimental setup details are outlined, while the results are presented and discussed in subsection 8.4.2.

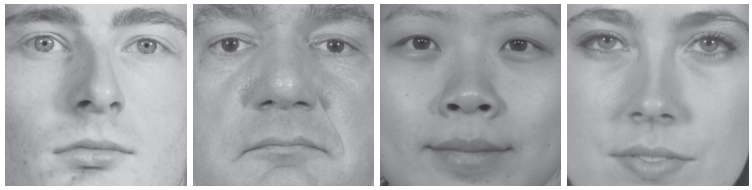
8.4.1 Experimental Setup

Three commonly used facial datasets, summarised in table 8.2, were chosen for experiments in this paper. From the FERET dataset, only frontal images were used, while from the AR Face dataset, frontal images without intentional obfuscations (such as sunglasses or scarves) were used. From the FRGC dataset, the complete "Fall2003" subset was used.

Table 8.2: Overview of the data used for experiments

Dataset	**Subjects**	**Images**	**Comparisons**	
			Genuine	Impostor
AR Face [14]	133	741	1757	8777
FERET [17]	994	2722	3649	493520
FRGC [16]	370	11358	219851	68264

(a) FERET



(b) FRGC

(c) AR Face

Figure 8.3: Example images after pre-processing

In the pre-processing stage, the face of a subject is detected and normalised according to eye coordinates detected by the *dlib* landmark detector [9]. Subsequently, the normalised region is cropped to 320×320 pixels and converted to grayscale. Example images from the used datasets (after pre-processing) are shown in figure 8.3. Thereafter, two state-of-the-art, open-source deep facial recognition frameworks (OpenFace [1] and FaceNet [20]) were used to extract features from the images. The resulting representation is a 1-dimensional feature vector containing 128 float values. The frameworks utilised pre-trained (on datasets disjoint from the ones used for the binarisation experiments in this paper) models, made available by their authors, were used.

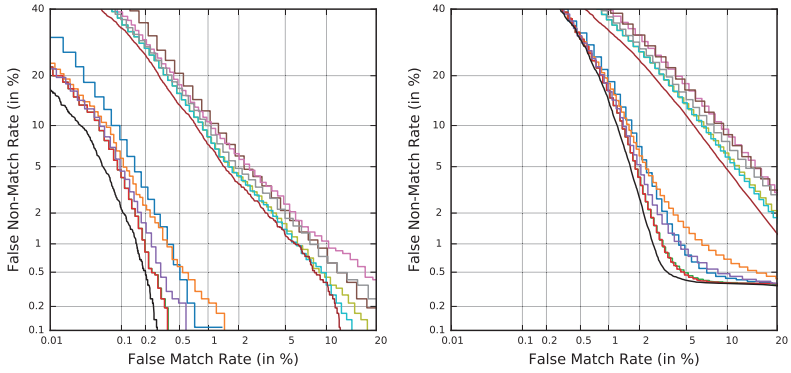
Suitable thresholds for quantisation intervals are determined via training on the feature space of the AR Face dataset and then used directly in tests on the remaining two datasets. For each binarisation method, all possible template comparisons (verification transactions) were performed to compute the biometric performance of the system. The baseline biometric performance was computed using the aforementioned original, float-based templates, which are compared using squared Euclidean distance.

The metrics used for evaluation were:

- Biometric performance: Detection error trade-off curve (DET) and equal error rate (EER)
- Template size: bits
- Computational workload: CPU instructions required to perform a single template comparison

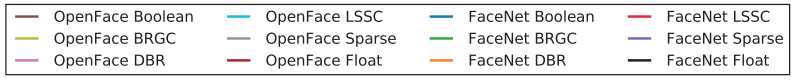
8.4.2 Results

Figure 8.4 shows DET curves for the performed experiments on the FERET (figure 8.4a) and FRGC (figure 8.4b) datasets. It can be seen, that using the FaceNet feature extractor yields results vastly superior to that of OpenFace. Furthermore, it can be seen, that the float-based representation performs only marginally better than the best binarisation schemes.



(a) FERET

(b) FRGC



(c) Legend

Figure 8.4: DET curves

In terms of EER, the baseline performance on the FERET dataset is 2.68% and 0.23% EER for OpenFace and FaceNet, respectively, while on the FRGC dataset, it is 7.35% and 2.13% EER for OpenFace and FaceNet, respectively. The experimental results for binarisation schemes are shown in table 8.3

with best result for each feature extractor/dataset pair marked in bold. Generally, the LSSC encoding has the best performance, which was to be expected, since it offers better separability than the remaining encodings. In most cases, the equal-width quantisation was better than the equal-probable quantisation. In summary, the best quantisation/encoding method pairs suffer only a negligible loss of biometric performance (in terms of EER) against the float-based baseline system: FaceNet loses 0.06 and 0.14 percentage points, while OpenFace loses 0.19 and 0.53 percentage points on FERET and FRGC datasets, respectively.

Table 8.3: Results (best one(s) for each dataset/extractor pair marked in bold)

Encoding	**Quantisation**	**Size (bits)**	**Performance (EER)**			
			FERET		FRGC	
			FaceNet	OpenFace	FaceNet	OpenFace
Boolean	eq. width	128	0.47%	3.34%	2.85%	9.10%
	eq. probable		0.49%	3.56%	2.80%	9.49%
DBR	eq. width	256	0.52%	3.46%	2.99%	9.31%
		384	0.98%	3.85%	3.72%	10.10%
	eq. probable	256	0.71%	3.65%	3.31%	9.30%
		384	0.76%	3.75%	3.60%	9.36%
BRGC	eq. width	256	**0.29%**	2.87%	**2.32%**	7.95%
		384	0.31%	2.99%	2.36%	8.17%
	eq. probable	256	0.35%	3.20%	2.61%	8.33%
		384	0.36%	3.30%	2.73%	8.42%
LSSC	eq. width	384	**0.29%**	**2.82%**	**2.32%**	**7.88%**
	eq. probable		**0.29%**	2.92%	2.40%	7.97%
Sparse	eq. width	512	0.34%	3.13%	2.54%	8.49%
	eq. probable		0.47%	3.36%	2.92%	8.63%

The binary templates are compared using Hamming distance, *i.e.* by using a binary xor followed by a $\mathrm{popcount}$, both of which are intrinsic operations on the vast majority of modern processors. By storing the binary vectors in arrays of unsigned integers, 64 bits at a time can be handled and then summed up using add operations. Hence, for comparing a binary vector of length $64 * n$, the required number of operations is: $3 * n - 1$. The float-based templates are in this case compared using squared Euclidean distance, which in one dimension corresponds to computing the dot product of the difference between the two feature vectors, *i.e.* a sum of element-wise multiplication between two copies of the difference vector. Table 8.4 summarises the required instruction numbers for the template representation types and sizes used in this paper's experiments. The original representation requires an order of magnitude more instructions; furthermore, those require floating point arithmetic instead of binary/integer arithmetic. It is therefore clear, that the binarised representation is vastly more efficient computationally.

Table 8.4: CPU instructions per template comparison

Representation	Instruction Type	Count
128 floats	float sub, mul and add	383
128 bits	binary xor and popcount	5
256 bits		11
384 bits		17
512 bits		23

8.5 Summary

In this paper, several methods for quantisation and encoding of float-valued deep (OpenFace and FaceNet) feature representation of facial images were benchmarked. In tests on commonly used large facial datasets, the binarised templates suffer only a negligible biometric performance loss against the original, float-valued representation of the deep facial templates, while vastly reducing the template size in bits. As a consequence of the more compact feature vector, and also by being able to use intrinsic CPU operations for template comparison, the computational and storage requirements are vastly reduced. This benchmark thus reveals that the existing binarisation methods can be readily applied to feature vectors produced by deep neural networks.

A promising future work avenue is using the binarised deep face templates to perform (multi-)biometric indexing for further workload reduction in large-scale biometric identification systems.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF), by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP), and the LOEWE-3 BioBiDa Project (594/18-17).

8.6 Bibliography

[1] AMOS, B., LUDWICZUK, B., AND SATYANARAYANAN, M. OpenFace: A general-purpose face recognition library with mobile applications. Tech. Rep. CMU-CS-16-118, CMU School of Computer Science, 2016.

[2] BRINGER, J., AND DESPIEGEL, V. Binary feature vector fingerprint representation from minutiae vicinities. In *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2010), IEEE, pp. 1–6.

[3] CAPPELLI, R., FERRARA, M., AND MALTONI, D. Minutia cylinder-code: A new representation and matching technique for fingerprint recognition. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 32*, 12 (December 2010), 2128–2141.

[4] CHEN, C., AND VELDHUIS, R. N. J. Binary biometric representation through pairwise adaptive phase quantization. *EURASIP Journal on Information Security 2011*, 1 (February 2011), 1–16.

[5] CHEN, C., VELDHUIS, R. N. J., AND AKKERMANS, T. A. M. K. A. H. M. Biometric quantization through detection rate optimized bit allocation. *EURASIP Journal on Advances in Signal Processing 2009*, 1 (May 2009), 1–16.

[6] GRAY, F. Pulse code communications, March 1953. U.S. Patent 2,632,058.

[7] JAIN, A. K., AND LI, S. Z. *Handbook of face recognition*. Springer, 2004.

[8] KEVENAAR, T. A. M., SCHRIJEN, G. J., VAN DER VEEN, M., AKKERMANS, A. H. M., AND ZUO, F. Face recognition with renewable and privacy preserving binary templates. In *Workshop on Automatic Identification Advanced Technologies (AutoID)* (October 2005), IEEE, pp. 21–26.

[9] KING, D. E. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research (JMLR) 10* (2009), 1755–1758.

[10] LEE, H., TEOH, A. B. J., JUNG, H. G., AND KIM, J. A secure biometric discretization scheme for face template protection. *Future Generation Computer Systems 28*, 1 (January 2012), 218–231.

[11] LIM, M.-H., AND TEOH, A. B. J. A novel encoding scheme for effective biometric discretization: Linearly separable subcode. *Trans. on Pattern Analysis and Machine Intelligence (TPAMI) 35*, 2 (February 2013), 300–313.

[12] LIM, M.-H., TEOH, A. B. J., AND KIM, J. Biometric feature-type transformation: Making templates compatible for secret protection. *Signal Processing Magazine 32*, 5 (September 2015), 77–87.

[13] LIM, M.-H., TEOH, A. B. J., AND TOH, K.-A. An efficient dynamic reliability-dependent bit allocation for biometric discretization. *Pattern Recognition 45*, 5 (May 2012), 1960–1971.

[14] MARTÍNEZ, A. M., AND BENAVENTE, R. The AR face database. Tech. Rep. 24, CVC, June 1998.

[15] PARKHI, O. M., VEDALDI, A., ZISSERMAN, A., ET AL. Deep face recognition. In *British Machine Vision Conference (BMVC)* (September 2015), BMVA Press, pp. 1–6.

[16] PHILLIPS, P. J., FLYNN, P. J., SCRUGGS, T., BOWYER, K. W., CHANG, J., ET AL. Overview of the face recognition grand challenge. In *Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2005), vol. 1, IEEE, pp. 947–954.

[17] PHILLIPS, P. J., MOON, H., RIZVI, S. A., AND RAUSS, P. J. The FERET evaluation methodology for face-recognition algorithms. *Transactions on pattern analysis and machine intelligence (TPAMI) 22*, 10 (October 2000), 1090–1104.

[18] RATHGEB, C., BUCHMANN, N., HOFBAUER, H., BAIER, H., UHL, A., AND BUSCH, C. Methods for accuracy-preserving acceleration of large-scale comparisons in CPU-based iris ecognition systems. *IET Biometrics 7*, 4 (July 2018), 356–364.

[19] SCHLETT, T., RATHGEB, C., AND BUSCH, C. A binarization scheme for face recognition based on multi-scale block local binary patterns. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2016), IEEE, pp. 1–4.

[20] SCHROFF, F., KALENICHENKO, D., AND PHILBIN, J. FaceNet: A unified embedding for face recognition and clustering. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2015), IEEE, pp. 815–823.

[21] SUN, Y., WANG, X., AND TANG, X. Deeply learned face representations are sparse, selective, and robust. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2015), IEEE, pp. 2892–2900.

[22] TAIGMAN, Y., YANG, M., RANZATO, M., AND WOLF, L. DeepFace: Closing the gap to human-level performance in face verification. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2014), IEEE, pp. 1701–1708.

[23] ZHAO, W., CHELLAPPA, R., PHILLIPS, P. J., AND ROSENFELD, A. Face recognition: A literature survey. *Computing surveys (CSUR) 35*, 4 (December 2003), 399–458.

Chapter 9

On the Application of Homomorphic Encryption to Face Identification

Abstract

The data security and privacy of enrolled subjects is a critical requirement expected from biometric systems. This paper addresses said topic in facial biometric identification. In order to fulfil the properties of unlinkability, irreversibility, and renewability of the templates required for biometric template protection schemes, homomorphic encryption is utilised. In addition to achieving the aforementioned objectives, the use of homomorphic encryption ensures that the biometric performance remains completely unaffected by the template protection scheme.

The main contributions of this paper are: It proposes an architecture of a system capable of performing biometric identification in the encrypted domain, as well as provides and evaluates an implementation using two existing homomorphic encryption schemes. Furthermore, it discusses the pertinent technical considerations and challenges in this context.

Addressed research question(s): RQ1, RQ5
Reference: DROZDOWSKI, P., BUCHMANN, N., RATHGEB, C., MARGRAF, M., AND BUSCH, C. On the application of homomorphic encryption to face identification. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2019), IEEE, pp. 173–180.

9.1 Introduction

Data exposure is a potential risk in biometric system deployments, which typically store their data secured using traditional encryption algorithms [19]. If this protection were to be compromised, serious problems would arise, including but not limited to, identity theft, cross-matching without consent, and severely limited renewability. Increasingly, the public and non-governmental organisations pay attention to those (and other) issues associated with centralised storage of sensitive personal and biometric data. In some areas, this contributes to the political process of widening legislation

against privacy violations (*e.g.* GDPR in Europe [11]), which entail significant responsibilities for the data controllers.

Recently, biometric *template protection* (see *e.g.* [22] for a survey) has been an active research field attempting to address said security and privacy challenges. The ISO/IEC IS 24745 [17] mandates several properties, which must be guaranteed by such schemes:

- **Unlinkability** referring to making it infeasible to determine if two or more protected templates were derived from the same instance. By fulfilling this property, cross-matching across different databases is prevented.
- **Irreversibility** referring to making it infeasible to reconstruct the original biometric data given a protected template and its corresponding secret. With this property fulfilled, the privacy of the users' data is increased, and additionally the security of the system is increased against presentation and replay attacks.
- **Renewability** referring to making it possible to revoke old protected templates and creating new ones from the same biometric instance and/or sample. With this property fulfilled, it is possible to revoke and reissue the templates in case of the database being compromised, thereby preventing misuse.
- **Performance preservation** referring to the requirement of the biometric performance not being significantly impaired by the protection scheme.

Three main biometric template protection approach classes can be distinguished: (1) biometric cryptosystems, which use the biometric data to bind or extract a key [7], (2) cancelable biometrics, which utilise irreversible transformations to the biometric samples or templates [20], and (3) (homomorphic) encryption of biometric data [2].

Homomorphic encryption (henceforth referred to as "HE") makes it possible to compute operations in the encrypted domain, which render the same result as those in the plaintext domain. Thus, provided that it is possible to implement a given biometric comparator to feasibly operate within the homomorphic domain, such a template protection scheme would operate without any loss of biometric performance, whereas some impairment is often inevitable in biometric cryptosystems and cancelable biometrics. In general, an encryption algorithm E has the homomorphic property for an operation $\odot$ if it holds $E(m_1) \odot E(m_2) = E(m_1 \odot m_2), \forall m_1, m_2 \in M$ where M is the set of all possible messages. HE schemes are classified depending on the number and type of supported $\odot$ operations (see *e.g.* [1] for a detailed survey). The following three HE scheme types exist today:

- **Partially Homomorphic Encryption (PHE)** schemes are defined as allowing only a single operation type an unlimited number of times. PHE schemes have been around for over 30 years and are the oldest HE schemes like the classical RSA scheme [23] and the El-Gamal scheme [10] supporting only either addition or multiplication.
- **Somewhat Homomorphic Encryption (SWHE)** schemes allow multiple operation types, but only a limited number of times. SWHE examples from literature are Yao's garbled circuit scheme [26], which supports arbitrary operations a limited number of times and the Boneh-Goh-Nissim (BGN) scheme [4], which supports unlimited number of additions and one multiplication.
- **Fully Homomorphic Encryption (FHE)** schemes support an unlimited number of operations. The first feasible FHE scheme was proposed by Gentry [13] in 2009 and many newer FHE schemes are based on Gentry's general FHE framework. Brakerski and Vaikuntanathan [6] utilise Gentry's framework to make their SWHE scheme a FHE scheme and introduce batching as an optimization [5]. The Brakerski's scheme was later optimised by Fan and Vercauteren [12].

While several authors investigated using FHE for biometric *verification* (see *e.g.* [3, 15, 16]) with promising results, the biometric *identification* scenario has not yet been addressed or merely considered a trivial extension of the proposed schemes. However, there exist several challenges and issues which must be dealt with for such schemes to be viable in the biometric identification scenario, especially if computational workload reduction (*i.e.* decreasing the computational complexity of the retrieval, see *e.g.* [18] for a survey) is to be employed. Accordingly, the contribution of this paper is twofold: (1) an example architecture of a system capable of performing biometric identification with homomorphically protected templates is described, implemented, and evaluated with a facial recognition system, and (2) the practical considerations and challenges relevant to the biometric identification scenario are discussed in the context of HE and potential solutions, along with the future research avenues being explored.

This paper is organised as follows: in section 9.2 the proposed system is described. Section 9.3 outlines the experimental setup and the results of the evaluations. The results and other relevant matters are discussed in section 9.4, while concluding remarks and a summary are given in section 9.5.

9.2 Proposed System

Figure 9.1 shows an overview of the proposed system. There are 3 entities in the system: A client, where the biometric features are extracted (not depicted) and encrypted; a database, where the encrypted references of the

enrolled subjects are stored and the distances between them and the probe computed in the encrypted domain and applies a decision threshold; and a trusted third party (TTP), which decrypts the thresholded scores and communicates a decision to the client. This description is an abstract one, the concrete HE schemes used in the implementation and evaluation are listed in section 9.3.

The proposed system ensures that the unencrypted, privacy-sensitive biometric features are only available to the client. All network transfers of the biometric probe, as well as the computations on the data happen in the encrypted domain, thereby preventing eavesdropping attacks on the network and malicious or rogue database system attacks. Since all database entries are stored encrypted, a database breach or insider threats yield no valuable attack vectors. To further strengthen the security of the proposed system, following measures are possible: 1) to curtail active attacks (*e.g.* Man-in-the-Middle or skimming) against the transferred feature vectors and the identification transaction decision, TLS can be deployed between the parties; 2) to prevent the attacker from conducting an analysis of a single entry or unsanctioned database audits, the order of the database entries could be randomly shuffled during or after each identification transaction. Both measures would have a negligible impact on the access time and no impact on the biometric performance.

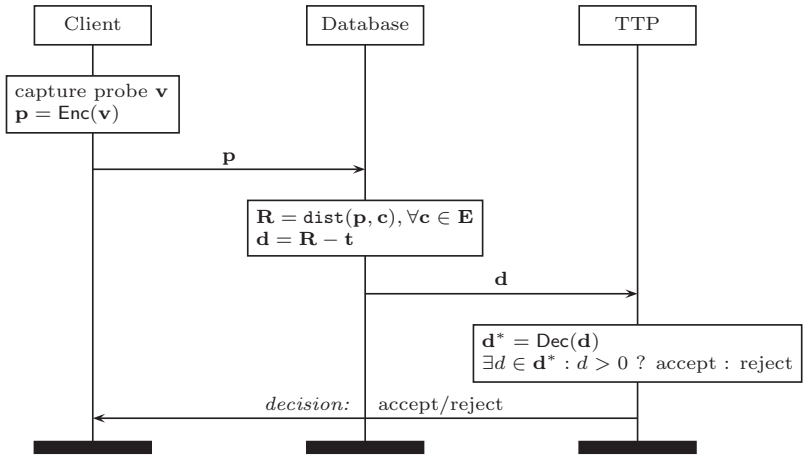


Figure 9.1: System overview sketch

1. Let $\mathbf{v}$ denote a biometric feature vector, with a constant number (n) of feature elements: $\mathbf{v} = (v_1, v_2, \ldots v_n)$ and $\mathbf{c}$ denote the element-wise encrypted version of said reference vector using a HE scheme with batching. Batching is a mechanism which allows to perform operations on vectors rather than individual numbers in the encrypted do-

main, thereby allowing vast (by several orders of magnitude) speed-ups on vectorisable operations. By design, the HE schemes used in this paper support a certain number of slots (s) for the elements in the encrypted vector. This number depends on the parameters of the encryption algorithm, including the encryption security. In this case (see subsection 9.3.1), $n < s$ for any reasonable (from the security perspective) combination of parameters, which means that $\mathbf{c}$ must be padded with zeroes beyond the n-th index, *i.e.* $\mathbf{c} = (c_1, c_2, \dots c_n, 0 \dots 0)$.

2. Let $\mathbf{E}$ denote an enrolment database consisting of N such encrypted vectors (subjects), *i.e.* $\mathbf{E} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_N \end{bmatrix}$. Let $\mathbf{p}$ denote an encrypted feature vector of a probe and $\odot$ an arbitrary function for distance computation between two biometric feature vectors. In this work, the squared Euclidean distance will be used, *i.e.* $\texttt{dist}(\mathbf{c}, \mathbf{p}) = \sum_{i=0}^{s} (\mathbf{c}_i - \mathbf{p}_i)^2$. While performing the subtraction and exponentiation in the encrypted domain is done trivially by directly using the operators in the encrypted domain, the summing up of the individual elements is not as straightforward, since batching techniques do not allow access to individual encrypted vector elements. To compute the sum, the observation made in [14] is utilised – the vector is successively circularly shifted and added onto itself s times. Afterwards, the sum of elements is present in the first element of the vector. The other elements of the vector are now irrelevant and are cleaned by multiplying with 0. Mathematically, a vector $\mathbf{k} = \{1, 0, \dots 0\}$ is defined and multiplied with the result.

3. The goal of a biometric identification is to first compute the comparison scores between the probe and all the items in the enrolment database, select the best one, and compare it against a predetermined threshold to make the final decision. Let $\mathbf{r}_i$ denote the result of comparison between the probe and the i'th entry in the database, *i.e.* $\mathbf{r}_i = \texttt{dist}(\mathbf{p}, \mathbf{c}_i) \cdot \mathbf{k}$. After processing the whole enrolment database, N such result vectors as separate ciphertexts are created: $\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_N \end{bmatrix} = \begin{bmatrix} (\mathbf{r}_{1,1}, 0, \dots 0) \\ (\mathbf{r}_{2,1}, 0, \dots 0) \\ \vdots \\ (\mathbf{r}_{N,1}, 0, \dots 0) \end{bmatrix}$ note, that each of $\mathbf{r}_i$ is encrypted in a separate ciphertext, rather than $\mathbf{R}$ being encrypted in one ciphertext. In the next step, those individual ciphertexts are combined. At this point, it is also possible to randomly shuffle the order of the ciphertexts, thereby preventing the

trusted third party from later learning which subject has been identified.

4. Through shifting the $\mathbf{r}_i$ vectors, a structure conceptually akin to a diagonal matrix is reached, *i.e.* $\mathbf{R} = \begin{bmatrix} (\mathbf{r}_{1,1}, 0, \dots 0) \\ (0, \mathbf{r}_{2,1}, \dots 0) \\ \vdots \\ (0, 0, \dots \mathbf{r}_{N,1}) \end{bmatrix}$. Those vectors are combined by adding them together, thereby producing a single encrypted vector holding the comparison scores of $\mathbf{p}$ against $\mathbf{E}$, *i.e.* a mapping was introduced so that $\mathbf{R} \mapsto (\mathbf{r}_{1,1}, \mathbf{r}_{2,1}, \dots \mathbf{r}_{N,1})$.

5. The next step is to compare the scores against a predefined threshold, *i.e.* to transform from the continuous spectrum of the comparison scores to the binary accept or reject decision. This is done by subtracting an encrypted vector storing at each element the threshold value (t), *i.e.* $\mathbf{t} = (t, t, \dots t)$ of length s, thus producing a decision vector $\mathbf{d} = \mathbf{R} - \mathbf{t}$, which is subsequently transmitted to the trusted third party.

6. In the last step, $\mathbf{d}$ is decrypted to determine the identification outcome (and communicate it to the client), *i.e.* decision $= \begin{cases} \text{accept} & \text{if } \exists d \in \mathbf{d} : d > 0 \\ \text{reject} & \text{if } \forall d \in \mathbf{d} : d \leq 0 \end{cases}$. If any of the elements in $\mathbf{d}$ is above 0, it is necessarily because the corresponding score in $\mathbf{R}$ has been greater than t, thus yielding an accept decision. Since due to batching access to individual elements in the encrypted vector is not possible, the whole vector must be decrypted by the trusted third party.

9.3 Experiments

9.3.1 Experimental Setup

The experimental evaluation was conducted on a frontal subset of the FERET database [21] in an open-set identification scenario with 500 enrolled data subjects, using 10-fold cross-validation. The features from the images were extracted using FaceNet with a pre-trained model provided by its authors [24]. FaceNet yields templates comprising of 512 floating-point feature elements; two such templates can then be compared using squared Euclidean distance. The open-source Microsoft SEAL HE library [25] was utilised to implement the identification protocol in the encrypted domain (see section 9.2). The choice was based on the presence of a high-level API and suitable HE schemes, namely: Brakerski/Fan-Vercauteren (henceforth referred to as "BFV") [12] and Cheon-Kim-Kim-Song (henceforth referred to as "CKKS")

[8] for integer and float based computations, respectively (see [1] for a detailed HE survey, including other available libraries). Thus, the templates produced by FaceNet can be encrypted directly using CKKS. To utilise BFV, a quantisation scheme is employed, whereby the continuously distributed values of the feature elements are mapped into discrete intervals (see *e.g.* [9] for more details). Although some information is lost through quantisation, the biometric performance should remain largely unaffected. Accordingly, following evaluations were conducted on commodity hardware (one 2.5GHz CPU, 8 GB RAM) in a virtualised Linux environment:

- Biometric performance (DET curve) with the original and quantised feature vectors.
- Time elapsed (in ms) for the computations in the HE domain.

9.3.2 Results

When applying the original float-based templates and CKKS scheme, the distance computation between two encrypted feature vectors was around 5000ms. However, by applying a quantisation scheme, which enables the use of BFV, significant speed-up was achieved – the distance computation only taking 850ms. The time consumed by the encryption and decryption operations was trivial in comparison to that of the distance computation, taking around 7ms and 2.5ms for CKKS and BFV, respectively. The space requirements for the generated keys are not excessive: $<$1MB for the public, secret, and relinearization keys (each) and around 10MB for the Galois keys. In figure 9.2, it can be seen that the biometric performance of the system was not degraded by the application of quantisation.

9.4 Technical Considerations

By utilising HE, the security objectives of a biometric template protection system (see section 9.1) are achieved. The unlinkability across different databases can be guaranteed, insofar they use a different set of keys for the encryption. The irreversibility of the templates is bound to the encryption strength, which in the used library, depending on the chosen parameters, can be 128, 192, or 256 bits. Renewability is ensured, since new protected templates can always be generated by changing the encryption keys. Finally, as the template comparator in the encrypted domain is functionally identical (yields the same comparison scores) to that of the plaintext domain, the biometric performance is not impacted.

The speed of the current implementation may be prohibitive for larger deployments. It should, however, be noted that the experiments were carried out using an ordinary set-up, *i.e.* without powerful CPUs, parallelisa-

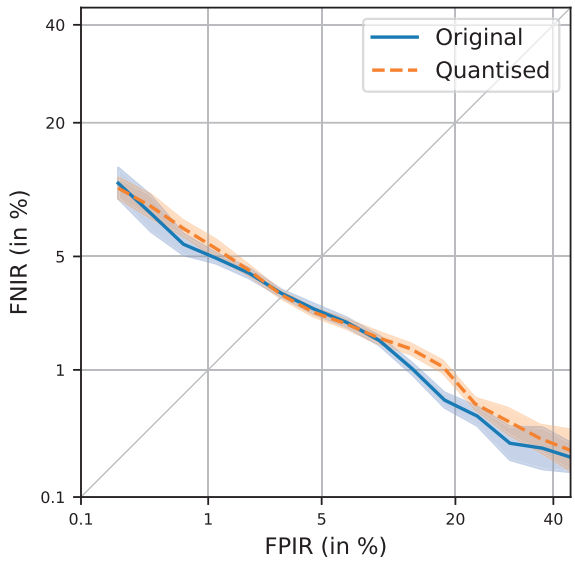


Figure 9.2: DET curves for biometric identification with original and quantised features (with 95% CI)

tion/distribution, *etc*. The hardware set-up notwithstanding, it would be beneficial to incorporate the concepts of computational workload reduction (see *e.g*. [18]) into such systems in order to narrow down the search space for each identification transaction. However, one drawback of using HE is that it limits the flexibility in the implementation – for instance, as previously mentioned, due to batching the feature vector elements cannot be accessed individually. This limitation makes *e.g*. the incremental recognition schemes (which facilitate early acceptance/rejection of likely/unlikely candidates) infeasible. The incorporation of more complicated schemes, such as indexing and binning, could be a potentially interesting future research avenue. On the other hand, a 1-to-first search strategy could already be implemented by slight alterations to the communication between the database and the trusted third party described in section 9.2, however likely at the cost of at least some information exposure.

In the evaluation of biometric systems, a multitude of factors need to be considered. Some of the most important properties are the biometric performance, computational workload, as well as data security and privacy. Those goals typically counterbalance each other, and a biometric system operator is inevitably faced with trade-offs. In the case of HE in the biometric identification scenario described in this paper, currently the goals of high biometric performance, as well as data security and privacy can be fulfilled, while reducing or accelerating the computational workload could be pursued in

future research to facilitate usage of such HE schemes in the practical, real-time applications.

9.5 Summary

In this paper, an architecture and an implementation thereof for a facial identification system in HE domain were presented and subsequently evaluated experimentally. The system fulfils the biometric template protection objectives defined in ISO/IEC IS 24745, namely unlinkability, irreversibility, renewability, and does not negatively affect the biometric performance. The paper also provided a discussion w.r.t. the technical considerations, challenges, and future work potential for utilisation of HE in the context of biometric face identification systems. While not yet practical for real-time applications, HE is definitely a promising avenue for future research and developments in this context.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF), by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP), and the LOEWE-3 BioBiDa Project (594/18-17). The authors acknowledge the financial support by the Federal Ministry of Education and Research of Germany in the framework of MEDIAN (FKZ 13N14798).

9.6 Bibliography

[1] ACAR, A., AKSU, H., ULUAGAC, A. S., AND CONTI, M. A survey on homomorphic encryption schemes: Theory and implementation. *Computing Surveys (CSUR) 51*, 4 (September 2018), 79:1–79:35.

[2] AGUILAR-MELCHOR, C., FAU, S., FONTAINE, C., GOGNIAT, G., AND SIRDEY, R. Recent advances in homomorphic encryption: A possible future for signal processing in the encrypted domain. *Signal Processing Magazine 30*, 2 (March 2013), 108–117.

[3] BODDETI, V. N. Secure face matching using fully homomorphic encryption. In *International Conference on Biometrics Theory, Applications and Systems (BTAS)* (2019), IEEE, pp. 1–10.

[4] BONEH, D., GOH, E.-J., AND NISSIM, K. Evaluating 2-DNF formulas on ciphertexts. In *Theory of Cryptography Conference* (February 2005), Springer, pp. 325–341.

[5] Brakerski, Z., Gentry, C., and Vaikuntanathan, V. (leveled) fully homomorphic encryption without bootstrapping. *Transactions on Computation Theory (TOCT) 6*, 3 (July 2014), 13.

[6] Brakerski, Z., and Vaikuntanathan, V. Fully homomorphic encryption from ring-LWE and security for key dependent messages. In *Annual Cryptology Conference* (August 2011), vol. 6841, Springer, pp. 505–524.

[7] Campisi, P. *Security and privacy in biometrics*, vol. 24. Springer, June 2013.

[8] Cheon, J. H., Kim, A., Kim, M., and Song, Y. Homomorphic encryption for arithmetic of approximate numbers. In *Advances in Cryptology – ASIACRYPT* (December 2017), Springer, pp. 409–437.

[9] Drozdowski, P., Struck, F., Rathgeb, C., and Busch, C. Benchmarking binarisation schemes for deep face templates. In *International Conference on Image Processing (ICIP)* (October 2018), IEEE, pp. 1–5.

[10] El-Gamal, T. A public key cryptosystem and a signature scheme based on discrete logarithms. *Transactions on information theory 31*, 4 (July 1985), 469–472.

[11] European Parliament. Regulation (EU) 2016/679. *Official Journal of the European Union L119* (April 2016), 1–88.

[12] Fan, J., and Vercauteren, F. Somewhat practical fully homomorphic encryption. *Cryptology ePrint Archive* (March 2012), 1–19.

[13] Gentry, C. *A fully homomorphic encryption scheme*. Ph.D. thesis, Stanford University, September 2009.

[14] Gentry, C., and Halevi, S. Implementing Gentry's fully-homomorphic encryption scheme. In *Annual international conference on the theory and applications of cryptographic techniques* (May 2011), Springer, pp. 129–148.

[15] Gomez-Barrero, M., Maiorana, E., Galbally, J., Campisi, P., and Fierrez, J. Multi-biometric template protection based on homomorphic encryption. *Pattern Recognition 67* (July 2017), 149–163.

[16] Imtiyazuddin, S., Rao, Y. V. S., and Rekha, N. R. Faster biometric authentication system using Fan and Vercauteren scheme. In *International Conference on Advances in Computing, Control and Communication Technology (IAC3T)* (September 2018), IEEE, pp. 48–53.

[17] ISO/IEC JTC1 SC27 IT SECURITY TECHNIQUES. *ISO/IEC 24745:2011. Information technology – Security techniques – Biometric information protection*. International Organization for Standardization and International Electrotechnical Committee, June 2011.

[18] KAVATI, I., PRASAD, M., AND BHAGVATI, C. Search space reduction in biometric databases: a review. In *Computer Vision: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2018, pp. 1600–1626.

[19] NANDAKUMAR, K., AND JAIN, A. K. Biometric template protection: Bridging the performance gap between theory and practice. *Signal Processing Magazine 32*, 5 (September 2015), 88–100.

[20] PATEL, V. M., RATHA, N. K., AND CHELLAPPA, R. Cancelable biometrics: A review. *Signal Processing Magazine 32*, 5 (September 2015), 54–65.

[21] PHILLIPS, P. J., MOON, H., RIZVI, S. A., AND RAUSS, P. J. The FERET evaluation methodology for face-recognition algorithms. *Transactions on pattern analysis and machine intelligence (TPAMI) 22*, 10 (October 2000), 1090–1104.

[22] RATHGEB, C., AND UHL, A. A survey on biometric cryptosystems and cancelable biometrics. *EURASIP Journal on Information Security 2011*, 1 (September 2011), 1–25.

[23] RIVEST, R. L., SHAMIR, A., AND ADLEMAN, L. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM 21*, 2 (February 1978), 120–126.

[24] SCHROFF, F., KALENICHENKO, D., AND PHILBIN, J. FaceNet: A unified embedding for face recognition and clustering. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2015), IEEE, pp. 815–823.

[25] Microsoft SEAL (release 3.2). `https://github.com/Microsoft/SEAL`, February 2019. Last accessed: 2020–03–11.

[26] YAO, A. C. Protocols for secure computations. In *Annual Symposium on Foundations of Computer Science* (November 1982), IEEE, pp. 160–164.

Chapter 10

Multi-biometric Identification with Cascading Database Filtering

Abstract

The growing scale and number of biometric deployments around the world necessitates research into technologies which facilitate fast identification queries and high discriminative power. In this context, this article presents a biometric identification system which relies on a successive pre-filtering of the potential candidate list using multiple biometric modalities, coupled with a weighted score-level information fusion. The proposed system is evaluated in a series of experiments using a compound dataset constructed from several publicly available datasets; an open-set identification scenario is considered with the enrolment database containing 1,000 chimeric instances. This evaluation shows that the proposed system exhibits a significantly increased biometric performance w.r.t. a weighted score-level or rank-level fusion based baseline, while simultaneously providing a consequential computational workload reduction in terms of penetration rate. Lastly, it is worth noting that the proposed system could be flexibly employed in any multi-biometric identification system, irrespective of the chosen types of biometric characteristics and the encoding of their extracted features.

Addressed research question(s): RQ1, RQ3, RQ4
Reference: DROZDOWSKI, P., RATHGEB, C., MOKROSS, B.-A., AND BUSCH, C. Multi-biometric identification with cascading database filtering. *Transactions on Biometrics, Behavior, and Identity Science (TBIOM)*, (March 2020), 1–14.

10.1 Introduction

Various market value studies (see *e.g.* [5, 54, 78]) evince the rapid growth of interest and investment in biometric technologies. Biometrics are being used by various governmental organisations around the world for purposes such as law enforcement and forensic investigations (see *e.g.* [26, 28, 56]), border

control (see *e.g.* [24, 25, 29, 60]), national ID systems (see *e.g.* [12, 81]), as well as during elections for voter registration (see *e.g.* [8, 11]). The largest of such deployments to date is located in India, where the Unique Identification Authority of India operates a national ID system (Aadhaar) which accommodates, at the time of this writing, almost 1.3 billion enrolled subjects (see *e.g.* the online dashboard [80]). Additionally, the prevalence and computing power of mobile devices (especially smartphones) has been steadily increasing. Together with the advances in embeddable high-quality sensors, those trends have sparked interest in (single and multi modal) mobile biometrics, which has become an active area of research and product development (see *e.g.* [4, 15, 32, 33, 71]).

With the aforementioned increase of the popularity and sizes of biometric systems in the governmental and commercial sectors alike, it is important to develop technologies which facilitate accurate and efficient processing of large amounts of biometric data. In particular, guaranteeing practical system response times by means of algorithmic solutions, rather than merely the scaling of the hardware architecture is of utmost interest. Those considerations are especially important for biometric identification (and duplicate enrolment check) scenarios, where the conventional biometric systems typically conduct an exhaustive search (entailing one-to-many comparison) to identify the biometric probes. Daugman, the pioneer of iris recognition, stated (in a recent interview) that performing accurate and efficient biometric identification (*i.e.* without an exhaustive search) is one of the most important, unsolved issues in biometrics in general. From the governmental side there exists a strong interest for computationally efficient biometric algorithms, as evidenced by multiple competitions and benchmarks (*e.g.* 1:N Evaluation under Face Recognition Vendor Test (FRVT) [57], one-to-many evaluations under Iris Exchange (IREX) [58], and Biometric Technology Rally [9]).

In recent years, a significant research effort has been devoted to addressing this topic by developing methods for computational workload reduction in biometric systems (see subsection 10.2.2 and a recent survey of Drozdowski *et al.* [20] for more details). The contribution of this work in this context is a proposal of an information fusion scheme, as well as an experimental evaluation thereof on a large compound dataset in the biometric open-set identification scenario. The scheme is based on a successive filtering of candidate shortlists coupled with information fusion on score level. It is shown that the proposed scheme increases the biometric performance w.r.t. the weighted score-level or rank-level fusion based baseline by an order of magnitude, while simultaneously significantly reducing the computational workload (in terms of penetration rate) of the biometric identification transactions. In related works, several authors utilised dimensionality reduction and/or binarisation to create short-length templates, which are

used to pre-filter the enrolment database in a two-stage framework (see *e.g.* Gentile *et al.* [30], Billeb *et al.* [6], and Pflug *et al.* [62]), whereas Drozdowski *et al.* [21] used biometric image morphing in a similar manner. All of those methods considered single-modal systems. A decision-based cascade operating on the principle of sequential fusion of fingerprint and iris recognition systems was presented by Elhoseny *et al.* [23]. Lastly, database pre-filtering based on demographic and geographic metadata is a known and widely used method of searching in large-scale databases (see *e.g.* Gehrmann *et al.* [27]). Soft biometrics (see *e.g.* Dantcheva *et al.* [14]) can also be used in an analogous manner. Best to the authors' knowledge, previous research has not considered a multi-modal fusion utilised in a cascading manner for the simultaneous purpose of computational workload reduction and biometric performance improvement.

The remainder of this article is organised as follows: section 10.2 provides a background overview of the two relevant related work areas – biometric information fusion and computational workload reduction. In section 10.3, the proposed system is described. The details of the experimental setup are outlined in section 10.4, while the results of the experiments are presented in section 10.5 and discussed in section 10.6. A summary and concluding remarks are given in section 10.7.

10.2 Background and Related Work

In this section, relevant background information and related work w.r.t. the two main topics of this article are outlined. Specifically, subsection 10.2.1 addresses biometric information fusion, while subsection 10.2.2 deals with computational workload reduction in biometric systems. Furthermore, the scope of this article and the proposed system is demonstrated within the overall overview of those research areas.

10.2.1 Biometric Information Fusion

One of the key goals of biometric information fusion is to increase the overall discriminative power of a biometric recognition system. Systems where biometric information fusion is utilised are referred to as multi-biometric systems. In such systems, multiple information sources are considered and combined (fused) with each other. In the context of biometrics, following main fusion categories can be distinguished (see *e.g.* Ross *et al.* [72] and ISO/IEC TR 24722 [42]):

Multi-type Information from multiple biometric characteristics (*e.g.* face and fingerprint) is used.

Multi-sensorial Biometric data is acquired using several complementary sensors (*e.g.* visible wavelength and near-infrared camera).

Multi-algorithm Biometric samples are processed using multiple complementary algorithms (*e.g.* texture and keypoint based image descriptors for feature extraction and/or different concepts for comparison).

Multi-instance Information from multiple instances of the same characteristic is used (*e.g.* left and right iris).

Multi-sample Multiple samples (acquisitions) of the same characteristic are used (*e.g.* for sample quality assurance or detection of reliable regions).

The system proposed in this article pertains to the first scenario. Specifically, three types of biometric characteristics are chosen and are subsequently used in a pre-filtering and fusion scheme. In addition to the coarse categories above, several levels of the biometric processing pipeline can be distinguished where information fusion can be performed (see *e.g.* Ross *et al.* [72]):

Sensor Information from multiple sensors or multiple samples (*e.g.* on the pixel level for images or phase level for audio/video signals) is combined prior to any other processing steps. See *e.g.* Jain *et al.* [45] and Kusuma *et al.* [52].

Feature Information from multiple extracted feature sets is consolidated. The data could come from the same biometric characteristic (*e.g.* multiple, complementary feature extractors are used) or different biometric characteristics (*e.g.* a common feature representation is used for the fusion). See *e.g.* Kanhangad *et al.* [47] and Yan *et al.* [84].

Score The comparison scores acquired from multiple information channels are combined (*e.g.* summed or averaged). Depending on the used biometric comparators, this often requires normalisation of the scores to a common domain. See *e.g.* Snelick *et al.* [76] and Jain *et al.* [46].

Rank First, the ranks (order) of potential matches of a probe against an enrolment database are established. Subsequently, heuristics (*e.g.* choosing the best rank or majority vote) are used to consolidate the information from multiple systems. See *e.g.* Abaza *et al.* [2] and Kumar *et al.* [50].

Decision The decisions (*i.e.* accept/reject) reached by multiple systems are combined using heuristics (*e.g.* majority voting or statistics-based rulesets). See *e.g.* Prabhakar *et al.* [63] and Paul *et al.* [59].

In the context of this work, information fusion on score and rank level is of most interest. This is partially because score-level fusion is amongst the most popular and best performing of the aforementioned methods (see Ross *et al*. [72]), and partially because the proposed system (see section 10.3) is designed to work at those levels of the biometric pipeline, *i.e*. irrespective of the chosen biometric characteristics, acquisition methods, and feature extraction algorithms.

Several extensive works and surveys on the topic of biometric information fusion have been published in the scientific literature. The interested reader is therefore referred to *e.g*. Ross *et al*. [72] for a comprehensive general introduction to this topic, Jain *et al*. [46] and Snelick *et al*. [76] for score-level fusion specifically, as well as Radu *et al*. [66], Dinca *et al*. [18], and ISO/IEC TR 24722 [42] for more recent works concerning the overall topic of biometric information fusion.

10.2.2 Computational Workload Reduction

There exists a broad variety of ways in which biometric systems can operate. The main two of them (quoted directly from the ISO/IEC international standards [41, 42, 43]) are:

Biometric verification Referring to the "process of confirming a biometric claim through biometric comparison".

Biometric identification Referring to the "process of searching against a biometric enrolment database to find and return the biometric reference identifier(s) attributable to a single individual".

In the context of biometric identification, two main scenarios can be distinguished, namely *closed-set* identification, where it is known that the enrolment database contains all the potential system users as data subjects, and *open-set* identification, where it is possible that some potential users (impostors) are not enrolled in the system.

Open-set biometric identification, which is, arguably, the most challenging from the practical point of view, is the focus of this article. Due to the necessity of protecting against impostors not enrolled with the system, as well as the lack of an identity claim during a transaction, in the worst case an exhaustive search (*i.e*. comparisons between the probe and the entire enrolment database) is required in order to make a decision. Unfortunately, two non-trivial problems are quickly encountered by this naïve approach:

Computational costs With an increasing size of the enrolment database, the response times become proportionally slower, hence requiring hardware investment and/or software optimisations to facilitate the growing number of the data subjects.

False positives costs Daugman [16] has pointed to a demanding relationship facing biometric identification systems:

$$P_N = 1 - (1 - P_1)^N \tag{10.1}$$

This equation denotes the probability (P_N) of at least one false positive occurrence in an identification transaction within a system which comprises N enrolled users and has a P_1 false positive probability of a one-to-one template comparison. Even when P_1 is very low (*i.e.* the system would exhibit good biometric performance in verification mode), P_N raises very quickly to unacceptable levels as N increases[1].

Since the overall computational costs in a biometric identification scenario are dominated by performing the biometric comparisons (see *e.g.* Drozdowski *et al.* [20]), most computational workload reduction approaches are aimed at that step in the system pipeline. It should be noted, that due to certain properties of biometric data (*i.e.* lack of inherent ordering, within-subject variability, and high dimensionality), many traditional approaches (such as normal database indexing) are often unsuitable or perform poorly (see Hao *et al.* [37]). Therefore, approaches specifically tailored to those properties have been developed. In particular, two main approach classes can be distinguished:

Pre-selection Approaches in this category concentrate on reduction of the potential search space, *i.e.* the number of necessary template comparisons (penetration rate) during a biometric identification transaction. Three principal sub-categories can be distinguished here:

Pre-filtering Multiple algorithms or feature representations are used. The idea is to first use computationally efficient (but somewhat inaccurate) methods to create a candidate shortlist, whereupon a computationally expensive (but accurate) method is used on this small, pre-filtered subset of the database (see *e.g.* Ratha *et al.* [67], Gentile *et al.* [30], and Billeb *et al.* [6]).

Binning The database is split into distinct bins/partitions based on some coarse auxiliary features. Examples include metadata (such as demographic and geographic attributes, see *e.g.* Gehrmann *et al.* [27]) or biometric characteristic specific features, such as fingerprint classes (see *e.g.* Drozdowski *et al.* [19]). During a biometric

[1] Although this equation ignores other system errors, such as the failure-to-acquire rate and also assumes that at a given threshold all subjects have the same false-match-rate (which likely is not the case), it nonetheless is a useful approximation through which the challenges of the biometric identification systems can be illustrated quantitatively.

identification only the bins corresponding to the sample are considered, thereby reducing the search space. As an alternative to such handcrafted features, unsupervised clustering can also used (see *e.g.* Ross *et al.* [73] and Pflug *et al.* [61]).

Datastructures The enrolment database is reorganised to take advantage of efficient ordering principles, for example based on search trees (see *e.g.* Proença [64] and Rathgeb *et al.* [68]) or fuzzy hashing (see *e.g.* Cappelli *et al.* [10] and Kaushik *et al.* [48]), thereby enabling sub-linear/logarithmic search time.

Feature transformation Approaches in this category concentrate on reducing the computational cost of the individual template comparisons. Typical approaches in this category accomplish this by reducing the dimensionality of the biometric templates by extracting the most discriminative parts (see *e.g.* Gentile *et al.* [31] and Rathgeb *et al.* [70]), utilising more efficient comparators such as integer/bit-based instead of float-based (see *e.g.* Lim *et al.* [53] and Drozdowski *et al.* [22]), or providing sample alignment invariance (see *e.g.* Rathgeb *et al.* [69] and Damer *et al.* [13]).

An exhaustive survey of this research area is out of scope for this article – for more details, the interested reader is referred to other works on this topic. Specifically, in a recently published work of Drozdowski *et al.* [20], a biometric characteristic-agnostic, concept-based taxonomy of computational workload reduction approaches in biometrics has been proposed. Additionally, the authors conducted a comprehensive survey of computational workload reduction in biometric identification systems in the context of said taxonomy. For biometric characteristic-specific works, surveys by Schuch [75] (fingerprint), Proença *et al.* [65] (iris), and Kavati *et al.* [49] (fingerprint, face, iris) are of interest.

In the context of the above categories of computational workload reduction approaches, the pre-selection (more specifically, pre-filtering) one is most relevant to this work. This is because, as previously mentioned, this article presents (see section 10.3) a method which relies on a successive candidate shortlist filtering, and works irrespective of the chosen type of biometric characteristics and their feature representations, thereby precluding any approaches which rely on specific feature transformations.

10.3 Proposed System

Consider a biometric enrolment database with references of N data subjects for K different biometric modalities (*i.e.* types of biometric characteristics). A standard approach for a biometric identification transaction would be to

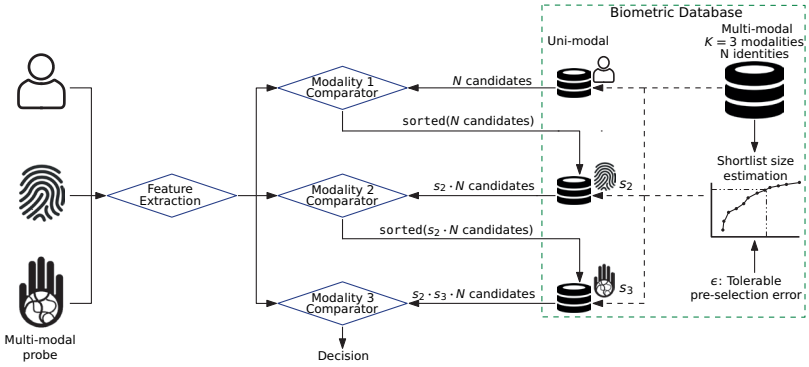


Figure 10.1: Overview of the proposed system

conduct the comparisons (C) exhaustively (*i.e.* $\#\mathrm{C}_{\mathrm{baseline}} = K \cdot N$ comparisons) for all the modalities and to fuse the scores using one of the traditional strategies (such as score or rank level fusion) described in subsection 10.2.1. This approach will serve as a baseline later on in the experiments. Here, an alternative method is proposed with the aim of improving the biometric performance and reducing the computational workload.

The conceptual overview of the proposed system (for $K = 3$) is shown in figure 10.1. On the biometric database side, the stored modalities are given a specific order (see subsection 10.3.2 and section 10.5 for more details on the ordering of the chosen types of biometric characteristics). The key idea is to successively filter the list of potential candidates based on the comparisons within the individual modalities, thus creating a multi-stage (K-stage), cascading filtering system. In the illustrated example, a biometric identification transaction would proceed as follows:

1. Features are extracted for each of the probe sample of each biometric type (*i.e.* modality).

2. Modality 1 (face) probe is compared exhaustively (N comparisons) against the enrolment database. Based on the sorted comparison scores, a certain fraction (denoted s_2) of the most promising candidates (a candidate shortlist) is passed onto the next level.

3. Modality 2 (fingerprint) probe is compared against the ($s_2 \cdot N$) most promising candidates. A fraction of those (s_3) is then passed onto the last level.

4. Modality 3 (fingervein) probe is compared against the ($s_2 \cdot s_3 \cdot N$) most promising candidates to reach the final identification decision.

The types of biometric characteristics (face, fingerprint, and fingervein) were chosen based on the criteria that the three types be not correlated, that they are widely deployed (in various operational systems), and exhibit desirable properties w.r.t. presentation attack detection (the latter especially concerning the fingervein). However, it should be noted that the system is not in any way reliant on those specific characteristics or this particular ordering thereof – the system design is applicable irrespective of the participating biometric characteristics and their feature representations. The order of the characteristics in the cascade is also flexible – more on this topic in subsection 10.3.2 and the experimental evaluation in section 10.5.

The sizes of the candidate lists passed between the levels of the cascade (values in s_i) are estimated empirically in a training step, see subsection 10.3.1 for more details. Since at the first level the whole database is used for the comparisons, s_1 would equal 1.0 and is not depicted in figure 10.1. The theoretical impact of the proposed system on the biometric performance and computational workload is described in subsection 10.3.2.

10.3.1 Shortlist Size Estimation

For each considered type of biometric characteristic, a tolerable pre-selection margin of error in terms of false negative identification rate is determined. This margin is denoted as $\epsilon \in [0\% \dots 100\%]$ and can be set arbitrarily low or high by the system operator depending on the system policy. The extreme values are unlikely in a practical scenario and are listed for the purposes of the mathematical definition only. This parameter is used for the purpose of shortlist size estimation for the pre-selection algorithm described in the previous subsection. The goal is to find the minimum fraction (denoted s) of candidate identities to pass between two levels of the cascade, so that the selected tolerable margin of error is not violated. Estimating s for a biometric type happens in a dedicated training step on a disjoint dataset, where a closed-set identification experiment is carried out and a cumulative match characteristic (CMC) curve is computed. Using the CMC curve, one first needs to calculate the minimum rank (r), so that $IR \geq 100\% - \epsilon$; then, r is expressed relative to the size (N_{train}) of the training enrolment database (E_{train}). An abstract, formal description of this concept is given in algorithm 10.1.

For a concrete example of the concept, see figure 10.2. There, a CMC curve for an example system (purely for illustrative purposes; the chosen type of biometric characteristic does not matter in this case) with 30 enrollees has been computed. Two operational points (with different ϵ values) are considered, expressing different system policies: a liberal one, wherein some pre-selection (false negative) errors are acceptable (depicted with the orange line), and a stringent one, which seeks to minimise pre-selection errors (depicted with the blue line). Consequently, in the case of the $\epsilon = 1\%$

Algorithm 10.1: Shortlist size estimation

Input: E_{train}, ϵ
Output: s_ϵ
1: $CMC \leftarrow \text{COMPUTECMC}(E_{train})$
2: $r_\epsilon \leftarrow \min\{r \in \{1 \dots N\} \mid CMC(r) \geq 100 - \epsilon\}$
3: $N_{train} \leftarrow \text{LENGTH}(E_{train})$
4: $s_\epsilon \leftarrow \frac{r_\epsilon}{N_{train}}$
5: **return** s_ϵ

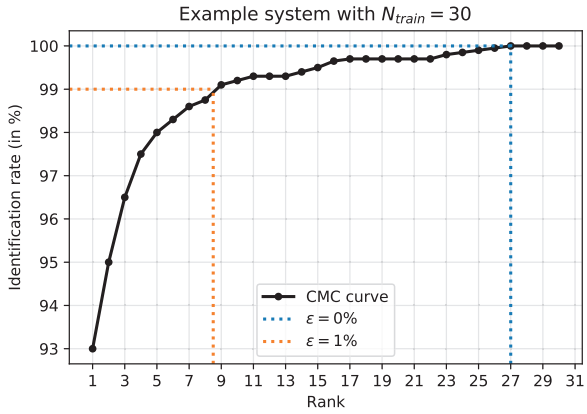


Figure 10.2: Determining s_ϵ based on a training CMC curve and ϵ

policy, the lowest rank satisfying the IR constraint (see line 2 in algorithm 10.1) is 9, meaning that the candidate shortlist passed onto the next level would be around $s_{\epsilon=1\%} = \frac{9}{30} = 30\%$ of the enrolment database. On the other hand, for the $\epsilon = 0\%$ policy, the constraint is satisfied only at rank 27, thus making the passed candidate shortlist $s_{\epsilon=0\%} = \frac{27}{30} = 90\%$ of the enrolment database.

Thus, for each considered modality and ϵ, a theoretically optimal candidate shortlist size, expressed as a fraction of the enrolment database (s, $s \in \{x \in \mathbb{R} \mid 0.0 < x \leq 1.0\}$), can be ascertained w.r.t. the system policy. A multi-stage system with K modalities would then have $S_\epsilon = [s_1 \dots s_K]$ shortlist sizes, with s_1 always equal to 1.0. Generally, more liberal (*i.e.* higher) values of ϵ mean smaller candidate shortlists (*i.e.* lower penetration rate), but increased potential of pre-selection (false negative) errors. On the other hand, the reduction of penetration rate contributes to reducing the false positive error rate. This is because the reduction in penetration rate corresponds to the factor N in equation (10.1) being reduced, *i.e.* there being fewer potential comparisons in a biometric identification transaction where an impostor

could, just by chance, get a better comparison score against a reference in the enrolment database. Note, that this is true insofar there exists no correlation between the comparison scores of the biometric characteristics in the system, *i.e*. that they be statistically independent. Certain biometric modalities can exhibit explicit or hidden symmetries and correlations (see *e.g*. Gomez-Barrero *et al*. [34] and *e.g*. Kumar *et al*. [51]). Such correlations have a non-trivial impact on the biometric performance of information fusion schemes (see *e.g*. Ulery *et al*. [79]). Generally speaking, the utility (in terms of increase of information entropy, see *e.g*. Adler *et al*. [3]) of correlated modalities may be lower than that of uncorrelated ones. Furthermore, specifically for the proposed scheme, using correlated modalities for the pre-filtering stage would be counter-productive, as computational workload would have to be expended on performing the comparisons, but little or no additional information would have been gained for the pre-selection of candidates. In other words, while the proposed scheme technically supports any combination of biometric modalities by the virtue of operating at the level of comparison scores, some attention is nevertheless required w.r.t. the choice of the modalities participating in the scheme. Ideally, completely uncorrelated modalities should be used. If correlated modalities are chosen, the results in terms of biometric performance and computational workload reduction may be degraded. Therefore, care and awareness is advised w.r.t. the choice of biometric modalities for the proposed scheme. Note, that this caveat of correlated data is also applicable to other existing biometric information fusion schemes. In this article, three uncorrelated biometric characteristics have been chosen for the experiments (see subsection 10.4.1).

10.3.2 System Ordering

The modalities participating in the cascade can be ordered arbitrarily – the number of possible permutations for a system with K modalities is $K!$. The ordering is expected to have a non-trivial impact on the computational workload and biometric performance. If the computational cost of individual template comparisons is also considered (see section 10.6), the system ordering has an impact not only on the biometric performance, but also the overall computational workload. The total number of comparisons in the proposed system is:

$$\#\mathrm{C}_{\text{proposed}} = N + \sum_{k=2}^{K} \prod_{i=1}^{k} s_i \cdot N \tag{10.2}$$

The key idea behind equation (10.2) being that $\#\mathrm{C}_{\text{proposed}} \ll \#\mathrm{C}_{\text{baseline}}$, *i.e*. reducing the penetration rate of the search. The lower bound of the penetration rate is then $p = \frac{1}{K} + \frac{K-1}{N}$, *i.e*. in the case of a 3-level system the

minimum penetration rate could be around 33.(3)%. This limit is due to exhaustive search always having to be be conducted for the first modality in the cascade. A potential extension of the proposed system could consider another scheme of computational workload reduction (*e.g.* binning) to be used prior to the first level of the cascade in order to avoid the necessity of conducting an exhaustive search there.

10.3.3 Combination with Weighted Score-level Fusion

The system proposed in the previous subsections uses the comparison scores in the shortlist from the final level of the cascade to make a decision. It is, however, also possible to combine the traditional weighted score-level fusion with the proposed scheme. Specifically, such a combined scheme would work as follows:

1. Conduct the cascading filtering with K modalities as described in the previous subsections.

2. Retrieve the identities of the subjects in the candidate shortlist produced at the last level of the cascade.

3. Retrieve the comparison scores corresponding to the candidate shortlist for the modality at the final level of the cascade *and* the previous levels of the cascade.

4. Fuse the scores.

In other words, the database is first filtered to find the most likely candidates using the individual modalities in the cascade, whereupon the comparison scores (for all the modalities) of the candidates in the shortlist are fused. In theory, such a system should exhibit a decreased computational workload, as well as an increased biometric performance. This idea is evaluated experimentally in addition to the system proposed in the previous subsections.

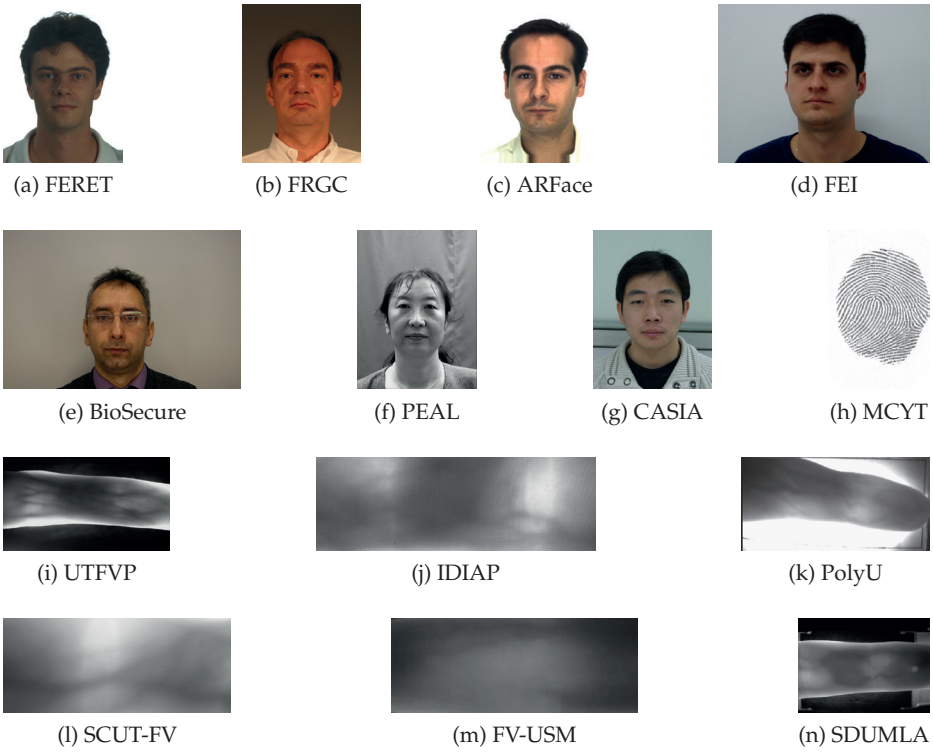
10.4 Experimental Setup

The following subsections outline the details of the experimental setup. The chosen datasets and processing pipelines are described in subsections 10.4.1 and 10.4.2, respectively. The baseline and proposed system configurations, as well as the evaluation metrics are described in subsections 10.4.3 and 10.4.4.

10.4.1 Datasets

The research conducted in this paper is aimed at cooperative systems, *i.e.* ones where biometric samples of reasonably good quality can be expected. Hence, in-the-wild, large time-scale, and occluded facial datasets (or parts thereof), as well as latent fingerprint datasets were not considered. Since none large-scale multi-modal datasets were available to the authors, it was decided to create a virtual dataset from existing single-modal ones.

While there exist several datasets with very large numbers of biometric samples, their size in terms of data subjects is typically much smaller. As such, several facial and fingervein datasets had to be considered in order to obtain a suitable number of data subjects. For the fingerprint and fingervein datasets, the individual instances (fingers) are not correlated and can therefore be treated as separate subjects.



(a) FERET (b) FRGC (c) ARFace (d) FEI

(e) BioSecure (f) PEAL (g) CASIA (h) MCYT

(i) UTFVP (j) IDIAP (k) PolyU

(l) SCUT-FV (m) FV-USM (n) SDUMLA

Figure 10.3: Example images from the selected datasets

The fingerprint database was used directly without any filtering. For the facial datasets, frontal images without intentional occlusions (*e.g.* scarves or sunglasses) were chosen, while some images with exceedingly poor quality

were removed from the fingervein and facial datasets (to facilitate reproducible research, the lists of chosen images and other experimental setup details will be made available online after this article is accepted for publication). It should be noted that the facial and especially fingervein data is extremely inhomogeneous across the chosen datasets. The images were acquired using different cameras/sensors, under varying lighting conditions, and the images have been saved in several distinct resolutions. The chosen datasets are listed in table 10.1 (the numbers given in the table are after the filtering was applied). Example images from the datasets are shown in figure 10.3.

Table 10.1: Used datasets

Characteristic	Dataset	Instances	Samples
Face	FERET	994	2,716
	FRGC	453	2,754
	ARFace	136	1,526
	FEI	200	600
	BioSecure	210	840
	PEAL	429	3,274
	CASIA	725	3,072
Fingerprint	MCYT	3,300	39,600
Fingervein	UTFVP	360	1,440
	IDIAP	220	440
	PolyU	312	3,132
	SCUT-FV	600	3,600
	FV-USM	492	5,904
	SDUMLA	636	3,816

Following the selection, the datasets of the same biometric characteristic have been merged and a compound dataset was constructed. This was done by repeatedly (ten times) shuffling the instances and samples from the original datasets to construct new chimeric instances. The ten copies of the dataset enable a tenfold cross-validation in the experimental evaluation. Each of the copies consists of 2,500 instances and approximately 15,000 samples (depending on availability, since different datasets contain different number of samples per instance). Finally, the resulting compound dataset has been split into two partitions:

Training Consists of 1,000 instances. Used for computing CMC curves in a closed-set identification scenario to approximate the appropriate shortlist sizes for each modality. Additionally used for computing information necessary for comparison score normalisation.

Testing Consists of 1,500 instances. Used for evaluating the baselines (for each modality individually and for several popular information fusion schemes) and the proposed system in an open-set identification scenario.

10.4.2 Processing Pipelines

The images were processed using exclusively open-source frameworks. While commercial frameworks may have offered a better or even errorless biometric performance on the dataset used in the experiments (see subsection 10.4.1), facilitating reproducible research has been deemed a higher priority, hence favouring the open-source frameworks. Furthermore, it has been shown both theoretically and in practice (see *e.g.* Daugman *et al.* [17] and Grother *et al.* [35]), that the biometric performance in the identification scenario decreases with the growing size of the enrolment database. In other words, in large-scale systems, optimal biometric performance is not to be expected, even from the commercial systems. Lastly, to evaluate the proposed fusion methods, the key metric is the *relative* biometric performance gain/loss w.r.t. to a baseline and not the *absolute* biometric performance achieved. Following tools and frameworks were used to extract features from the images and compare the resulting templates:

Face A neural-network based approach is used. Specifically, the FaceNet CNN of Schroff *et al.* [74] is used with a pre-trained model made available by the authors[2]. The network learns to map facial images to Euclidean space, whereby the produced templates (embeddings) can be directly compared using Euclidean distance.

Fingerprint The features (minutiae triplets, *i.e.* 2-D location and angle) are extracted using a neural-network based approach. In particular, the FingerNet CNN of Tang *et al.* [77] is used with a pre-trained model made available by the authors[3]. To compare such templates, a minutiae pairing and scoring algorithm of the sourceAFIS system of Važan [82] is used[4].

Fingervein A minutiae based approach is used. Specifically, the maximum curvature algorithm of Miura *et al.* [55] is used to extract the skeleton of the fingervein patterns, which is subsequently thinned using the

[2] `https://github.com/davidsandberg/facenet`

[3] `https://github.com/felixTY/FingerNet`

[4] The original algorithm uses minutiae quadruplets, *i.e.* additionally considers the minutiae type (*e.g.* ridge ending or bifurcation). Since minutiae triplets are extracted by FingerNet, the algorithm has been modified to ignore the type information. Using FingerNet instead of the native minutiae extractor provided by sourceAFIS is preferred, as it has yielded higher biometric performance.

method presented by Guo *et al*. [36]. The minutiae are retrieved from the vein skeleton with a convolution kernel proposed by Olsen *et al*. [1]. Lastly, using the method of Xu *et al*. [83] and Hartung *et al*. [38, 39], the variable-sized minutiae vector is translated into the Spectral Minutiae Representation (SMR), which is fixed-length and additionally offers certain implicit rotation and scaling invariance. Such templates can be compared using a simple correlation measure (likewise presented in [83]), which is a common approach in image processing.

Table 10.2 summarises the information about the utilised data processing pipelines.

Table 10.2: Data processing pipelines

Characteristic	Extraction	Representation	Size	Comparison
Face	FaceNet	1-D embedding	512 floats	Euclidean distance
Fingerprint	FingerNet	Minutiae triplets set	Variable	Minutiae pairing
Fingervein	Spectral minutiae	2-D matrix	256×128 floats	Correlation

For the biometric fusion, two scenarios were considered (see *e.g*. Jain *et al*. [46], Snelick *et al*. [76], and Ho *et al*. [40] for details):

Score level The scores were normalised using Z-score method, which is one of the most commonly used score normalisation methods and relies on the arithmetic mean and standard deviation of the scores data. This method is expected to perform well when prior knowledge about the score distributions is available – which is the case in this experimental setup (see subsection 10.4.1). Subsequently, the normalised scores were fused with a sum-rule method (using those methods, very good biometric performance has been observed in general, see *e.g*. Jain *et al*. [44] and ISO/IEC TR 24722 [42]).

Rank level A Borda count based method (see Black [7]), which is a group consensus function and a generalisation of the majority vote, was used. The method relies on summing the ranks assigned to the probe-reference pairs based on the comparison scores during a biometric identification transaction and requires no prior training.

In both cases, a weighted variant was also included, whereby the individual types of biometric characteristics are assigned relative weights, which are multiplied with the normalised scores prior to the fusion. The optimal weights' combinations were estimated experimentally (see subsection 10.4.3).

10.4.3 Baseline and Proposed System Configurations

To establish the baseline, against which the results of the proposed methods can be benchmarked, the following experiments were conducted on the testing subset of the compound dataset in an open-set identification scenario:

- Each of the 3 modalities individually.
- Weighted score-level and rank-level fusion (see subsection 10.4.2) of all possible combinations of 2 modalities and of all 3 modalities.

Pairs of weights in the interval $[0.05 \ldots 0.95]$ with a step size of 0.05 were considered for the score and rank level fusion, thus yielding a total of 19 and 171 weights combinations for the fusion of 2 and 3 modalities, respectively.

Two versions of the proposed system were evaluated:

- Cascading filtering.
- Cascading filtering + weighted score-level fusion with a sum-rule.

For the second item above, same combinations of weights as in the baseline were used. Furthermore, all possible orderings of the modalities in the cascade were evaluated. All the experiments with the baseline and the proposed system were conducted using a tenfold cross-validation, as mentioned in subsection 10.4.1. Table 10.3 lists the number of configurations in each of the experiment types.

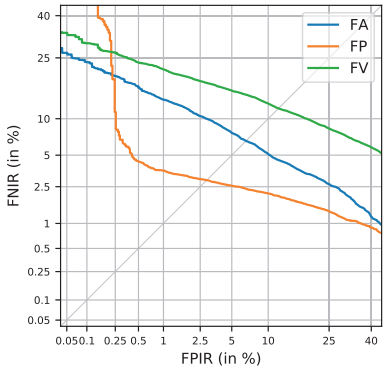
Table 10.3: Configurations per experiment

Experiment	Modalities	Orderings	Weights	Epsilons	Total
Individual baseline	1	3	—	—	3
Weighted fusion baseline	2	1	19	—	19
Weighted fusion baseline	3	1	171	—	171
Cascading filtering	2	6	—	7	42
Cascading filtering	3	6	—	7	42
Cascading filtering + weighted score fusion	2	6	19	7	798
Cascading filtering + weighted score fusion	3	6	171	7	7,182

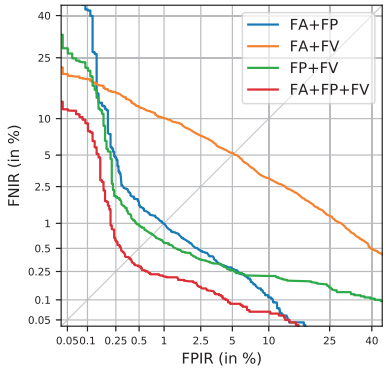
10.4.4 Evaluation Metrics

The systems were evaluated on two key aspects, using ISO/IEC standard methods and metrics [41] as well as additional, commonly used ones:

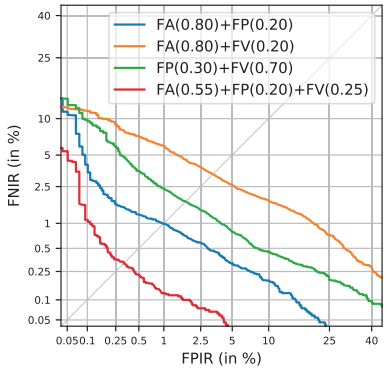
Biometric performance DET curves, equal-error-rate (EER), and false negative identification rate at a certain (here 0.1%) false positive identification rate (denoted FPIR0.1). Additionally, the decidability index over the genuine and impostor score distributions (defined as: $d' =$



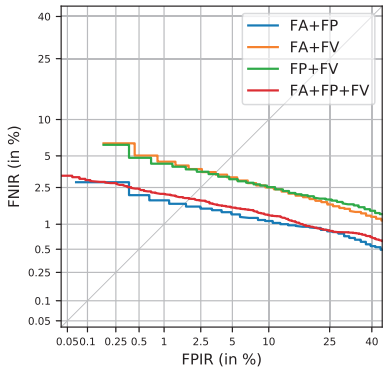
(a) Single modality



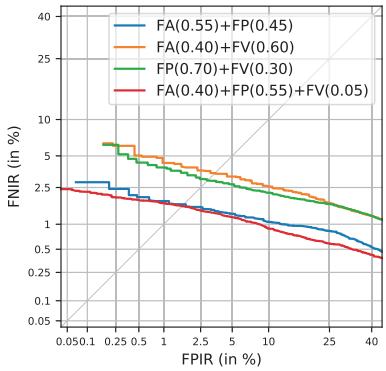
(b) Score fusion



(c) Weighted score fusion



(d) Rank fusion



(e) Weighted rank fusion

Figure 10.4: Baseline results

Table 10.4: Baseline results (with 95% CI)

Method	Modality	PR	τ	EER (in %)	FPIR0.1 (in %)	d'
Individual	FA	1.000	22.240 ± 0.773	6.648 ± 0.473	22.216 ± 0.583	3.153 ± 0.079
	FP	1.000	47.234 ± 3.435	2.864 ± 0.238	47.202 ± 2.592	3.638 ± 0.066
	FV	1.000	29.204 ± 0.401	11.958 ± 0.508	29.186 ± 0.302	2.218 ± 0.044
Rank fusion	FA+FP	1.000	2.784 ± 0.373	1.491 ± 0.199	2.594 ± 0.300	2.540 ± 0.038
	FA+FV	1.000	6.360 ± 0.544	3.081 ± 0.250	6.280 ± 0.416	2.404 ± 0.047
	FP+FV	1.000	6.332 ± 0.501	3.212 ± 0.314	6.252 ± 0.383	2.478 ± 0.034
	FA+FP+FV	1.000	3.127 ± 0.374	1.862 ± 0.200	2.960 ± 0.297	3.502 ± 0.059
Score fusion	FA+FP	1.000	25.340 ± 2.157	0.959 ± 0.152	25.253 ± 1.632	4.561 ± 0.082
	FA+FV	1.000	17.830 ± 0.336	5.049 ± 0.303	17.802 ± 0.254	3.033 ± 0.053
	FP+FV	1.000	14.587 ± 1.347	0.732 ± 0.092	14.430 ± 1.026	4.267 ± 0.048
	FA+FP+FV	1.000	5.615 ± 0.481	0.379 ± 0.112	5.295 ± 0.382	4.851 ± 0.054
Rank fusion weighted	FA(0.55)+FP(0.45)	1.000	2.743 ± 0.393	1.582 ± 0.195	2.550 ± 0.317	2.540 ± 0.039
	FA(0.40)+FV(0.60)	1.000	6.153 ± 0.573	3.342 ± 0.265	6.070 ± 0.438	2.391 ± 0.051
	FP(0.70)+FV(0.30)	1.000	5.970 ± 0.835	2.921 ± 0.294	5.884 ± 0.640	2.510 ± 0.043
	FA(0.40)+FP(0.55)+FV(0.05)	1.000	2.349 ± 0.303	1.564 ± 0.117	2.121 ± 0.254	3.532 ± 0.049
Score fusion weighted	FA(0.80)+FP(0.20)	1.000	5.107 ± 0.574	0.992 ± 0.146	4.901 ± 0.440	4.739 ± 0.079
	FA(0.80)+FV(0.20)	1.000	10.511 ± 0.309	3.311 ± 0.367	10.459 ± 0.234	3.632 ± 0.072
	FP(0.30)+FV(0.70)	1.000	8.533 ± 0.394	1.704 ± 0.103	8.459 ± 0.299	3.541 ± 0.049
	FA(0.55)+FP(0.20)+FV(0.25)	1.000	1.986 ± 0.160	0.324 ± 0.063	1.504 ± 0.136	4.985 ± 0.063

$\frac{|\mu_g - \mu_i|}{\sqrt{\frac{1}{2}(\sigma_g^2 + \sigma_i^2)}}$, where μ and σ stand for the means and standard deviations of the genuine and impostor score distributions, respectively) is reported.

Computational workload Penetration rate (PR), *i.e.* the number of the preselected candidate templates as a fraction of the total number of templates in the enrolment database.

Additionally, a metric which brings the two aspects together (adapted from Proença *et al.* [65]) is used. The metric (τ) calculates the Euclidean distance from the optimal operating point (*i.e.* FPIR0.1 = 0 and PR $\approx$ 0) and is defined as follows: $\tau = \sqrt{(\text{FPIR0.1})^2 + \text{PR}^2}$.

10.5 Results

In this section, the experimental results are presented. First, in subsection 10.5.1, the baseline is established. Subsequently, subsection 10.5.2 shows the empirical shortlist sizes estimation for the proposed system, while its results are presented in subsection 10.5.3. All the tables and figures in this section use a short notation for the biometric characteristics: FA (face), FP (fingerprint), and FV (fingervein). For the weighted fusion variants, the relative weights are written in parentheses immediately following their corresponding biometric characteristics.

10.5.1 Baseline

The results of the baseline experiments are shown in figure 10.4 and table 10.4. All possible combinations of modalities are shown; whereas for the

weighted scenario, the results of the configuration with the lowest FPIR0.1 value for each modality combination are given. Looking at the baseline results, following conclusions can be reached:

- The biometric performance of the individual modalities is only moderate. This is to be expected due to open-source tools being used, as well as the high degree of homogeneity and sometimes poor quality of the facial and fingervein data. However, it is also demonstrated that the biometric performance can be improved to useful levels by applying information fusion.

- The score-level fusion performs better than the rank-level fusion.

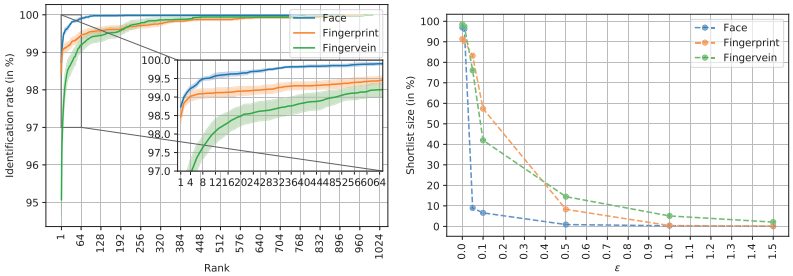
- The results can be further improved by applying relative weighting of the modalities. This is especially the case in terms of FPIR0.1 for the score-level fusion and less so for the rank-level fusion. It should be noted that the exact optimal weights are only pertinent for a particular experimental setup (*i.e.* the specific databases, algorithms, *etc.*) and should not be used to reach general conclusions about weighted biometric fusion.

- The biometric performance in terms of EER of the best combination of 2 modalities and weights is around 1%, whereas using all 3 modalities reduces the EER down to around 0.35%. However, it should be noted that the FPIR0.1 is relatively high in both cases – around 5% and 1.5% for 2 and 3 modalities, respectively.

- Since the baseline setup relies on an exhaustive search method, the penetration rate is 1.0 and τ depends solely on the values of FPIR0.1.

10.5.2 Shortlist Size Estimation

To estimate the shortlist sizes, the methodology outlined in subsection 10.3.1 is followed. Accordingly, CMC curves are computed on the training partition of the dataset and then used to estimate the shortlist size for several ϵ values. In figure 10.5, the CMC curves are shown, along with the relation between the ϵ value and the shortlist size. It can be seen, that relatively high identification rate is achieved at very low ranks; however, it takes a while before 100% is reached, especially for the fingerprint and fingervein modalities. It should be noted, that those CMC curves do not provide a general statement w.r.t. to the relative strength of the chosen types of biometric characteristics; they merely provide a benchmark and overview relevant to the particular experimental setup (*i.e.* the specific databases, algorithms, *etc.*) used in this work.



(a) CMC curves (with 95% CI)

(b) The relation between ϵ value and the shortlist size as a percentage of the enrolment DB size

Figure 10.5: Estimation of the shortlist sizes

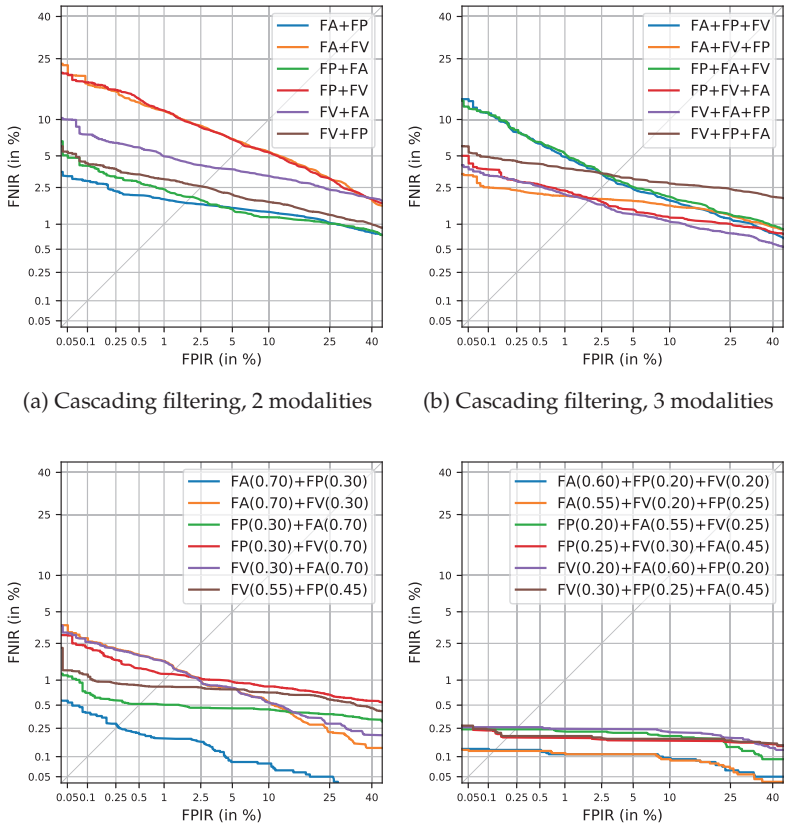
10.5.3 Proposed

The results of the experiments with the proposed system are shown in figure 10.6 and table 10.5. All possible modality combinations and orderings are shown; whereas for the weighted scenario, the results of the configuration with the lowest τ value for each modality combination and ordering are given. Looking at the results of the proposed system, following conclusions can be reached:

- Using the proposed technique alone improves the biometric performance for 2 modalities. For 3 modalities, a relatively good biometric performance is reached, albeit it is somewhat lower than the baseline.

- By combining the proposed technique with a weighted score-level fusion, the biometric performance is significantly improved (by an order of magnitude in some cases, *cf*. table 10.6). The best baseline weighted score-level fusion configuration achieves approximately 0.992% and 0.324% EER for 2 and 3 modalities, respectively. The best configuration of the proposed scheme achieves approximately 0.254% and 0.109% EER for 2 and 3 modalities, respectively. Even more significant improvements can be seen in the higher security region of the error curves. The best baseline weighted score-level fusion configuration achieves approximately 4.901% and 1.504% FPIR0.1 for 2 and 3 modalities, respectively. The best configuration of the proposed scheme achieves approximately 0.333% and 0.125%. It should be noted that the exact optimal weights are only pertinent for a particular experimental setup (*i.e*. the specific databases, algorithms, *etc*.) and should not be used to reach general conclusions about weighted biometric fusion.

- In all the cases, the penetration rate (and hence the computational workload) is significantly reduced – down to 0.545 and 0.388 for 2 and 3 modalities, respectively. Those results are close to the theoretical maximum reduction (*i.e.* down to $\sim\frac{1}{K}$) for the proposed scheme as described in subsection 10.3.2.



(a) Cascading filtering, 2 modalities

(b) Cascading filtering, 3 modalities

(c) Cascading filtering and weighted score-level fusion, 2 modalities

(d) Cascading filtering and weighted score-level fusion, 3 modalities

Figure 10.6: Proposed system's results

10.6 Discussion

This section expands on the discussion items provided directly with the results in the previous section. Specifically, the results in terms of biometric

Table 10.5: Proposed system's results (with 95% CI)

Method	Modality	ϵ	PR	τ	EER (in %)	FPIR0.1 (in %)	d'
Cascading	FA+FP	1.0	0.501	2.769 ± 0.602	1.719 ± 0.172	2.721 ± 0.093	4.408 ± 0.093
	FA+FV	1.5	0.500	18.015 ± 2.582	6.365 ± 0.446	18.008 ± 0.390	2.777 ± 0.050
	FP+FA	1.5	0.501	4.333 ± 1.387	1.983 ± 0.162	4.302 ± 0.210	4.929 ± 0.118
	FP+FV	1.5	0.501	17.518 ± 1.622	6.226 ± 0.356	17.511 ± 0.245	2.787 ± 0.044
	FV+FA	1.0	0.525	7.252 ± 1.176	4.308 ± 0.348	7.232 ± 0.178	3.747 ± 0.141
	FV+FP	1.0	0.525	5.015 ± 0.749	3.816 ± 0.342	4.987 ± 0.114	3.941 ± 0.082
	FA+FP+FV	1.0	0.335	10.461 ± 2.695	3.011 ± 0.220	10.456 ± 2.033	3.426 ± 0.054
	FA+FV+FP	1.5	0.334	2.652 ± 0.524	1.920 ± 0.177	2.630 ± 0.398	4.513 ± 0.093
	FP+FA+FV	1.5	0.334	9.829 ± 1.967	3.069 ± 0.212	9.823 ± 1.484	3.409 ± 0.047
	FP+FV+FA	1.0	0.335	3.673 ± 0.990	1.913 ± 0.125	3.657 ± 0.749	4.951 ± 0.098
	FV+FA+FP	0.5	0.382	3.017 ± 0.424	2.172 ± 0.215	2.992 ± 0.322	4.444 ± 0.094
	FV+FP+FA	0.5	0.386	5.028 ± 0.809	3.269 ± 0.267	5.013 ± 0.611	4.243 ± 0.140
Cascading + score fusion weighted	FA(0.70)+FP(0.30)	0.05	0.545	0.648 ± 0.079	0.254 ± 0.051	0.333 ± 0.102	5.702 ± 0.083
	FA(0.70)+FV(0.30)	0.05	0.545	2.713 ± 0.770	1.335 ± 0.171	2.654 ± 0.591	4.293 ± 0.076
	FP(0.30)+FA(0.70)	0.5	0.541	0.838 ± 0.168	0.509 ± 0.101	0.626 ± 0.159	5.640 ± 0.099
	FP(0.30)+FV(0.70)	1.5	0.501	2.440 ± 1.035	1.163 ± 0.144	2.380 ± 0.794	4.415 ± 0.061
	FV(0.30)+FA(0.70)	0.1	0.710	2.685 ± 0.742	1.327 ± 0.172	2.583 ± 0.576	4.275 ± 0.080
	FV(0.55)+FP(0.45)	0.5	0.573	1.482 ± 0.881	0.832 ± 0.134	1.338 ± 0.697	4.583 ± 0.064
	FA(0.60)+FP(0.20)+FV(0.20)	0.05	0.388	0.409 ± 0.012	0.109 ± 0.022	0.125 ± 0.027	6.356 ± 0.076
	FA(0.55)+FV(0.20)+FP(0.25)	0.05	0.386	0.407 ± 0.013	0.111 ± 0.027	0.121 ± 0.036	6.522 ± 0.073
	FP(0.20)+FA(0.55)+FV(0.25)	0.1	0.537	0.592 ± 0.022	0.215 ± 0.041	0.242 ± 0.042	6.215 ± 0.071
	FP(0.25)+FV(0.30)+FA(0.45)	0.1	0.605	0.642 ± 0.020	0.176 ± 0.047	0.204 ± 0.047	6.186 ± 0.077
	FV(0.20)+FA(0.60)+FP(0.20)	0.1	0.483	0.553 ± 0.036	0.239 ± 0.065	0.259 ± 0.064	6.243 ± 0.085
	FV(0.30)+FP(0.25)+FA(0.45)	0.1	0.554	0.597 ± 0.026	0.185 ± 0.048	0.213 ± 0.053	6.184 ± 0.077

performance and computational workload reduction are addressed in subsections 10.6.1 and 10.6.2, respectively. Lastly, subsection 10.6.3 outlines and discusses the potential limitations of this work.

10.6.1 Biometric Performance

It appears that the most successful system ordering follows the training CMC curves, *i.e*. preferring the type of biometric characteristic with the highest identification rate at low ranks to be used first. Accordingly, the best orderings (in terms of the τ metric) in the experiments were Face-Fingerprint and Face-Fingerprint-Fingervein for 2 and 3 modalities, respectively. In general, as has been demonstrated in the previous section, the proposed system increases the biometric performance when benchmarked against the baseline. This increase happens both in terms of FPIR0.1, as well as EER. Although the FNIR can be somewhat higher than that of the baseline (*cf*. figure 10.7), this happens at values of FPIR which are considered impractical for operational systems. Those errors occur due to the pre-filtering – if, for example, at the first level of the cascade a sample of bad quality is filtered out, the proposed system cannot recover, whereas a score-fusion based system might, provided excellent scores for the other modalities. On the other hand, by the act of pre-filtering the database, the potential for false positives is decreased (recall subsection 10.3.2), thus yielding better results in terms of FPIR0.1. In other words, the proposed scheme can be used to increase the security of biometric identification systems which already employ information fusion of multiple biometric modalities.

10.6.2 Computational Workload Reduction

In addition to the aforementioned biometric performance improvement, the proposed system has an impact on the computational complexity of a biometric identification transaction. In this context, two scenarios can be distinguished depending on the *cost of template comparisons* for the used modalities:

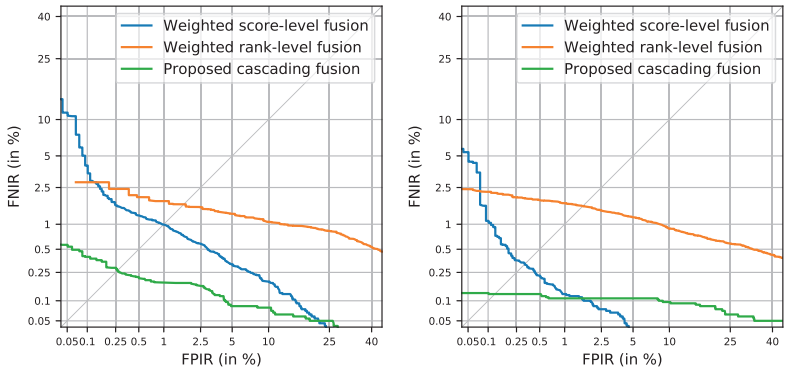
Same cost irrespective of the modality In this case, the computational workload depends exclusively on the penetration rate (recall equation (10.2)). To minimise it, the modalities should be ordered corresponding to the ascending order of their respective shortlist sizes, *i.e.* $S = \{s_1 \dots s_K \mid s_i \leq s_j, \forall i < j\}$. The computational workload (W) of an identification transaction in such a setup would be equal to the total number of comparisons, *i.e.* $W = \#\mathrm{C}_{\mathrm{proposed}}$.

Different cost This case adds an extra factor (w_k) in the equations, representing the cost of the template comparison for the k'th modality, to be multiplied with the shortlist and enrolment database sizes. To minimise the computational workload, the ordering of the system would be $S' = \{s'_1 \dots s'_K \mid s'_i * w_i \leq s'_j * w_j, \forall i < j\}$, and the total computational workload for a biometric identification transaction $W = N \cdot w_1 + \sum_{k=2}^{K} \prod_{i=1}^{k} S'_i \cdot N \cdot w_k$.

In this work, exclusively the first scenario was considered, due to the difficulty of consistently estimating the computational cost of individual template comparisons (see *e.g.* Drozdowski *et al.* [20] for a more detailed discussion on this topic). The main reason for this are the different feature representations and comparators across the modalities. One could, in theory, measure the execution time; however, this effectively amounts to measuring the efficiency of the software implementation and/or the underlying hardware architecture. This limited general use notwithstanding, such experiments would be useful for a specific system implementation (*e.g.* a commercial deployment).

Table 10.6: Summary of the results – best configuration for each of the tested fusion methods (with 95% CI)

Method	Modality	ϵ	PR	τ	EER (in %)	FPIR0.1 (in %)	d′
Rank fusion weighted	FA(0.55)+FP(0.45)	—	1.000	2.743 ± 0.393	1.582 ± 0.195	2.550 ± 0.317	2.540 ± 0.039
	FA(0.40)+FP(0.55)+FV(0.05)	—	1.000	2.349 ± 0.303	1.564 ± 0.117	2.121 ± 0.254	3.532 ± 0.049
Score fusion weighted	FA(0.80)+FP(0.20)	—	1.000	5.107 ± 0.574	0.992 ± 0.146	4.901 ± 0.440	4.739 ± 0.079
	FA(0.55)+FP(0.20)+FV(0.25)	—	1.000	1.986 ± 0.160	0.324 ± 0.063	1.504 ± 0.136	4.985 ± 0.063
Proposed cascading fusion	FA(0.70)+FP(0.30)	0.05	0.545	0.648 ± 0.079	0.254 ± 0.051	0.333 ± 0.102	5.702 ± 0.083
	FA(0.60)+FP(0.20)+FV(0.20)	0.05	0.388	0.409 ± 0.012	0.109 ± 0.022	0.125 ± 0.027	6.356 ± 0.076



(a) 2 modalities

(b) 3 modalities

Figure 10.7: Summary of the results – best configuration for each of the tested fusion methods

10.6.3 Limitations

In terms of computational workload reduction, the main limitation of the proposed system is a hard limit of the potential penetration rate reduction, as described in subsection 10.3.2. Specifically, the biometric comparisons need to be conducted exhaustively on the first level of the cascade, thereby effectively limiting the minimum penetration rate to $\frac{1}{K}$, where K is the number of modalities in the cascade. Indeed, as reported in subsection 10.5.3, the results of the proposed system closely approach this maximal penetration rate reduction, while simultaneously improving the biometric performance. The proposed scheme could, however, be extended by considering another method of computational workload reduction (*e.g.* binning) prior to the cascade in order to further reduce the penetration rate and avoid the exhaustive search at the first level of the cascade.

Another potential limitation is the necessity of the training step, in order to facilitate the shortlist sizes estimation, as well as score normalisation. This, however, is a common property of many (if not most) effective biometric information fusion systems.

In terms of a practical implementation, it should be noted that fully parallelised computations of the comparison scores across all the cascade levels are not possible. Specifically, while the computations on the individual cascade levels are, naturally, trivially parallelisable, it is not possible to compute all the cascade levels simultaneously. This is because the computations at each subsequent level of the cascade need to wait for the completion of the previous level, *i.e.* the creation of the candidate shortlist.

10.7 Summary

This article presented a biometric information fusion-based system which addresses two of the main challenges associated with biometric identification: biometric performance and computational workload. By successively filtering the candidate lists using the individual modalities and subsequently fusing the remaining comparison scores, the biometric performance in the region of the DET curve which is relevant for security sensitive applications, can be significantly improved, while simultaneously reducing the penetration rate (computational workload) of a biometric identification transaction. The proposed method could be seamlessly integrated into many operational multi-modal biometric identification systems, as it is designed to work irrespective of the chosen biometric characteristics or their respective feature representations, and only requires a straightforward training step for the purpose of parameter estimation.

A summary of the results (best configurations in terms of τ) for each of the fusion methods is shown in figure 10.7 and table 10.6. It can be seen that, w.r.t. using the weighted score-level or rank-level fusion alone, the proposed system has the following effects:

Biometric performance is improved in terms of of EER and FPIR0.1 – by an order of magnitude.

Computational workload is reduced in terms of penetration rate – down to around 55% and 39% for 2 and 3 modal system, respectively.

Operational flexibility is retained due to lack of dependence on specific biometric characteristics or template representations.

Future work in this area could consist of, for example, testing the proposed system with an even larger database (albeit those are difficult to come by in the research context), as well as using commercial off-the-shelf biometric recognition systems to assess the practicability of the proposed concept in the context of real biometric applications and operational (not virtual) datasets.

Acknowledgements

This research work has been funded by the German Federal Ministry of Education and Research and the Hessen State Ministry for Higher Education, Research and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE, and the LOEWE-3 BioBiDa Project (594/18-17).

10.8 Bibliography

[1] AASTRUP OLSEN, M., HARTUNG, D., BUSCH, C., AND LARSEN, R. Convolution approach for feature detection in topological skeletons obtained from vascular patterns. In *Workshop on Computational Intelligence in Biometrics and Identity Management (CIBIM)* (April 2011), IEEE, pp. 163–167.

[2] ABAZA, A., AND ROSS, A. Quality based rank-level fusion in multibiometric systems. In *International Conference on Biometrics: Theory, Applications, and Systems (BTAS)* (September 2009), IEEE, pp. 1–6.

[3] ADLER, A., YOUMARAN, R., AND LOYKA, S. Towards a measure of biometric feature information. *Pattern Analysis and Applications 12*, 3 (September 2009), 261–270.

[4] ARONOWITZ, H., LI, M., TOLEDO-RONEN, O., HARARY, S., GEVA, A., BEN-DAVID, S., RENDEL, A., HOORY, R., RATHA, N., PANKANTI, S., AND NAHAMOO, D. Multi-modal biometrics for mobile authentication. In *IEEE International Joint Conference on Biometrics* (September 2014), IEEE, pp. 1–8.

[5] BHUTANI, A., AND BHARDWAJ, P. Biometrics market size by application. Tech. Rep. GMI493, Global Market Insights, August 2017.

[6] BILLEB, S., RATHGEB, C., BUSCHBECK, M., REININGER, H., AND KASPER, K. Efficient two-stage speaker identification based on universal background models. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2014), IEEE, pp. 1–6.

[7] BLACK, D. *The theory of committees and elections*. Cambridge University Press, 1958.

[8] BOWYER, K. W., ORTIZ, E., AND SGROI, A. Iris recognition technology evaluated for voter registration in Somaliland. *Biometric Technology Today 2015*, 2 (February 2015), 5–8.

[9] BURT, C. DHS S&T biometric technology rally results suggest face best for fast processing. `https://www.biometricupdate.com/201909/dhs-st-biometric-technology-rally-results-suggest-face-best-for-fast-processing`, September 2019. Last accessed: 2020–03–11.

[10] CAPPELLI, R., FERRARA, M., AND MALTONI, D. Fingerprint indexing based on minutia cylinder-code. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 33*, 5 (May 2011), 1051–1057.

[11] Consortium for Elections and Political Process Strengthening. Assessment of electoral preparations in the Democratic Republic of the Congo. Tech. rep., CEPPS, May 2018.

[12] Dalwai, A. Aadhaar technology and architecture: principles, design, best practices and key lessons. Tech. rep., Unique Identification Authority of India (UIDAI), March 2014.

[13] Damer, N., Terhörst, P., Braun, A., and Kuijper, A. Efficient, accurate, and rotation-invariant iris code. *Signal Processing Letters 24*, 8 (August 2017), 1233–1237.

[14] Dantcheva, A., Elia, P., and Ross, A. What else does your biometric data reveal? A survey on soft biometrics. *Transactions on Information Forensics and Security (TIFS) 11*, 3 (March 2016), 441–467.

[15] Das, A., Galdi, C., Han, H., Ramachandra, R., Dugelay, J.-L., and Dantcheva, A. Recent advances in biometric technology for mobile devices. In *International Conference on Biometrics Theory, Applications and Systems (BTAS)* (October 2018), IEEE, pp. 1–11.

[16] Daugman, J. Biometric decision landscapes. Tech. Rep. UCAM-CL-TR-482, University of Cambridge - Computer Laboratory, January 2000.

[17] Daugman, J., and Downing, C. Searching for doppelgängers: assessing the universality of the IrisCode impostors distribution. *IET Biometrics 5*, 2 (June 2016), 65–75.

[18] Dinca, L. M., and Hancke, G. P. The fall of one, the rise of many: A survey on multi-biometric fusion methods. *IEEE Access 5* (April 2017), 6247–6289.

[19] Drozdowski, P., Fischer, D., Rathgeb, C., Schiel, C., and Busch, C. Database binning and retrieval in multi-fingerprint identification systems. In *International Workshop on Information Forensics and Security (WIFS)* (December 2018), IEEE, pp. 1–7.

[20] Drozdowski, P., Rathgeb, C., and Busch, C. Computational workload in biometric identification systems: An overview. *IET Biometrics 8*, 6 (November 2019), 351–368.

[21] Drozdowski, P., Rathgeb, C., and Busch, C. Turning a vulnerability into an asset: Accelerating facial identification with morphing. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (May 2019), IEEE, pp. 2582–2586.

[22] DROZDOWSKI, P., STRUCK, F., RATHGEB, C., AND BUSCH, C. Benchmarking binarisation schemes for deep face templates. In *International Conference on Image Processing (ICIP)* (October 2018), IEEE, pp. 191–195.

[23] ELHOSENY, M., ESSA, E., ELKHATEB, A., HASSANIEN, A. E., AND HAMAD, A. Cascade multimodal biometric system using fingerprint and iris patterns. In *International Conference on Advanced Intelligent Systems and Informatics (AISI)* (August 2017), Springer, pp. 590–599.

[24] EUROPEAN COMMISSION. Smart borders. `https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/smart-borders_en`, 2018. Last accessed: 2020–03–11.

[25] EUROPEAN UNION AGENCY FOR THE OPERATIONAL MANAGEMENT OF LARGE-SCALE IT SYSTEMS IN THE AREA OF FREEDOM, SECURITY AND JUSTICE. Eurodac storage capacity increased. `https://www.eulisa.europa.eu/Newsroom/News/Pages/Eurodac-storage-capacity-increased.aspx`, April 2016. Last accessed: 2020–03–11.

[26] FEDERAL BUREAU OF INVESTIGATION. CODIS - NDIS statistics. `https://www.fbi.gov/services/laboratory/biometric-analysis/codis/ndis-statistics`, June 2018. Last accessed: 2020–03–11.

[27] GEHRMANN, C., RODAN, M., AND JÖNSSON, N. Metadata filtering for user-friendly centralized biometric authentication. *EURASIP Journal on Information Security 2019*, 1 (June 2019), 7.

[28] GEMALTO. Automated Fingerprint Identification System (AFIS) - a short history. `https://www.gemalto.com/govt/biometrics/afis-history`, April 2019. Last accessed: 2020–03–11.

[29] GEMALTO. DHS's automated biometric identification system IDENT - the heart of biometric visitor identification in the USA. `https://www.gemalto.com/govt/customer-cases/ident-automated-biometric-identification-system`, March 2019. Last accessed: 2020–03–11.

[30] GENTILE, J. E., RATHA, N., AND CONNELL, J. An efficient, two-stage iris recognition system. In *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2009), IEEE, pp. 211–215.

[31] GENTILE, J. E., RATHA, N., AND CONNELL, J. SLIC: short-length iris codes. In *International Conference on Biometrics: Theory, Applications, and Systems (BTAS)* (2009), IEEE, pp. 1–5.

[32] GOFMAN, M., MITRA, S., CHENG, K., AND SMITH, N. Quality-based score-level fusion for secure and robust multimodal biometrics-based authentication on consumer mobile devices. In *International Conference on Software Engineering Advances (ICSEA)* (November 2015), IARIA, pp. 274–276.

[33] GOFMAN, M. I., MITRA, S., CHENG, T.-H. K., AND SMITH, N. T. Multimodal biometrics for enhanced mobile device security. *Communications of the ACM 59*, 4 (April 2016), 58–65.

[34] GOMEZ-BARRERO, M., RATHGEB, C., RAJA, K. B., RAGHAVENDRA, R., AND BUSCH, C. Biometric symmetry: Implications on template protection. In *European Signal Processing Conference (EUSIPCO)* (August 2017), IEEE, pp. 941–945.

[35] GROTHER, P., NGAN, M., AND HANAOKA, K. Ongoing face recognition vendor test (FRVT) part 2: Identification. Tech. Rep. NISTIR 8238, National Institute of Standards and Technology, November 2018.

[36] GUO, Z., AND HALL, R. W. Parallel thinning with two-subiteration algorithms. *Communications of the ACM 32*, 3 (March 1989), 359–373.

[37] HAO, F., DAUGMAN, J., AND ZIELINSKI, P. A fast search algorithm for a large fuzzy database. *Transactions on Information Forensics and Security (TIFS) 3*, 2 (June 2008), 203–212.

[38] HARTUNG, D., OLSEN, M. A., XU, H., AND BUSCH, C. Spectral minutiae for vein pattern recognition. In *International Joint Conference on Biometrics (IJCB)* (October 2011), IEEE, pp. 1–7.

[39] HARTUNG, D., OLSEN, M. A., XU, H., NGUYEN, H. T., AND BUSCH, C. Comprehensive analysis of spectral minutiae for vein pattern recognition. *IET Biometrics 1*, 1 (March 2012), 25–36.

[40] HO, T. K., HULL, J. J., AND SRIHARI, S. N. Decision combination in multiple classifier systems. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 1*, 1 (January 1994), 66–75.

[41] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 19795-1:2006. Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*. International Organization for Standardization and International Electrotechnical Committee, April 2006.

[42] ISO/IEC JTC1 SC37 BIOMETRICS. Iso/iec tr 24722:2015. information technology – biometrics – multimodal and other multibiometric fusion. Tech. rep., International Organization for Standardization, December 2015.

[43] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 2382-37:2017. Information technology – Vocabulary – Part 37: Biometrics*, 2 ed. International Organization for Standardization and International Electrotechnical Committee, February 2017.

[44] JAIN, A., KLARE, B., AND ROSS, A. Guidelines for best practices in biometrics research. In *International Conference on Biometrics (ICB)* (May 2015), IEEE, pp. 541–545.

[45] JAIN, A., AND ROSS, A. Fingerprint mosaicking. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (May 2002), vol. 4, IEEE, pp. IV–4064–IV–4067.

[46] JAIN, A. K., NANDAKUMAR, K., AND ROSS, A. Score normalization in multimodal biometric systems. *Pattern recognition 38*, 12 (December 2005), 2270–2285.

[47] KANHANGAD, V., KUMAR, A., AND ZHANG, D. Contactless and pose invariant biometric identification using hand surface. *Transactions on Image Processing (TIP) 20*, 5 (May 2011), 1415–1424.

[48] KAUSHIK, V. D., UMARANI, J., GUPTA, A. K., AND GUPTA, P. An efficient indexing scheme for face database using modified geometric hashing. *Neurocomputing 116* (2013), 208–221.

[49] KAVATI, I., PRASAD, M., AND BHAGVATI, C. Search space reduction in biometric databases: a review. In *Computer Vision: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2018, pp. 1600–1626.

[50] KUMAR, A., AND SHEKHAR, S. Personal identification using multibiometrics rank-level fusion. *Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 41*, 5 (September 2011), 743–752.

[51] KUMAR, A., AND WANG, K. Identifying humans by matching their left palmprint with right palmprint images using convolutional neural network. In *International Workshop on Deep Learning and Pattern Recognition* (July 2016), SPIE.

[52] KUSUMA, G. P., AND CHUA, C.-S. PCA-based image recombination for multimodal 2D + 3D face recognition. *Image and Vision Computing 29*, 5 (April 2011), 306–316.

[53] LIM, M.-H., TEOH, A. B. J., AND KIM, J. Biometric feature-type transformation: Making templates compatible for secret protection. *Signal Processing Magazine 32*, 5 (September 2015), 77–87.

[54] MARKETS AND MARKETS. Biometric system market by authentication type - global forecast to 2023. Tech. Rep. SE 3449, Markets and Markets, July 2018.

[55] MIURA, N., NAGASAKA, A., AND MIYATAKE, T. Extraction of finger-vein patterns using maximum curvature points in image profiles. *Transactions on Information and Systems 90*, 8 (August 2007), 1185–1194.

[56] MOSES, K. R., HIGGINS, P., MCCABE, M., PROBHAKAR, S., AND SWANN, S. *Fingerprint Sourcebook*. US Department of Justice, 2010, ch. Automated Fingerprint Identification System (AFIS), pp. 1–33.

[57] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. Face Recognition Vendor Test (FRVT) 1:N Evaluation. https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-1n-2018-evaluation, October 2017. Last accessed: 2020–03–11.

[58] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST). Iris Exchange (IREX). https://www.nist.gov/programs-projects/iris-exchange-irex-overview, 2018. Last accessed: 2020–03–11.

[59] PAUL, P. P., GAVRILOVA, M. L., AND ALHAJJ, R. Decision fusion for multimodal biometrics using social network analysis. *Transactions on Systems, Man, and Cybernetics: Systems 44*, 11 (November 2014), 1522–1533.

[60] PAYNTER, T. Northrop Grumman wins $95 million award from Department of Homeland Security to develop next-generation biometric identification services system. https://news.northropgrumman.com/news/releases/northrop-grumman-wins-95-million-award-from-department-of-homeland-security-to-develop-next-generation-biometric-identification-services-system, February 2018. Last accessed: 2020–03–11.

[61] PFLUG, A., BUSCH, C., AND ROSS, A. 2D ear classification based on unsupervised clustering. In *International Joint Conference on Biometrics (IJCB)* (September 2014), IEEE, pp. 1–8.

[62] PFLUG, A., RATHGEB, C., SCHERHAG, U., AND BUSCH, C. Binarization of spectral histogram models: An application to efficient biometric identification. In *International Conference on Cybernetics (CYBCONF)* (June 2015), IEEE, pp. 501–506.

[63] PRABHAKAR, S., AND JAIN, A. K. Decision-level fusion in fingerprint verification. *Pattern Recognition 35*, 4 (April 2002), 861–874.

[64] PROENÇA, H. Iris biometrics: Indexing and retrieving heavily degraded data. *Transactions on Information Forensics and Security (TIFS) 8*, 12 (December 2013), 1975–1985.

[65] PROENÇA, H., AND NEVES, J. Iris biometric indexing. In *Iris and Periocular Biometric Recognition*. Institution of Engineering and Technology, July 2017, pp. 101–124.

[66] RADU, P., SIRLANTZIS, K., HOWELLS, G., DERAVI, F., AND HOQUE, S. A review of information fusion techniques employed in iris recognition systems. *International Journal of Advanced Intelligence Paradigms 4*, 3/4 (February 2012), 211–240.

[67] RATHA, N. K., KARU, K., CHEN, S., AND JAIN, A. K. A real-time matching system for large fingerprint databases. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 18*, 8 (August 1996), 799–813.

[68] RATHGEB, C., BREITINGER, F., BAIER, H., AND BUSCH, C. Towards bloom filter-based indexing of iris biometric data. In *International Conference on Biometrics (ICB)* (May 2015), IEEE, pp. 422–429.

[69] RATHGEB, C., BREITINGER, F., BUSCH, C., AND BAIER, H. On application of Bloom filters to iris biometrics. *IET Biometrics 3*, 4 (December 2014), 207–218.

[70] RATHGEB, C., UHL, A., AND WILD, P. On combining selective best bits of iris-codes. In *European Workshop on Biometrics and Identity Management (BioID)* (March 2011), Springer, pp. 227–237.

[71] RATTANI, A., DERAKHSHANI, R., AND ROSS, A. *Selfie Biometrics: Advances and Challenges*. Springer, 2019.

[72] ROSS, A., NANDAKUMAR, K., AND JAIN, A. K. *Handbook of multibiometrics*. Springer, 2006.

[73] ROSS, A., AND SUNDER, M. S. Block based texture analysis for iris classification and matching. In *Conference on Computer Vision and Pattern Recognition - Workshops (CVPRW)* (June 2010), IEEE, pp. 30–37.

[74] SCHROFF, F., KALENICHENKO, D., AND PHILBIN, J. FaceNet: A unified embedding for face recognition and clustering. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2015), IEEE, pp. 815–823.

[75] SCHUCH, P. Survey on features for fingerprint indexing. *IET Biometrics 8*, 1 (January 2019), 1–13.

[76] SNELICK, R., INDOVINA, M., YEN, J., AND MINK, A. Multimodal biometrics: issues in design and testing. In *International Conference on Multimodal Interfaces* (November 2003), ACM, pp. 68–72.

[77] TANG, Y., GAO, F., FENG, J., AND LIU, Y. FingerNet: An unified deep network for fingerprint minutiae extraction. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 108–116.

[78] THAKKAR, D. Global biometric market analysis: Trends and future prospects. https://www.bayometric.com/global-biometric-market-analysis/, August 2018. Last accessed: 2020–03–11.

[79] ULERY, B., HICKLIN, A., WATSON, C., FELLNER, W., AND HALLINAN, P. Studies of biometric fusion. Tech. Rep. NISTIR 7346, National Institute of Standards and Technology, September 2006.

[80] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Aadhaar dashboard. https://www.uidai.gov.in/aadhaar_dashboard/. Last accessed: 2020–03–11.

[81] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Role of biometric technology in Aadhaar enrollment. Tech. rep., UIDAI, January 2012.

[82] VAŽAN, R. SourceAFIS – opensource fingerprint matcher. https://sourceafis.machinezoo.com/, 2019. Last accessed: 2020–03–11.

[83] XU, H., VELDHUIS, R. N. J., KEVENAAR, T. A. M., AKKERMANS, A. H. M., AND BAZEN, A. M. Spectral minutiae: A fixed-length representation of a minutiae set. In *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (June 2008), IEEE, pp. 1–6.

[84] YAN, X., KANG, W., DENG, F., AND WU, Q. Palm vein recognition based on multi-sampling and feature-level fusion. *Neurocomputing 151*, 2 (March 2015), 798–807.

Chapter 11

Turning a Vulnerability into an Asset: Accelerating Facial Identification with Morphing

Abstract

In recent years, morphing of facial images has arisen as an important attack vector on biometric systems. Detection of morphed images has proven challenging for automated systems and human experts alike. Likewise, in recent years, the importance of efficient (fast) biometric identification has been emphasised by the rapid rise and growth of large-scale biometric systems around the world.

In this paper, the aforementioned, hitherto unrelated, topics within the biometrics domain are combined: the properties of morphed images are exploited for the purpose of improving the transaction times of a biometric identification system. Specifically, morphs of two or more samples are used in the pre-selection step of a two-stage biometric identification system. In a proof-of-concept experimental evaluation using two state-of-the-art open-source facial recognition frameworks it is shown, that the proposed system achieves hit rates comparable to that of an exhaustive search-based baseline, while significantly reducing the penetration rate (and thus the computational workload) associated with the biometric identification transactions.

Addressed research question(s): RQ1, RQ3, RQ4
Reference: DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Turning a vulnerability into an asset: Accelerating facial identification with morphing. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (May 2019), IEEE, pp. 2582–2586.

11.1 Introduction

In recent years, the interest around biometric technologies has been growing steadily. This is evidenced by various market value studies (see *e.g.* [1, 20]), as well as flourishing deployments of national and international systems for

purposes of, among others, personal identification, law enforcement, and facilitating elections (see *e.g.* [2, 6, 7, 27]).

In this paper, two hitherto unrelated areas of biometric research are combined:

1. Computational workload reduction in biometric identification.
2. Facial image morphing.

Specifically, facial image morphing, a crucial vulnerability of operational biometric systems is turned into an advantage through which the penetration rate (computational workload) of biometric identification transactions can be significantly reduced. This is achieved by employing a two-stage retrieval approach, which exploits certain properties of morphed facial images.

The remainder of this paper is organised as follows: section 11.2 introduces the relevant background concepts and the related work. In section 11.3, the proposed system is described and visualised conceptually. Section 11.4 presents the experimental setup and the achieved results, while a summary and concluding remarks are given in section 11.5.

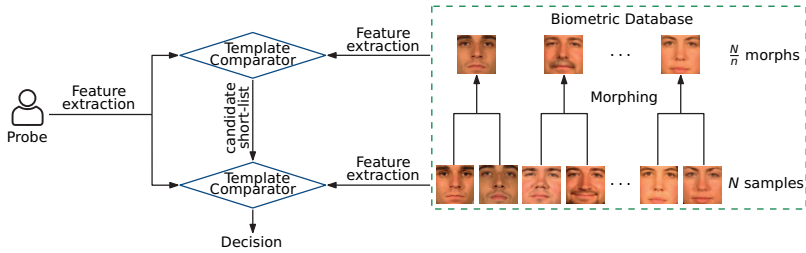


Figure 11.1: Proposed system overview (here, $n = 2$)

11.2 Background and Related Work

In this section, the research fields relevant to this paper are briefly introduced: the operation modes of a biometric system and challenges associated with biometric identification (subsection 11.2.1), and facial images morphing (subsection 11.2.2).

11.2.1 Operation Modes of a Biometric System

Biometric systems generally operate in one of two modes:

Verification Resolved in a 1:1 comparison between a biometric probe and the biometric reference of a claimed identity.

Identification No identity claim is made. Thus, in the worst case, an exhaustive linear search is required in order to find a candidate list or to reach a decision with the rank one on the list.

The second case is obviously more challenging from the practical point of view. However, the naïve approach of the exhaustive search suffers from two key issues:

Computational cost The growing number of enrolled subjects, gradually slows down the response times, which in turn requires investment into optimisations and/or hardware architecture.

False positives costs The probability of at least one false positive (P_N) occurring in a identification scenario is: $P_N = 1 - (1 - P_1)^N$, where N is the number of enrolled subjects and P_1 the false positive probability of a one-to-one template comparison (see Daugman [3]). This relationship is very demanding – even for systems which perform extremely well in verification mode (*i.e.* have low P_1), the value of P_N very quickly becomes unacceptably high, as the number of enrolled subjects N increases.

Performing accurate and efficient biometric identification (*i.e.* not by an exhaustive search) has been stated to be one of the important, unsolved issues in the biometrics field in general by Daugman, the inventor of iris recognition in a recent interview [12]. Over time, many approaches have been developed in this field; for more insights, the reader is referred to surveys by *e.g.* Proença *et al.* [22], Schuch *et al.* [25], and Kavati *et al.* [15].

11.2.2 Morphing of Facial Images

By using image morphing methods, it is possible to create biometric samples which *contain biometric information from two or more distinct data subjects*. The resulting artificial sample resembles the two (or more) original samples in the image and feature domain; thus, breaking the unique link between data subjects and their biometric reference data (*i.e.* the enrolment record). In other words, the subjects whose biometric samples were used to create the morphed image can both be matched (accepted) during subsequent biometric recognition transactions with the morphed reference image. This vulnerability was first introduced by Ferrara *et al.* [8] (the so-called "magic passport") and shown to be a feasible attack vector against automated systems and human experts alike [9]. A typical morphing process includes:

1. Facial landmark detection and triangulation in two or more images.
2. Landmark averaging to a single set of landmarks.

3. Image warping and alpha blending.

The process is quite simple, and even non-experts can generate realistic looking morphed face images with a variety of inexpensive or even free software tools. Figure 11.2 shows an example of facial image morphing.

(a) Subject 1 (b) Morph (c) Subject 2

Figure 11.2: Morphing example (from Scherhag *et al*. [23])

In recent years, significant research effort has been devoted to development of methods capable of automatically detecting morphed images. Among others, methods based on general purpose texture descriptors (*e.g*. Scherhag *et al*. [23]), deep learning (*e.g*. Seibold *et al*. [26]), media forensics (*e.g*. Hildebrandt *et al*. [11]), and camera noise (*e.g*. Debiasi *et al*. [4]) have been proposed. For more detailed treatment of morph creation and detection methods, the reader is referred to a recent survey by Makrushin *et al*. [18]. It is not the intention of this paper to develop morphing detection algorithms; instead, the goal is to take advantage of morphing in the context of a biometric identification system.

11.3 Proposed System

Figure 11.1 shows a conceptual view of the proposed system. Following symbols are used:

N the number of enrolled subjects.

n the number of samples contributing to a morph.

k the number of morphs in the selected candidate short-list.

The key idea is to perform a fusion of the enrolled samples on image level through morphing. Each thus created image contains biometric information

from multiple subjects (recall subsection 11.2.2). The morphed images are expected to retain enough discriminative power, so that upon retrieval the comparison score of a biometric against the correct (mated) morphed image will tend to be better than scores against other (non-mated) morphs. It would then possible to select a candidate short-list based on the comparison scores between the biometric probe and the morphs. Hence, a biometric identification transaction proceeds in a two-stage process (conceptually similar to *e.g.* Gentile *et al.* [10]):

1. Perform template comparisons between the biometric probe and the enrolled morphed samples exhaustively. Based on the comparison scores, pre-select a short-list of the most likely candidates.
2. Within the candidate short-list, perform template comparisons between the biometric probe and the normal enrolled samples.

In order for the system to reduce the computational workload associated with an identification transaction, the following relation must be satisfied: $\frac{N}{n} + k * n < N$. Figure 11.3 visualises this relation between the parameters n and k, and the number of necessary comparisons for a biometric identification transaction for $N = 400$ subjects. The baseline (exhaustive search), which is not dependent on those parameters is plotted as a horizontal line for reference.

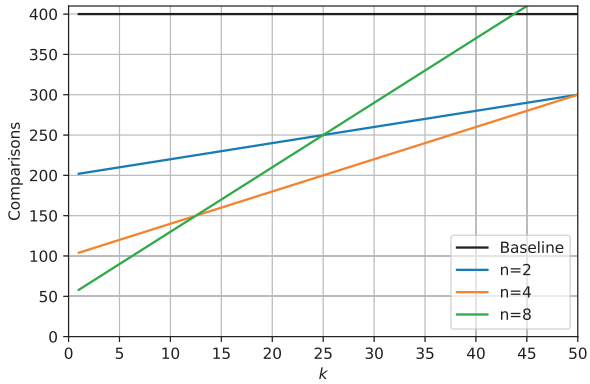


Figure 11.3: Template comparisons per identification transaction

11.4 Experiments

In this section, the experimental setup and the used dataset are described (subsection 11.4.1), along with the results of the experiments (subsection 11.4.2).

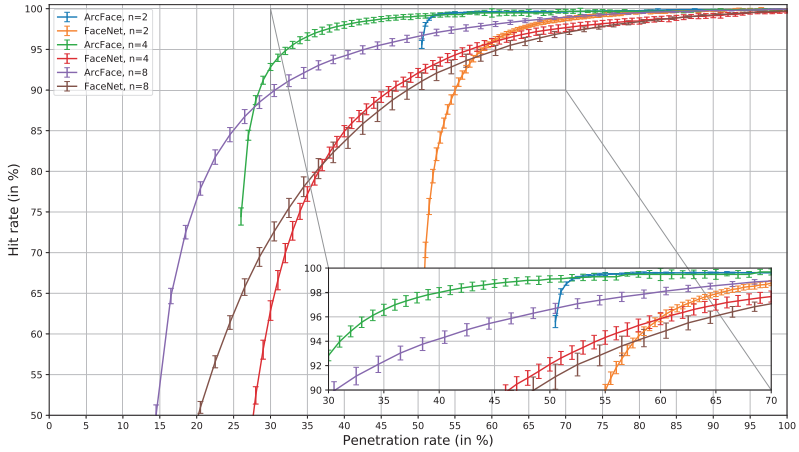


Figure 11.4: Results (with errorbars denoting the 95% confidence interval)

11.4.1 Experimental Setup

The experiments are conducted using the FERET facial image database [21], which contains 14126 images from 1199 data subjects (with varying number of images per subject). Specifically, frontal images compliant with the ISO/IEC requirements [14] for high quality facial images have been selected, resulting in a subset consisting of 6963 images from 573 subjects. The morphed images were automatically generated from pairs of images using landmark detection by dlib [16], Delaunay triangulation [17], and image warping and alpha blending as described in [19]. The alpha value is always set to $\frac{1}{n}$, *i.e.* the samples contribute equally to the morph.

Two state-of-the-art open-source facial recognition frameworks based on deep neural networks were used: FaceNet [24] and ArcFace [5], both with the pre-trained models provided by their authors. The frameworks extract feature vectors (embeddings) from the non-morphed and morphed facial images consisting of 512 float values, which can be subsequently compared using metrics such as the cosine distance.

The experiments were conducted in the biometric identification mode, utilising cross-validation over 10 folds. The variables mentioned in section 11.3 are as follows: $N = 400$ (200 men and 200 women; morphs were created within the same gender only, but no other heuristics were used when deciding which samples to morph, *i.e.* they were selected at random), $n \in \{2, 4, 8\}$, and $k \in \{1 \dots \frac{N}{n}\}$. Thus, for each experiment, several thousand identification transactions are performed (depending on the enrolled subjects, since the number of samples per subject in the dataset varies). The biometric performance is evaluated in terms of metrics defined by ISO/IEC [13]: hit rate

(HR)[1], penetration rate (PR), and rank-1 identification rate (RR-1).

11.4.2 Results

The accuracies achieved by various configurations of the proposed system in the pre-selection step are shown in table 11.1 and figure 11.4, where the trade-off between hit rate and penetration rate is plotted. Best results for the high hit rates (≥99%) are achieved using the ArcFace feature extractor, whereby the penetration rate is approximately halved. The FaceNet feature extractor achieves significantly poorer results, albeit it still manages to reduce the penetration rate somewhat. It can also be observed, that enough biometric information for high hit rates and penetration rate reduction is retained even when morphing $n = 8$ subjects together, although the best results occur when $n = 2$ or $n = 4$.

Table 11.1: Pre-selection results

Feature Extractor	n	**PR at**		
		95% HR	**99% HR**	**99.5% HR**
ArcFace	2	50.5%	52.0%	55.0%
	4	32.0%	48.0%	57.0%
	8	42.5%	70.5%	80.5%
FaceNet	2	57.0%	86.0%	95.0%
	4	42.5%	70.5%	80.5%
	8	60.5%	86.5%	94.5%

The results achieved by the baseline and the two-stage system (with optimal k values) are shown in tables 11.2 and 11.3, respectively. All the results are reported with a 95% confidence interval. It is observed, that particularly for the ArcFace feature extractor virtually no biometric performance loss occurs at $n = 2$, while the penetration rate is significantly reduced.

Table 11.2: Baseline results

Feature Extractor	**RR-1**	**PR**
ArcFace	99.18% ± 0.11%	1.0
FaceNet	98.84% ± 0.16%	

11.5 Summary

In this paper, two heretofore unrelated fields within the domain of biometrics have been combined. Specifically, the properties of morphed facial im-

[1] *i.e.* 100% minus the pre-selection error rate.

Table 11.3: Two-stage system results

Feature Extractor	n	k	RR-1	PR
ArcFace	2	5	98.82% $\pm$ 0.12%	52.5%
	4	20	97.57% $\pm$ 0.26%	45.0%
	8	25	96.09% $\pm$ 0.19%	50.0%
FaceNet	2	30	96.97% $\pm$ 0.31%	65.0%
	4	30	93.61% $\pm$ 0.56%	55.0%
	8	30	96.51% $\pm$ 0.26%	72.5%

ages have been used at the pre-selection step of a two-stage biometric identification system. It has been shown, that through the morphing process of two or even more data subjects, enough biometric information is retained to facilitate an accurate pre-selection of a candidate short-list. The experiments with two state-of-the-art open-source deep facial recognition frameworks show high ($\geq$99%) hit rates, while reducing the associated penetration rates (and thus the computational workload) down to around 50% of the baseline system. Future work could consist of a more comprehensive evaluation extending this proof-of-concept study, for example utilising:

- Other feature extractors and recognition frameworks, particularly commercial off-the-shelf systems.
- Additional morphing techniques and tools.
- Larger datasets.

Furthermore, improvements to the morphing process of the enrolled samples could be attempted – for instance, morphing most similar subjects together, rather than doing so randomly. Lastly, the morphed images could possibly be organised into a tree-like hierarchical search structure with the aim of further reducing the search space. One technical limitation of the proposed method is that it requires frontal images of good quality, albeit in practice such images are already being captured in data acquisition with controlled environment and cooperative data subjects (*e.g.* passport issuance).

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF), by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP), and the LOEWE-3 BioBiDa Project (594/18-17).

11.6 Bibliography

[1] BHUTANI, A., AND BHARDWAJ, P. Biometrics market size by application. Tech. Rep. GMI493, Global Market Insights, August 2017.

[2] CONSORTIUM FOR ELECTIONS AND POLITICAL PROCESS STRENGTHENING. Assessment of electoral preparations in the Democratic Republic of the Congo. Tech. rep., CEPPS, May 2018.

[3] DAUGMAN, J. Biometric decision landscapes. Tech. Rep. UCAM-CL-TR-482, University of Cambridge - Computer Laboratory, January 2000.

[4] DEBIASI, L., SCHERHAG, U., RATHGEB, C., UHL, A., AND BUSCH, C. PRNU-based detection of morphed face images. In *International Workshop on Biometrics and Forensics (IWBF)* (June 2018), IEEE, pp. 1–7.

[5] DENG, J., GUO, J., AND ZAFEIRIOU, S. ArcFace: Additive angular margin loss for deep face recognition. *Computing Research Repository (CoRR)* (January 2018), 1–11.

[6] EUROPEAN UNION AGENCY FOR THE OPERATIONAL MANAGEMENT OF LARGE-SCALE IT SYSTEMS IN THE AREA OF FREEDOM, S., AND JUSTICE. *Biometrics in Large-Scale IT*. eu-LISA, 2015.

[7] FEDERAL BUREAU OF INVESTIGATION. CODIS - NDIS statistics. `https://www.fbi.gov/services/laboratory/biometric-analysis/codis/ndis-statistics`, June 2018. Last accessed: 2020–03–11.

[8] FERRARA, M., FRANCO, A., AND MALTONI, D. The magic passport. In *International Joint Conference on Biometrics (IJCB)* (September 2014), IEEE, pp. 1–7.

[9] FERRARA, M., FRANCO, A., AND MALTONI, D. On the effects of image alterations on face recognition accuracy. In *Face Recognition Across the Imaging Spectrum*. Springer, February 2016, pp. 195–222.

[10] GENTILE, J. E., RATHA, N., AND CONNELL, J. An efficient, two-stage iris recognition system. In *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2009), IEEE, pp. 211–215.

[11] HILDEBRANDT, M., NEUBERT, T., MAKRUSHIN, A., AND DITTMANN, J. Benchmarking face morphing forgery detection: Application of StirTrace for impact simulation of different processing steps. In *International Workshop on Biometrics and Forensics (IWBF)* (April 2017), IEEE, pp. 1–6.

[12] Institute of Electrical and Electronics Engineers. Biometric council newsletter. http://ieee-biometrics.org/images/pdf/Newsletter_Nov_2015_corrected.pdf, November 2015. Last accessed: 2020–03–11.

[13] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC 19795-1:2006. Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*. International Organization for Standardization and International Electrotechnical Committee, April 2006.

[14] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC 19794-5:2011. Information technology – Biometric data interchange formats – Part 5: Face image data*. International Organization for Standardization and International Electrotechnical Committee, October 2011.

[15] Kavati, I., Prasad, M., and Bhagvati, C. Search space reduction in biometric databases: a review. In *Computer Vision: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2018, pp. 1600–1626.

[16] King, D. E. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research (JMLR) 10* (2009), 1755–1758.

[17] Lee, D.-T., and Schachter, B. J. Two algorithms for constructing a Delaunay triangulation. *International Journal of Computer & Information Sciences 9*, 3 (June 1980), 219–242.

[18] Makrushin, A., and Wolf, A. An overview of recent advances in assessing and mitigating the face morphing attack. In *European Signal Processing Conference (EUSIPCO)* (September 2018), IEEE, pp. 1017–1021.

[19] Mallick, S. Face morph using OpenCV – C++ / Python. https://www.learnopencv.com/face-morph-using-opencv-cpp-python/, March 2016. Last accessed: 2020–03–11.

[20] Markets and Markets. Biometric system market by authentication type - global forecast to 2023. Tech. Rep. SE 3449, Markets and Markets, July 2018.

[21] Phillips, P. J., Moon, H., Rizvi, S. A., and Rauss, P. J. The FERET evaluation methodology for face-recognition algorithms. *Transactions on pattern analysis and machine intelligence (TPAMI) 22*, 10 (October 2000), 1090–1104.

[22] Proença, H., and Neves, J. Iris biometric indexing. In *Iris and Periocular Biometric Recognition*. Institution of Engineering and Technology, July 2017, pp. 101–124.

[23] SCHERHAG, U., RAGHAVENDRA, R., RAJA, K. B., GOMEZ-BARRERO, M., RATHGEB, C., AND BUSCH, C. On the vulnerability of face recognition systems towards morphed face attacks. In *International Workshop on Biometrics and Forensics (IWBF)* (April 2017), IEEE, pp. 1–6.

[24] SCHROFF, F., KALENICHENKO, D., AND PHILBIN, J. FaceNet: A unified embedding for face recognition and clustering. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2015), IEEE, pp. 815–823.

[25] SCHUCH, P. Survey on features for fingerprint indexing. *IET Biometrics 8*, 1 (January 2019), 1–13.

[26] SEIBOLD, C., SAMEK, W., HILSMANN, A., AND EISERT, P. Detection of face morphing attacks by deep learning. In *International Workshop on Digital Watermarking (IWDW)* (August 2017), Springer, pp. 107–120.

[27] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Role of biometric technology in Aadhaar enrollment. Tech. rep., UIDAI, January 2012.

Chapter 12

Towards Pre-alignment of Near-infrared Iris Images

Abstract

The necessity of biometric template alignment imposes a significant computational load and increases the probability of false positive occurrences in biometric systems. While for some modalities, automatic pre-alignment of biometric samples is utilised, this topic has not yet been explored for systems based on the iris.

This paper presents a method for pre-alignment of iris images based on the positions of automatically detected eye corners. Existing work in the area of automatic eye corner detection has hitherto only involved visible wavelength images; for the near-infrared images, used in the vast majority of current iris recognition systems, this task is significantly more challenging and as of yet unexplored. A comparative study of two methods for solving this problem is presented in this paper. The eye corners detected by the two methods are then used for the pre-alignment and biometric performance evaluation experiments. The system utilising image pre-alignment is benchmarked against a baseline iris recognition system on the iris subset of the BioSecure database. In the benchmark, the workload associated with alignment compensation is significantly reduced, while the biometric performance remains unchanged or even improves slightly.

Addressed research question(s): RQ1, RQ4
Reference: DROZDOWSKI, P., RATHGEB, C., HOFBAUER, H., WAGNER, J., UHL, A., AND BUSCH, C. Towards pre-alignment of near-infrared iris images. In *International Joint Conference on Biometrics (IJCB)* (October 2017), IEEE, pp. 359–366.

12.1 Introduction

The iris is one of the main biometric characteristics used in biometric systems around the world. At the time of this writing, the Indian Aadhaar system has enrolled over 1 billion subjects' multi-modal (including iris) biomet-

ric data [27]. The border control system of United Arab Emirates checks every traveller against a growing blacklist consisting of hundreds of thousands of subjects [1]. The deployments of this size and importance face strenuous requirements in terms of, among other matters, biometric performance and computational efficiency.

Following Daugman's approach [5], which is the core of most public operational systems, four major modules constitute an iris recognition system: (1) acquisition of the near-infrared image, where most current deployments require subjects to fully cooperate with the capture device in order to capture images of sufficient quality; (2) pre-processing, which involves a detection of inner and outer iris boundaries, a detection of eyelids, an exclusion of eyelashes as well as contact lens rings, a scrubbing of specular reflections and an estimation of quality factors [14]. Subsequently, the iris is mapped to dimensionless coordinates, *i.e.* a normalized rectangular texture, and an according noise mask is stored; (3) feature extraction, in which a two-dimensional binary feature vector, *i.e.* iris-code, is generated by applying adequate filters to the pre-processed iris texture. This binary data representation enables compact storage and rapid (4) comparison, which is based on the estimation of Hamming distance (HD) scores between pairs of iris-codes. In the comparison stage circular bit shifts are applied to iris-codes and HD scores are estimated at $\pm K$ different shifting positions, *i.e.* relative tilt angles, in order to compensate the biometric sample misalignment. The minimal obtained HD, which corresponds to an optimal alignment, represents the final score.

Considering multiple shifting positions during a template comparison increases the computational workload of the system and the probability of a false match with K [6]. This is especially important for identification systems, where an exhaustive search of the reference database is performed during an authentication attempt. By pre-aligning the eye images, the aforementioned cost (in terms of computational workload and biometric performance degradation) could be significantly reduced, thus partially alleviating the issues created by the necessity of alignment compensation. For the biometric references, the pre-alignment could be performed at enrollment stage, while any additional computational cost of pre-alignment of the biometric probes would be inconsequential in relation to the template comparison costs, since in any sizeable biometric identification system, the computational costs are dominated by the template comparisons [9]. Although image pre-alignment has been utilised in, for instance, fingerprint and face based biometric systems (see *e.g.* [26] and [4]), as of yet it has not been explored in the context of iris recognition systems.

The remainder of this paper is organised as follows: In section 12.2, the related work is presented. Section 12.2.2 explains the usage of eye corners in eye images pre-alignment and outlines the proposed approaches to au-

tomatic detection of eye corners in near-infrared images. The experimental set-up and obtained results are presented and discussed in section 12.3, while concluding remarks are given in section 12.4.

12.2 Related Work

The work presented in this paper combines two areas of research - automatic detection of eye corners and reduction of the alignment costs in iris identification systems. This section is accordingly divided into two subsections.

12.2.1 Eye Corner Detection

Facial landmark detection represents a well-studied area in computer vision. It forms the basis for numerous types of applications, such as face recognition or emotion estimation. Facial landmarks detected by state-of-the-art methods tend to vary in number and type; however, the vast majority of approaches extracts eye corner positions as specified in ISO/IEC 19794-5 [12].

In the context of iris recognition, automatic eye corner detection approaches for visible spectrum images have been presented by a number of researchers. Xu *et al*. [29] base their approach on the semantic features of the inner and outer eye corners, an angle model based on the eyelids and utilise a logistic regression classifier for the detection. Xia and Yan [28] use weighted variance projection function to detect first the regions of interest and then the eye corners themeselves. Erdogmus and Dugelay [8] use the Hough transform to detect the eyelid contours and subsequently establish the eye corners at the intersection of polynomials fitted to said contours. Santos and Proença [24] perform experiments on low-quality data, in which they utilise sclera segmentation and eyelid contours to generate a set of candidate points, from which the final eye corner locations are chosen based on a fusion of a number of metrics calculated for all the points in the candidate set. More recently, Zhang *et al*. [31] used a two-step process in which the rough locations of the eye corners are estimated and refined using image texture information.

All of the above report excellent results, ranging between 90% and 100% correct detection of eye corners - depending on how the groundtruth was established and what metrics and parameters were used to measure the detection accuracy. However, it is important to reiterate, that all of the mentioned approaches use *visible wavelength* eye images (or regions of interest extracted from facial images). Eye corner detection in *near-infrared* images, which are currently used in operational (large-scale) iris recognition systems, is a significantly more challenging task. In contrast to iris images acquired at visible wavelengths, near-infrared images exhibit a low contrast between sclera and skin, *cf*. figure 12.4. Hence, a proper sclera segmentation, which is re-

quired in some of the mentioned approaches, is not feasible for near-infrared images.

12.2.2 Alignment Cost Reduction

As has been mentioned in section 12.1, the traditional iris-code based iris identification systems require significant workload to be put into alignment compensation. In recent years, some interest has been exhibited towards developing methods to reduce (or even eliminate) the number of relative alignment positions that need to be considered in order to achieve an acceptable biometric performance. Du *et al*. [7] have presented a feature extractor for iris recognition based on one-dimensional signatures and showed that such an approach does not require an alignment of extracted templates. Alonso-Fernandez *et al*. [2] suggested to apply scale invariant feature transform to extract iris texture features prior to the normalisation step, where a comparison of keypoint-based feature vectors does not require the traditional alignment procedure. A partial alignment-compensating representation of the commonly used iris-code matrix was proposed by Rathgeb and Busch [20]. However, published rotation-invariant feature representations either require a more complex comparison process or reveal unpractical biometric performance. In the latter case, these may still be applied in a pre-selection step of a biometric identification scenario, see *e.g*. work of Konrad *et al*. [17]. Recently, Rathgeb *et al*. [21] introduced a method based on an analysis of the nature of iris-code and comparison scores between those. In a two-step process, the number of relative positions that need to be considered for two biometric samples was significantly decreased.

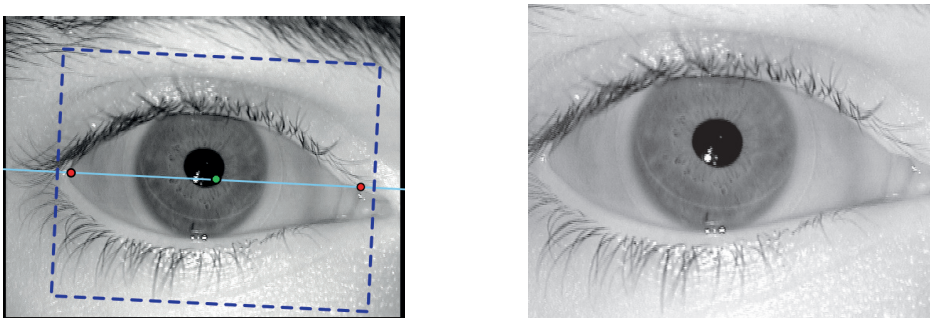
In this paper, two methods for eye corner detection in near-infrared eye images are presented. Before describing these, a brief outline of how the eye corners are used to pre-align an eye image is given below. Based on the two – left (L) and right (R) – eye corner points, the angle of a line through the two points is calculated as shown in equation (12.1).

$$\angle(L,R) = \arctan\left(\frac{R_y - L_y}{R_x - L_x}\right). \tag{12.1}$$

The image is rotated by the given angle, such that a line drawn between L and R is horizontal. The image is subsequently cropped in order to remove boundary artefacts resulting from the rotation. Those artefacts generate strong edges, which might negatively influence the segmentation process. The center of rotation C, which serves as the center of the cropped area, is based on the corner points as well:

$$C_x = \frac{L_x + R_x}{2},\ C_y = \frac{L_y + R_y}{2}. \tag{12.2}$$

The size of the cropped image is set to 512×400 pixels. Figure 12.1 shows the eye corner landmarks, the line between the landmarks and the framing of the resulting cropping and rotation. As can be seen in the image, the inner eye corner is hard to define due to missing color information.



(a) (b)

Figure 12.1: Iris image with eye corner landmarks (red), the rotation center (green), the horizon line and the frame for cropping and rotation (a) as well as the resulting image (b).

This method of absolute image pre-alignment is used with the eye corner locations produced with the methods outlined in following subsections. Note, that the aim is to align iris images prior to the segmentation stage. Alternatively, eye corners could be detected as part of the segmentation process, which might allow for an application of geometrical constraints, *e.g.* based on the detected pupil center.

12.2.3 Adapting Facial Landmark Detectors (FaceLD)

There exist many facial landmark detectors, which are made available in open-source toolboxes, *e.g.* dlib [16] and Bob [3] with menpofit [11], which were used in experiments performed for this paper. Those frameworks include pre-trained machine learning models, which are capable of detecting a large number of specific landmarks on a human face, among which are the eye corners. Naturally, these systems require an entire or at least a large part of a face to be present in an image. The eye images captured for the iris recognition systems only include a small part of the periocular area or are cropped (those two image formats are standardised by ISO/IEC 19794-6 [13]). A surprisingly effective idea is to utilise a high quality, noiseless facial image and insert the eye images into it, as shown in figure 12.2, so that the left and right face halves together with the inserted eye images are mirror reflections of each other. As an optional post-processing step, a semi-transparent smoothing transformation can be applied along the borders of

the eye images. The two methods of inserting the eye image into the face image are referred to as *basic* (12.2a) and *smooth* (12.2b). The facial landmarks are then detected, as with processing a normal face image; the landmark positions from the left and right side of the face are averaged, expecting more robustness. The last step is to translate the eye corners positions from the coordinate system of the face image to that of the eye image and use them in the pre-alignment experiments.

(a) Eye image plainly inserted upon the facial image.

(b) A semi-transparent transform applied around the eye image edges.

Figure 12.2: Insertion of an iris image to a high resolution frontal face image.

12.2.4 Landmark Detection for Eye Images (EyeLD)

A logical next step is to train a model dedicated for eye images alone. The open-source dlib package [16] implements a landmarking model presented by Kazemi and Sullivan [15], which relies on an ensemble of randomized regression trees. For the training, a groundtruth of 9 landmarks marked by a single operator is used; it contains the eye corners themselves, pupil center and points along the lower and upper eyelid arches, as shown in figure 12.3. In the pre-alignment experiments, the detected eye corners are used directly or computed using the intersection between the polynomials or circles fitted (least-squares sense) to the eyelid landmarks.

12.3 Results

The evaluation of the proposed approaches is focused on following matters:

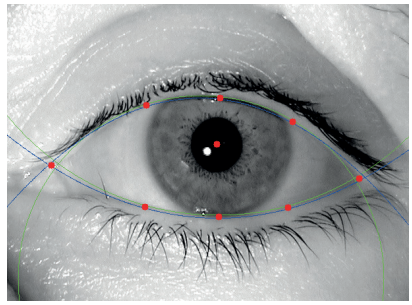


Figure 12.3: The 9 landmarks automatically detected by the model on a sample image. The curves show locating the eye corners by fitting circles (green) and polynomials (blue) to the eyelid landmarks.

Biometric performance By pre-aligning the images, the number of shifting positions considered at the comparison stage is changed. This obviously affects the biometric performance of the system, which is evaluated by calculating the equal-error rate (EER) and the false non-match rate measured at false match rate of 0.01% ($\text{FNMR}_{0.01}$). We are interested in the minimum EER found and also define diminishing returns (DR) of the EER, where we allow the EER to be up to 10% over the minimum EER. This usually results in a drastically reduced remaining rotation. The cause for this are the outliers, which allow to slightly improve the EER at a much higher cost of alignment compensation.

Workload The required alignment compensation ($\pm K$) after the pre-alignment step.

Pre-alignment accuracy How far are the results yielded by the pre-alignment step from the objectively optimal alignment.

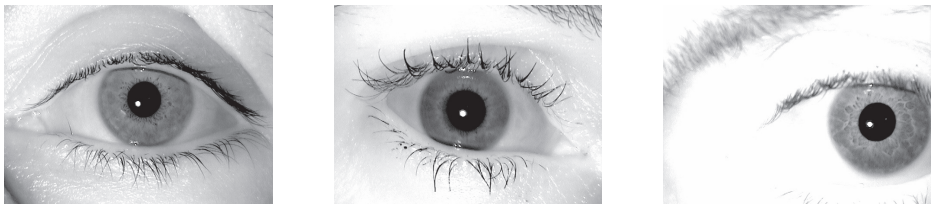


Figure 12.4: Example images from the BioSecure database.

The dataset chosen for the evaluation of the proposed approaches is the iris subset of the BioSecure database [19]. It contains 1680 left and right eye images from 210 subjects; the images of size 640×480 pixels were captured

using a near-infrared camera. Most of the publicly available iris datasets come in the cropped image format, which makes them unsuitable for our experiments; the images in the BioSecure dataset are uncropped. Additionally, the quality of images varies in terms of eye position, rotation and illumination conditions, as shown in figure 12.4. For the model training (see subsection 12.2.4), the dataset is divided into 5 subsets, each containing 1344 training images and 336 test images. This allows to generate landmarks for the whole dataset, while ensuring that the training and test sets are always disjoint.

In the employed iris recognition system, the iris of a given sample image is detected and transformed to a normalised rectangular texture of 512×64 pixels. The normalised iris texture is divided into texture stripes to obtain 10 one-dimensional signals, each one averaged from adjacent texture rows. A row-wise convolution with a Log-Gabor wavelet is performed on each signal and the two bits of phase information are used to generate a 512×20 bits iris-code. During alignment compensation, the rotation per bit corresponds to $\frac{360}{512} \approx 0.7°$. We have employed the algorithm that was made available in [23] and described in detail in [22]. For the biometric performance evaluation, all possible template comparisons are considered. This results in a total of 2520 genuine comparisons and almost 1.4 million impostor comparisons. It should, however, be noted, that the results presented in the following sections can be achieved irrespective of the chosen feature extraction algorithm.

12.3.1 Baseline and Groundtruth

First, in order to create a reference point for the proposed methods, baseline and groundtruth results are established. The *baseline* is a normal, iris-code based system, which performs $K = \pm 24$ bit shifts during a template comparison. The *groundtruth* consists of the manually marked landmark types shown in figure 12.3; the eye corners for pre-alignment calculations are used directly or computed as the intersection of polynomials or circles fitted to eyelid landmarks. Those results are listed in table 12.1.

Table 12.1: Baseline and groundtruth results (in %).

Method		**Minimum**			**DR**		
		K	EER	$FNMR_{0.01}$	K	EER	$FNMR_{0.01}$
Baseline		±20	2.506	4.296	±10	2.705	4.795
Groundtruth	Eye corners	±7	2.148	3.690	±5	2.336	4.288
	Polynomial fitting	±5	2.188	3.496	±3	2.347	4.171
	Circle fitting	±15	2.347	3.699	±5	2.506	4.178

When benchmarked against the baseline system, the proposed pre-alignment technique allows to significantly reduce the required remaining alignment compensation (K). This verifies the conceptual soundness of the approach

with manually marked landmark points. The diminishing returns (DR) allows for a trade-off between biometric performance improvement and workload reduction. As can be seen in table 12.1, the diminishing returns EER for the groundtruth is the same (or better) as for the baseline with full alignment compensation (minimum EER). In other words, by pre-aligning the samples, the required workload can be dramatically decreased without negatively affecting the biometric performance of the system.

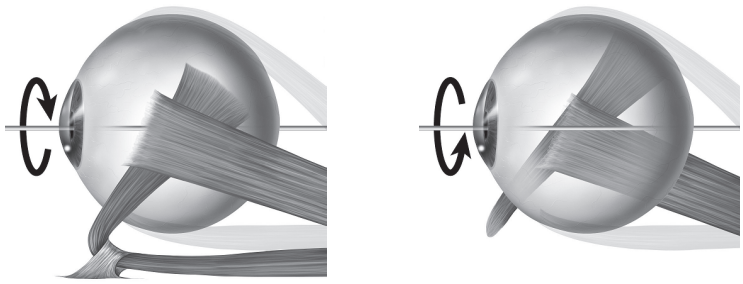


Figure 12.5: Eye with muscles responsible for torsional movement in the eye socket highlighted. Images by Patrick J. Lynch, medical illustrator (CC BY 2.5).

It is also important to address, why the pre-alignment does not fully eliminate the need for further alignment compensation at the iris-code template comparison stage, *i.e.* why $K \neq 0$. This remaining rotation of up to $\pm$7Bit $\approx \pm 4.92°$ is to be expected, since landmarks from the periocular region and not from the eye itself are used. The eye can rotate in the eye socket; this includes torsional movement induced by the superior/inferior rectus and superior/inferior oblique muscles [25] (see figure 12.5), with a range of motion that is "generally limited to angles of less than 10°" [30]. In recent years, methods for eye alignment during refractive surgery have been developed [10]. While extremely accurate, they depend on either continuous, active tracking (video) or static tracking based on a set of points marked in a reference image. The methods presented in this paper, however, perform the pre-alignment based on a single sample image.

12.3.2 Algorithmic Landmark Detection

Figure 12.6 shows various example images with landmarks detected by the proposed approaches marked. The results for the biometric performance and workload evaluation of the two approaches (subsections 12.2.3 and 12.2.4) are shown in table 12.2 and figure 12.7. Of interest are the benchmark against the baseline, *i.e.* by how much K decreased in an automated setting and the benchmark against the groundtruth (especially in case of EyeLD), *i.e.* by how much the automated approaches could still be improved.

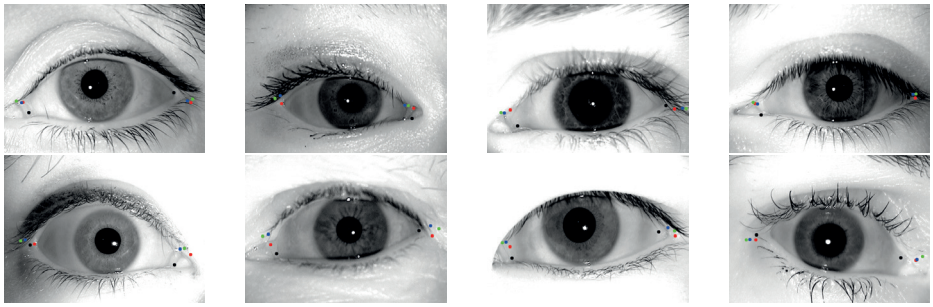
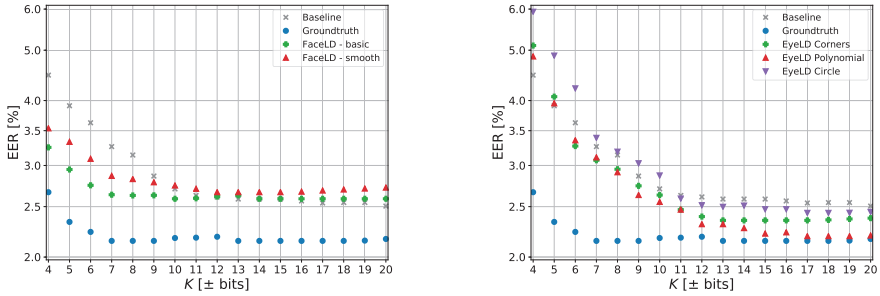


Figure 12.6: Example images with landmarks detected by the proposed approaches: FaceLD - basic (black), EyeLD - Corners (red), EyeLD - Polynomial (green), EyeLD - Circle (blue).



(a) Adaptation of facial landmarks detection.

(b) Dedicated eye landmarks detection.

Figure 12.7: Biometric performance comparison for the evaluated approaches (note the logarithmic scale of the y-axis).

Table 12.2: Algorithmic results (in %).

Method		Minimum			DR		
		K	EER	$\text{FNMR}_{0.01}$	K	EER	$\text{FNMR}_{0.01}$
FaceLD	Basic	±10	2.589	3.961	±6	2.748	4.625
	Smooth	±12	2.665	4.280	±7	2.864	4.654
EyeLD	Eye corners	±13	2.352	4.027	±11	2.467	4.142
	Polynomial fitting	±17	2.193	3.558	±12	2.313	3.868
	Circle fitting	±23	2.396	3.836	±11	2.592	4.214

For both approach classes, we observe an improvement over the baseline in both the minimum and diminishing returns EER setting. In all cases (except for circle fitting), K is significantly reduced (up to being halved), while the biometric performance in terms of EER and $\text{FNMR}_{0.01}$ remains unchanged or is improved. The FaceLD approach based on the Bob frame-

work [3] and menpofit [11] model performs well. The model offered by the dlib [16] package was also tried, but was left out due to very poor results. The smoothing transform around the eye image edges in FaceLD approach performs worse than the basic version of FaceLD approach. This could be due to the eye corners being blurred out when they are located near the border of the eye image. Thus, potentially a more sophisticated approach would have to be applied. In terms of alignment workload reduction, the FaceLD approaches outperform the EyeLD approaches. On the other hand, the biometric performance of EyeLD approaches is better than both the FaceLD approaches and the baseline. It is also worth noting, that while in terms of K reduction, the EyeLD approaches do not match the results achieved by the groundtruth, one could safely assume that with a large enough training corpus, the results of the groundtruth and the automated method would converge. One idea for future work is to mirror and rotate the available images in each training set fold, thus dramatically enlarging the training set and thereby the landmark detection accuracy.

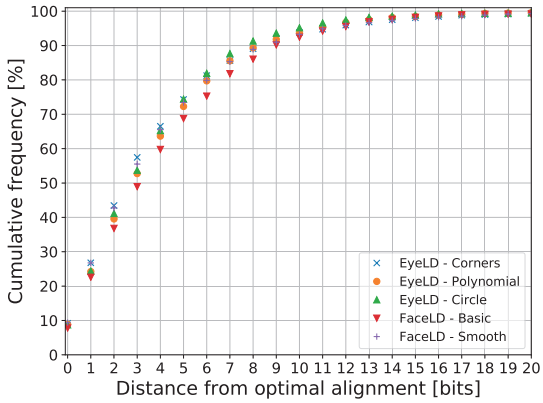


Figure 12.8: Cumulative distributions of the distance from the optimal alignment achieved by the presented pre-alignment approaches.

Table 12.3: Parameters of the impostor score distributions.

K	0	1	2	3	4	8	16	24
mean	0.498	0.495	0.492	0.489	0.486	0.478	0.469	0.466
st. deviation	0.024	0.023	0.023	0.022	0.021	0.018	0.016	0.014
skewness	-0.026	-0.034	-0.052	-0.083	-0.127	-0.351	-0.598	-0.717
ex. kurtosis	0.291	0.378	0.466	0.550	0.650	1.202	2.225	2.842

Figure 12.8 shows a cumulative distribution of the distance from optimal alignment after the pre-alignment step performed by the proposed ap-

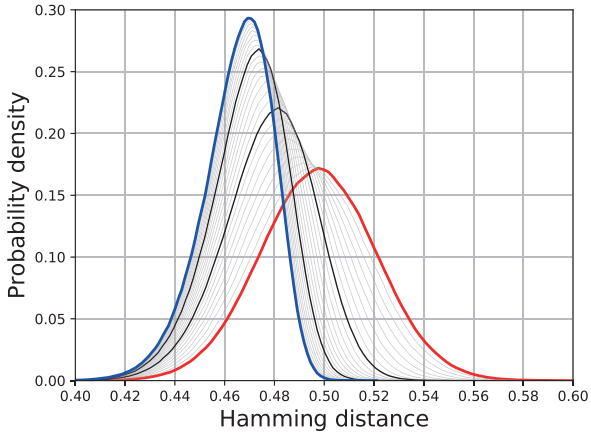


Figure 12.9: Kernel density estimate of impostor scores from no (red) to $K = \pm 24$ bits (blue) rotation compensation.

proaches. For example, it can be seen, that after a pre-alignment step around 80% of the images are less than 8 bits from the optimal alignment. Note, that this figure does not necessarily reflect the EER scores, since optimal alignment is not a guarantee for an optimal score (bad quality images can have a high HD score even at the optimal alignment position). In other words, the distance from the optimal alignment would only be a good predictor of biometric performance, if and only if the quality of images (apart from rotational variation) was very high.

While the pre-alignment is not expected to have a significant positive impact on the genuine scores, it affects the impostor scores significantly. As can be seen in figure 12.9 and table 12.3, when no alignment compensation is applied (*i.e.* $K = 0$), the impostor scores approximate a normal distribution around $HD = 0.5$. However, with the growing K value, the distribution moves towards left (*i.e.* towards genuine scores distribution). As has been mentioned in section 12.1, this increases the probability of false positives due to larger overlap between the genuine and impostor distributions. By pre-aligning and decreasing K, this effect is counteracted, thereby slightly improving the biometric performance in addition to reducing the workload.

12.4 Conclusion

In this paper, a software-based approach to alignment cost reduction in iris recognition systems has been introduced. Experiments conducted on the iris subset of the BioSecure database have lead to following key findings:

- Pre-alignment improves the biometric performance in terms of EER

and $FNMR_{0.01}$ when benchmarked against a baseline system.

- Pre-alignment reduces the required alignment compensation workload in terms of K when benchmarked against a baseline system.
- Proposed landmark detection approaches work, but as the groundtruth experiments demonstrate, there is still room for improvement.

While there exists a number of approaches for automatic corner detection in visible spectrum images, the authors are not aware of such work in the near-infrared spectrum. In this paper, two methods for achieving this task were presented with the resulting landmarks used for image pre-alignment. In addition to eye landmark detection accuracy refinement, a potentially interesting area for future work is investigating the possibility of application of the presented approaches to cropped eye images, as defined in ISO/IEC 19794-6 [13]. It is worth noting, that many areas of biometric research could benefit from iris image pre-alignment. This pertains in particular to privacy-enhancing technologies, *i.e*. biometric template protection [18], in which comparisons are performed in an encrypted domain, such that a proper alignment is (in many cases) not feasible.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF), as well as by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP) and by the Austrian Science Fund (FWF), project no. P27776.

12.5 Bibliography

[1] AL-RAISI, A. N., AND AL-KHOURI, A. M. Iris recognition and the challenge of homeland and border control security in UAE. *Telematics and Informatics 25*, 2 (May 2008), 117–132.

[2] ALONSO-FERNANDEZ, F., TOME-GONZALEZ, P., RUIZ-ALBACETE, V., AND ORTEGA-GARCIA, J. Iris recognition based on SIFT features. In *International Conference on Biometrics, Identity and Security (BIdS)* (September 2009), IEEE, pp. 1–8.

[3] ANJOS, A., SHAFEY, L. E., WALLACE, R., GÜNTHER, M., MCCOOL, C., AND MARCEL, S. Bob: a free signal processing and machine learning toolbox for researchers. In *Conference on Multimedia Systems (ACMMM)* (October 2012), ACM, pp. 1449–1452.

[4] CAO, X., WEI, Y., WEN, F., AND SUN, J. Face alignment by explicit shape regression. *Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2014), 2887–2894.

[5] DAUGMAN, J. How iris recognition works. *Transactions on Circuits and Systems for Video Technology (TCSVT) 14*, 1 (January 2004), 21–30.

[6] DAUGMAN, J. Information theory and the IrisCode. *Transactions on Information Forensics and Security 11*, 2 (February 2016), 400–409.

[7] DU, Y., IVES, R. W., ETTER, D. M., AND WELCH, T. B. Use of one-dimensional iris signatures to rank iris pattern similarities. *Optical Engineering 45*, 3 (March 2006), 1–10.

[8] ERDOGMUS, N., AND DUGELAY, J.-L. An efficient iris and eye corners extraction method. In *Structural, Syntactic, and Statistical Pattern Recognition: Joint IAPR International Workshop* (August 2010), Springer, pp. 549–558.

[9] HAO, F., DAUGMAN, J., AND ZIELINSKI, P. A fast search algorithm for a large fuzzy database. *Transactions on Information Forensics and Security (TIFS) 3*, 2 (June 2008), 203–212.

[10] HOLMQVIST, K., NYSTRÖM, M., ANDERSSON, R., DEWHURST, R., JARODZKA, H., AND DE WEIJER, J. V. *Eye tracking: A comprehensive guide to methods and measures*, 1 ed. Oxford University Press, September 2011.

[11] I MEDINA, J. A., ANTONAKOS, E., BOOTH, J., SNAPE, P., AND ZAFEIRIOU, S. Menpo: A comprehensive platform for parametric image alignment and visual deformable models. In *Proceedings of the 22nd ACM International Conference on Multimedia* (November 2014), MM '14, ACM, pp. 679–682.

[12] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 19794-5:2011. Information technology – Biometric data interchange formats – Part 5: Face image data*. International Organization for Standardization and International Electrotechnical Committee, October 2011.

[13] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 19794-6:2011. Information technology – Biometric data interchange formats – Part 6: Iris image data*. International Organization for Standardization and International Electrotechnical Committee, October 2011.

[14] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 29794-6:2015. Information technology – Biometric sample quality – Part 6: Iris image data*. International Organization for Standardization and International Electrotechnical Committee, July 2015.

[15] KAZEMI, V., AND SULLIVAN, J. One millisecond face alignment with an ensemble of regression trees. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2014), IEEE, pp. 1867–1874.

[16] KING, D. E. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research (JMLR) 10* (2009), 1755–1758.

[17] KONRAD, M., STÖGNER, H., UHL, A., AND WILD, P. Computationally efficient serial combination of rotation-invariant and rotation compensating iris recognition algorithms. In *International Conference on Computer Vision Theory and Applications (VISAPP)* (May 2010), SciTePress, pp. 85–90.

[18] NANDAKUMAR, K., AND JAIN, A. K. Biometric template protection: Bridging the performance gap between theory and practice. *Signal Processing Magazine 32*, 5 (September 2015), 88–100.

[19] ORTEGA-GARCIA, J., FIERREZ, J., ALONSO-FERNANDEZ, F., GALBALLY, J., FREIRE, M. R., ET AL. The multiscenario multienvironment BioSecure multimodal database (BMDB). *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 32*, 6 (June 2010), 1097–1111.

[20] RATHGEB, C., AND BUSCH, C. Comparing binary iris biometric templates based on counting Bloom filters. In *Iberoamerican Congress on Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications (CIARP)* (November 2013), Springer, pp. 262–269.

[21] RATHGEB, C., HOFBAUER, H., UHL, A., AND BUSCH, C. TripleA: Accelerated accuracy-preserving alignment for iris-codes. In *International Conference on Biometrics (ICB)* (June 2016), IEEE, pp. 1–8.

[22] RATHGEB, C., UHL, A., AND WILD, P. *Iris Recognition: From Segmentation to Template Security*, 1 ed., vol. 59 of *Advances in Information Security*. Springer, December 2013.

[23] RATHGEB, C., UHL, A., WILD, P., AND HOFBAUER, H. Design decisions for an iris recognition SDK. In *Handbook of Iris Recognition*, K. Bowyer and M. J. Burge, Eds., 2 ed., Advances in Computer Vision and Pattern Recognition. Springer, July 2016, pp. 359–396.

[24] SANTOS, G., AND PROENÇA, H. A robust eye-corner detection method for real-world data. In *International Joint Conference on Biometrics (IJCB)* (October 2011), IEEE, pp. 1–7.

[25] SPARKS, D. L. The brainstem control of saccadic eye movements. *Nature Reviews Neuroscience 3*, 12 (December 2002), 952–964.

[26] TAMS, B. Absolute fingerprint pre-alignment in minutiae-based cryptosystems. In *International Conference of the BIOSIG Special Interest Group (BIOSIG)* (September 2013), IEEE, pp. 1–12.

[27] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Aadhaar dashboard. `https://www.uidai.gov.in/aadhaar_dashboard/`. Last accessed: 2020–03–11.

[28] XIA, H., AND YAN, G. A novel method for eye corner detection based on weighted variance projection function. In *International Congress on Image and Signal Processing (ICSIP)* (October 2009), IEEE, pp. 1–4.

[29] XU, C., ZHENG, Y., AND WANG, Z. Semantic feature extraction for accurate eye corner detection. In *International Conference on Pattern Recognition (ICPR)* (December 2008), IEEE, pp. 1–4.

[30] YOUNG, L. R., AND SHEENA, D. Survey of eye movement recording methods. *Behavior Research Methods & Instrumentation 7*, 5 (September 1975), 397–429.

[31] ZHANG, Z., SHEN, Y., LIN, W., AND ZHOU, B. Eye corner detection with texture image fusion. In *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)* (December 2015), IEEE, pp. 992–995.

Chapter 13

Detection of Glasses in Near-infrared Ocular Images

Abstract

Eyeglasses change the appearance and visual perception of facial images. Moreover, under objective metrics, glasses generally deteriorate the sample quality of near-infrared ocular images and as a consequence can worsen the biometric performance of iris recognition systems. Automatic detection of glasses is therefore one of the prerequisites for a sufficient quality, interactive sample acquisition process in an automatic iris recognition system. In this paper, three approaches (i.e. a statistical method, a deep learning based method and an algorithmic method based on detection of edges and reflections) for automatic detection of glasses in near-infrared iris images are presented. Those approaches are evaluated using cross-validation on the CASIA-IrisV4-Thousand dataset, which contains 20000 images from 1000 subjects. Individually, they are capable of correctly classifying 95-98% of images, while a majority vote based fusion of the three approaches achieves a correct classification rate (CCR) of 99.54%.

Addressed research question(s): RQ2
Reference: DROZDOWSKI, P., STRUCK, F., RATHGEB, C., AND BUSCH, C. Detection of glasses in near-infrared ocular images. In *International Conference on Biometrics (ICB)* (February 2018), IEEE, pp. 202–208.

13.1 Introduction

In recent years, iris recognition has become a popular modality for biometric systems and is used in many large-scale deployments (e.g. the Indian National ID project [22]). The technology is also increasingly being used in automatic (without human operator supervision) systems, such as smart border/airport gates and mobile devices [14]. Operational systems typically capture iris images in the near-infrared light spectrum, in which the iris patterns are much more pronounced than in the visible light spectrum, even for darkly pigmented irides [6]. According to recent reports [20, 21], over 50%

of adult population in the developed world wear eyeglasses. The pervasiveness of short-sightedness (myopia) has been on an extreme rise in Eastern Asia and around the world in general; a recent report in Nature News [7] states:

> East Asia has been gripped by an unprecedented rise in myopia, also known as short-sightedness. Sixty years ago, 10-20% of the Chinese population was short-sighted. Today, up to 90% of teenagers and young adults are. In Seoul, a whopping 96.5% of 19-year-old men are short-sighted. Other parts of the world have also seen a dramatic increase in the condition, which now affects around half of young adults in the United States and Europe - double the prevalence of half a century ago. By some estimates, one-third of the world's population - 2.5 billion people - could be affected by short-sightedness by the end of this decade.

Due to specular reflections, blur, scratches and other factors, glasses tend to decrease the biometric sample quality and consequently often the biometric performance of the systems. While several researchers have investigated the impact of glasses on face recognition systems, the scientific literature on iris recognition contains very little related work on this subject, except for a paper in which a small-scale quantification of the effects of glasses on iris image pre-processing is presented [13] and glasses being mentioned as a significant noise factor (e.g. [1, 3, 9]). ISO/IEC 29794-6 biometric sample quality standard [10] specifically recommends to instruct data subjects to remove glasses during acquisition or to perform the acquisition with additional care.

Therefore, and due to the prevalence of glasses in the world population, automatic detection of glasses is an important matter in iris recognition (as will also be substantiated by the experiments described in section 13.3). It is of particular interest for automatic sample acquisition systems, where such a detection module would enable an interactive sample acquisition and thus facilitate higher sample quality. While this is a well-researched topic in systems working with images of the facial region (e.g. [2, 23]), doing so in images of ocular region alone has not received enough attention. In this paper, three methods for accomplishing said task are presented and benchmarked.

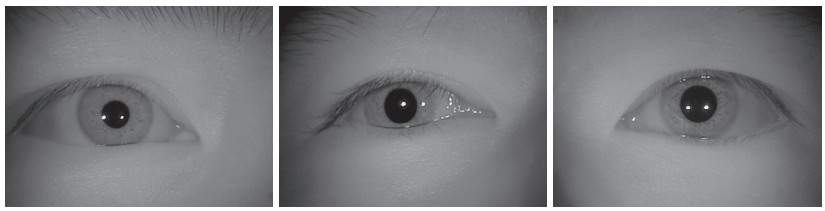
This paper is organised as follows: in section 13.2, the used dataset and experimental setup are described. Section 13.3 provides an overview of the impact of glasses on iris recognition. In section 13.4 the three proposed automatic glasses detection approaches are presented and evaluated. Concluding remarks are given in section 13.5.

13.2 Experimental Setup

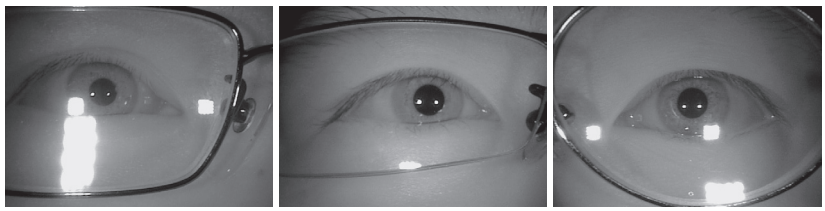
The Thousand subset of the CASIA-IrisV4 database [5] (henceforth referred to as "CASIA-Thousand dataset") was chosen for the experiments performed for this paper. Said dataset contains near-infrared iris images of size 640×480 pixels and, due to its size, is suitable for large-scale testing. Additionally, for subjects who are glass-wearers, it contains images both with and without glasses, thus enabling a direct biometric performance benchmark. Figure 13.1 shows example images from the dataset, while table 13.1 summarises its properties[1]. Observe the high fraction of subjects who are glass-wearers coinciding with the statistics mentioned in section 13.1. The groundtruth labels (with/without glasses) had to be assigned to all the images, which was done manually by a single researcher via visual inspection.

Table 13.1: Overview of the CASIA-Thousand dataset

	Samples	**Subjects**	**Instances**
Total	20000	1000	2000
Without glasses	14664	1000	2000
With glasses	5336	617	1193



(a) Without glasses



(b) With glasses

Figure 13.1: Example images from the CASIA-Thousand dataset. Samples (a) and (b) are captured from the same eye instance.

[1] Observe, that since for every subject/instance there is at least one sample without glasses in the dataset, the numbers for subjects/instances *seemingly* do not add up.

The images were processed with commonly used methods (specifically, Viterbi algorithm for segmentation [19], Daugman's rubber sheet model for normalisation, LogGabor wavelet for feature encoding and fractional Hamming distance for template comparison [6]) implemented by the open-source OSIRIS [15] and USIT [16] frameworks. Subsequently, two evaluations took place:

- The impact of glasses on sample quality (some metrics from ISO/IEC 29794-6 standard [10]) and thereby on iris recognition in terms of biometric performance measured in equal error rate (EER). (section 13.3)
- The classification accuracy of the proposed detection approaches using cross-validation over 4 folds (i.e. 15000 training and 5000 test images), measured in correct classification rate (CCR). (section 13.4)

13.3 Impact of Glasses on Iris Recognition

The topic of glasses in iris recognition systems has often been mentioned in the scientific literature (e.g. [1, 3, 9]) and presentations [18]. It is commonly agreed that they can have detrimental effect on sample quality due to specular reflections, dirt, optical distortions and shadows. A decrease in sample quality in turn negatively affects the segmentation accuracy and/or biometric performance. Furthermore, as shown in figure 13.2, they introduce potential for explicit failures, where the reflections or frame can be misunderstood as pupilliary or limbic boundaries by the segmentation algorithm (the red blobs in the images represent areas masked out by the algorithm as eyelids and noise). Those assertions notwithstanding, with an exception of a small investigation [13], studies quantifying the effects glasses have on the biometric performance of iris recognition systems are lacking in the scientific literature.

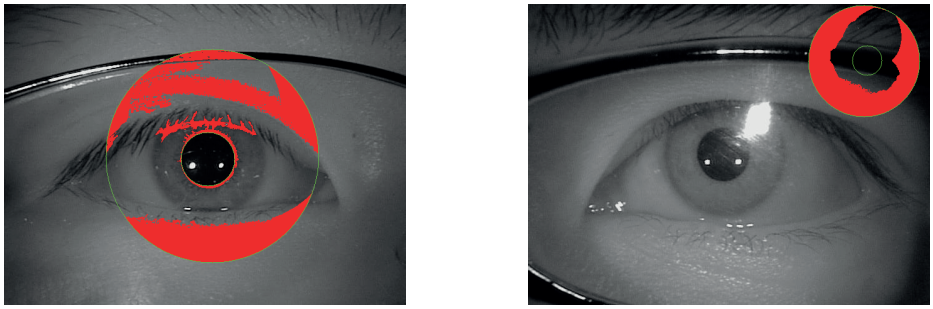


Figure 13.2: Segmentation failures caused by glasses

The results of a biometric verification experiment on the CASIA-Thousand dataset, shown in table 13.2, demonstrate the negative impact of glasses on

an iris recognition system. In addition to the data shown in the table, the motion blur in images with glasses was calculated to be twice as high as in images without glasses, which in turn can negatively affect other iris image quality metrics, such as the iris-pupil and iris-sclera contrast.

Table 13.2: Impact of glasses on iris recognition

Metric	Without glasses	With glasses
EER, all images	6.86%	12.16%
EER, no segmentation failures	3.79%	10.67%
Images with usable iris area $\geq 70\%$	59.19%	51.63%

The aforementioned issues are also mentioned in the ISO/IEC 29794-6 biometric sample quality standard [10], where it is recommended to perform data acquisition so that the specular reflections on the iris are minimised or even to instruct the data subject to remove their glasses. In some, particularly automatic systems, doing so would require automatically detecting the glasses. In the next section, methods of automatic detection of glasses in near-infrared iris images are described and evaluated.

13.4 Automatic Detection Approaches

As discussed earlier, automatic detection of glasses appears to be an overseen or underappreciated issue in the scientific literature. However, based on the sheer numbers of glass-wearers in the population (section 13.1) and the significant impact of glasses on the biometric performance (section 13.3), it is abundantly clear that methods for automatic glasses detection are beneficial for iris recognition systems. With it in place, such systems would be enabled to provide actionable feedback to the capture subject - meaning to ask the subject to take off the glasses and to subsequently initiate a recapture. In this section, three such approaches are presented and evaluated on near-infrared iris image data.

13.4.1 Texture Descriptor

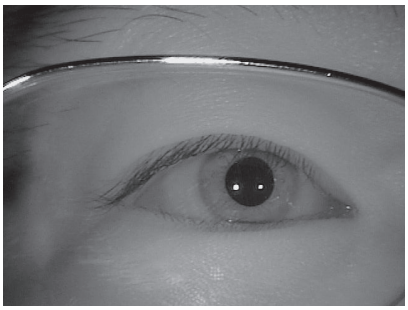
Binarized statistical image features (BSIF) [12] is a generic texture descriptor, which uses filters learned from patches of natural images. Pre-trained filters made available as part of the above publication are used. The process of using BSIF to detect glasses in iris images is as follows:

1. An input image (figure 13.3a), is convolved with 8 stacked linear filters of size 15×15 pixels; the sign of each filter response is used to binarise it (such that negative responses become 0 and positive responses become 1), resulting in a binary string of length 8 for each pixel of the

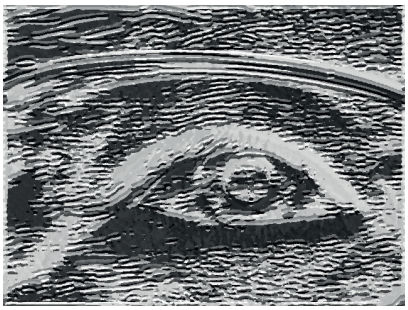
image. The integer representation of those binary strings lies in range (in range 0 to $2^8 - 1$), and can be thus displayed as a 256-bit grayscale image, as shown in figure 13.3b.

2. The aforementioned integer values for the whole image are stored in a histogram, as shown in figure 13.3c.

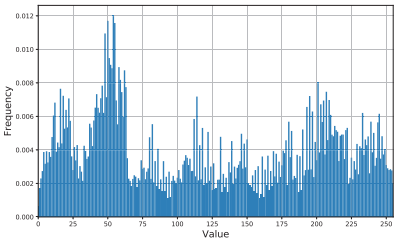
3. Using the previously (section 13.2) mentioned cross-validation loop for training and testing, the classification decision is obtained by passing the histogram as an input to a support vector machine (SVM). It uses a linear kernel, which is suitable for high-dimensional vectors. A lightweight implementation provided by the libsvm (version 3.22) [4] library was used; for training the SVM, 1000000 was used as cost parameter and 0.001 was used for termination tolerance. The parameters were estimated empirically on a small, disjoin training set.



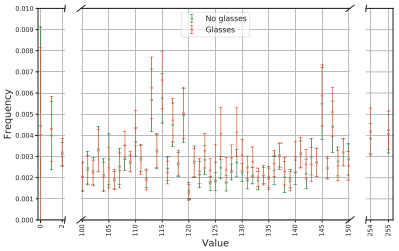
(a) Image



(b) BSIF applied



(c) BSIF grayscale values histogram for one image



(d) BSIF histogram value frequencies (mean and one standard deviation as errorbars) for the entire CASIA-Thousand dataset

Figure 13.3: BSIF-based approach

As shown in figure 13.3d, subtle differences in the BSIF-histogram values' frequency distribution are perceivable. The SVM is capable of using

those to distinguish between and correctly classify images with and without glasses. To facilitate reproducible research, the trained SVM is released publicly at [8].

13.4.2 Deep Learning

Deep neural networks convince by successful application with a huge variety of tasks [17], including classification specifically [24]. The problem of classifying images with and without glasses falls well within the areas in which deep neural networks are commonly applied.

Using the Caffe framework (version 1.0) [11], a deep convolutional neural network for classification of images has been created; its topology can be seen in table 13.3. The neural network is trained and tested using the previously (section 13.2) mentioned cross-validation loop. The images are resized to 320×240 pixels and the training is run over 20000 iterations, with batch size of 32 images and 20000 as step size. Using 15000 input images and 5 steps, the network was trained for about 213 epochs. The learning rate is set to 0.0001 and gets multiplied by 0.25 after every step.

Multiple other architectures, which differed mostly in the input dimensions and the size of the convolution layers, were tested. It turned out that input dimensions larger than 320×240 (e.g. 640×480) are not necessary to achieve good classification results. Thus, for computational performance reasons the relatively small network was chosen in order to attain an acceptable trade-off between classification accuracy and throughput. The dimensions of the convolution layers were determined by the size of potential feature blocks, which are effected by glasses being present in an image. To facilitate reproducible research, the trained network is released publicly at [8].

Table 13.3: Topology of the DNN-based approach

Part	Layer	Iterations	Details
Feature Extraction	Convolution	2	1. 17×17 pixels, 48 filters 2. 7×7 pixels, 96 filters
	ReLu		—
	Pooling		Max, 3×3 pixels, 1 filter, stride 2
Classification	Fully connected	2	96 neurons
	ReLu		—
	Dropout		—
	Fully connected	1	2 neurons
	Decision		linear classifier

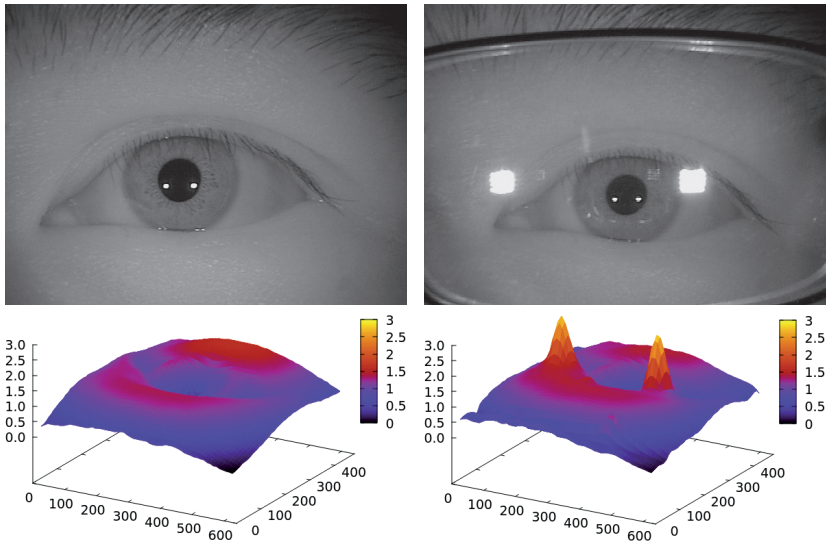
13.4.3 Edge and Reflection Detection based Algorithm

Two key differences between images with and without glasses are more pronounced specular reflections and stronger edges due to the frames of

glasses. The classification approach described in this subsection is based on detection and quantification of those image features.

13.4.3.1 Reflections

The image is divided into blocks of equal size (chosen empirically to be 30×30 pixels and the brightness of each block is computed relative to the brightness of the entire image, thus producing a map of relative brightness deviation. The block size filters out small, natural reflections (figure 13.4a), whereas large, artificial reflections are very well pronounced (figure 13.4b).



(a) Without glasses

(b) With glasses

Figure 13.4: Reflection detection with a relative brightness measure. The two specular reflections caused by the glasses are clearly observed by this proposed metric.

13.4.3.2 Edges

The process of detecting and measuring edges for glasses detection in an iris image is described below and shown in figure 13.5 and described below.

1. The image is convolved with a simple kernel which detects horizontal edges. This process is independent of the average brightness of the image, since only the local brightness gradients are computed. (figure 13.5b)

2. The grayscale image is transformed into a black and white image. This is done by applying a brightness threshold (usually between 128 and 129, estimated empirically on a small disjoint training set), which only accepts sharp brightness transitions and ignores blurred edges. (figure 13.5c)

3. Due to illumination artefacts or image compression many edges have small gaps. A dilation filter of size 7×7 pixels (estimated empirically on a small disjoint training set) is used to fill those gaps. (figure 13.5d)

4. The edges in the middle of the image are masked out, since they tend to be natural eye edges. (figure 13.5e)

5. To distinguish between individual edges, the flood fill algorithm with 8 directions is applied. This algorithm finds connected pixels and represents them with different colours. (figure 13.5f)

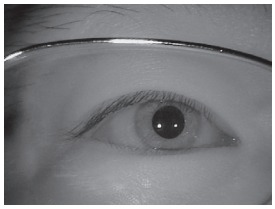
6. The width and height of the found edges is calculated using the leftmost and rightmost, and topmost and bottommost pixels. Very small edges are discarded (e.q. the small points on the right top corner in figure 13.5f) are discarded because they do not contain information. Subsequently, then the ratio between widths and heights of the remaining edges is computed. (figure 13.5g)

13.4.3.3 Classification

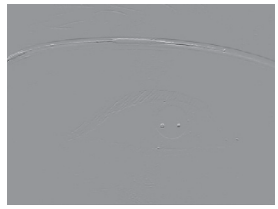
The reflection and edge detection methods described in subsections 13.4.3.1 and 13.4.3.2, respectively, are applied to an iris sample. Using the previously (section 13.2) mentioned cross-validation loop, tuples containing the values of largest relative brightness block and the edge with the highest width-to-height ratio are passed to a SVM, which performs the classification decisions. A radial basis function (RBF) kernel was chosen, since it is well suited for low-dimensional vectors. For training the SVM, 10000 was used as cost parameter, which was estimated using a small, disjoint training set. As shown in figure 13.6, the two metrics (reflection and edge scores) contain sufficient discriminative power to distinguish quite accurately between images with and without glasses, albeit some overlap (and thereby classification errors) is still present.

13.4.4 Results

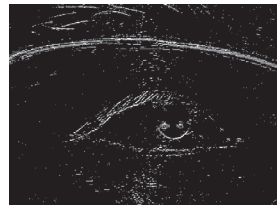
The classification accuracy of the proposed methods is estimated by performing cross-validation over 4 folds. The results are shown in table 13.4. All three proposed approaches perform well, with overall accuracy ranging between 95-99%. Notice, however, that the CCR for images with and



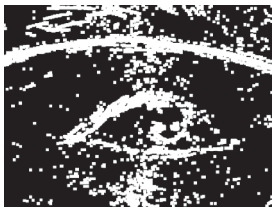
(a) Image



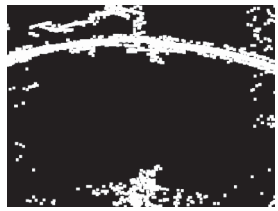
(b) Detection



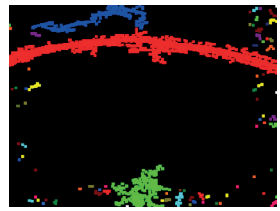
(c) Binarisation



(d) Dilation



(e) Outmasking



(f) Assignment

Edge colour	Width	Height	Ratio
Red	639	111	5.76
Blue	279	66	4.23
Green	142	100	1.42
⋮	⋮	⋮	⋮

(g) Measurement

Figure 13.5: Edge detection and measurement

without glasses vary - for instance, the neural network classifies more images without glasses correctly, whereas the statistical approach does so for images with glasses. This suggests a possibility of fusing the decisions of the approaches, so that their individual weaknesses are compensated for. Performing a majority vote of all three approaches was able to significantly increase the CCR. A conjunction based fusion of all three or different configurations of two approaches was also tried, but was found to be less successful than the majority vote (albeit still improving upon the accuracy of the individual approaches).

13.4.5 Classification Errors

It is of interest to investigate what types of images were incorrectly classified by the proposed approaches. Figure 13.7 shows such example images and corresponding error reasons. With a larger dataset and hence more training data, the classification errors could potentially be further reduced.

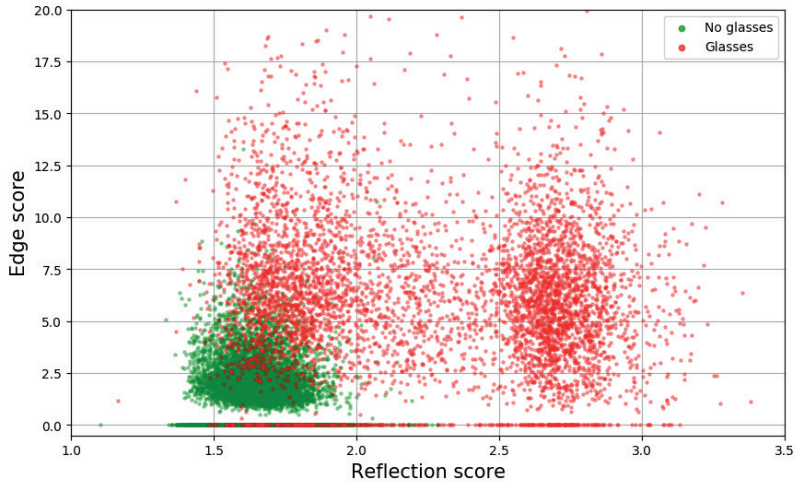


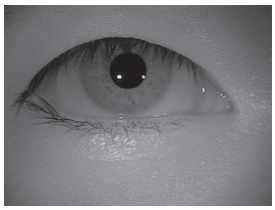
Figure 13.6: A scatter plot of edge and reflection scores for all images from the CASIA-Thousand dataset, which shows significant separation between the two image classes

Table 13.4: Results of the evaluation (with 95% CI)

Approach	CCR (in%)		
	Without Glasses	With glasses	Overall
Texture Descriptor (13.4.1)	97.79 ± 0.95	98.54 ± 0.69	98.08 ± 0.44
Deep Learning (13.4.2)	99.28 ± 0.22	97.33 ± 1.60	98.97 ± 0.29
Edges and Reflections (13.4.3)	97.18 ± 0.38	92.37 ± 2.23	95.43 ± 0.36
Majority vote	99.72 ± 0.08	98.79 ± 0.66	99.54 ± 0.12

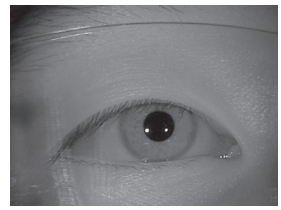
13.5 Conclusion

Glasses make iris recognition more challenging, since they can have a detrimental effect on sample quality and thereby biometric performance of a system. In section 13.3, it has been shown that on the CASIA-Thousand dataset, the equal error rate on the subset of images with glasses is twice that of the subset of images without glasses. It is therefore of interest to automatically detect glasses in iris images in order to handle such images separately or reacquire once the data subject has been asked to remove their glasses. In this paper, three approaches for automatically detecting glasses in near-infrared ocular images have been presented. They achieve classification accuracy in range of 95-98%, which can be further improved on by a decision-level fusion. A majority vote of all three approaches achieved an overall 99.54% correct classification rate, whereas slightly lower (but still above 99%) correct classification rate was achieved with an conjunction-based fusion of two approaches. In contrast to other approaches for glasses detection, the pro-

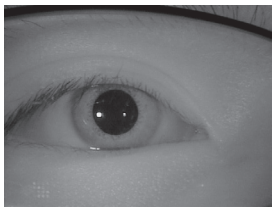


(a) Strong eye edge

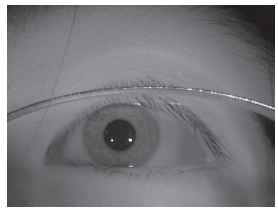
(b) Strong make-up



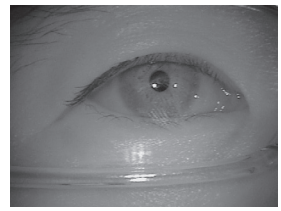
(c) Thin/no frame



(d) Frame at upper margin



(e) Frame position



(f) Transparent frame

Figure 13.7: Examples of incorrectly classified images from all 3 methods. Figures (a)-(b) falsely classified as glasses, figures (c)-(f) falsely classified as non-glasses.

posed methods require only a single-frame image and work with the ocular area alone instead of whole face. They could be seamlessly integrated into operational automatic systems, for instance to facilitate interactive image acquisition, where the data subjects would be required to take off glasses if detected. Furthermore, such systems often capture images of both eyes simultaneously, thus the accuracy of glasses detection could be further improved by performing a multi-instance fusion, i.e. a conjunction of the decisions from both eyes.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF) as well as by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP).

13.6 Bibliography

[1] BHARADWAJ, S., BHATT, H. S., VATSA, M., AND SINGH, R. Periocular biometrics: When iris recognition fails. In *International Conference on*

Biometrics: Theory, Applications and Systems (BTAS) (September 2010), IEEE, pp. 1–6.

[2] BORZA, D., DARABANT, A. S., AND DANESCU, R. Eyeglasses lens contour extraction from facial images using an efficient shape description. *Sensors 13*, 10 (October 2013), 13638–13658.

[3] BOWYER, K. W., HOLLINGSWORTH, K., AND FLYNN, P. J. Image understanding for iris biometrics: A survey. *Computer Vision and Image Understanding 110*, 2 (May 2007), 281–307.

[4] CHANG, C., AND LIN, C. LIBSVM: A library for support vector machines. *Transactions on Intelligent Systems and Technology (TIST) 2* (May 2011), 27:1–27:27. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[5] CHINESE ACADEMY OF SCIENCES' INSTITUTE OF AUTOMATION. CASIA iris image database. `http://biometrics.idealtest.org/`, December 2010. Last accessed: 2020–03–11.

[6] DAUGMAN, J. How iris recognition works. *Transactions on Circuits and Systems for Video Technology (TCSVT) 14*, 1 (January 2004), 21–30.

[7] DOLGIN, E. The myopia boom. *Nature 519*, 7543 (March 2015), 276–278.

[8] DROZDOWSKI, P., STRUCK, F., RATHGEB, C., AND BUSCH, C. Glasses detection models. `https://github.com/dasec/glasses-detection-models`, December 2017.

[9] H.PROENÇA, AND ALEXANDRE, L. A. Iris recognition: Analysis of the error rates regarding the accuracy of the segmentation stage. *Image and vision computing 28*, 1 (January 2010), 202–206.

[10] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 29794-6:2015. Information technology – Biometric sample quality – Part 6: Iris image data*. International Organization for Standardization and International Electrotechnical Committee, July 2015.

[11] JIA, Y., SHELHAMER, E., DONAHUE, J., KARAYEV, S., LONG, J., ET AL. Caffe: Convolutional architecture for fast feature embedding. In *International Conference on Multimedia* (November 2014), ACM, pp. 675–678.

[12] KANNALA, J., AND RAHTU, E. BSIF: Binarized statistical image features. In *International Conference on Pattern Recognition (ICPR)* (November 2012), IEEE, pp. 1363–1366.

[13] LIM, S., LEE, K., BYEON, O., AND KIM, T. Efficient iris recognition through improvement of feature vector and classifier. *ETRI Journal 23*, 2 (June 2001), 61–70.

[14] NIGAM, I., VATSA, M., AND SINGH, R. Ocular biometrics: A survey of modalities and fusion approaches. *Information Fusion 26* (November 2015), 1–35.

[15] OTHMAN, N., DORIZZI, B., AND GARCIA-SALICETTI, S. OSIRIS: An open source iris recognition software. *Pattern Recognition Letters 82*, 2 (September 2016), 124–131.

[16] RATHGEB, C., UHL, A., WILD, P., AND HOFBAUER, H. Design decisions for an iris recognition SDK. In *Handbook of Iris Recognition*, K. Bowyer and M. J. Burge, Eds., 2 ed., Advances in Computer Vision and Pattern Recognition. Springer, July 2016, pp. 359–396.

[17] SCHMIDHUBER, J. Deep learning in neural networks: An overview. *Neural networks 61* (January 2015), 85–117.

[18] SUN, Z. Lecture: Recent progress of iris recognition. `http://www.comp.hkbu.edu.hk/wsb17/lecturer_details.php?lect_id=8`, January 2017. Last accessed: 2020–03–11.

[19] SUTRA, G., GARCIA-SALICETTI, S., AND DORIZZI, B. The Viterbi algorithm at different resolutions for enhanced iris segmentation. In *International Conference on Biometrics (ICB)* (March 2012), IEEE, pp. 310–316.

[20] THE EUROPEAN COUNCIL OF OPTOMETRY AND OPTICS. ECOO Blue Book: Data on optometry and optics in Europe. `https://www.ecoo.info/2017/05/08/ecoo-publishes-blue-book-2017/`, May 2017.

[21] THE VISION COUNCIL. VisionWatch: The Vision Council member benefit report, September 2016.

[22] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Aadhaar dashboard. `https://www.uidai.gov.in/aadhaar_dashboard/`. Last accessed: 2020–03–11.

[23] WU, C., LIU, C., SHUM, H., XY, Y., AND ZHANG, Z. Automatic eyeglasses removal from face images. *Transactions on pattern analysis and machine intelligence (TPAMI) 26*, 3 (March 2004), 322–336.

[24] ZHANG, G. P. Neural networks for classification: A survey. *Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 30*, 4 (November 2000), 451–462.

Chapter 14

SIC-Gen: A Synthetic Iris-Code Generator

Abstract

Nowadays large-scale identity management systems enrol more than one billion data subjects. In order to limit transaction times, biometric indexing is a suitable method to reduce the search space in biometric identifications. Effective testing of such biometric identification systems and biometric indexing approaches requires large datasets of biometric data. Currently, the size of the publicly available iris datasets is insufficient, especially for system scalability assessments. Synthetic data generation offers a potential solution to this issue; however, it is challenging to generate data that is both statistically sound and visually realistic - for the iris, the currently available approaches prove unsatisfactory.

In this paper, we present a method for generation of synthetic binary iris-based templates, *i.e.* Iris-Codes, which are the *de facto* standard used throughout major biometric deployments around the world. We validate the statistical properties of the synthetic templates and show that they closely resemble ones produced from real ocular images. With the proposed approach, large databases of synthetic Iris-Codes with flexibly adjustable properties can be generated.

Addressed research question(s): RQ2
Reference: DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. SIC-Gen: A synthetic Iris-Code generator. In *International Conference of the Biometrics Special Interest Group (BIOSIG)* (September 2017), IEEE, pp. 61–69.

14.1 Introduction

The iris is one of the most widely applied biometric modalities. In recent years, several large-scale deployments have been created, most notably the Indian National ID program [14], which has, at the time of this writing, enrolled over one billion subjects with biometric data including the irides. Despite using efficient comparators (e.g. Hamming distance for the iris) and parallelism, the computational load faced by such deployments in the identification scenario is extremely high. With biometric workload reduction as

a motivation, many approaches for indexing of iris data have been developed [11]. However, evaluation of such approaches and their scalability is often questionable due to lack of large test datasets. While various publicly available iris databases with near-infrared (NIR) data exist, they are relatively small. At the time of this writing, some of the largest publicly available datasets, CASIA-IrisV4-Thousand and ND-CrossSensor-Iris-2013, contain merely 20.000 images from 1000 subjects and 146.550 images from 676 subjects, respectively. This is several orders of magnitude smaller than some of the large-scale deployments nowadays.

Synthetic data generation is one possible way of dealing with the issue of testing efficient indexing methods. Most of the existing approaches for synthetic iris generation attempt to synthesise an entire iris image or texture [1, 7, 8, 12, 13, 15, 16, 17, 18]. The main issues with such approaches include the computational costs and the difficulty in guaranteeing the statistical properties of the real data. The vast majority of operational iris biometric systems are based on the Iris-Code [2], making it a *de facto* standard. Generating Iris-Codes (feature vectors) directly is therefore also viable and may offer better control over the statistical properties of the synthetic data. Recently, two such approaches have been proposed. Proença and Neves [12] provide a method of Iris-Code synthesis based on bit correlations; the method is shown to attain some of the desired statistical properties (the shapes of the genuine and impostor distributions). It is also somewhat flexible with adjustable parameters; however, it does not allow to generate a set of templates following a desired score distribution. Furthermore, the filter response resulting from the typical feature extraction process is not modelled (in other words, the produced synthetic Iris-Codes scantily resemble the ones produced from real iris images through the commonly used iris processing pipeline). Lastly, typical error patterns between two mated templates are not modelled. Daugman [3] proposed to use a simple hidden Markov model to generate a stream of bits and showed that it can be adjusted, so that the produced templates mimic the impostor distribution of real iris templates. However, the produced streams are 1-dimensional (*i.e.* do not model the correlation between the Iris-Code rows); furthermore, the method does not offer a way to generate more than one template per subject (*i.e.* it is not possible to use it for simulating genuine comparisons). As such, it might only be useful for stress-testing of iris identification systems.

In this paper, we present a synthetic Iris-Code generator, which both reflects the statistical properties of the real Iris-Codes and resembles the real templates visually. An important feature of the proposed approach is its flexibility, in that it allows to generate Iris-Codes with an arbitrary resolution and an arbitrary score distribution of mated templates, unlike any of the approaches currently in the literature. To facilitate reproducible research, the software written in Python3 programming language, is released to the

scientific community under a permissive license.

The remainder of this paper is organised as follows: section 14.2 describes the proposed method of synthetic Iris-Code generation. In section 14.3 the properties of the generated templates are validated, while section 14.4 contains concluding remarks.

14.2 Proposed Method

When generating synthetic Iris-Codes, several matters have to be taken into account:

- Dataset

  Score distributions The distributions of Hamming distance scores must closely resemble the ones produced by real data.

  Degrees of freedom Based on a large number of comparison scores from non-mated templates, the effective number of independent bits (degrees of freedom) can be calculated. Degrees of freedom can be seen as discrimination entropy as a measure of information content in iris images and has to be close to that of the real data.

- Individual templates

  Bit correlation The bits in an Iris-Code are far from independent. There exist correlations between both rows and columns, which result in long sequences of identical consecutive bits. The reason for this is partially the anatomy of iris patterns, as well as the nature of the commonly used feature extractors [3]. Those correlations have to be reflected in the synthetic data.

  Error patterns The majority of bit mismatches between two mated Iris-Codes occurs for bits resulting from wavelet response close to 0 (*i.e.* where the response phase changes). Those occur mostly on the edges of the bit sequences, and are called the ”fragile” bits [5]. They have to be present in the synthetic data. Additional noise sources, such as the occlusions resulting from the eyelids, have to be modelled as well.

  Rotation In the real data, rotations of the eye, which are mainly caused by head tilts (*i.e.* roll pose), potentially result in misalignment between two mated samples. In Iris-Codes, this is represented by circular horizontal shifts of the matrix columns, which have to be modelled in the synthetic data.

The proposed generator synthesises Iris-Codes as pairs of mated templates, referred to as Iris-Codes $IC1$ and $IC2$ in the algorithm description and figure 14.1 below. The bold-filled arrows denote the changes to the template throughout the process, while the thin arrows denote the system parameters.

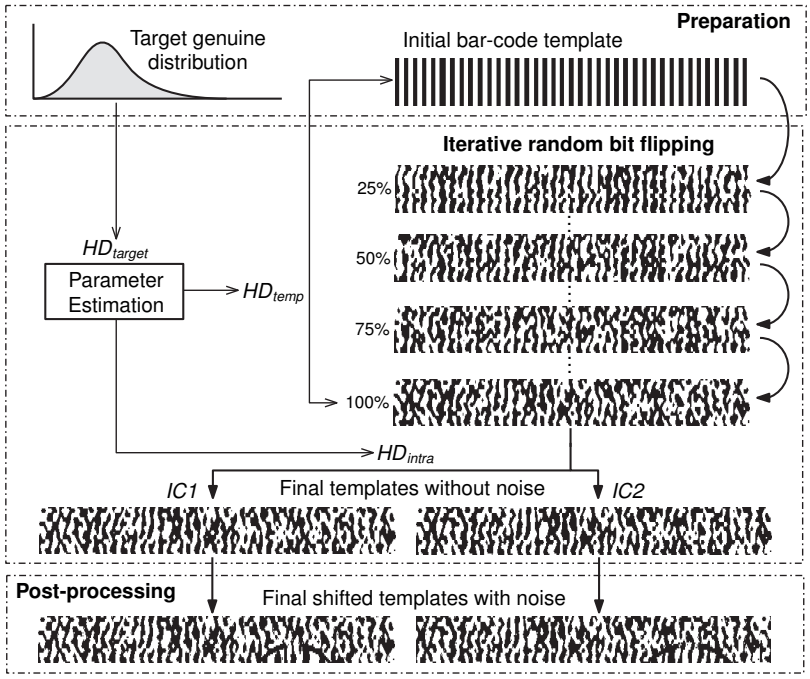


Figure 14.1: The process of generating an Iris-Code pair with SIC-Gen

1. **Preparation**, during which a base Iris-Code matrix is created as follows:

   - The first row is created by generating alternating sequences of 0's and 1's with lengths drawn from a normal distribution. The distribution parameters can be estimated empirically, by measuring the sequence lengths in real Iris-Codes.
   - By duplicating that row, a simple bar-code pattern is generated.

2. **Parameter Estimation**, during which system configuration variables are calculated based on the user input.

   - A target Hamming distance (HD_{target}) between $IC1$ and $IC2$ is drawn from a random distribution.

- HD_{temp} and HD_{intra} (see figure 14.1 and next step of the process description), are estimated based on HD_{target}. Following relations are satisfied: $HD_{temp} + HD_{intra} = C$ and $HD_{target} = 2HD_{intra} - O$, where O is the expected overlap of bit mismatches introduced by the process described in the next step; C remains constant for a batch of generated templates, and affects the effective number of independent bits (degrees of freedom) in the synthetic data.

3. **Iterative bit flipping**, during which a pair of mated Iris-Code templates is created from the base Iris-Code.

   - The bits at the edges of consecutive bit sequences (*i.e.* where sequences of 1's turn to 0's and vice versa) are randomly flipped. After HD_{temp} from the original bar-code template is reached, the template is split into $IC1$ and $IC2$. Subsequently, bit flipping occurs until HD_{intra} between them is reached.
   - Additionally, majority voting and median filtering are applied to make the patterns visually smoother. Furthermore, the chances of bit flips are adjusted on per-row basis to simulate the collarette and furrow structures in real irides.
   - This step can be accelerated by applying an initial shifting pattern to the bar-code template produced in step 1.

4. **Post-processing**, during which additional noise factors are accounted for. Those include:

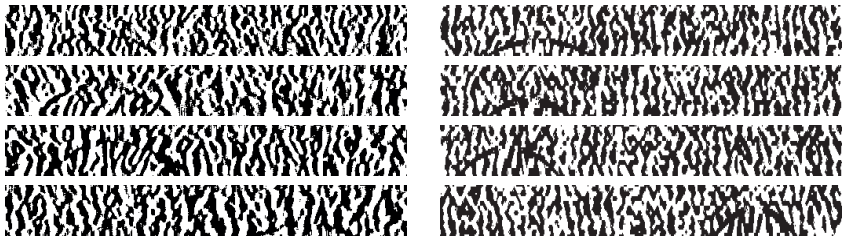
   - Adding the characteristic pattern resulting from an eyelid, as well as the noise beneath it.
   - Adding additional noise in the row near the pupil and simulating occlusions.
   - Storing the noise masks.
   - Applying circular shifts to the Iris-Code to simulate sample roll pose.

The process generates Iris-Codes of a default size; smaller sizes, if desired, are sampled from this size. The default dimension is motivated by the ISO/IEC international standard on Biometric sample quality [6]. There, the minimum iris radius is recommended to be at least 80 pixels (for the smallest reported human iris), which corresponds to a texture width of $80 * 2\pi \approx 502$ pixels when unrolled. The recommended optimal iris-pupil ratio is 0.2, which corresponds to a pupil of $80 * 0.2 = 16$ pixels, and thus an iris texture of 64 rows. Thus, the default size of the generated Iris-Codes is 512×64 bits. There are numerous adjustable parameters, which allow to mimic different

properties of the Iris-Code (e.g. the correlations between rows and columns, noise). Notably, it is also possible to *guarantee* an arbitrary distribution of genuine scores and thereby simulate sample quality. For the data generated in this paper, the HDs are drawn from a Weibull distribution, due to its close resemblance to real data; another candidate could be the Gamma distribution. Yet another approach could be to empirically estimate a distribution from real data and use it instead.

14.3 Validation

In this section, the properties of the synthetically generated data are validated with respect to the requirements outlined in section 14.2. The visual comparison between real and synthetic Iris-Codes can be seen in figure 14.2. The real Iris-Codes were produced by using the OSIRIS toolkit [10] to process the near-infrared images from the iris subset of the BioSecure [9] database. The toolkit provides the commonly used 2D-Gabor feature extraction algorithm to produce the Iris-Codes. The synthetic Iris-Codes bear an excellent resemblance to the real ones.

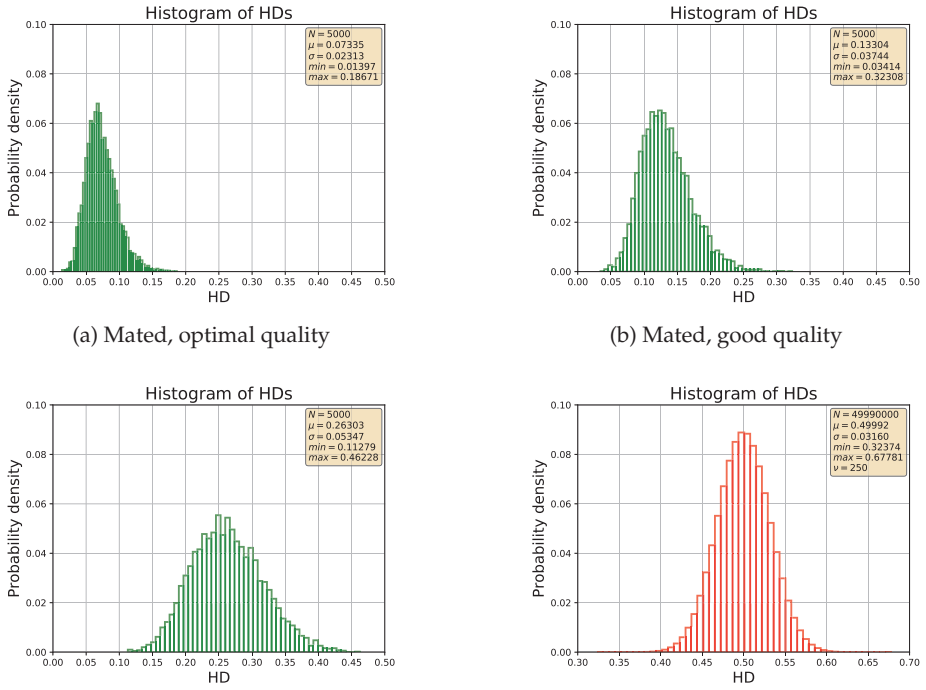


(a) Real (b) Synthetic

Figure 14.2: Example Iris-Codes produced from real eye images and generated by the proposed method

After confirming the visual appearance of the synthetic Iris-Codes to closely resemble that of the real data, their statistical properties are validated. Figure 14.3d shows the distribution of scores for non-mated templates for a large number of comparisons (N). The resulting distribution and its statistical properties (the yellow box in the image), including degrees of freedom (ν), are identical to that exhibited by the real data, shown by Daugman in [2]. In figures 14.3a, 14.3b and 14.3c, example distributions of comparison scores for mated templates are shown, representing simulating of optimal, good and non-optimal quality data, respectively. As mentioned earlier, the mated distributions can be specified arbitrarily due to the nature of the template generation process (see section 14.2). The score distributions

in figure 14.3 were produced using Iris-Codes of size 256×8 bits (same as used by Daugman in the paper cited above), sampled from the default size Iris-Codes generated by the process described in the previous section.



(a) Mated, optimal quality

(b) Mated, good quality

(c) Mated, non-optimal quality

(d) Non-mated

Figure 14.3: Distributions of Hamming distances for a large number of comparisons between synthetic templates

Due to correlations between bits in an Iris-Code, its rows comprise of sequences of consecutive identical bits. It is of interest to verify, that the synthetic data follows that property. As real data reference, sequence lengths for all templates from the iris subset of the BioSecure database were computed. In figure 14.4, those distributions are shown, along with sequence lengths produced by Daugman's HMM from [3]. The distribution for the synthetic data generated by SIC-Gen closely follows the one exhibited by the real data.

Figure 14.5 shows example error patterns for comparisons between mated and non-mated templates. For the mated template pairs, the bit mismatches occur at the edges of sequences of consecutive identical bits, resulting in the pattern akin to that shown in real data by Hollingsworth *et al*. [5].

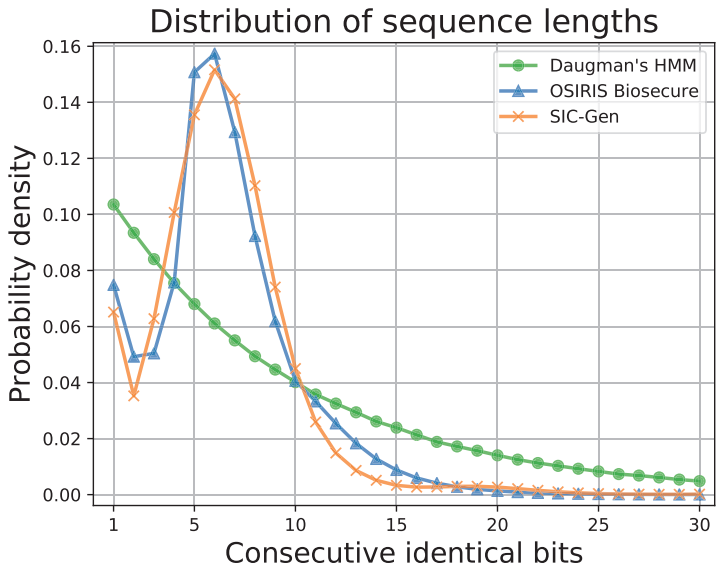
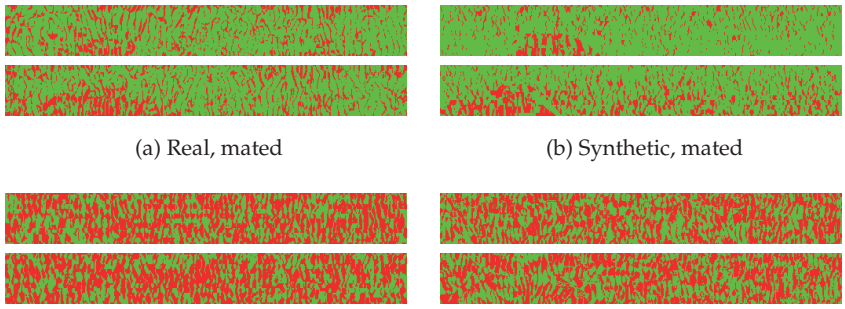


Figure 14.4: Visualisation of lengths of sequences of consecutive bits in real data from BioSecure database, SIC-Gen synthetic templates and synthetic templates generated with Daugmann's HMM



(a) Real, mated

(b) Synthetic, mated

(c) Real, non-mated

(d) Synthetic, non-mated

Figure 14.5: Example error patterns for comparisons between the real Iris-Codes from the BioSecure dataset and between the synthetic Iris-Codes

14.4 Conclusion and Future Work

In this paper, a method for generating synthetic Iris-Codes has been presented. The proposed method allows for a flexible specification of the score distribution between mated templates, to allow simulating different sam-

ple quality, acquisition environments etc.; the bit mismatches between two mated templates follow the so-called "fragile bits" patterns observed in real data. Simultaneously, the important statistical properties (*e.g.* degrees of freedom) of the distribution of non-mated comparison scores are maintained. Additionally, the synthetic Iris-Codes resemble the real ones visually. They reflect the correlations between Iris-Code bits resulting in long sequences of consecutive identical bits, as well as the typical noise sources, such as the eyelid pattern, circular shifts, wavelet noise and additional noise near the pupil. By accounting for all the aforementioned statistical and visual properties of real iris data, the proposed method represents a significant improvement over the current state-of-the-art and can be used in research cases where large iris datasets are needed, but unavailable. In future work, the authors intend to employ the synthetic Iris-Codes in large-scale testing of biometric indexing approaches, as well as to attempt to generate iris textures and/or images from the synthetic data using learning-based methods, *e.g.* Galbally *et al*. [4].

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF) as well as by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP).

14.5 Bibliography

[1] CUI, J., WANG, Y., HUANG, J., TAN, T., AND SUN, Z. An iris image synthesis method based on PCA and super-resolution. In *International Conference on Pattern Recognition (ICPR)* (August 2004), vol. 4, IEEE, pp. 471–474.

[2] DAUGMAN, J. How iris recognition works. *Transactions on Circuits and Systems for Video Technology (TCSVT) 14*, 1 (January 2004), 21–30.

[3] DAUGMAN, J. Information theory and the IrisCode. *Transactions on Information Forensics and Security 11*, 2 (February 2016), 400–409.

[4] GALBALLY, J., ROSS, A., GOMEZ-BARRERO, M., FIERREZ, J., AND ORTEGA-GARCIA, J. Iris image reconstruction from binary templates: An efficient probabilistic approach based on genetic algorithms. *Computer Vision and Image Understanding 117*, 10 (October 2013), 1512–1525.

[5] HOLLINGSWORTH, K. P., BOWYER, K. W., AND FLYNN, P. J. The best bits in an iris code. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 31*, 6 (June 2009), 964–973.

[6] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 29794-6:2015. Information technology – Biometric sample quality – Part 6: Iris image data*. International Organization for Standardization and International Electrotechnical Committee, July 2015.

[7] LEFOHN, A., BUDGE, B., SHIRLEY, P., CARUSO, R., AND REINHARD, E. An ocularist's approach to human iris synthesis. *Computer Graphics and Applications 23*, 6 (November 2003), 70–75.

[8] MAKTHAL, S., AND ROSS, A. Synthesis of iris images using Markov random fields. In *European Signal Processing Conference (EUSIPCO)* (September 2005), IEEE, pp. 1–4.

[9] ORTEGA-GARCIA, J., FIERREZ, J., ALONSO-FERNANDEZ, F., GALBALLY, J., FREIRE, M. R., ET AL. The multiscenario multienvironment BioSecure multimodal database (BMDB). *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 32*, 6 (June 2010), 1097–1111.

[10] OTHMAN, N., DORIZZI, B., AND GARCIA-SALICETTI, S. OSIRIS: An open source iris recognition software. *Pattern Recognition Letters 82*, 2 (September 2016), 124–131.

[11] PROENÇA, H., AND NEVES, J. Iris biometric indexing. In *Iris and Periocular Biometric Recognition*. Institution of Engineering and Technology, July 2017, pp. 101–124.

[12] PROENÇA, H., AND NEVES, J. C. Creating synthetic IrisCodes to feed biometrics experiments. In *Workshop on Biometric Measurements and Systems for Security and Medical Applications* (September 2013), IEEE, pp. 8–12.

[13] SHAH, S., AND ROSS, A. Generating synthetic irises by feature agglomeration. In *International Conference on Image Processing (ICIP)* (October 2006), IEEE, pp. 317–320.

[14] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Aadhaar dashboard. `https://www.uidai.gov.in/aadhaar_dashboard/`. Last accessed: 2020–03–11.

[15] WECKER, L., SAMAVATI, F., AND GAVRILOVA, M. Iris synthesis: a reverse subdivision application. In *International Conference on computer graphics and interactive techniques in Australasia and South East Asia* (November 2005), ACM, pp. 121–125.

[16] WEI, Z., TAN, T., AND SUN, Z. Synthesis of large realistic iris databases using patch-based sampling. In *International Conference on Pattern Recognition (ICPR)* (December 2008), IEEE, pp. 1–4.

[17] ZUO, J., AND SCHMID, N. A. A model based, anatomy based method for synthesizing iris images. In *International Conference on Biometrics (ICB)* (January 2005), Springer, pp. 428–435.

[18] ZUO, J., SCHMID, N. A., AND CHEN, X. On generation and analysis of synthetic iris images. *Transactions on Information Forensics and Security (TIFS) 2*, 1 (March 2007), 77–90.

Chapter 15

Score Fusion Strategies in Single-Iris Dual-Probe Recognition Systems

Abstract

Multiple samples can be utilised at the comparison stage of a biometric system in order to increase its biometric performance via information fusion or decision heuristics. It has been shown, that in a single-instance dual-probe setup, fusing the probe scores yields significant biometric performance increase over the single-probe baseline. Additionally, using the probe-probe comparison score was demonstrated to further improve the biometric performance of a fingerprint recognition system in a study by Cheng et al. In this paper, through a benchmark on the CASIA-IrisV4-Interval dataset and on the iris corpus of the BioSecure dataset, the aforementioned method is shown to be viable for an iris recognition system. However, since it requires an additional parameter, which must be estimated empirically, we propose a simpler method which exhibits similar biometric performance, while requiring no additional parametrisation.

Addressed research question(s): RQ3
Reference: DROZDOWSKI, P., WIEGAND, N., RATHGEB, C., AND BUSCH, C. Score fusion strategies in single-iris dual-probe recognition systems. In *International Conference on Biometric Engineering and Applications (ICBEA)* (May 2018), ACM, pp. 13–17.

15.1 Introduction

In past years, several multi-biometric iris recognition systems have been proposed [14, 16], some of which consolidate information from multiple samples of a single eye instance during enrolment. Some of these single-instance multi-sample fusion approaches have been found to significantly improve the recognition accuracy of iris recognition systems. The vast majority of proposed iris-based multi-sample fusion schemes process multiple extracted feature vectors, *i.e.* binary iris-codes, at the time of enrolment. The

first conceptual scheme of this kind was presented in [6], in which a majority vote-based coding is applied for each bit position of an odd number of iris-codes, with the goal of reducing the intra-class variation between the resulting reference and probe iris-codes. In [19], a weighted majority voting was proposed to improve the accuracy of an iris recognition system. A weight map, which indicates the stability of iris-code bits, is obtained from several iris-codes at enrolment. Comparison scores are then estimated as a weighted sum of mis-matching bits. A similar approach based on personalized weight maps has been presented in [7]. In [10], so-called "fragile" bits, *i.e.* bits which exhibit a higher probability than others to flip their value during a genuine comparison, are detected by comparing several iris-codes obtained from a single eye. Incorporating those bits into noise masks extracted in the iris segmentation stage was shown to improve the overall biometric performance of the iris recognition system. In contrast to the aforementioned approaches, a signal-level fusion of iris texture information extracted from multiple frames of a video was proposed in [9]. Based on a pixel-wise averaging, a single normalised iris texture is obtained. Such textures exhibit higher quality/reliabiltiy, and have been shown to improve the biometric performance of an iris recognition system. This scheme has been derived from a concept which was first introduced for face recognition [1]. Similar schemes have been proposed for fingerprint recognition systems [11, 17]. In [2], a score fusion of single-fingerprint dual-probe is proposed, where in addition to utilising the reference-probe comparison scores, the probe-probe comparison score is incorporated into a score fusion. In this paper, said score fusion method, along with proposal of further heuristics are applied in an iris-based system and benchmarked.

The remainder of this paper is organised as follows: in section 15.2, the employed fusion strategies for single-iris dual-probe iris recognition are described. In section 15.3, the experimental setup and results are presented, while section 15.4 contains a summary of the paper.

15.2 Fusion Strategies

State-of-the-art iris recognition systems capture multiple samples during acquisition stage for the purpose of supporting compensation of pose or gaze variations or for providing some fundamental presentation attack detection (PAD) [8]. Those additional samples can then be utilised at comparison stage. Specifically, in a system where two probe samples are present at comparison stage, three comparison scores can be computed as shown in figure 15.1: two (HD_1 and HD_2) between the reference and each probe and one (HD_3) between the two probes themselves. It is then possible to fuse the scores, for example, in following ways:

- Using only the scores between the reference and probes, an Average

Reference-Probe score (referred to as "ARP"): $(HD_1 + HD_2)/2$. Observe, that for fusing the scores no normalisation is required, since the experimental scores stem from a single biometric system (same modality and same comparison algorithm).

- Using all three scores, an Average Reference-Probe score weighted by the probe-probe score (referred to as "w-ARP"): $(HD_1 - a * HD_3 + HD_2 - a * HD_3)/(2 - a)$, where a is estimated on a training set, so that it maximises the biometric performance.

- Based on the probe-probe score, the reference-probe scores are either fused using ARP or only the minimum is used (referred to as "Min-or-ARP"). Here, the probe-probe score functions as a quality check – if one (or both) probes are of bad quality, then HD_3 is likely to be high. In this case, if HD_3 exceeds the acceptance threshold of the biometric system, it will therefore be better, instead of ARP, to simply use the minimum of HD_1 and HD_2. Doing so will disproportionally favour genuine transactions, by providing better chances of acceptance even in case of one sample being of bad quality; whereas the impact on impostor scores is expected to be negligible.

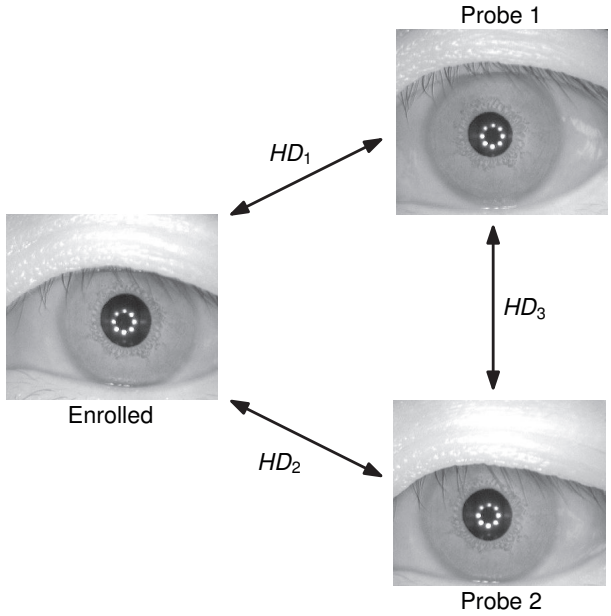


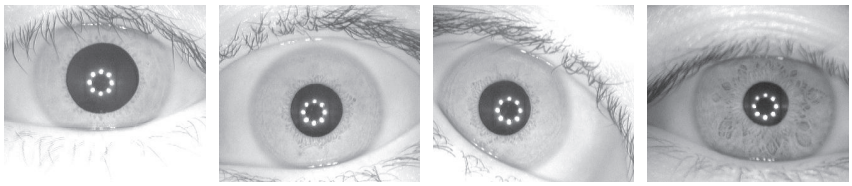
Figure 15.1: Single-iris dual-sample iris recognition

15.3 Performance Evaluation

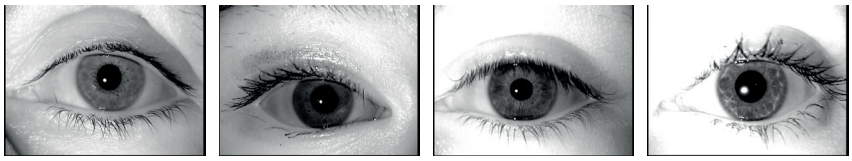
This section contains the evaluation of the dual-sample fusion schemes described in section 15.2. In subsection 15.3.1, the used dataset and the experimental setup details are outlined, while the results are presented and discussed in subsection 15.3.2.

15.3.1 Dataset

The experiments were performed on the CASIA-IrisV4-Interval database [3] (henceforth referred to as "CASIA") and the iris corpus of the BioSecure database [12] (henceforth referred to as "BioSecure"), both containing images captured in near-infrared light spectrum. An overview of the datasets is shown in table 15.1, while example images are shown in figure 15.2. Several subjects had to be removed from the BioSecure dataset due to labelling errors.



(a) CASIA



(b) BioSecure

Figure 15.2: Example images from the datasets

Table 15.1: Dataset overview

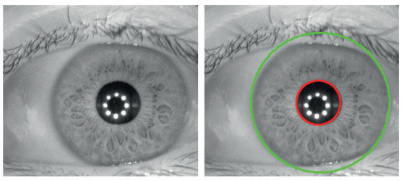
Dataset	Subjects	Instances	Images	Resolution
CASIA	249	395	2639	320×280 px
BioSecure	210	420	1680	640×480 px

The raw images were processed with the commonly used methods, as shown in figure 15.3. After segmentation using the Viterbi algorithm [18], where the iris and pupil boundaries are located, the iris textures were normalised according to the rubbersheet model [5] and subsequently enhanced

Table 15.2: Numbers of comparisons performed during experiments. ("Fusion" refers to all three fusion experiments, *i.e.* ARP, w-ARP and Min-or-ARP, since for each one of those the transactions numbers are identical)

Dataset	Experiment	Genuine	Impostor
CASIA	Baseline	41594	7993888
	Fusion	20797	3996944
BioSecure	Baseline	10080	2111632
	Fusion	5040	1055816

by applying Contrast Limited Adaptive Histogram Equalization (CLAHE). Feature extraction was performed with the Daugman-like 1D-LogGabor algorithm (LG), generating iris-codes of size $512 \times 20 = 10240$ bits. Such templates are compared using fractional Hamming distance with circular shifts applied to account for sample misalignment. The implementations of the aforementioned algorithms were provided by open-source frameworks OSIRIS [13] and USIT [15]. The evaluation of the methods described in section 15.2 along with a single-sample baseline were performed in verification mode. In the experiments, all possible transactions were performed; table 15.2 shows the numbers of transactions for each experiment.



(a) Iris detection in a raw image

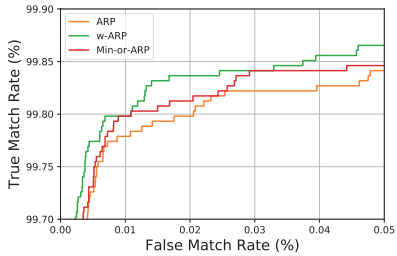
(b) Normalised and enhanced texture

(c) 1D Log-Gabor Iris-Code

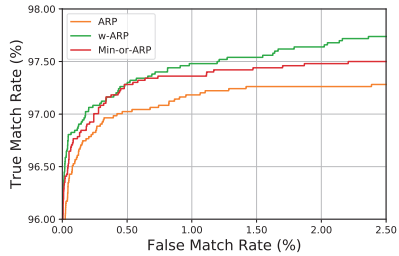
Figure 15.3: Iris recognition processing chain

15.3.2 Results

Figure 15.4 shows the receiver operating characteristic (ROC) curves of the benchmarked approaches. The baseline is not shown, since its biometric performance is well below that of the fusion approaches (see table 15.3). It can be seen, that by incorporating the third comparison score (between the two probes – w-ARP) into the score fusion, biometric performance can be improved over that of a simple score fusion of the two reference-probe



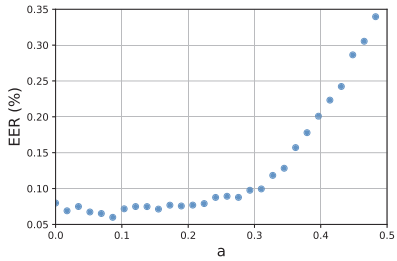
(a) CASIA



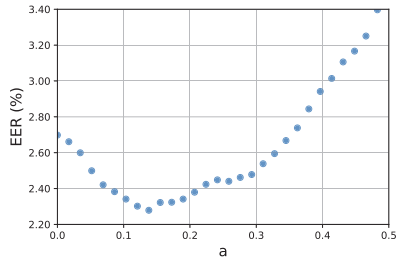
(b) BioSecure

Figure 15.4: ROC curves

scores (ARP). It also appears that said third comparison score can be effectively used as a quality check, since the Min-or-ARP algorithm slightly outperforms the plain score fusion strategy (ARP) and has a biometric performance comparable to that of w-ARP. In figure 15.5, it can be seen that the biometric performance of w-ARP varies strongly with the value of the a parameter. This provides strong motivation for introducing the Min-or-ARP scheme, as a parameter-free alternative.



(a) CASIA



(b) BioSecure

Figure 15.5: Scatter plots for w-ARP scheme showing the dependence of biometric performance on the a parameter

In table 15.3, additional metrics for benchmarking the strategies are listed. Those are: equal-error-rate (EER), area under ROC curve (AUC) and decidability (d'). The decidability is computed using the means and standard deviation of the genuine and impostor score distributions: $d' = \frac{|\mu_1 - \mu_2|}{\sqrt{\frac{1}{2} * (\sigma_1^2 + \sigma_2^2)}}$ (higher values are better). This metric is useful in assessing the intrinsic decidability of a biometric decision problem, although with the limitation of ignoring statistical moments higher than second-order [4]. It can be observed, that the dual-sample set-ups all outperform the baseline significantly, while

the benefits of the additional heuristics (w-ARP, Min-or-ARP) over ARP are noticeable, especially in the significantly higher decidability values.

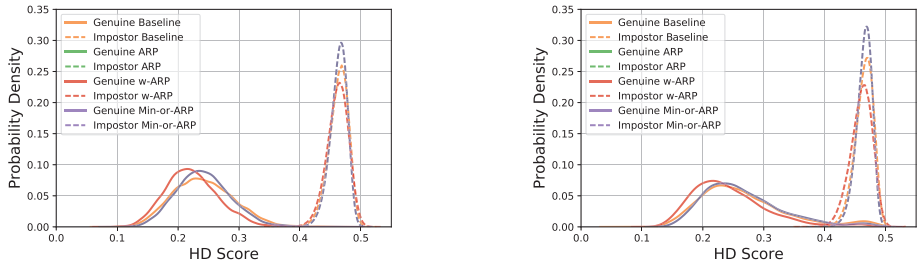
Table 15.3: Results

Dataset	System	EER	AUC	d'
CASIA	Baseline	0.00334	0.99938	5.88276
	ARP	0.00130	0.99968	6.69575
	w-ARP	0.00110	0.99969	7.07860
	Min-or-ARP	0.00115	0.99968	6.74932
BioSecure	Baseline	0.04563	0.97855	3.80808
	ARP	0.02698	0.98806	4.25278
	w-ARP	0.02317	0.99103	4.60225
	Min-or-ARP	0.02455	0.98919	4.49606

In addition to the biometric performance and decidability metrics, it is interesting to take a look at the statistical and visual properties of the genuine and impostor score distributions produced by the algorithms described in section 15.2. Those are listed in table 15.4, while figure 15.6 shows kernel density estimates of the distributions. Most noticeable is that the genuine distribution for w-ARP has significantly shifted to the left, while its corresponding distribution has only done so slightly, which explains the improved biometric performance.

Table 15.4: Distribution statistics

Dataset	Type	System	Min	Max	Mean	Std	Skew	Ex. kurt.
CASIA	Genuine	Baseline	0.076	0.484	0.242	0.051	0.377	0.390
		ARP	0.090	0.484	0.242	0.045	0.383	0.506
		w-ARP	0.079	0.487	0.223	0.044	0.532	0.821
		Min-or-ARP	0.090	0.484	0.242	0.044	0.363	0.496
	Impostor	Baseline	0.354	0.524	0.463	0.016	-0.568	0.479
		ARP	0.371	0.520	0.463	0.014	-0.599	0.587
		w-ARP	0.355	0.530	0.461	0.018	-0.397	0.277
		Min-or-ARP	0.355	0.520	0.463	0.014	-0.607	0.616
BioSecure	Genuine	Baseline	0.064	0.497	0.266	0.072	0.972	0.846
		ARP	0.142	0.491	0.266	0.065	0.959	0.875
		w-ARP	0.120	0.497	0.248	0.063	1.078	1.415
		Min-or-ARP	0.142	0.491	0.262	0.062	1.031	1.304
	Impostor	Baseline	0.351	0.526	0.465	0.015	-0.591	0.499
		ARP	0.370	0.511	0.465	0.013	-0.631	0.669
		w-ARP	0.356	0.522	0.460	0.017	-0.395	0.087
		Min-or-ARP	0.370	0.511	0.464	0.013	-0.639	0.667



(a) CASIA

(b) BioSecure

Figure 15.6: Kernel density estimates for the score distributions

15.4 Summary

In this paper, several methods for fusing information in single-iris dual-probe authentication scenario were benchmarked. It has been shown that using two probe samples can yield significant biometric performance improvements over the single probe sample baseline. Specifically, aside from a simple score fusion of the two reference-probe scores, a third score – between the two probes – can be utilised. Here, two methods were tested: one a direct re-implementation of the idea from single-fingerprint dual-probe system of Cheng et al., where the third score is directly incorporated into the score fusion. A second method was proposed, where the third score acts as a probe quality check, based on which the reference-probe scores are either fused or only the minimum is used. Both methods yield slight improvements over the simple score fusion method in terms of biometric performance (ROC curves) and decidability (d'). The advantage of the proposed method (Min-or-ARP) over the existing weighted fusion method (w-ARP) is that it does not require additional parametrisation (in w-ARP, the a parameter has to be estimated on a training set to minimise the EER).

For the systems operating in verification mode, the additional computational workload of the dual-sample approach is negligible – 2 or 3 template comparisons instead of 1, while in the identification mode, the workload would be doubled (the score between the two probes need only be calculated once). Lastly, the dual-sample approach could be effortlessly incorporated into some operational systems, since they might already capture multiple samples, *e.g.* for PAD.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF) as well as by the Hessen State Ministry for

Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP).

15.5 Bibliography

[1] CHELLAPPA, R., KRUGER, V., AND ZHOU, S. Probabilistic recognition of human faces from video. In *International Conference on Image Processing (ICIP)* (September 2002), vol. 1, IEEE, pp. 1–4.

[2] CHENG, X., TULYAKOV, S., AND GOVINDARAJU, V. Multiple-sample fusion of matching scores in biometric systems. In *Computer Vision and Pattern Recognition Workshops (CVPRW)* (June 2011), IEEE, pp. 120–125.

[3] CHINESE ACADEMY OF SCIENCES' INSTITUTE OF AUTOMATION. CASIA iris image database. `http://biometrics.idealtest.org/`, December 2010. Last accessed: 2020–03–11.

[4] DAUGMAN, J. Biometric decision landscapes. Tech. Rep. UCAM-CL-TR-482, University of Cambridge - Computer Laboratory, January 2000.

[5] DAUGMAN, J. How iris recognition works. *Transactions on Circuits and Systems for Video Technology (TCSVT) 14*, 1 (January 2004), 21–30.

[6] DAVIDA, G., FRANKEL, Y., AND MATT, B. On enabling secure applications through off-line biometric identification. In *Symposium on Security and Privacy* (May 1998), IEEE, pp. 148–157.

[7] DONG, W., SUN, Z., AND TAN, T. Iris matching based on personalized weight map. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 33*, 9 (September 2011), 1744–1757.

[8] GALBALLY, J., AND GOMEZ-BARRERO, M. A review of iris antispoofing. In *International Conference on Biometrics and Forensics (IWBF)* (March 2016), IEEE, pp. 1–6.

[9] HOLLINGSWORTH, K., PETERS, T., BOWYER, K. W., AND FLYNN, P. J. Iris recognition using signal-level fusion of frames from video. *Transactions on Information Forensics and Security (TIFS) 4*, 4 (December 2009), 837–848.

[10] HOLLINGSWORTH, K. P., BOWYER, K. W., AND FLYNN, P. J. The best bits in an iris code. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 31*, 6 (June 2009), 964–973.

[11] JAIN, A., AND ROSS, A. Fingerprint mosaicking. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (May 2002), vol. 4, IEEE, pp. 4064–4067.

[12] ORTEGA-GARCIA, J., FIERREZ, J., ALONSO-FERNANDEZ, F., GALBALLY, J., FREIRE, M. R., ET AL. The multiscenario multienvironment BioSecure multimodal database (BMDB). *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 32*, 6 (June 2010), 1097–1111.

[13] OTHMAN, N., DORIZZI, B., AND GARCIA-SALICETTI, S. OSIRIS: An open source iris recognition software. *Pattern Recognition Letters 82*, 2 (September 2016), 124–131.

[14] RADU, P., SIRLANTZIS, K., HOWELLS, G., DERAVI, F., AND HOQUE, S. A review of information fusion techniques employed in iris recognition systems. *International Journal of Advanced Intelligence Paradigms 4*, 3/4 (February 2012), 211–240.

[15] RATHGEB, C., UHL, A., WILD, P., AND HOFBAUER, H. Design decisions for an iris recognition SDK. In *Handbook of Iris Recognition*, K. Bowyer and M. J. Burge, Eds., 2 ed., Advances in Computer Vision and Pattern Recognition. Springer, July 2016, pp. 359–396.

[16] ROSS, A., NANDAKUMAR, K., AND JAIN, A. K. *Handbook of multibiometrics*. Springer, 2006.

[17] RYU, C., HAN, Y., AND KIM, H. Super-template generation using successive Bayesian estimation for fingerprint enrollment. In *International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)* (July 2005), Springer, pp. 710–719.

[18] SUTRA, G., GARCIA-SALICETTI, S., AND DORIZZI, B. The Viterbi algorithm at different resolutions for enhanced iris segmentation. In *International Conference on Biometrics (ICB)* (March 2012), IEEE, pp. 310–316.

[19] ZIAUDDIN, S., AND DAILEY, M. N. Iris recognition performance enhancement using weighted majority voting. In *International Conference on Image Processing (ICIP)* (October 2008), IEEE, pp. 277–280.

Part IV

Conclusions

Chapter 16

Summary of Results

Abstract

In this chapter, the research questions (see section 2.1) are answered by summarising the insights and results from the research articles contained in this thesis.

16.1 Research Question 1

Most of the research articles in this thesis concern themselves with this question. It is indeed possible to vastly reduce the computational workload in biometric identification through various methods at different levels of the biometric data processing pipeline. For the truly large gains in computational efficiency there usually (albeit not always, see *e.g.* chapter 10) follows a slight trade-off with biometric performance. Nevertheless, the results show that reduction down to less than 1% of the baseline (exhaustive search) computational workload is feasible before the impact on the biometric performance becomes prohibitive.

16.2 Research Question 2

The current state-of-the-art in this thesis' research area has been described and harmonised:

- In chapter 4, a biometric characteristic independent taxonomy of approaches for computational workload reduction in biometric identification systems has been developed based on a comprehensive survey of the scientific literature.
- Metrics, as well as experimental protocol prerequisites for modality and method agnostic reporting of computational workload reduction have been developed (see chapter 4 and appendix A) and they are currently being considered for inclusion in the revision project of ISO/IEC 19795-1 [2].

Some other aspects relevant to the operation of large-scale identification systems have also been investigated:

- In chapter 13 and [3], the impact of eyeglasses on iris recognition systems is assessed for near-infrared and visual wavelength images. Furthermore, methods for detection of glasses in such images have been developed.
- In chapter 14, a generator of visually and statistically realistic synthetic iris templates (Iris-Codes) has been conceptualised and implemented in order to facilitate creation of large synthetic databases which can be used *e.g.* for system stress testing and evaluation.

16.3 Research Question 3

Most of the existing methods surveyed in chapter 4 do not incorporate information fusion. Several research articles contained in this thesis address this matter:

- In chapter 5, an auxiliary feature (fingerprint type) has been utilised to perform multi-instance binning.
- In chapter 6, a multi-instance fusion (left and right iris) has been performed on the feature level, prior to organising the templates into a hierarchical search structure (see also appendix A).
- In chapter 10, a method of candidate short-list filtering which works with an arbitrary number and type of biometric characteristics and representations has been presented.
- In chapter 11, information fusion on the signal level has been performed, whereby biometric information from multiple data subjects has been fused into one image to enable a pre-selection step using the fused data.

16.4 Research Question 4

Although many methods of computational workload reduction are tied to a specific feature representation (*i.e.* they take advantage of its intrinsic properties), it is possible to create methods which can be used prior to the feature extraction or even irrespective of the chosen feature representation. This has been demonstrated in several research articles contained in this thesis:

- The methods presented in chapter 5 and 11 operate with raw samples (images – *i.e.* irrespective of the feature representation) to facilitate the computational workload reduction.

- The method presented in chapter 10 requires, irrespective of the chosen type of biometric characteristic and feature representations thereof, only the comparison scores and ranked candidate lists. The method filters the candidate lists in a cascading framework, thereby resulting in successively smaller candidate short-lists and a reduced computational workload.
- In chapter 12, the iris images have been pre-aligned (*i.e.* prior to segmentation and feature extraction), which results in a smaller computational workload associated with the sample alignment typically performed on the feature vectors (*i.e.* Iris-Codes).

16.5 Research Question 5

Data security and privacy are important in the context of biometric systems. Previous state-of-the-art work on computational workload reduction has not addressed this aspect. Two research articles in this thesis are devoted to this topic:

- In chapter 7, a row-based permutation has been applied to the feature vectors, which have subsequently been organised into a hierarchical search structure (see also appendix A). The resulting system fulfils the privacy and security objectives of ISO/IEC 24745 [1] by ensuring the unlinkability, irreversibility, and renewability of the protected templates (*i.e.* biometric references). Furthermore, the proposed system significantly reduces the computational workload, while maintaining a comparable biometric performance w.r.t. to the baseline system.
- In chapter 9, a general purpose method (homomorphic encryption) has been used to design a protocol which fulfils the aforementioned privacy and security objectives in a biometric identification system.

16.6 Summary

Computational workload constitutes one of the key challenges and design considerations w.r.t. biometric identification systems. With the increasing size of a biometric enrolment database, the naïve retrieval method (*i.e.* exhaustive search) becomes impractical, especially in terms of the needed computational effort, which in this case grows linearly with the number of enrollees. As the number, size, and scope of the worldwide deployments of biometric identification systems steadily increases, it necessitates research into the topic of computationally efficient biometric identification. A number of methods exist in this field; they can be coarsely divided into two categories (see chapter 4 for a more fine-grained taxonomy):

Workload reduction: Methods which seek to minimise the required computational effort (*e.g.* by reducing the search space) associated with the biometric identification transactions.

Acceleration: Methods which seek to optimise the software (*e.g.* implementations or level of abstraction of the algorithms) and/or hardware (*e.g.* distributed processing or reconfigurable computing), thereby speeding up the biometric identification transactions, but not actually reducing the required computational effort.

This thesis concerned itself with the first category. Specifically, a number of pre-selection and feature transformation methods were developed, concentrating on incorporating additional constraints and extensions, which have not previously been sufficiently addressed in the scientific literature, such as:

- Being agnostic w.r.t. type of biometric characteristic and/or their feature representations.
- Being applied prior to feature extraction.
- Facilitating or directly utilising biometric information fusion.
- Including biometric template protection.

Additionally, the previous research in this field has been surveyed and systematised into a biometric characteristic agnostic taxonomy. This, as well as the propositions listed above, formed the core of the research questions addressed in this thesis. The results of the individual research articles in this context have been outlined in sections (16.1–16.5) dedicated to each of the five individual research questions, respectively. Summarising the points made in those sections, all five research questions posed in section 2.1 have been answered positively.

The issues pertaining to computational workload in biometric identification systems can be expected to remain an active focus of future research. A number of promising avenues exist in this area; several of those are briefly outlined and discussed in chapter 17.

16.7 Bibliography

[1] ISO/IEC JTC1 SC27 IT SECURITY TECHNIQUES. *ISO/IEC 24745:2011. Information technology – Security techniques – Biometric information protection*. International Organization for Standardization and International Electrotechnical Committee, June 2011.

[2] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 19795-1:2006. Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*. International Organization for Standardization and International Electrotechnical Committee, April 2006.

[3] OSORIO-ROIG, D., DROZDOWSKI, P., RATHGEB, C., MORALES-GONZÁLEZ, A., GAREA-LLANO, E., AND BUSCH, C. Iris recognition in visible wavelength: Impact and automated detection of glasses. In *International Conference on Signal-Image Technology Internet-Based Systems (SITIS)* (November 2018), IEEE, pp. 542–546.

Chapter 17

Future Work

Abstract

This chapter outlines some avenues of potential future research in the areas addressed by this thesis.

17.1 Scalability

The publicly available datasets for biometric research are relatively small (*cf*. tables 2.1 and 4.1). Furthermore, the acquisition environments in such datasets may or may not correspond to the real conditions in the operational systems. Therefore, there exists a need to validate (*e.g*. in terms of scalability) the academic approaches with real, large-scale, operational data. As personal biometric data is currently legally (*i.e*. due to GDPR [2] in the European Union) categorised within "special categories" of personal data (formerly, "sensitive personal data"), it is subject to numerous protections, thereby often making data sharing legally and practically inconvenient. Those difficulties notwithstanding, for one of the research articles in this thesis (chapter 5), anonymised operational data (albeit not the images themselves) was kindly provided by the German Federal Police (Bundeskriminalamt)[1] and used to validate the approach proposed in the aforementioned research article.

17.2 Unconstrained Data

The research in this thesis concentrated on computational workload reduction for cooperative data, *i.e*. data with relatively high quality (with a possible/arguable exception of chapter 5, which used a scanned/rolled-ink fingerprint dataset), such as that used in the national identity registries or biometric passports. However, depending on the operational scenario, such data may not be available, *e.g*. in surveillance (see *e.g*. [4]) or mobile data acquisition (see *e.g*. [1]). The development of computational workload reduction methods specifically tailored to such data would definitely be of

[1] `https://www.bka.de/EN`

interest. An example of a practical application, where computationally efficient biometric identification would be important could be a quick (real-time) criminal blacklist check for images acquired by a police officer in the field using a mobile device (*e.g.* during a routine traffic control).

17.3 Deep Learning

Deep learning has been an extremely active field of research in recent years, with many impressive achievements in various domains, including computer vision and pattern recognition. Biometrics has also already experienced the benefits of deep learning methods, *e.g.* the breakthrough biometric performances in facial recognition. Although certain inherent challenges exist (*e.g.* model and decision transparency, explainability, and interpretability – see *e.g.* [6]), the deep learning methods are being embraced in many research areas within biometrics.

In the context of the biometric identification, deep learning could also be of interest. For example, one might investigate challenges such as learning to map variable-length feature vectors into fixed-length ones and learning compact (*e.g.* binary and/or short) encodings from the extracted features. Those could then be used directly for the template comparisons with efficient comparators or as a basis for indexing (or other) algorithms for search space reduction. The recent thesis of Schuch [5] has addressed some of those topics for the fingerprint-based biometric identification systems. The results presented there clearly show the potential of deep learning in this context and invite research in this field (also for other types of biometric characteristics).

17.4 Standardisation

First steps towards standardising the topic of computational workload in biometric systems have been described in appendix A, while several pertinent issues have also been discussed in chapter 4. Based on those, contributions to ISO/IEC 19795-1 [3] have been made, thereby making the SC37 committee aware of the issue and providing a platform for subsequent discussions. Parts of the comments have already been accepted for inclusion in the standard, while others are still under consideration. As such, and due to standardisation being an ongoing process, a continuous engagement with the works of the committee is necessary to ensure that the matter of computational workload (and the reduction thereof) in biometric systems is adequately represented in the standard.

17.5 Bibliography

[1] DAS, A., GALDI, C., HAN, H., RAMACHANDRA, R., DUGELAY, J.-L., AND DANTCHEVA, A. Recent advances in biometric technology for mobile devices. In *International Conference on Biometrics Theory, Applications and Systems (BTAS)* (October 2018), IEEE, pp. 1–11.

[2] EUROPEAN PARLIAMENT. Regulation (EU) 2016/679. *Official Journal of the European Union L119* (April 2016), 1–88.

[3] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 19795-1:2006. Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*. International Organization for Standardization and International Electrotechnical Committee, April 2006.

[4] NEVES, J., NARDUCCI, F., BARRA, S., AND PROENÇA, H. Biometric recognition in surveillance scenarios: a survey. *Artificial Intelligence Review 46*, 4 (December 2016), 515–541.

[5] SCHUCH, P. *Deep Learning for Fingerprint Recognition Systems*. Ph.D. thesis, Norwegian University of Science and Technology, October 2019.

[6] ZHANG, Q.-S., AND ZHU, S.-C. Visual interpretability for deep learning: a survey. *Frontiers of Information Technology & Electronic Engineering 19*, 1 (January 2018), 27–39.

Part V

Appendix

Appendix A

Bloom Filter-based Search Structures for Indexing and Retrieving Iris-Codes

Abstract

Large-scale biometric deployments are becoming ubiquitous. The computational workload of the conventional retrieval method, which requires performing 1:N comparisons in the identification mode, quickly becomes impractical for large systems. This issue necessitates the research into algorithms for efficient biometric identification. In recent years, many such approaches have been proposed, but the scalability of the proposed systems is often questionable. Furthermore, the lack of a unified methodology for biometric workload reduction reporting often makes a direct benchmark or even a thorough evaluation of the proposed schemes cumbersome.

In this article, we propose an iris indexing scheme based on Bloom filters and binary search trees. With the help of a statistical model, the system is shown to be theoretically scalable for an arbitrary number of enrollees. We evaluate this system empirically on a combined database from numerous publicly available datasets, containing a total of 11,936 iris images from 1477 instances. The system tested in an open-set identification scenario maintains the biometric performance yielded by an iris-code 1:N baseline - a true positive identification rate (TPIR) of approximately 98%, measured at 0.1% false positive identification rate (FPIR). These results are achieved at less than 10% of the baseline workload. In a proof-of-concept multi-iris indexing experiment, the TPIR measured at 0.1% FPIR is increased to over 99%, while the workload remains the same as in the case of the single-iris system above. We seek to report our experimental results clearly and exhaustively; in order to do so, we define a number of prerequisites necessary for a transparent and comprehensive methodology of reporting biometric workload reduction results.

Addressed research question(s): RQ1, RQ2
Reference: DROZDOWSKI, P., RATHGEB, C., AND BUSCH, C. Bloom filter-based search structures for indexing and retrieving Iris-Codes. *IET Biometrics 7* (May 2018), 260–268.

A.1 Introduction

In recent years, several large-scale biometric systems have been introduced worldwide. By far the largest of these is the Indian National ID project, which at the time of this writing has successfully enrolled over 1 billion subjects [28] with biometric data from iris, face and fingerprints. Two main challenges associated with large-scale biometric identification are the computational cost and the risk of false positives. A naïve, brute-force approach is to perform template comparisons between the probe and all enrolled reference templates (i.e. 1:N comparisons). Even with excellent hardware and reliance on parallelism, the computational cost quickly becomes prohibitive. Similarly, the possibility of false positive occurrences quickly becomes unacceptable. In [4], Daugman shows the probability of at least one false positive (P_N) occurring in a identification scenario to be calculated using equation (A.1), where N is the number of enrolled subjects and P_1 the false positive probability of a one-to-one template comparison.

$$P_N = 1 - (1 - P_1)^N \tag{A.1}$$

A biometric system which performs well in the verification mode (i.e. has a low P_1) is not necessarily suitable for the much more demanding identification mode. Observe, that for values of P_1 which are acceptable for biometric verification, the value of P_N might very quickly become unacceptably high as the number of enrolled subjects N increases. The equation ignores other system errors (e.g. FTA).

Following Daugman's approach [5], which is the core of most public operational systems, four major modules constitute an iris recognition system: (1) image acquisition, where most current deployments require subjects to fully cooperate with the capture device in order to capture images of sufficient quality; (2) pre-processing, which involves a detection of inner and outer iris boundaries, a detection of eyelids, an exclusion of eyelashes as well as contact lens rings, a scrubbing of specular reflections and an estimation of quality factors. Subsequently, the iris is mapped to dimensionless coordinates, i.e. a normalised rectangular texture, and an according noise mask is stored; (3) feature extraction, in which a two-dimensional binary feature vector consisting of $W_{IC} \times H_{IC}$ bits, i.e. iris-code, is generated by applying adequate filters to the pre-processed iris texture. This binary data representation enables compact storage and rapid (4) comparison, which is based on the estimation of Hamming distance (HD) scores between pairs of iris-codes. In the comparison stage circular bit shifts are applied to iris-codes and HD scores are estimated at K different shifting positions, i.e. relative tilt angles, resulting in $W_{IC} * H_{IC} * K$ bit comparisons. The minimal obtained HD, which corresponds to an optimal alignment, represents the final score.

Despite the rapid comparison and high resilience against false matches offered by the iris-code representation, the sheer scale of the major biometric deployments makes biometric workload reduction a very relevant and important research topic. Recently, a promising workload reduction approach employing a Bloom filter-based representation of iris-codes and binary search trees has been proposed in a proof-of-concept study [23]. In this article, we further analyse and expand upon said idea.

A.1.1 Contribution of Work and Article Organisation

The remainder of this article is organised as follows:

1. The related works in the area of biometric workload reduction for iris recognition are outlined. (section A.2)

2. A proposal for a standardised way of biometric workload reduction reporting is made. (section A.2)

3. A basic Bloom filter-based system for biometric indexing is described in detail. This system forms the basis of the work performed for this article. (section A.3)

4. A general model for Bloom filter-based indexing is introduced along with a number of improvements for the basic system; those include scalability to an arbitrary number of enrollees and a first of its kind attempt of multi-iris indexing. (section A.3)

5. The experiment methodology and results are presented and discussed, along with future work items and concluding remarks. (sections A.4 - A.6)

A.2 Workload Reduction in Iris Biometric Systems

A.2.1 Categories

In this work, we distinguish between three main categories of workload reduction approaches for iris biometric identification, as listed below. Other biometric modalities often use similar approach types; however, as of this writing, there exists no generalised, modality-agnostic, categorisation of workload reduction methods.

- **Cascading Algorithms**: a computationally efficient algorithm finds a short-list of most likely candidates. A slower, more accurate algorithm then runs over only this short-list.

- **Binning**: extraction of one or more distinguishing features from the biometric references and organising the enrolled database into "bins". During an identification transaction, template comparisons are performed only in the bin to which the probe has been classified.
- **Indexing**: approaches that strive for space search reduction. Often, hierarchical and/or probabilistic data structures are utilised.

A.2.2 Related Works

Table A.1 provides a summary of related works on the topic of biometric workload reduction for the iris modality.

Table A.1: Related works (results as reported by the authors, or if unavailable, extracted from the presented plots)

Category	Method	Dataset	Biometric performance	Workload reduction	Remarks
Cascading Algorithms	Gentile *et al.* [9]	MMU	TPIR 93%	12-fold reduction	high pre-selection error rate (limited TPIR)
	Konrad *et al.* [16]	CASIA-V1 CASIA-V3-Interval MMU	92% IR, 0% FAR 89% IR, 0.85% FAR 79% IR, 0.85% FAR	70-80% time reduction	low alignment cost
	Rathgeb *et al.* [25]	CASIA-V3-Interval	97.2-99.2% RR-1	5% bit comparisons	early rejection of unlikely templates
Binning	Qiu *et al.* [22]	CASIA-V2, UPOL, UBIRIS	86% CCR	number of classes	
	Ross *et al.* [26]	UPOL	0% EER	30% bits used	tested on very high quality images
	Sun *et al.* [27]	CASIA-V4 ND Clarkson	0% EER 0.9% EER 0.54% EER	number of classes	
Indexing	Mukherjee *et al.* [19]	CASIA-V3	80-84% hit rate	8-30% penetration rate	first work on iris indexing
	Gadde *et al.* [8]	CASIA-V3	99.8% hit rate	17.2% penetration rate	
	Hao *et al.* [10]	UAE	0% FAR, 0.64% FRR	0.006% penetration rate	twice the original storage required; tested on very high quality images
	Jayaraman *et al.* [14]	UBIRIS	98.7% hit rate	7.1-8.3% penetration rate	low quality images
	Mehrotra *et al.* [18]	CASIA BATH IITK	1.6-43.6% bin miss rate 4-72% bin miss rate 1.5-44% bin miss rate	39.96-0.63% penetration rate 26.14-0.06% penetration rate 41.4-0.2% penetration rate	
	Proença [21]	CASIA-V4-Thousand UBIRIS	94% TPIR at 0.1% FPIR 85% TPIR at 10.0% FPIR	$2.5-7.5\%$ av. penetration rate $38-65\%$ av. penetration rate	low quality images
	Rathgeb *et al.* [23]	IITDv1	same or better as baseline	$O(\log N)$ penetration rate	small test dataset (~200 subjects)

Despite being a relatively new topic, many approaches for reducing workload in iris identification systems have been proposed. While the results often appear promising, in many cases the way the results are reported is not following an uniform, standard methodology. Specifically, one can observe many different metrics in which the authors report the workload reduction. Furthermore, many of the schemes have hidden costs or limitations (especially in terms of scalability), which are sometimes not explicitly stated or explored. Overall, these issues often make the direct benchmark and scalability assessment impractical. This leads us to presenting, in the next subsection, a proposal for standardising the way of reporting workload reduction for biometric identification systems.

A.2.3 Workload Reduction Reporting

Unlike the biometric performance reporting (see [12]), the current methodology for biometric workload reduction reporting is not standardised. The

aforementioned standard does contain metrics related to workload reduction (e.g. penetration rate), however, their definitions do not reflect the total workload faced by an identification system. This has lead the researchers to adapt many different ways of reporting the workload reduction achieved by their system (table A.1). It is therefore evident, that the biometric research community could greatly benefit from a clear, transparent and unified methodology. Below, we present a set of abstract, high-level prerequisites (in bold typeface) necessary for a clear and unambiguous reporting of workload reduction in biometric identification systems. Those prerequisites are universal and modality-agnostic. Using the regular typeface, the rationale for each of the prerequisites is provided and, where applicable (particularly in P1 and P6), possible concrete metrics for an iris system are proposed. This topic is of interest to the research community – we thus create a solid starting point for further discussions and deliberations in the community and at ISO/IEC JTC 1/SC 37 Biometrics.

P1 The baseline workload should be explicitly stated. This is to be expressed in terms of template (e.g., the iris-code) size in bits, with the alignment compensation and other costs accounted for, the number of enrolled subjects and the penetration rate (defined as a fraction of the number of necessary template comparisons and the number of enrolled templates - in case of the baseline, the penetration rate is 1.0, since no optimisations occur) in an open-set identification scenario. Otherwise, there is no clear and direct point of reference for the workload of the proposed system.

P2 The baseline biometric performance of a state-of-the-art algorithm should be explicitly stated, in a manner described in the ISO/IEC international standard 19795-1 on biometric performance testing and reporting [12]. In particular, of interest is the biometric performance of an open-set identification scenario, expressed with the true positive identification rate (TPIR) and false positive identification rate (FPIR) plotted as an ROC curve. Otherwise, akin to **P1**, it will not be possible to establish potential biometric performance costs incurred by the workload reduction of the proposed scheme.

P3 The workload of the proposed scheme is to be stated in the manner described in P1. If these parameters vary (e.g. due to different scheme configurations or non-determinism), then a range or an upper bound should be given. If a pre-selection step is involved, then it should be accounted for within the above parameters; if that is not feasible, then its cost should be stated separately.

P4 The biometric performance of the proposed scheme should be reported according to the ISO/IEC 19795-1 standard [12]. In particular, of inter-

est is the biometric performance of an open-set identification scenario, expressed with the true positive identification rate (TPIR) and false positive identification rate (FPIR) plotted as an ROC curve. This is necessary, because without regard for biometric performance, arbitrarily high workload reduction can be claimed. A scheme will, for the most part, only be viable if the biometric performance does not become significantly lowered; in any case, the trade-offs should be mentioned.

P5 The additional costs and benefits of the proposed scheme should be listed (e.g. offline costs, storage requirements, alignment invariance). It should also be stated whether or not the template comparisons can be performed using fast CPU instructions (bitwise operators in particular). This is important to allow a general, well-informed evaluation of the system and the trade-offs associated with the workload reduction.

P6 The total workload for both the baseline and the optimised system should be computed. By applying equation (A.2), the total workload reduction of the proposed system can be succinctly and precisely stated as a fraction (F) of the workload of the baseline (e.g. "$F = 0.4$ of the baseline workload") in the worst and average case. Using this metric to summarise the results is advantageous, as it provides the readers with a single value, with which they can immediately and reliably assess the workload reduction conferred by the proposed system. The reasoning behind this requirement is including all the workload related variables in for the sake of accuracy and transparency.

Formula (equation (A.2)) to describe the system workload in a single lookup during an identification scenario ($\mathcal{W}$) is derived from the parameters stated in the requirements above: N - the number of subjects enrolled, p - the penetration rate and C - the cost of a single step (i.e. cost of a one-to-one template comparison). In case of iris recognition, the templates are represented as binary vectors; the cost of a single step can be then expressed in terms of bit comparisons, so $C = W_{IC} * H_{IC} * K$.

$$\mathcal{W} = N * p * C \tag{A.2}$$

The results in this article will be presented in conformance with the methodology proposed in this section.

A.3 Methodology

In [23], a proof-of-concept study of an iris identification system based on Bloom filters and binary search trees is presented. Its basics are briefly described in subsections A.3.1 and A.3.2, since that system forms the founda-

tion for the work performed for this article. In subsection A.3.3, a general model for Bloom filter-based indexing is introduced.

A.3.1 Bloom Filter

A Bloom filter [1] is a probabilistic data structure for the purpose of efficient membership queries in one against set retrievals. Bloom filters convince by successful application in a variety of scenarios [2]. Recently in [24], the applicability of Bloom filters to iris-based biometrics has been assessed with a proposed template protection scheme. In our article, the Bloom filter-based representation is used for the purpose of a scalable and efficient biometric indexing scheme for open-set identification, vastly extending on the proof-of-concept work in [23].

A Bloom filter is a binary vector of fixed length. This can also be represented as a set of integers (i.e. activated indexes). Data addition proceeds by applying transformation/mapping function(s) to the data and inserting the resultant items into the filter (i.e. activating bits at the according positions). The retrieval is performed by applying the same function(s) to the probe data and comparing if the produced items match with the ones stored in the reference filter. In its original form, this concept considers only full matches (binary yes/no decision); however, it can be extended for fuzzy matching by employing a (dis)similarity metric between two filters. Finally, it is worth noting, that two or more Bloom filters can be seamlessly combined using a set union. This useful property along with the sparseness of the data representation form the basis of the system described in the next subsection.

A.3.2 Bloom Filter-based Indexing

The process of transforming the iris-code templates to a Bloom filter-based representation is as follows: First, the two-dimensional iris-code is divided into l blocks of equal size, $W_B \times H_B$, which are then inserted into Bloom filters. Instead of using multiple hash functions as in the original Bloom filter concept, a single transformation function is applied. It interprets the columns $(c_1, \ldots, c_{W_B})$ in a block as binary numbers, converts them to base-10 integers and inserts those into the Bloom filter corresponding to that block (i.e. $\mathbf{b}[\text{int}(c_i)] = 1$). The column values are obviously always in range $0 \leq \text{int}(c_i) < 2^{H_B}$. Thereby, the resulting biometric template (denoted $\mathbf{B}$) is a fixed-length (l) sequence of Bloom filters $(\mathbf{b}_1, \ldots, \mathbf{b}_l)$. A dissimilarity score of two such biometric templates is calculated as an average of pairwise Hamming distances between corresponding Bloom filters in the sequences, as shown in equation (A.3), where $\mathbf{B}$ and $\mathbf{B}'$ denote the reference and probe template, respectively.

$$DS(\mathbf{B}, \mathbf{B}') = \frac{1}{l} \sum_{i=1}^{l} ds(\mathbf{b}_i, \mathbf{b}'_i)$$
$$ds(\mathbf{b}, \mathbf{b}') = \frac{|\mathbf{b} \oplus \mathbf{b}'|}{(|\mathbf{b}| + |\mathbf{b}'|)} \tag{A.3}$$

Where $|\cdot|$ represents the population count, i.e. Hamming weight. Observe, that like in the case of the iris-code, the comparator utilises efficient bitwise instructions and can be trivially parallelised. It is also worth noting, that this data representation is, to a certain degree, rotation invariant, at the cost of loss of local information and loss of information about the number of identical columns in a block. The block size determines the sparseness of the representation; we define the level to which a filter is filled as the number of values present in the filter as a fraction of the number of possible values in the filter (i.e. $|\mathbf{b}| / 2^{H_B}$). In order to identify suitable configurations of $W_B \times H_B$ for tree construction, we estimate the level to which a filter is expected to be filled when holding just one iris-code template. The results are shown in table A.2.

Table A.2: Approximation of filling a Bloom filter resulting from a block of height H_B and width W_B with random data (lower values reflect higher data representation sparseness and fewer potential collisions)

H_B \ W_B	**8**	**16**	**32**	**64**
5	0.22	0.4	0.64	0.87
6	0.12	0.22	0.4	0.64
7	0.061	0.12	0.22	0.39
8	0.031	0.061	0.12	0.22
9	0.016	0.031	0.061	0.12
10	0.0078	0.016	0.031	0.061
11	0.0039	0.0078	0.016	0.031
12	0.0020	0.0039	0.0078	0.016

The presented values come from random data and thus represent an upper bound. A model accounting for the characteristics of an iris-code is presented in the next subsection. Nevertheless, this simple estimation provides a good intuition regarding the relation between block size, filter size and its sparseness. As will be demonstrated later, the data representation sparseness is a crucial factor for the system accuracy. Before expanding upon this matter, we describe the system operation in the identification mode.

A binary tree data-structure is constructed from all the enrolled templates. First, the tree root is created from all enrolled templates (i.e. $\bigcup_{i=1}^{N} \mathbf{B}_i$),

while the children of the root node each contain half of the enrolled templates (i.e. $\bigcup_{i=1}^{N/2} \mathbf{B}_i$ and $\bigcup_{i=N/2+1}^{N} \mathbf{B}_i$). The union of templates corresponds to ORing the individual binary filters. This process is repeated for node creation at subsequent tree levels, until at the end, the individual templates $(\mathbf{B}_1, \ldots, \mathbf{B}_N)$ are inserted as tree leaves. Since the scheme inherently requires no specific insertion ordering, the templates are always inserted into the next available position in the tree. In other words, the insertion does not occur at a deeper level until the previous one is completely filled, thus ensuring that the tree always remains balanced and its height is $\lceil \log_2(N) + 1 \rceil$.

The lookup in an identification scenario begins at the tree root. The tree is traversed by calculating the dissimilarity scores (equation (A.3)) between the probe template and two nodes at the next tree level; subsequently choosing the one with lower score until a leaf is reached. The key idea is to take advantage of a sparse data representation in the nodes - in that case, for the probe comparisons against the tree $DS_{genuine} \ll DS_{impostor}$ generally holds true. In other words, the genuine probes will be able to traverse the tree using the correct path to reach a matching leaf template. At the leaf, a final decision is made based on an acceptance threshold. The complexity class of a single lookup is $O(\log N)$. Note, however, that there is also a non-trivial constant factor: for each step, the scores at both child nodes have to be computed in order to make a traversal decision. Thus, while still of logarithmic complexity, the actual number of template comparisons per lookup is $2 * \log N - 1$. The tree construction and item retrieval processes are shown in figure A.1.

In a proof-of-concept study [23], this system has shown promising results both in terms of biometric performance and workload reduction. However, the scalability of the scheme has, until now, not been assessed. The obvious limitation of the basic scheme is lack of support for a large number of enrollees. As more subjects are added to the single tree, the data representation at the top tree levels will lose its sparseness, thus decreasing the difference between the genuine and impostor scores (i.e. eventually $DS_{genuine} \not\ll DS_{impostor}$ at the top tree levels). The lack of large difference between genuine and impostor scores impairs the ability to make correct traversal direction decisions for genuine probes, thus yielding a severe negative impact on the rate of true positives. In other words, it becomes increasingly more difficult to distinguish between genuine and impostor probes as the sparseness of the data representation diminishes.

To summarise, we are interested in being able to pinpoint when a Bloom filter reaches its capacity and is unable to accommodate more templates without accuracy loss and how the system design can be altered to be able to accommodate more enrollees. These two matters are explored in the following sections. They introduce a statistical model and implementation improvements for Bloom filter-based indexing, which ensure that the scheme

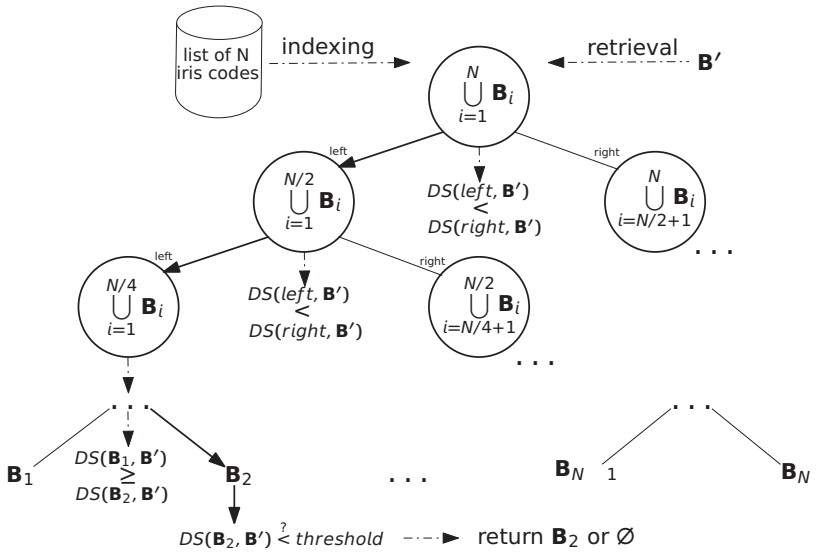


Figure A.1: Indexing and retrieval in the Bloom filter-based system. In this case, the retrieval follows the bold path down to a leaf, where the final decision is made.

can be theoretically applied to an arbitrary number of enrolled subjects.

A.3.3 Generic Model for Bloom Filter-based Indexing

A Bloom filter-based template can be equivalently represented as a sequence of sets of unique integer values within a certain range. The integer values then correspond to activated indices in the filters, which are obtained via a simple transformation shown earlier in this section. For the purpose of the model, the set-based representation is used in order to simplify the argument and notation. During an authentication, a pairwise comparison of the sets from probe and reference templates takes place. Given the assumption that all sets for a given $W_B \times H_B$ configuration exhibit similar characteristics, for the purposes of the model, the discourse is simplified to looking at a single set of integers (i.e. one member of the sequence of sets).

Let $\mathbf{b}$ denote a Bloom filter created from an iris-code block of size $W_B \times H_B$. Assuming that all values in the block are *mutually independent* and drawn from a *uniform distribution*, then:

$$\mathbf{b} = \left\{ x \in \mathbb{N}_0 \mid 0 \leq x < 2^{H_B} \right\}, |\mathbf{b}| = 1 - \left(1 - \frac{1}{2^{H_B}} \right)^{W_B} \tag{A.4}$$

However, iris-code columns are *not mutually independent* - neighbouring columns have a high probability of being equal. This is partly due to the nat-

ural correlation in the iris (e.g. radial furrows) and partly due to the nature of the commonly used iris feature extractors, and comprehensibly analysed in the recent work "Information Theory and the IrisCode" [6]. Due to said data correlation, a block of data from an iris-code will have fewer unique values than shown in the equations above. Let ϵ denote the difference between expected number of duplicate values in an iris-code and randomly generated, mutually independent values (i.e. ϵ will be subtracted from the value of W_B in the above equation). ϵ varies depending on parameters such as the dataset itself, feature extractor and block sizes. It can be readily approximated using a training set, or potentially by a more elaborate analysis of the nature of an iris-code (see e.g. [15]).

The equations above can be used for estimation of a filter containing a single template (recall table A.2). By extension of the above reasoning, a root of a Bloom filter template tree can also be modelled as a set of unique integers. The root ($\mathbf{R}$) consists of $\mathbf{r}_1 \ldots \mathbf{r}_l$ filters and is created by taking the union of multiple Bloom filter templates. As the overall model is simplified to consider only a single item in a template, a tree root model can be denoted as shown in equation (A.5).

$$\mathbf{r}_i = \bigcup_{i=1}^{N} \mathbf{b}_i \tag{A.5}$$

It follows trivially, that $\sum_{i=1}^{N} |\mathbf{b}_i|$ is the expected number of *non-unique* items in $\mathbf{r}_i$. However, it is necessary to account for the collisions; the expected number of *unique* items in a root can be estimated as shown in equation (A.6).

$$|\mathbf{r}_i| = \left(1 - \left(1 - \frac{1}{2^{H_B}}\right)^{N*(W_B-\epsilon)}\right) * 2^{H_B} \tag{A.6}$$

As a concrete example, consider the following system configuration: $W_B = 16$, $H_B = 8$, $\epsilon = 8$ and $N = 25$. Using above equation, the expected number of unique items is $|\mathbf{r}_i| \approx 139$, corresponding to the filter being around 54% full.

The final step is estimating the overlap between $\mathbf{r}_i$ and an arbitrary, random (impostor) Bloom filter $\mathbf{b}$. This simply means computing the expected cardinality of a set intersection of these two. Let O denote this overlap:

$$O = |\mathbf{r}_i \cap \mathbf{b}| \tag{A.7}$$

The probability of the expected overlap outcome follows a hypergeometric distribution:

$$P(O{=}o) = \frac{\binom{|\mathbf{b}|}{o}\binom{2^{H_B}-|\mathbf{b}|}{|\mathbf{r}_i|-o}}{\binom{2^{H_B}}{|\mathbf{r}_i|}} \tag{A.8}$$

This distribution can be used to validate the fit of the model to real data. The mean of above distribution will be used where a single number metric of tree root filling is needed instead of an entire distribution. Let Θ denote said metric.

$$\Theta = |\mathbf{r}_i| * \frac{|\mathbf{b}|}{2^{H_B}} \tag{A.9}$$

Later on in this article, the model's correspondence with real data will be shown and it will be demonstrated how the model can be used to identify viable configurations of the Bloom filter-based system.

A.3.4 Scalable System

As has been shown earlier, the number of enrolled subjects that can be accommodated by the basic, single-tree scheme is severely limited due to overfilling of the nodes near the top of the tree. The single-tree scheme can be trivially extended to store the enrolled templates in multiple (T) trees instead of one, thereby alleviating the overfilling issue. By doing so, an arbitrarily large number of subjects can be accommodated. This system can be operated in two modes:

1. **Simple** During a lookup, the constructed trees are successively traversed. The final decision can be made either once the first candidate identity is found or once all the trees are traversed.

2. **Selective** The tree roots are utilised in a pre-selection step. From the T constructed trees, a subset of t most promising trees ($t \ll T$) is selected based on the dissimilarity scores between the probe and tree roots - $DS(\mathbf{R}_1, \mathbf{B}'), \ldots, DS(\mathbf{R}_T, \mathbf{B}')$. The t chosen trees are then traversed as described in the "simple" mode.

The penetration rate of the proposed system in worst case is deterministic as shown in equations (A.10) and (A.11). On average, a match will be found after considering approximately half of the templates/trees, thus reducing the results of below equations by a factor of 2. Observe also, that in the selective traversal scheme, this factor is likely to be larger, since trees are traversed in descending order of likelihood of finding a match.

$$p_{simple} = \begin{cases} \frac{T*(2*(\log\frac{N}{T}-1))}{N} & \text{if } T < \frac{N}{2} \\ 1.0 & \text{otherwise} \end{cases} \tag{A.10}$$

$$p_{selective} = \begin{cases} \frac{T+t*(2*(\log \frac{N}{T}-1))}{N} & \text{if } T < \frac{N}{2} \\ 1.0 & \text{otherwise} \end{cases} \tag{A.11}$$

Figure A.2 visualises above equations. The x-axis is normalised w.r.t. the number of enrolled templates. Note, that values for $^{T}/_{N}$ larger than $^{1}/_{4}$ are only plotted for completeness - they do not make much sense as a system configuration, since they correspond to performing a brute-force search among the individual Bloom filter-based templates. The theoretical penetration rates for the proposed system are very low - compare that to the naïve iris-code baseline with a penetration rate of 1.0 and 0.5 in worst and average case, respectively.

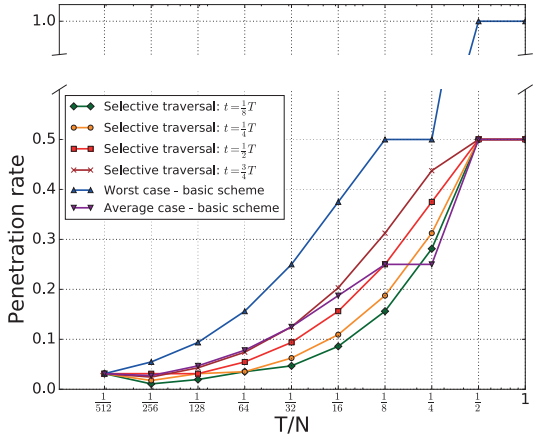


Figure A.2: Lookup in the Bloom filter-based system

Finally, recall (subsection A.3.2) that when a tree is traversed, *both* child node scores need to be calculated and compared at every level. One can take advantage of the fact that genuine score sequences are expected to decrease as the tree is traversed and make a quick decision about the traversal direction. A simple heuristic is proposed: calculate the score for *one* child node first and if the score of the first child node is lower than its parent, then this node is immediately selected as the correct traversal direction. While the factor of 2 is not entirely eliminated, its non-deterministic decrease is substantial.

A.3.5 Multi-Iris Indexing

In this section, an early study into feasibility of multi-iris indexing is presented. Although, in itself, multi-iris biometrics (possible due to the mutual

independence of irides from the left/right eye of the same subject) is not a new idea, best to the authors' knowledge, it is the first such attempt in the scientific literature for biometric indexing schemes. The proposed approach is based on the Bloom filter template representation and a feature level fusion of two irides from a single subject upon enrolment and prior to an authentication transaction. Let $\mathbf{B}_M$ denote a multi-iris template of a subject. It consists of the left eye template ($\mathbf{B}_L$) and the right eye template ($\mathbf{B}_R$) fused together, as shown in equation (A.12). In the concrete implementation, this is simply an element-wise union of two Bloom filter sets.

$$\mathbf{B}_M = \mathbf{B}_L \cup \mathbf{B}_R \tag{A.12}$$

With this scheme, one can expect a drop in the false positive identification rate and an increase in the true positive identification rate - overall a biometric performance superior to that of a standard Bloom filter-based system, which uses only a single iris per subject. Observe also, that the multi-iris system works in precisely the same way as described earlier in section A.3; the only change and additional, negligible, computational cost is the template fusion.

A.4 Experimental Setup

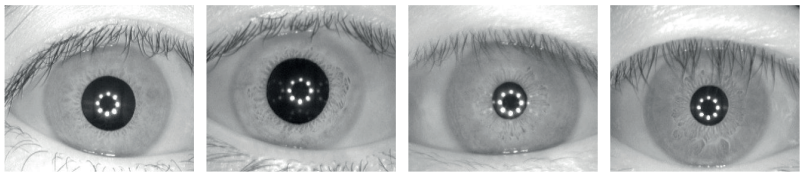
Table A.3 shows the datasets selected for the system evaluation in this work, while figure A.3 shows example images from the chosen datasets.

Table A.3: Evaluation dataset overview

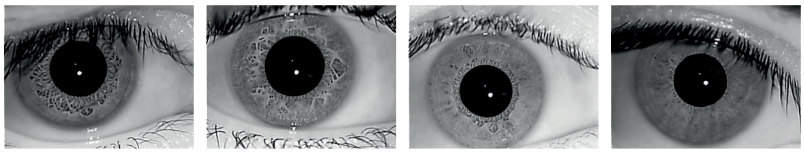
Dataset	Instances	Images	Resolution	Av. iris diameter	Quality
CASIAv4-Interval [3]	395	2639	320x280 px	∼210 px	High
IITDv1 [17]	448	2240	320x240 px	∼205 px	High
Biosecure [20]	420	1680	640x480px	∼225 px	Medium
ND-Iris-Template-Aging[1] [7]	214	5377	640x480px	∼250 px	High

Above data has been merged into one large dataset. The purpose of this is a higher number of enrolled subjects and a larger number of impostor transactions compared to the proof-of-concept study [23]. Henceforth, this dataset will be referred to as "Combined". The CASIAv4-Thousand dataset was also considered, but was not used, since only just over half of the segmented images had more than 70% usable iris area (metric defined in [13]); furthermore, in many images, the subjects are wearing glasses. Thus, in lieu of the feasibility of high-quality data acquisition nowadays and a specialised work concerning degraded data (e.g. [21]), it was deemed more interesting to work with biometric data of high quality. Nevertheless, some images had

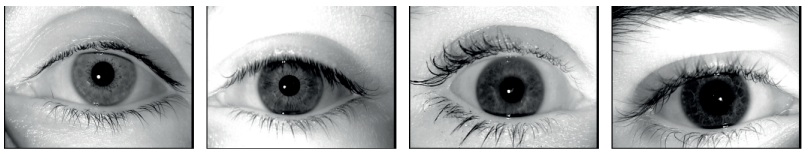
[1]Only a single point in time from the dataset was used.



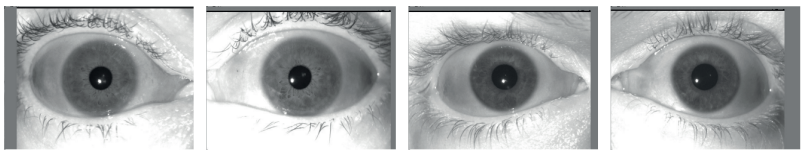
(a) CASIAv4-Interval



(b) IITDv1



(c) Biosecure



(d) ND-Iris-TemplateAging

Figure A.3: Example images from the datasets

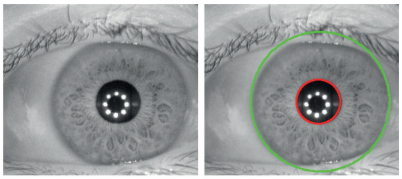
to be excluded from the experiments due to the datasets containing duplicate, identical images and labelling errors. The dataset was then split into disjoint groups as shown in table A.4. The IITD and CASIA images were used for enrolment and genuine comparisons as well as impostor comparisons, while the Biosecure and Ageing images were used exclusively for impostor comparisons.

Table A.4: Dataset split (templates) for the experiments

Dataset	Enrolled	Genuine	Impostor
Combined	512	2240	9073
Combined multi-iris	256	824	3534

The images were processed using the commonly used method, whose

steps are illustrated in figure A.4. The technique generates iris-codes of size $512 \times 20 = 10{,}240$ bits.



(a) Detection

(b) Enhanced texture

(c) 1D Log-Gabor

(d) Quadratic Spline Wavelet

Figure A.4: Iris recognition processing chain: (a) iris detection in the raw image, (b) normalized pre-processed iris texture, and (c)-(d) iris-codes of applied feature extractor. Image taken from CASIA-v4-Interval iris database [3].

Using thus produced iris-codes, the conducted performance-related experiments, in which the maximum possible number of genuine and impostor transactions were performed to ensure robustness of the results, were as follows:

- The baseline - an iris-code based system performing a brute-force search in the identification scenario.
- The basic Bloom filter system as described in subsection A.3.2.
- The proposed improvements to the basic Bloom filter scheme presented in subsection A.3.4.

The metrics used for system evaluation were:

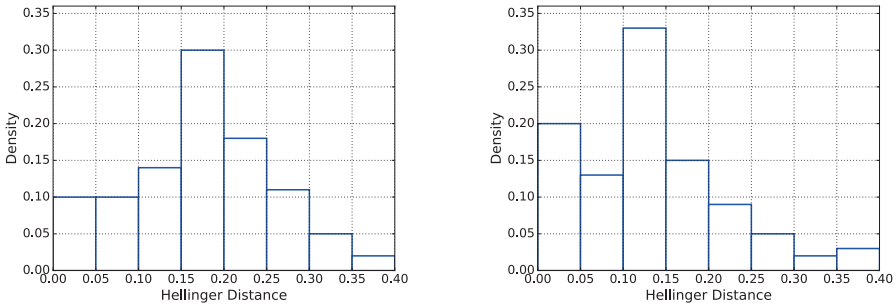
- Biometric performance: True positive identification rate (TPIR) and false positive identification rate (FPIR) ROC curve. Additionally, a fixed point on the ROC curve corresponding to TPIR at 0.1% FPIR (denoted $TP_{0.1}$).
- Workload: $\mathcal{W}$ and $\bar{F}$ as defined in subsection A.2.3.

A.5 Results

In this section, the proposed statistical model and Bloom filter system improvements are evaluated.

A.5.1 Parameter Estimation

A simple metric is used to measure whether or not the model is a reasonable representation of the real iris data. The Hellinger distance quantifies the dissimilarity between two distributions with a single value between 0 and 1, where small values indicate a good fit. To assess the general applicability of the model, the value of this metric was computed for all relevant system configurations. As figure A.5 reveals, the model corresponds quite well to the real data.



(a) LG (b) QSW

Figure A.5: Fit between the model and real data for all the relevant system configurations

With the model accuracy established, we turn attention to the Θ metric described earlier. Figure A.6 shows the correlation between true positive identification rate and the Θ metric from the model for all relevant system configurations (block size, number of constructed and traversed trees). Observe that $TP_{0.1}$ generally declines when Θ increases. Extrapolating from these results, the model could be used to instantly assess viability of an arbitrary system configuration by merely calculating Θ. Doing so could potentially be immensely time-saving, since only the likely viable configurations pointed out by the model would have to be tested empirically.

In other words, the true positive accuracy depends on the degree to which the Bloom filters are filled at the top levels of the tree. Table A.5, created from empirical data, illustrates the Bloom filter filling in the *single-tree* system for the Combined dataset. Observe that at the top tree levels, the

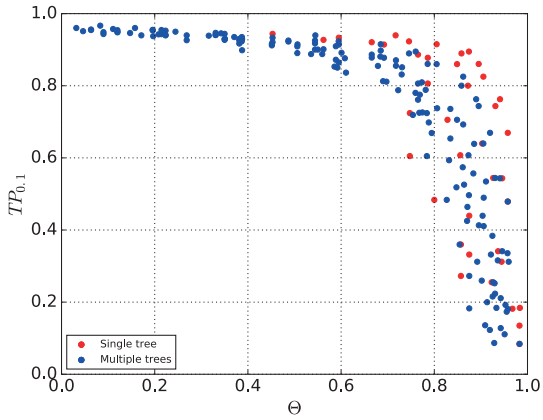


Figure A.6: The correlation between biometric performance and top tree level node filling

filters are nearly full for almost all relevant system configurations. This thus severely affects the ability to make correct tree traversal direction decisions and has a negative impact on the system performance in terms of TPIR (see next subsection).

Table A.5: % of bits set to 1 at the top levels of the basic, single-tree system (level 0 is the tree root)

Configuration		**LG**		**QSW**	
		Tree level			
H_B	W_B	0	1	0	1
8	8	98.10	95.38	99.63	98.48
8	16	99.67	98.74	99.97	99.79
8	32	99.99	99.79	100.00	99.99
10	8	78.39	69.20	87.51	78.81
10	16	90.73	84.38	96.57	92.08
10	32	97.29	94.21	99.42	98.06
12	8	40.01	32.40	48.36	39.29
12	16	57.86	48.73	68.78	58.58
12	32	75.85	67.06	86.15	77.79

A.5.2 Performance Evaluation

The baseline score for the datasets is established with a brute-force, iris-code identification system with alignment compensation of ± 8 bits, correspond-

ing to approximately $\pm 5.625°$. The total workload for a lookup in the baseline system is $\mathcal{W} = 8.91 * 10^7$ and $\mathcal{W} = 4.46 * 10^7$ in the worst and average case, respectively. As expected with high-quality data, the system achieves excellent $TP_{0.1}$ rates: 98.03% and 97.90% for the LG and QSW feature extractor, respectively.

Subsequently, the proposed Bloom filter-based system was tested in identification mode with varying configurations and improvements (block width, height, number of trees constructed, number of trees traversed etc.). Table A.6 shows results achieved by the basic Bloom filter system, with a single and multiple trees constructed. Observe, how construction of multiple trees increases the biometric performance, albeit at the cost of heavier workload in relation to the single-tree system.

Table A.6: The results of the 3 configurations with best performance in the single and multiple tree schemes

Configuration			**Workload per lookup**				**Extractor**	
H_B	W_B	T	**Worst case**		**Average case**		**LG**	**QSW**
			$\mathcal{W}$	$\bar{F}$	$\mathcal{W}$	$\bar{F}$		
10	8	1	$2.10 * 10^6$	0.0235	$2.10 * 10^6$	0.0471	91.19%	89.97%
12	8	1	$8.39 * 10^6$	0.0941	$8.39 * 10^6$	0.1882	94.68%	93.64%
12	16	1	$4.19 * 10^6$	0.0471	$4.19 * 10^6$	0.0941	94.32%	93.43%
8	16	16	$2.10 * 10^6$	0.0235	$1.05 * 10^6$	0.0235	96.87%	97.18%
10	16	16	$8.39 * 10^6$	0.0941	$4.19 * 10^6$	0.0941	97.19%	97.16%
12	16	16	$3.36 * 10^7$	0.3765	$1.68 * 10^7$	0.3765	97.48%	97.50%

To resolve the matter of increased workload due to construction of multiple trees, the pre-selection of t promising trees for traversal is applied (see subsection A.3.4). In table A.7, it can be seen that this technique allows to decrease the workload tremendously - for instance, traversing $^1/_4$ of the constructed trees essentially maintains the performance of a full traversal. Further workload reduction is possible (even traversing only a single most promising tree), albeit for most system configurations the biometric performance decrease is then significant.

The workload reduction effects of the quick traversal direction selection heuristic (abbreviated "qd") can be taken advantage of without loss in biometric performance. For the viable system configurations (i.e. with low Θ) the biometric performance of the scheme with and without the heuristic applied are virtually indistinguishable. Further biometric performance improvement can be obtained using the multi-iris indexing scheme with selective tree traversal. In experiments, we observe for many system configurations $TP_{0.1} > 99\%$, which is better than the single-iris baseline and comparable to a multi-iris baseline with a score level fusion scheme.

A summary of the results of well-performing configurations of each proposed system version is presented in figure A.7 and table A.8. It can be seen,

Table A.7: The results of the Bloom filter scheme with selective tree traversal

Configuration				Workload per lookup				Extractor	
				Worst case		Average case		LG	QSW
H_B	W_B	T	t	$\mathcal{W}$	F	$\mathcal{W}$	F		
8	16	16	8	$1.31 * 10^6$	0.0147	$7.86 * 10^5$	0.0176	96.76%	96.95%
10	16	16	8	$5.24 * 10^6$	0.0588	$3.15 * 10^6$	0.0706	97.17%	97.50%
12	16	16	8	$2.10 * 10^7$	0.2353	$1.26 * 10^7$	0.2824	97.50%	97.31%
8	16	16	4	$7.86 * 10^5$	0.0088	$5.24 * 10^5$	0.0118	96.22%	95.53%
10	16	16	4	$3.15 * 10^6$	0.0353	$2.10 * 10^6$	0.0471	97.05%	97.45%
12	16	16	4	$1.26 * 10^7$	0.1412	$8.39 * 10^6$	0.1882	97.42%	97.45%
8	16	16	2	$5.24 * 10^5$	0.0059	$3.93 * 10^5$	0.0088	94.05%	93.21%
10	16	16	2	$2.10 * 10^6$	0.0235	$1.57 * 10^6$	0.0353	96.79%	96.91%
12	16	16	2	$8.39 * 10^6$	0.0941	$6.29 * 10^6$	0.1412	97.32%	97.31%
8	16	16	1	$3.93 * 10^5$	0.0044	$3.93 * 10^5$	0.0088	90.58%	88.15%
10	16	16	1	$1.57 * 10^6$	0.0176	$1.57 * 10^6$	0.0353	95.95%	95.62%
12	16	16	1	$6.29 * 10^6$	0.0706	$6.29 * 10^6$	0.1412	97.15%	96.56%

that when benchmarked with the iris-code baseline, the proposed schemes greatly reduce the necessary lookup workload in the identification mode, while maintaining a very high biometric performance.

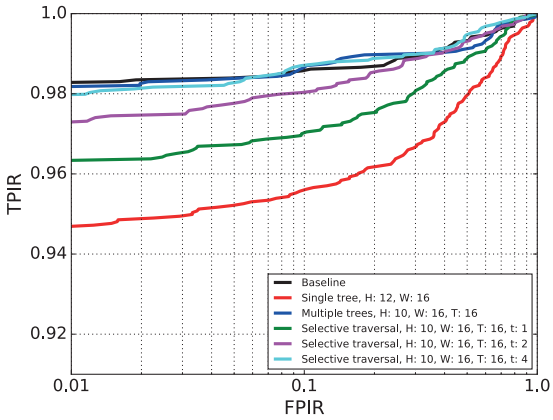
Table A.8: A summary of the results for various system improvements

System/Improvement	Biometric performance ($TP_{0.1}$)	Worst case workload per lookup (F)
Baseline	$97.90 - 98.03\%$	1.0
Single tree	$89.97 - 94.68\%$	$0.0235 - 0.0941$
Multiple trees	$96.87 - 97.50\%$	$0.0235 - 0.3765$
Selective traversal	$93.21 - 97.50\%$	$0.0059 - 0.1412$
Multi-iris	$> 99\%$	Same as selective traversal
Quick traversal decision	Same as selective traversal	$0.0043 - 0.1264$

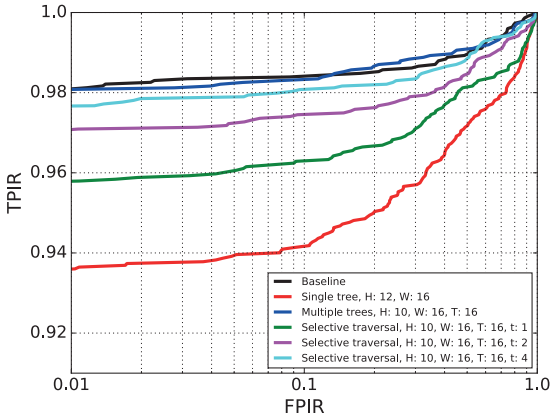
A.5.3 Discussion

In order to accommodate a larger number of enrolled templates, the basic Bloom filter-based approach was expanded with construction of multiple search trees and efficient pre-selection of the trees to traverse upon lookup.

The empirical experiments on the Combined dataset with 512 enrollees show, that the system is capable of achieving biometric performance comparable with the naïve baseline implementation and better than the basic, single-tree Bloom filter-based implementation. This performance has been achieved by having to traverse only $^1/_4$ or less of the constructed search trees. It can be reasonably assumed, that this property will hold when more subjects are enrolled (and thereby more trees built), thus maintaining the low workload requirements for larger systems. This assumption is reasonable, since in order to be in the worst $^3/_4$ of the trees, a genuine probe tem-



(a) LG



(b) QSW

Figure A.7: ROC curves comparison for different system versions

plate would have to be of very bad quality - an issue, which can be effectively eliminated by a sampling quality check.

The overall workload in the identification mode is vastly reduced: in several configurations, $F < 0.05$ with only a minor impact on the biometric performance. The performance and workload reduction achieved by our system are on par with, or exceed the current state of the art (table A.1).

The numerous available system configurations offer significant flexibility in adjusting the biometric performance and workload to the individual needs of a given application. This is beneficial in terms of feasibility for real-world deployments, since their varying requirements can be accommodated

by fine-tuning the system parameters. In addition to the empirical results, the statistical model can instantly assess whether or not a given system configuration is likely to perform well, thus decreasing the number of required empirical tests.

Most importantly, however, we have shown that the improved Bloom filter-based indexing scheme is scalable both in terms of the biometric performance and the workload. In other words, the low workload and high biometric performance can be expected for arbitrarily sized systems. This is a key point differentiating our approach from the majority of the related works (table A.1), which for most part offer no scalability analysis. A direct benchmark in terms of biometric performance and workload is infeasible, partly due to different datasets used and partly due to the wildly varying result reporting methodologies across the current literature. This is the chief reason for our proposal of a standardisation of biometric workload reduction reporting methodology (subsection A.2.3).

A.6 Conclusion and Future Research

The global interest for biometrics has been steadily growing and several large-scale deployments have appeared around the world. The current trends make biometric workload reduction a relevant and attractive research area. In a recent interview, Daugman, the inventor of iris recognition, has stated that performing accurate and efficient biometric identification (e.g. by means of indexing, rather than exhaustive search) is one of the important, unsolved issues in the biometrics field in general [11].

In this article, we expanded on a recently proposed biometric indexing approach based on Bloom filters and binary search trees. Several improvements were proposed; with these in place, the system is capable of supporting an arbitrarily large number of enrollees. The biometric workload in identification scenario is greatly reduced (below 10%), while a high (¿98% TPIR at 0.1% FPIR) biometric performance of a baseline system is maintained. Additionally, in a proof-of-concept study, the Bloom filter-based system is shown to be capable of accurate and efficient multi-iris indexing, which shows promise for future research. Experiments on a larger dataset, as well as a study into multi-biometric capabilities of the system and combination with soft biometric traits are planned.

Furthermore, due to differences in methods for workload reduction reporting in the scientific community, we proposed a set of prerequisites for a more transparent and unified reporting methodology. We encourage ISO/IEC JTC 1/SC 37 Biometrics to take up work on a standardised framework for biometric workload reduction reporting, using the prerequisites proposed by us as a starting point for discussions and deliberations.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF) as well as by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP).

A.7 Bibliography

[1] BLOOM, B. H. Space/time trade-offs in hash coding with allowable errors. *Communications of the ACM 13*, 7 (July 1970), 422–426.

[2] BRODER, A., AND MITZENMACHER, M. Network applications of Bloom filters: A survey. *Internet Mathematics 1*, 4 (2004), 485–509.

[3] CHINESE ACADEMY OF SCIENCES' INSTITUTE OF AUTOMATION. CASIA iris image database. `http://biometrics.idealtest.org/`, December 2010. Last accessed: 2020–03–11.

[4] DAUGMAN, J. Biometric decision landscapes. Tech. Rep. UCAM-CL-TR-482, University of Cambridge - Computer Laboratory, January 2000.

[5] DAUGMAN, J. How iris recognition works. *Transactions on Circuits and Systems for Video Technology (TCSVT) 14*, 1 (January 2004), 21–30.

[6] DAUGMAN, J. Information theory and the IrisCode. *Transactions on Information Forensics and Security (TIFS) 11*, 2 (February 2016), 400–409.

[7] FENKER, S. P., AND BOWYER, K. W. Analysis of template aging in iris biometrics. In *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (June 2012), IEEE, pp. 45–51.

[8] GADDE, R. B., ADJEROH, D., AND ROSS, A. Indexing iris images using the Burrows-Wheeler transform. In *International Workshop on Information Forensics and Security (WIFS)* (December 2010), IEEE, pp. 1–6.

[9] GENTILE, J. E., RATHA, N., AND CONNELL, J. An efficient, two-stage iris recognition system. In *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2009), IEEE, pp. 211–215.

[10] HAO, F., DAUGMAN, J., AND ZIELINSKI, P. A fast search algorithm for a large fuzzy database. *Transactions on Information Forensics and Security (TIFS) 3*, 2 (June 2008), 203–212.

[11] INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. Biometric council newsletter. `http://ieee-biometrics.org/images/pdf/Newsletter_Nov_2015_corrected.pdf`, November 2015. Last accessed: 2020–03–11.

[12] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 19795-1:2006. Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*. International Organization for Standardization and International Electrotechnical Committee, April 2006.

[13] ISO/IEC JTC1 SC37 BIOMETRICS. *ISO/IEC 29794-6:2015. Information technology – Biometric sample quality – Part 6: Iris image data*. International Organization for Standardization and International Electrotechnical Committee, July 2015.

[14] JAYARAMAN, U., PRAKASH, S., AND GUPTA, P. An efficient color and texture based iris image retrieval technique. *Expert Systems with Applications 39*, 5 (April 2012), 4915–4926.

[15] KONG, A. W. K., ZHANG, D., AND KAMEL, M. S. An analysis of IrisCode. *Transactions on Image Processing (TIP) 19*, 2 (February 2010), 522–532.

[16] KONRAD, M., STÖGNER, H., UHL, A., AND WILD, P. Computationally efficient serial combination of rotation-invariant and rotation compensating iris recognition algorithms. In *International Conference on Computer Vision Theory and Applications (VISAPP)* (May 2010), SciTePress, pp. 85–90.

[17] KUMAR, A., AND PASSI, A. Comparison and combination of iris matchers for reliable personal authentication. *Pattern Recognition 43*, 3 (March 2010), 1016–1026.

[18] MEHROTRA, H., SRINIVAS, B. G., AND MAJHI, B. Indexing iris biometric database using energy histogram of DCT subbands. In *International Conference on Contemporary Computing (IC3)* (August 2009), vol. 40, Springer, pp. 194–204.

[19] MUKHERJEE, R., AND ROSS, A. Indexing iris images. In *International Conference on Pattern Recognition (ICPR)* (December 2008), IEEE, pp. 1–3.

[20] ORTEGA-GARCIA, J., FIERREZ, J., ALONSO-FERNANDEZ, F., GALBALLY, J., FREIRE, M. R., ET AL. The multiscenario multienvironment BioSecure multimodal database (BMDB). *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 32*, 6 (June 2010), 1097–1111.

[21] PROENÇA, H. Iris biometrics: Indexing and retrieving heavily degraded data. *Transactions on Information Forensics and Security (TIFS) 8*, 12 (December 2013), 1975–1985.

[22] QIU, X., SUN, Z., AND TAN, T. Global texture analysis of iris images for ethnic classification. *International Conference on Biometrics (ICB) 3832* (January 2006), 411–418.

[23] RATHGEB, C., BREITINGER, F., BAIER, H., AND BUSCH, C. Towards bloom filter-based indexing of iris biometric data. In *International Conference on Biometrics (ICB)* (May 2015), IEEE, pp. 422–429.

[24] RATHGEB, C., BREITINGER, F., BUSCH, C., AND BAIER, H. On application of Bloom filters to iris biometrics. *IET Biometrics 3*, 4 (December 2014), 207–218.

[25] RATHGEB, C., UHL, A., AND WILD, P. Incremental iris recognition: A single-algorithm serial fusion strategy to optimize time complexity. *International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (September 2010), 1–6.

[26] ROSS, A., AND SUNDER, M. S. Block based texture analysis for iris classification and matching. In *Conference on Computer Vision and Pattern Recognition - Workshops (CVPRW)* (June 2010), IEEE, pp. 30–37.

[27] SUN, Z., ZHANG, H., TAN, T., AND WANG, J. Iris image classification based on hierarchical visual codebook. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 36*, 6 (June 2014), 1120–1133.

[28] UNIQUE IDENTIFICATION AUTHORITY OF INDIA. Aadhaar dashboard. `https://www.uidai.gov.in/aadhaar_dashboard/`. Last accessed: 2020–03–11.

Nomenclature

d'	Sensitivity/Decidability index
CCR	Correct classification rate
CMC	Cumulative match characteristic
DET	Detection error trade-off
DS	Dissimilarity
EER	Equal-error rate
FMR	False match rate
FNIR	False-negative identification-error rate
FNMR	False non-match rate
FPIR	False-positive identification-error rate
FTA	Failure-to-acquire rate
HD	Hamming distance
HR	Hit rate
IS	International standard
ISO/IEC	International Organization for Standardization and the International Electrotechnical Commission
PR	Penetration rate
ROC	Receiver operating characteristic
RR-1	Rank-1 identification rate
TPIR	True-positive identification rate

ISBN 978-82-326-4578-7 (printed version)
ISBN 978-82-326-4579-4 (electronic version)
ISSN 1503-8181

NTNU
Norwegian University of
Science and Technology

Pawel Drozdowski

Doctoral Thesis

Doctoral theses at NTNU, 2020:115

Pawel Drozdowski

Efficient privacy-preserving biometric identification in large-scale multibiometric systems

NTNU
Norwegian University of
Science and Technology
Faculty of Information Technology
and Electrical Engineering
Department of Information Security
and Communication Technology

Doctoral theses at NTNU, 2020:115

NTNU

NTNU
Norwegian University of
Science and Technology