

Doctoral thesis

Doctoral theses at NTNU, 2021:111

Ruven Pillay

Multispectral and Hyperspectral Imaging of Art

Quality, Calibration and Visualization

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Information Technology and Electrical
Engineering
Department of Computer Science



Norwegian University of
Science and Technology

Ruven Pillay

Multispectral and Hyperspectral Imaging of Art

Quality, Calibration and Visualization

Thesis for the Degree of Philosophiae Doctor

Gjøvik, April 2021

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science



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 Computer Science

© Ruven Pillay

ISBN 978-82-326-6787-1 (printed ver.)

ISBN 978-82-326-6204-3 (electronic ver.)

ISSN 1503-8181 (printed ver.)

ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2021:111

Printed by NTNU Grafisk senter

Abstract

Multispectral and hyperspectral imaging can be powerful tools for analyzing and documenting works of art due to their ability to simultaneously capture both accurate spectral and spatial information. The data can be used for a wide range of diagnostic and analytical purposes, including materials identification, pigment mapping, the detection of hidden features or areas of lost material, for colorimetric analysis or for precise quantitative documentation.

However, a number of technical challenges exist which have prevented multispectral and hyperspectral imaging from realizing their full potential and which have prevented the technologies from becoming more widely used and routine analytical tools. Both multispectral and hyperspectral imaging systems require careful and precise acquisition workflows in order to produce useful data. In addition, processing and calibration of the acquired data can be a challenge for many cultural heritage users. Hyperspectral imaging, in particular, can produce vast quantities of raw data that require complex processing and the ability to manage the large resulting volumes of data. Moreover, the final high resolution and multidimensional data that is produced can be difficult to use or to visualize.

This thesis, therefore, seeks to address some of these issues and seeks to analyze and quantify potential problems and then propose tools, workflows and methodologies to resolve and mitigate them. The research presented here focuses on two main areas. The first research area concerns the quality of spectral data and how to measure, quantify and improve it. To do so, it is necessary to first establish exactly what spectral quality is and what methods can be used to quantify it. These methods are then applied to ascertain the levels of quality seen in data acquired under routine operating conditions with an evaluation of data from an extensive round-robin test of hyperspectral imaging systems.

In order to improve the quality of spectral data, the various elements that contribute to and affect spectral quality within a system are then analyzed. Ac-

quisition and calibration pipelines are then defined for both multispectral and hyperspectral equipment with practical guidelines and workflows provided that aim to help users produce the best quality data possible.

The second research area concerns the visualization of such data and examines ways to facilitate and make large and complex image data accessible online. For this, an architecture, visualization techniques and a full software platform are presented for the efficient distribution and visualization of high resolution multi-modal and multispectral or hyperspectral image data. This work is then extended in order to push the technology to the limits and to apply the techniques to the field of astronomy where image sizes are at their most extreme.

Translation of the Abstract in Norwegian

Multispektral og hyperspektral avbildning kan være kraftige verktøy for å analysere og dokumentere kunstverk, på grunn av dets evne til samtidig å fange både nøyaktig spektral og romlig informasjon. Dataene kan brukes til et bredt spekter av diagnostiske og analytiske formål, inkludert materialidentifikasjon, pigmentkartlegging, påvisning av skjulte egenskaper eller områder med tappt materiale, for kolorimetrisk analyse eller for presis kvantitativ dokumentasjon.

Imidlertid eksisterer det en rekke tekniske utfordringer som har forhindret multispektral og hyperspektral avbildning fra å utnytte sitt fulle potensiale, og som har forhindret teknologiene fra å bli mer brukt og fra å bli standard analytiske verktøy. Både multispektrale og hyperspektrale bildesystemer krever nøyte og presise avbildningsarbeidsflyter for å produsere nyttige data. I tillegg kan behandling og kalibrering av innhentede data være en utfordring for mange brukere i kulturarvsektoren. Spesielt hyperspektral avbildning kan produsere store mengder rådata som krever kompleks behandling og muligheten til å håndtere de store resulterende datamengdene. Videre kan det endelige høyoppløselige og flerdimensjonale datasettet som produseres være vanskelig å bruke eller å visualisere.

Denne avhandlingen søker derfor å adressere noen av disse problemene og søker å analysere og kvantifisere potensielle problemer, og deretter foreslå verktøy, arbeidsflyter og metoder for å redusere og løse dem. Forskingen som presenteres her fokuserer på to hovedområder. Det første forskningsområdet gjelder selve kvaliteten på spektraldataene og hvordan man måler, kvantifiserer og forbedrer denne kvaliteten. For å gjøre dette er det nødvendig å først fastslå nøyaktig hva spektral kvalitet er og hvilke metoder som kan brukes til å kvantifisere den. Disse metodene blir deretter brukt for å fastslå kvalitetsnivåene til i datasett innhentet under rutinemessige driftsforhold, gjennom en evaluering av data fra en omfattende «round-robin test» av hyperspektrale bildesystemer.

For å forbedre kvaliteten på spektraldata analyseres de forskjellige elementene som bidrar til og påvirker spektralkvaliteten i et system. Avbildnings- og kalibreringssarbeidsflyter defineres deretter for både multispektralt og hyperspektralt utstyr, med foreslåtte praktiske retningslinjer og arbeidsflyter for at brukerne skal produsere data med best mulig kvalitet.

Det andre forskningsområdet gjelder visualisering av slike data og undersøker måter å legge til rette for og gjøre store og komplekse bildedata tilgjengelige online. For dette presenteres en arkitektur, visualiseringsteknikker og en full programvareplattform for effektiv distribusjon og visualisering av høyoppløselige

multimodale og multispektrale eller hyperspektrale bildedata. Dette arbeidet er også utvidet for å presse teknologien til det ytterste og til å anvende teknikkene på astronomi-feltet der bildestørrelsene er på sitt mest ekstreme.

Foreword

The research presented in this thesis was carried out after having already worked for many years on spectral imaging within the museum sector, first at the National Gallery in London and, subsequently at the C2RMF, the national research and restoration laboratories housed within the Louvre museum in Paris. The undertaking of this PhD allowed strands from this work to be brought together in a coherent manner and enabled the research to be extended and pushed in a number of new directions.

The space afforded to carry out this work at the C2RMF is gratefully acknowledged and I would like to thank those who helped make this research possible and who gave their support. In particular, I would like to thank Christian Lahanier whose visionary and inspirational work allowed multispectral imaging to be first brought to the laboratory and Morwena Joly whose belief and persistence later made hyperspectral imaging a reality there.

I would also like to acknowledge the various projects that helped provide data for this thesis and the co-authors with whom the papers within this thesis were written. I also gratefully acknowledge the support of NTNU Gjøvik and of the Norwegian Colour and Visual Computing Laboratory.

I would especially like to express my gratitude to my PhD supervisors. To my co-supervisor, Sony George and, in particular, to Jon Y. Hardeberg, who showed remarkable and determined faith in my work and in this PhD. This PhD would not have been completed without his stoic perseverance and steadfast support throughout.

Finally I would like to thank my wife for her boundless patience and wholehearted encouragement during this undertaking.

Contents

Abstract	i
Translation of the Abstract in Norwegian	iii
Foreword	v
Contents	x
List of Abbreviations	xi
I Introduction	1
1 Introduction	3
1.1 Motivation	3
1.2 Research Aims & Questions	4
1.3 List of Contributing Papers	5
2 Background	9
2.1 Introduction	9
2.2 Scientific Imaging	10

2.3	Spectral Imaging	11
2.3.1	Multispectral Imaging	12
2.3.2	Hyperspectral Imaging	12
2.3.3	Ultraspectral Imaging	14
2.3.4	Scanning Methods	14
2.3.5	Detector Technologies	16
2.4	Spectral Imaging within Cultural Heritage	17
2.4.1	Multispectral Imaging of Art	17
2.4.2	Hyperspectral Imaging of Art	20
2.5	Spectral Image Quality	22
2.5.1	Noise	22
2.5.2	Measuring Spectral Differences	24
2.5.3	Round-Robin Tests	27
2.6	Spectral Image Calibration	30
2.6.1	Radiometric Calibration	31
2.6.2	Geometric Calibration	32
2.7	Classification	33
2.8	Data Structures & Visualization	34
2.8.1	Spectral Data Structures	34
2.8.2	Visualization of Very Large Image Data	37
2.8.3	Optimizing Performance for Very Large Images	38
2.8.4	Compression	39
3	Summary of Papers	41
3.1	Introduction	41
3.2	Paper 1: Quality Evaluation in Spectral Imaging - Quality Factors and Metrics	43

3.3	Paper 2: Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems	44
3.4	Paper 3: Studying That Smile: A Tutorial on Multispectral Imaging of Paintings	46
3.5	Paper 4: Hyperspectral Imaging of Art: Acquisition and Calibration Workflows	48
3.6	Paper 5: Multi-Modal Data Visualization and Analysis of “The Bedroom at Arles” by Vincent van Gogh	50
3.7	Paper 6: Web-Based Visualization of Very Large Scientific Astronomy Imagery	52
4	Discussion	55
4.1	Research Contributions	56
4.1.1	How can the quality of spectral data be measured? What is the quality of acquired experimental spectral data? . .	56
4.1.2	What are the factors that influence multispectral and hyperspectral data quality and how can acquisition and calibration workflows be improved?	58
4.1.3	How to make high resolution quantitative image data accessible? How can large volumes of spectral and other scientific imaging data be visualized?	59
4.2	Scope of the Research	61
4.2.1	Breadth of the Research	61
4.2.2	Chronology and Timeline	62
4.2.3	Research Impact	63
5	Conclusion	65
5.1	Perspectives	66
	Bibliography	67

II Original Papers	87
Paper 1: Quality Evaluation in Spectral Imaging - Quality Factors and Metrics	89
Paper 2: Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems	105
Paper 3: Studying That Smile: A Tutorial on Multispectral Imaging of Paintings	133
Paper 4: Hyperspectral Imaging of Art: Acquisition and Calibration Workflows	147
Paper 5: Multi-Modal Data Visualization and Analysis of “The Bedroom at Arles” by Vincent van Gogh	159
Paper 6: Web-Based Visualization of Very Large Scientific Astronomy Imagery	167

List of Abbreviations

2D Two Dimensional. 15, 34, 35

3D Three Dimensional. 35, 50, 67

BIL Band Interleaved Line. 36

BIP Band Interleaved Pixel. 36

BRDF Bidirectional Reflectance Distribution Function. 27

BSQ Band Sequential. 35, 36

C2RMF Centre de Recherche et de Restauration des Musées de France. v, 20

CCD Charge-Coupled Device. 16, 19

CIE Commission Internationale de l'Eclairage. 36

CMOS Complementary Metal Oxide Semiconductor. 16

IEEE Institute of Electrical and Electronics Engineers. 36, 66

IIIF International Image Interoperability Framework. 63

InSb Indium Antimonide. 16

LCTF Liquid Crystal Tunable Filter. 14, 18

LED Light Emitting Diode. 19

LWIR Long Wave Infrared. 11, 12

- MCT** Mercury Cadmium Telluride. 16, 21
- MIR** Mid Infrared. 12, 21
- MNF** Maximum Noise Fraction. 34
- MSE** Mean Square Error. 24, 25
- NIR** Near Infrared. 12, 16, 18, 19, 23
- PCA** Principal Component Analysis. 33, 34
- PPI** Pixel Purity Index. 34
- PSNR** Peak Signal-to-Noise Ratio. 25
- PTFE** Polytetrafluoroethylene. 29
- SAM** Spectral Angle Mapper. 26
- sCMOS** Scientific CMOS. 16
- SNR** Signal to Noise Ratio. 23–25, 44
- SWIR** Short Wave Infrared. 11, 12, 16, 19, 21, 23, 44
- UV** Ultra-Violet. 12, 29
- VIS** Visible. 12
- VNIR** Visible - Near Infrared. 11, 12, 16, 21, 23, 44

Part I

Introduction

Chapter 1

Introduction

1.1 Motivation

Spectral imaging techniques such as multispectral and hyperspectral imaging have become established analytical techniques in a wide range of fields including remote sensing [9], food science [10], astronomy [11], mineralogy [12], agriculture [13], medicine [14], archaeology [15] and many others. Both multispectral and hyperspectral imaging are able to simultaneously capture both spectral and spatial information and this combination can provide valuable and unique information that can be used for a number of purposes. Spectral imaging is able to produce data that is *quantitative* and that can be applied to a number of analytical tasks. The technique is sometimes referred to as “imaging spectroscopy” or “chemical imaging” due to its ability to distinguish or identify materials and is commonly used to map the distribution and abundance of materials as well as for other purposes such as measuring change or for performing accurate colorimetric measurements.

Thanks to this versatility, spectral imaging has become an increasingly used technique in the field of cultural heritage and considerable progress has been made since the introduction of filter-based multispectral imaging systems able to capture a handful of spectral bands in the mid-1990’s [16]. The development of hyperspectral imaging enables far higher spectral resolutions to be acquired with typical pushbroom hyperspectral cameras able to capture hundreds of narrow contiguous spectral bands and this technology was soon applied to the study of paintings [17]. The rich combination of spatial and spectral information can be a powerful tool for analyzing and documenting works of art and

can be applied to a wide range of diagnostic and analytical purposes within the field. These include materials identification, pigment mapping, the detection of hidden features or areas of lost material, for colorimetric analysis or for precise quantitative documentation.

Despite the wider availability and use of multispectral and hyperspectral imaging, they have yet to become routine analytical tools within the field of cultural heritage and a number of technical challenges have prevented the technology from realizing its full potential. Both multispectral and hyperspectral imaging require careful and precise acquisition workflows. Processing and calibration of the acquired data requires complex software along with the expertise and computing infrastructure that goes with it. Hyperspectral imaging, in particular, can produce vast quantities of data that require complex processing and the ability to manage the large resulting volumes of data. Moreover, the final high resolution and multidimensional data that is produced can be difficult to manipulate or to visualize.

This thesis, therefore, seeks to address some of these issues and aims to analyze and quantify potential problems and then propose tools, workflows and methodologies to resolve and mitigate them.

1.2 Research Aims & Questions

The research presented in this thesis covers two core areas encompassing three research questions. The first concerns the quality of spectral data and how to measure, quantify and improve it. The second research area concerns the visualization of such data and examines ways to facilitate and make large and complex image data accessible online.

From this, three specific research questions can be defined that the work presented in this thesis is intended to answer:

Research Question: How can the quality of spectral data be measured? What is the quality of acquired experimental spectral data?

The question of the quality of spectral data is a central one and forms an essential basis for fully exploiting the results produced by multispectral or hyperspectral imaging. To understand quality, however, it is necessary to first establish and define how differences in spectra can be measured and how the spectral quality of a system can be quantified.

Once a framework for evaluating spectral quality has been established, it's possible to look both qualitatively and quantitatively at the kind of problems and errors that often occur during the acquisition and processing of multispectral and hyperspectral data. By understanding and quantifying such errors, it is, therefore, possible to propose methodologies and techniques to improve the quality of the acquired data with a goal of defining full acquisition and calibration pipelines that are adapted to the acquisition of works of art.

Research Question: What are the factors that influence multispectral and hyperspectral data quality and how can acquisition and calibration workflows be improved?

Multispectral or hyperspectral imaging of works of art, such as paintings, requires specific care and the use of protocols that are adapted to this type of target. One of the core goals of the research carried out for this thesis, therefore, is to examine in detail the specific acquisition and calibration pipelines necessary for both multispectral and hyperspectral imaging of works of art and to propose workflows and protocols that allow the highest possible quality data to be obtained.

Research Question: How to make high resolution quantitative image data accessible? How can large volumes of spectral and other scientific imaging data be visualized?

Once full calibrated data has been acquired and calibrated, it is necessary to then exploit the data in some way. An often neglected aspect of multispectral and hyperspectral imaging is the issue of access and visualization of the final results. Multi and hyperspectral imaging systems can produce extremely large volumes of quantitative and complex data that are difficult to exploit and make available to end-users. This data is often also acquired along with other imaging modalities that need to be made available simultaneously. The final research question of this thesis, therefore, addresses this issue and proposes solutions to enable high resolution spectral and quantitative image data to be made available online.

1.3 List of Contributing Papers

The research that constitutes the core of this thesis is represented in six core papers. These papers cover the various research areas included in the thesis and

address issues such as how spectral quality can be defined; what is the spectral quality found in real acquired hyperspectral data; how to improve the quality of acquired multispectral and hyperspectral data; and finally how to efficiently access and visualize the kinds of massive quantitative data sets produced by spectral imaging.

Core Papers:

- Paper 1: *Quality Evaluation in Spectral Imaging - Quality Factors and Metrics*
Shrestha, R., **Pillay, R.**, George, S. & Hardeberg, J. Y.
Journal of the International Colour Association, 12, 22–35 (2014)
- Paper 2: *Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems*
Pillay, R., Picollo, M., Hardeberg, J. Y. & George, S.
Sensors, 20, 381 (2020)
- Paper 3: *Studying That Smile: A Tutorial on Multispectral Imaging of Paintings*
Ribés, A., **Pillay, R.**, Schmitt, F., Lahanier, C.
IEEE Signal Processing Magazine, 25, 4, 14-26 (2008)
- Paper 4: *Hyperspectral Imaging of Art: Acquisition and Calibration Workflows*
Pillay, R., Hardeberg, J. Y. & George, S.
Journal of the American Institute of Conservation, 58, 3-15 (2019)
- Paper 5: *Multi-Modal Data Visualization and Analysis of “The Bedroom at Arles” by Vincent van Gogh*
Pillay, R.
Proceedings of the 22nd CIPA Symposium, October 11-15, 2009, Kyoto, Japan
- Paper 6: *Web-Based Visualization of Very Large Scientific Astronomy Imagery*
Bertin, E., **Pillay, R.** & Marmo, C.
Astronomy and Computing Journal, 10, 43–53 (2015)

Paper 1 provides a survey of the various metrics that can be used to quantify multispectral or hyperspectral image quality and discusses possible global frameworks for spectral image quality. **Paper 2** presents the detailed results from a unique round-robin test of hyperspectral imaging systems and provides an in-depth analysis of the data, examining the types of errors and variability seen in the data.

Paper 3 and **Paper 4** both look in detail at the issue of spectral acquisition and calibration and define methodologies for improving the final quality of spectral data. The first, **Paper 3** focuses on multispectral imaging and provides a tutorial for the calibration of multispectral data, whereas **Paper 4** focuses on hyperspectral imaging and proposes an end-to-end workflow for acquisition and calibration that ensures accurate and reliable spectral data.

Paper 5 addresses the technology required to visualize high resolution multi-modal imagery and offers a case study on the visualization of such data acquired from a painting. Finally **Paper 6** looks at the visualization of extremely large quantitative scientific imaging data and presents a case study from the field of astronomy where the volumes and complexity of imaging data are at their most extreme.

The core papers are included in full in Part II of this thesis and summaries of each of them can be found in Chapter 3.

Supporting Papers:

Additionally, there are two supporting papers that provide related research that do not form part of the core thesis, but which are closely associated to the work presented here.

- S1 *A Study of Spectral Imaging Acquisition and Processing for Cultural Heritage*
George S, Hardeberg J.Y., Linhares J, MacDonald L.W., Montagner C, Nascimento S, Picollo M, **Pillay R**, Vittorino T & Webb E.K.
Chapter 8, *Digital Techniques for Documenting and Preserving Cultural Heritage*, 141-158, Arc Humanities Press (2018)
- S2 *Assessment of Multispectral and Hyperspectral Imaging Systems for Digitisation of a Russian Icon*
MacDonald L.W., Vitorino T, Picollo M, **Pillay R**, Obarzanowski M, Sobczyk J, Nascimento S & Linhares J
Heritage Science, 5, 41 (2017)

Both supporting papers are concerned with the round-robin test described in core **Paper 2**. Paper **S1** is a book chapter and provides an introduction and overview of the round-robin test as well as preliminary results. Paper **S2** provides results and an analysis from one of the other test targets used in the round-robin test.

Chapter 2

Background

2.1 Introduction

This chapter provides background material on the core topics covered in this thesis. Although each of the papers included here also provides significant background material, it is useful to provide here a consolidated and expanded version of this background material in order to give an overview and facilitate the reading and understanding of the work presented within this thesis.

The chapter is aimed at readers who are familiar with digital imaging and who have some experience of multispectral or hyperspectral imaging, but who may not necessarily be aware of the details of the technologies used or the details of the processing required for such data.

The chapter begins by looking at the essential concepts behind these imaging techniques and provides an overview of the various technologies and methodologies employed. This is followed by a historical overview of the use of multispectral and hyperspectral imaging within the field of cultural heritage and how the technique has evolved since the first uses of multispectral imaging. The topic of spectral quality is then covered with an overview of the topics of noise, spectral distance metrics and round-robin tests. The calibration of spectral imaging systems is then discussed including both radiometric and geometric calibration. Finally, background information on the accessibility, visualization and storage of high resolution spectral data is provided.

2.2 Scientific Imaging

Imaging has a long history within cultural heritage and has been used since the very early days of analogue film to record and document works of art. The earliest form of imaging, black & white photography, was long used to document art and is able to provide a high resolution and accurate record of the surface of a work of art. The introduction of color film made it possible to additionally record the colors within a painting. However, analogue film is prone to instability and change over time, making it unreliable as a means of accurately and durably recording color.

Measuring and mapping the colors within a work of art is, nevertheless, an important activity in art conservation and curatorship. Very early colorimetric analyses for conservation were carried out by eye with the use of painted or printed color palettes, which were used as references and compared to the colors within the work of art to document them. The introduction of equipment such as colorimeters enabled more objective and accurate measurements to be made of the color.

It was not until the introduction of digital imaging that reliable color imaging become a real possibility. Accurate colorimetric imaging was indeed one of early goals of digital imaging in the field of cultural heritage and many attempts were made to use the quantitative nature of digital imaging to provide colorimetric results [18, 19, 20].

Scientific imaging techniques that use light or electromagnetic radiation at wavelengths that are invisible to the human eye also have a long history within cultural heritage. The application of techniques such as X-ray radiography [21], infrared [22, 23] and ultra-violet [24] photography to the study of art enabled imaging to be used as a scientific tool and as a means to analyze and study works of art in new and unique ways. These different imaging techniques use different parts of the electromagnetic spectrum (Figure 2.1) to illuminate the work of art and each part of the spectrum is able to provide different and complementary information.

The combination of these techniques with digital imaging has provided a number of important opportunities with regards to the analysis, visualization and documentation of works of art. If correctly acquired, processed and stored, digital imaging is capable of producing scientific data that provides measurements that are both highly accurate and reliable. Such quantitative data allows imaging to be used as a powerful and non-destructive tool for scientific analysis. Combining this data with image processing, statistical analysis and machine

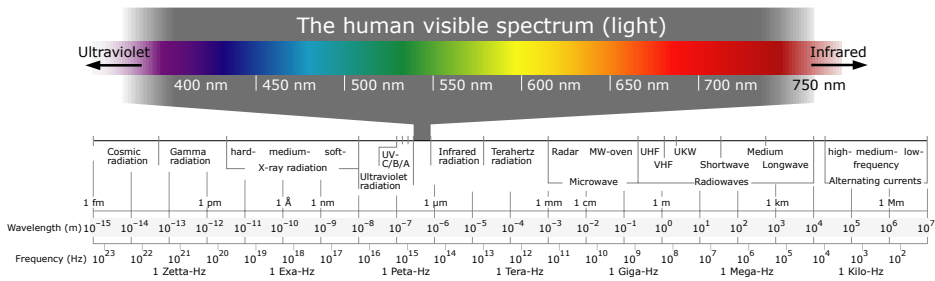


Figure 2.1: Electromagnetic spectrum. (Image credit: Horst Frank, CC BY-SA 4.0, via Wikimedia Commons)

learning techniques, allows image data to be used in more sophisticated ways, enabling uses such as materials analysis, change detection, hidden feature detection and many more.

2.3 Spectral Imaging

Spectral imaging is an imaging technique that has been made possible by the ability to acquire quantitative data through the use of digital technologies. It extends traditional scientific imaging by acquiring not just a single image, but many images at different wavelengths to obtain a reflectance (or emission) spectra at each point.

Due to the limits of the technologies used, spectral imaging systems are able to only acquire spectra within certain regions of the electromagnetic spectrum and within a specific range of wavelengths. These regions can be classified into several distinct wavelength ranges, which are listed in Table 2.1.

The most widely used spectral imaging cameras employed within the field of cultural heritage operate in the VNIR spectral region of 400–1000 nm. These allow spectra to be collected within the visible region, allowing accurate colorimetry to be performed. Spectra are also obtained within the near infrared region where it is possible to distinguish between certain pigments or materials that may be similar in color. SWIR spectral imaging systems that are able to see further into the infrared are also used for cultural heritage, but are relatively rare due to their higher costs. MIR and LWIR spectral systems, however, are orders of magnitude more expensive and are, therefore, extremely rare in the field of cultural heritage.

Description	Acronym	Wavelength Range (nm)
Ultra-Violet	UV	100 - 400
Visible	VIS	400 - 700
Near infrared	NIR	700 - 1000
Combined visible - near infrared	VNIR	400 - 1000
Short wave infrared	SWIR	1000 - 2500
Mid-infrared	MIR	3000 - 5000
Long-wave infrared	LWIR	5000 - 12000

Table 2.1: Common categories of spectral imaging and the range of wavelengths associated

Various acquisition technologies and techniques exist for spectral imaging, which are described in the following sections.

2.3.1 Multispectral Imaging

Standard three channel RGB camera technology is limited in terms of colorimetric accuracy and is prone to metamer error as the color recorded is dependent on the illuminant [25]. The desire to improve on this led to the idea of using multispectral imaging. Instead of three color channels, multispectral imaging uses filters to capture multiple channels spanning the spectral range of the camera, as can be seen in Figure 2.2(a). Acquiring multiple channels in this way allows for far higher color accuracy and reduces the potential for metamer errors. Most importantly, however, it allows the reflectance spectra to be estimated at each pixel. However, the limited number of and sometimes overlapping channels obtained through multispectral imaging can require spectral estimation methods to be used [26, 27].

2.3.2 Hyperspectral Imaging

Hyperspectral imaging, also known as imaging spectroscopy, bridges the domains of imaging and spectroscopy and was initially developed for and used primarily in the fields of satellite and space imaging [9] where it has been used to analyze, for example, planetary material composition or for the quantitative and qualitative analysis of terrestrial vegetation [28, 29].

Hyperspectral imaging provides a powerful combination of high spectral resolution and dense spatial mapping and has become a valuable analytical tool

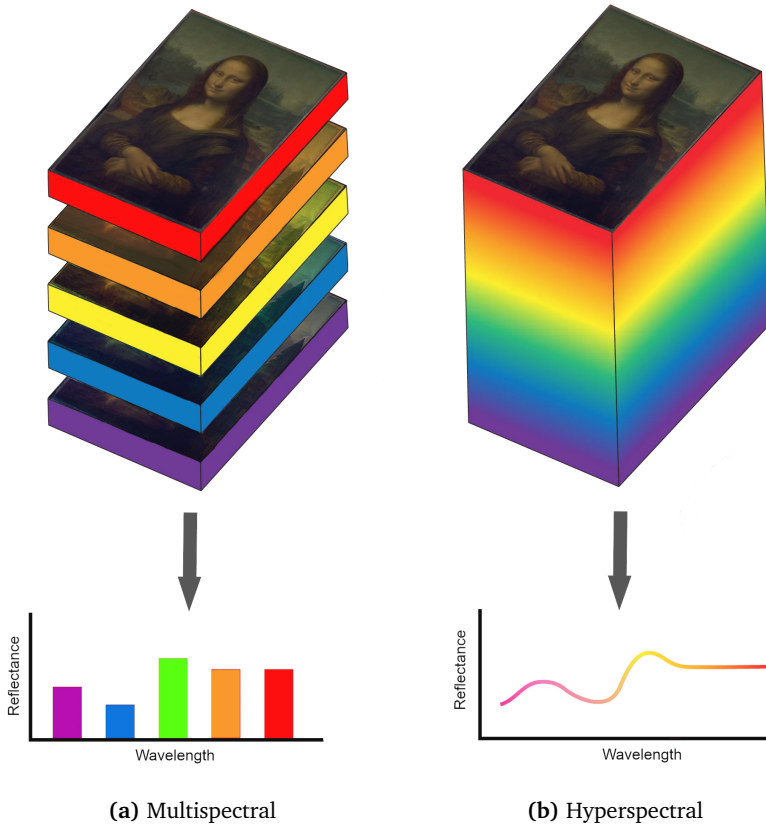


Figure 2.2: Illustration of the differences between the spectra acquired by multispectral and hyperspectral imaging

in a wide range of fields, including remote sensing [9], food science [10], astronomy [11], mineralogy [12], agriculture [13], medicine [14], archaeology [15] and many others. Hyperspectral data can be used in a number of ways, including, for example, materials mapping and identification, the detection of hidden features, change monitoring and many other applications.

Hyperspectral systems typically acquire up to hundreds of bands that are both continuous and contiguous over the spectral range of the system as can be seen in Figure 2.2(b). Thus, unlike multispectral imaging, hyperspectral imaging captures the full reflectance spectra without the need to perform spectral reconstruction. The distinction between multispectral and hyperspectral is, however, not always a clear one. Although attempts have been made to formalize a definition of each technology [30, 31], some acquisition technologies can be placed in either category.

2.3.3 Ultraspectral Imaging

A further category of “ultraspectral imaging” exists for very high resolution spectral imaging. These systems are capable of spectral resolutions of less than 1 nm and are able to distinguish fine molecular adsorption or emission line features [32, 33]. However, such instruments have yet to be applied to the imaging of works of art.

2.3.4 Scanning Methods

Hyperspectral imaging itself encompasses a range of different technologies. In addition, there are several scanning techniques generally employed, which are widely used to distinguish between hyperspectral systems according to how they physically scan a target. These scanning techniques are usually classified as either *snapshot*, *whiskbroom* or *pushbroom* hyperspectral imaging and can be seen in schematic form in Figure 2.3.

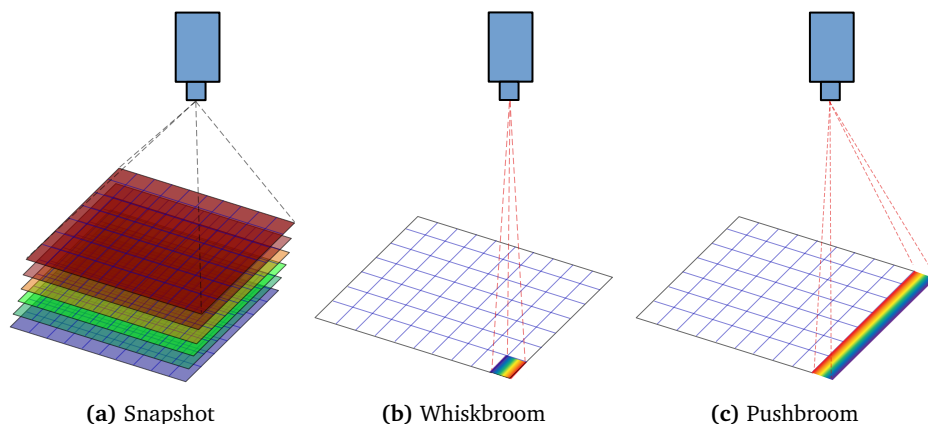


Figure 2.3: Illustration of the differences between the different scanning methods used for hyperspectral imaging. Snapshot scanning acquires the entire image at each spectral band. Whiskbroom acquires the entire spectra at a single pixel. Pushbroom acquires the entire spectra for a single line of pixels

Snapshot (or staring) hyperspectral imaging works in the same way as multispectral imaging, involving the use of a monochrome detector array and a filter-based system that captures a series of monochrome images at each band, as is illustrated in Figure 2.3(a). In order to obtain sufficiently finely spaced and contiguous spectra, tunable filters such as LCTFs (Liquid Crystal Tunable Filters) or Bragg tunable filters are commonly used. Snapshot imaging, there-

fore, captures an entire 2D image of the object in one shot with each spectral band acquired separately in subsequent shots.

An alternative to the use of the kind of filters used in multispectral or snapshot hyperspectral imaging is the use of dispersion optics. Dispersion optics (or spectrographs) are components consisting of prisms or diffraction gratings which split the incident light into its spectral components with each wavelength projected onto a different spatial location. By capturing this light with a linear or array detector, the entire spectra can be captured simultaneously. This technique allows very fine spectral resolutions to be acquired and dispersion optics of this sort are the most commonly used optics in current hyperspectral imaging systems.

Dispersion optics are employed in both whiskbroom and pushbroom systems. Unlike snapshot systems, these scanning techniques acquire the full spectral information in a single step. As a consequence, instead of capturing an entire image, only a single line is acquired at a time and the full spectral image must, therefore, be built up line by line. These systems function essentially as line scanners and, therefore, require the camera system to physically move at constant speed along the axis perpendicular to the scan line in order to acquire spectral data over a 2D area.

Whiskbroom scanning employs linear detectors which allow the dispersed spectra to be acquired along the line of pixels of the detector. These systems, therefore, capture the spectra from only a single point at a time, as is illustrated in Figure 2.3(b). In order to obtain a line spatially, whiskbroom systems include a moving mirror that sequentially scans across a line and reflects light into the line detector, thereby collecting data one pixel at a time across the line of sight. Whiskbroom scanning was often used for early spectrometers on orbital satellite systems [34]. Line by line scanning of this sort is perfectly suited to use on satellites where the camera moves smoothly at a constant speed along its orbital path.

Pushbroom scanning is similar to whiskbroom imaging in that it also scans an image line by line. However, instead of a line detector, a 2D detector array is used, which allows both spatial and spectral information to be captured simultaneously along the different axes of the detector. This allows an entire spatial line to be captured simultaneously to the entire spectra for each pixel across that line (Figure 2.3(c)). As with whiskbroom scanning, pushbroom systems require mechanical translation perpendicular to the orientation of the line being scanned. For ground or laboratory-based use, mechanical translation stages can be used to supply the necessary scan motion. Pushbroom systems are the

most common type of hyperspectral imaging used both generally and within the field of cultural heritage [8].

For a more complete description of the techniques and technologies behind hyperspectral imaging, see [35] or [36].

2.3.5 Detector Technologies

As mentioned in the previous sections, different camera technologies are required for different spectral regions. The detectors found in spectral cameras use various semiconductor materials that have different optimal operating ranges and sensitivities. This sensitivity (or quantum efficiency) is dependent on the wavelength of the incident light being measured and is different for each semiconductor material.

The most common detectors used for visible or NIR spectral imaging are Silicon-based detectors. These are sensitive over a wavelength range of approximately 400–1000 nm and offer good performance as well as being cost-effective. The silicon-based detectors used in spectral imaging are mostly packaged in CCD (Charge-Coupled Device) form, though CMOS (Complementary Metal Oxide Semiconductor) and sCMOS (Scientific CMOS) detectors are also sometimes used for VNIR spectral imaging.

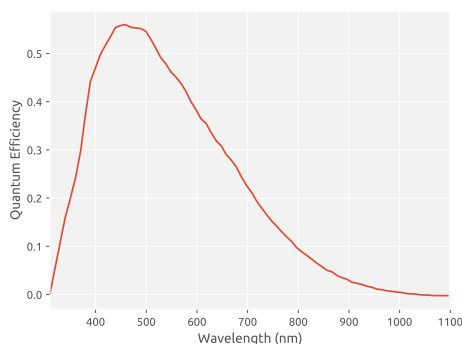


Figure 2.4: Quantum efficiency of the Kodak KAI-2020 CCD over the visible and near infra-red spectral region. (Image source: [4])

For detectors for the SWIR spectral range of 1000–2500 nm, other semi-conductor materials must be used. In-GaAs (Indium Gallium Arsenide) detectors are sensitive over a spectral range of around 1000–1600 nm, but in order to go further into the infra-red, MCT (Mercury Cadmium Telluride) or InSb (Indium Antimonide) based detectors are usually employed. These can only to be manufactured with smaller numbers of pixels than Silicon detectors and MCT cameras have at most around 320 pixels of spatial resolution for models that are currently commercially available.

The sensitivity of all detectors varies as a function of wavelength and a typical sensitivity curve for a CCD found in several common VNIR hyperspectral imaging systems can be seen in Figure 2.4. The sensitivity of the detector can drop

significantly at the edges of the range of spectral sensitivity and this variation in sensitivity needs to be corrected for during calibration. This can be compensated for in a snapshot system or in multispectral imaging where each spectral band is acquired independently. This allows the integration time to be set individually for each filter allowing the full dynamic range of the detector to be used for each band. However, for spectral imaging systems based on dispersion optics, the integration time must be set globally for the entire spectral range. Only a small part of the dynamic range of the detectors will, therefore, be used in areas of low sensitivity resulting in data quality that can vary considerably across the spectral range of the system.

2.4 Spectral Imaging within Cultural Heritage

In this section, we will look in more detail at the history of spectral imaging within the field of cultural heritage and describe the technologies that were adopted and how these evolved.

2.4.1 Multispectral Imaging of Art

The development and introduction of multispectral imaging technologies was of great interest to the museum sector due to its ability to capture simultaneously both accurate quantitative and colorimetric data. The first attempt to apply such technologies within a major museum was the VASARI project [16], which sought to capture both high resolution imagery and produce accurate scientific colorimetric data that could be used to monitor the changes in a painting over time. The system, which can be seen in Figure 2.5, used seven 50 nm wide interferential bandpass filters in the visible wavelength range of 400–700 nm coupled with a stabilized halogen-based light source that was used to illuminate the painting with a spectral power distribution that was continuous and smooth over the entire spectral range of the camera. Unusually, the interferential filters were placed in front of the light source rather than in front of the camera lens, as is the case with most multispectral imaging systems, in order to reduce distortions due to chromatic aberration and especially to limit the amount of incident light on the painting.

This enabled the system to obtain accuracies of better than $1\Delta E$ in the CIELAB color space [37]. The aim was also to match the spatial resolution of the 10x8 inch film typically used in museum photography and a scanning and mosaic

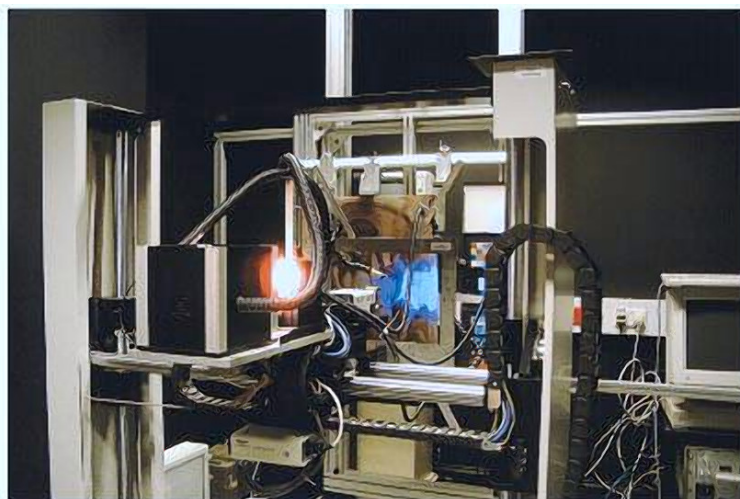


Figure 2.5: VASARI multispectral imaging system at the National Gallery, London: 7 band filter-based system mounted on a moving XY translation stage with a spectral range of 400–700 nm. (Image source: [37])

technique was necessary to obtain the necessary image sizes of up to 10,000 x 10,000 pixels per painting.

The system was, however, not portable and further development resulted in multispectral imaging systems that were not only portable (Figure 2.6), but which also extended the acquired spectral range into the NIR [38]. Many other multispectral systems were also later developed using different filters or sensors [39]. However, the use of a limited set of filters requires spectral estimation (or reconstruction) to be carried out [40] and attempts were made to design an optimal set of filters for paintings [41].

In order to obtain more accurate spectra without the need for spectral reconstruction, extra interferential filters were sometimes used [42]. However, to acquire even greater numbers of bands in a practical way, it is necessary to use tunable filters, such as LCTFs and various attempts were made using such filters with 16 [43] and 31 spectral bands [44] in the visible region and even up to 70 bands by means of tunable light sources [45]. An alternative approach was to use LCTFs, but select a subset of wavelengths according to the target [46].

The other major development in multispectral imaging was its use in the NIR spectral region over the wavelength range of 700–1000 nm. Early attempts using LCTF-based NIR multispectral imaging had been tested on soft media such as scrolls and ostraca [47]. For paintings, the extension of filter sets into

the NIR made it possible to see under-drawing and pentimenti and to begin to perform basic materials identification through spectral classification [48].



Figure 2.6: 13 filter portable multispectral CRISATEL camera. (Image source: [38])

The SWIR spectral region can provide valuable and complementary information for materials identification that allows pigments with similar colors to be clearly distinguished. This extra information can be used to perform more sophisticated materials classification and identification. Early attempts at this included studies coupling multispectral filter-based NIR imaging with multispectral filter-based SWIR [49], which were used to successfully distinguish similar pigments.

An alternative approach to the use of an image sensor array (such as a CCD) was the use of single-pixel photo-multipliers to acquire the multispectral data point by point rather than the whole image at once. In the

visible region, this method was able to acquire 32 bands through mechanical movement of the spectrophotometers and a point-by-point acquisition of the surface [50]. This system was later extended into the SWIR region [51, 52].

These systems were essentially used for multispectral reflectance imaging. However fluorescence can also provide useful information for the study of works of art. The emitted fluorescence can vary with respect to wavelength and can provide important information relating to the chemical composition of the work of art. Early attempts to use multispectral imaging to map the spectra of induced visible fluorescence [53] were able to provide useful materials identification.

The development of LED (Light Emitting Diode) based lighting and improvements in their stability and colorimetric content also allowed LED lighting to be used for multispectral imaging. LED lighting has significantly reduced thermal content and enabled multispectral imaging to be simplified and made less expensive. It also allowed multispectral imaging to be applied to fragile documents [54] and palimpsests [55, 56].

Multispectral imaging systems, however, can be very costly, narrowing their use to a limited set of institutions and making acquisition of such equipment often dependent on external sources of funding. In order to address this issue, simplified versions using off-the-shelf color cameras have also been proposed and used within museums [57, 58] and more recently using low-cost equipment [59].

2.4.2 Hyperspectral Imaging of Art

VNIR Imaging

The earliest attempts to use pushbroom hyperspectral imaging for the study of cultural heritage date to around 2004 with systems put together at the University of Parma [60] and, especially, at the Institute of Applied Optics (IFAC) in Florence for the analysis of paintings [17, 61]. Since then a number of other cultural heritage institutions around the world have acquired hyperspectral imaging systems, including notably, the NGA in Washington [62, 63]. A review of the work of these two institutions on hyperspectral imaging can be found in [64].

Although the use of hyperspectral imaging within cultural heritage institutions remains relatively rare, an increasing number of studies have made use of such equipment and hyperspectral imaging has been successfully used for a variety of applications including the revealing of otherwise hidden features and areas of lost material [61], for materials identification and pigment mapping [65, 66] where methods and algorithms from the field of remote sensing have been applied as well as for documentation and visualization [67].



Figure 2.7: Hyperspectral imaging system at the C2RMF, Paris: 160-band pushbroom camera with wavelength range of 400–1000 nm mounted on XY translation stage

A number of system designs have been implemented for the imaging of art using push-broom hyperspectral cameras. Standard laboratory systems for industrial use are usually horizontally mounted with the target lying flat on a horizontal surface. This set-up works well for manuscripts, books or drawings. However, for easel paintings, a vertical alignment is better suited, which therefore requires a custom-made set-up, such as that in Figure 2.7. Various push-broom based hyperspectral imaging systems have been put in place for cultural heritage to analyze paintings [61, 62], codices [68, 69], wall paintings [70], tapestries [71] and many others. In almost all of these systems, the push-broom hyperspectral camera moves linearly along one or more axes, while the painting remains mounted on a static easel. It is also possible, however, to mount the hyperspectral camera on a motorized rotating tripod [72] or use a rotating scan mirror in front of the hyperspectral camera to enable a form of snapshot imaging of the work of art [63].

SWIR Imaging

The majority of spectral imaging carried out for cultural heritage has involved the VNIR spectral region due to its lower cost and greater general availability. The SWIR spectral region can, however, provide valuable and complementary information for materials identification [73] and has been used with success in a wide range of studies including pigment mapping [66], the detection of egg yolk and animal skin glue paint binders [74], the identification of iron gall ink [75], the materials mapping of Chinese paintings [76] and many others.

MIR and LWIR Imaging

The use of equipment able to detect in the infrared region beyond the SWIR range has been very rare. Nevertheless attempts have been made at using the MIR spectral region (2500–5500 nm), where pigment analysis was carried out using an MCT-based system [77].

Luminescence and Fluorescence Imaging

The above examples of hyperspectral imaging have involved spectral reflectance imaging. But hyperspectral imaging can also be used to measure luminescence and fluorescence. As the light emitted is orders of magnitude weaker, this requires very sensitive spectral equipment. Examples of this include luminescence to identify pigments [78], of Egyptian Fayum portraits [79] and the fluorescence of medieval miniatures [80].

2.5 Spectral Image Quality

2.5.1 Noise

Noise can be a significant problem in spectral imaging systems and excessive noise can be detrimental to the use of hyperspectral data for tasks such as classification, anomaly detection or materials identification [81]. The level of noise in the spectral data is, therefore, an important criterion for defining the quality of a spectral imaging system.

Noise Sources

Several types of noise are generated in a digital imaging system and these sources of noise can originate from within the detector itself as well as from within the associated electronic subsystems. The most significant of these noise sources are fixed pattern noise (dark current) and shot (photon) noise.

Fixed pattern noise can be generated in various parts of an electronic system and is characterized by a noise pattern that is independent of the incident light. This form of noise, which is also known as Johnson (or Nyquist) noise, is essentially thermal in nature and will be different for each of the pixels in a detector. This thermal noise increases with increasing temperature and is proportional to the integration time used. It exists even in the absence of any light and so is also often referred to as the “dark current”. This form of noise is essentially added to the image data and in order to calibrate for this, the noise pattern needs simply to be subtracted from the raw spectral image data on a per pixel basis.

The other major form of noise is shot noise, which, unlike fixed pattern noise, is entirely dependent on the level of incident light. Shot noise is random in nature and can be modeled as a Poisson process where the noise is proportional in magnitude to the square root of the number of photons incident on the detector.

For a more detailed description and analysis of the various types of noise and their origins within a detector, see section 5.3 of [82].

Hyperspectral systems are particularly prone to noise, given the narrow wavelengths used and the wide variability in quantum efficiency across the spectral range of the detector (see Section 2.3.5 and the example in Figure 2.4). Even though the measured photon signal may be excellent through the central wavelengths of the operating range of the detector, the signal at the extremes of the operating range can be very weak with respect to the magnitude of fixed pat-

tern noise and, therefore, very noisy. The result is data quality that is highly variable between spectral bands.

Reducing Noise

There are, however, several techniques that can be used to reduce noise within a system and to, therefore, improve the SNR (Signal to Noise Ratio).

Often the most problematic source of noise is fixed pattern noise. As this form of noise is dependent on temperature, cooling the detector will reduce the magnitude of this noise. NIR and SWIR detectors are particularly prone to this form of noise [83] and cameras for these wavelength ranges are usually cooled to reduce the amount of it.

The level of fixed pattern noise forms in effect a lower threshold at which level any signal is equal in magnitude to the noise. Increasing the amount of incident light on the detector will, therefore, reduce the proportion of fixed pattern noise relative to the signal. Spectral cameras are, therefore, designed to maximize the transmission of light through to the detector and use optimized optics and coatings along with lenses with large apertures and, therefore, low f -numbers in order to achieve this.

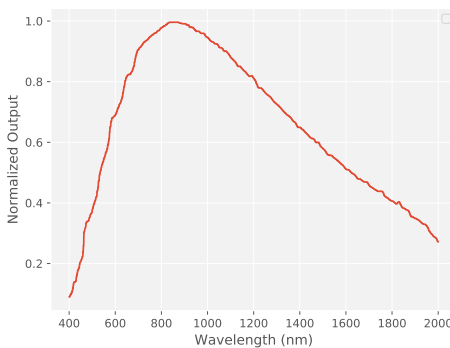


Figure 2.8: Spectral content of a typical Halogen illuminant used for VNIR spectral imaging. (Image source: [4])

When using the cameras, the relative level of fixed pattern noise can be reduced by increasing the global level of illumination and, in particular, by increasing the spectral content of the light sources in those wavelengths where the illuminant is weak and the detector has low sensitivity. The spectral power distribution for a Halogen illuminant typically used for spectral imaging is shown in Figure 2.8, where the large difference in spectral content between wavelengths is clearly visible. For a VNIR hyperspectral system, the combina-

tion of illuminant and quantum efficiency results in very low amounts of light at the lower end of the wavelength range at around 400 nm.

For shot noise, various signal processing techniques can be used to improve SNR. Frame averaging, for example, is a technique whereby a number of co-

incident scans of the same image (for a snapshot system) or line (for a whiskbroom or pushbroom system) on the target are averaged together to improve SNR [84]. This technique improves SNR by a factor of \sqrt{N} , where N is the number of co-incident scans.

Other techniques are aimed specifically at line-scanning hyperspectral systems and seek to reduce the *differences* in SNR across the spectral range. These techniques improve the *relative* SNR between bands and as a consequence reduce the SNR at the central wavelengths.

One example of this is the use of an equalization filter. Equalization filters are designed to reduce the sensitivity in the central most sensitive region of the detector's spectral range, thereby improving the relative SNR at the extremes. This can greatly improve the overall SNR of the data, but necessarily requires longer integration times and reduces the SNR for the central wavelengths as well as the average SNR globally. The ideal equalization filter would be the exact inverse of the combination of quantum efficiency and illuminant. However, obtaining an exact inverse is difficult to obtain in practice.

One further method for reducing noise is through the use of denoising algorithms and various image denoising techniques have been adapted for use with spectral images [85, 86, 87].

Although the combination of all these techniques can significantly improve the quality of spectral data, noise will invariably exist, particularly at the extremes of the spectral range of the detector.

2.5.2 Measuring Spectral Differences

The level of error globally within an image is often defined by measures such as the Mean Square Error (MSE) [88], which is the sum of the squared differences between the values of each pixel and its reference value divided by the number of pixels (Equation 2.1, where M and N are the pixel dimensions of the image and x and \bar{x} are the reference and measured pixel values respectively).

$$MSE = \frac{1}{M \times N} \sum_{i=0}^{i=M \times N} (x_i - \bar{x}_i)^2 \quad (2.1)$$

This is effectively a measure of the statistical variance within the accumulated errors. The quality of an image can, therefore, be defined as a value derived from the inverse of the MSE. This is often defined by the signal to noise ratio

(SNR), which is the ratio of the mean signal to the MSE. The PSNR (Peak Signal-to-Noise Ratio) is also often used, which is the ratio of the maximum possible pixel value to the MSE.

For spectral images, consisting of multiple images at different wavelengths, the MSE will vary with wavelength and so is usually calculated separately for each spectral band. Although, these provide a good generic measure for image quality, the quality of the spectra themselves obtained at each pixel location is more difficult to assess.

Determining what the quality of spectral data is is not something that is straightforward. There are a number of ways to quantify spectral accuracy and the appropriate measure of accuracy can depend in large part on the use to which that data will be put.

Spectroscopy

A common way of studying spectra in the field of traditional point spectroscopy is through the identification and localization of peaks and troughs within the spectra. Many materials or pigments have characteristic spectral curves that can be distinguished by their shapes and by the precise location of any peak. An extension of this principle is through the use of derivative analysis, which uses the derivative of the spectra to facilitate the identification of peaks and troughs and elucidate subtle spectral features that are not readily apparent otherwise.

Derivative analysis can be a useful tool for the identification of pigments [89] as well as for material identification more generally [90]. An example of the use of the derivative to distinguish two similar pigments can be seen in Figure 2.9, where the two red pigments carmine and vermilion are shown along with their derivatives. These derivatives have different forms with peaks in distinct locations, allowing the two different pigments to be clearly distinguished.

Spectral Distance

When using spectral data for classification tasks, the notion of the “spectral distance” between the measured spectra and a reference spectra is generally used and there are a number of ways to quantify this spectral distance. The MSE measure defined earlier can be also be applied to individual spectra and this provides a value for the “distance” between two spectra. However, this measure weights all differences between spectra equally, whereas more often than not, we are more interested in differences in *shape* rather than in the *absolute values* of the spectra. In other words, two spectra which have identical

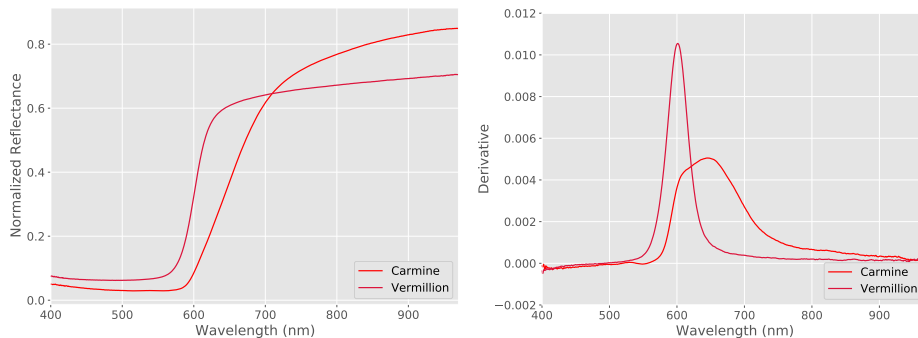


Figure 2.9: Reflectance spectra (left) and derivative of the reflectance spectra (right) for the pigments carmine and vermillion. Although these pigments are both red, the spectra can be used to distinguish between them. The derivative allows this distinction to be more clearly made with different shapes and peaks at distinct locations

shapes, but which differ in amplitude should be considered more similar than two spectra with similar amplitudes and very different shapes.

A number of spectral distance measures have been developed to attempt to do just that. The most widely used of these distance metrics is the Spectral Angle used in Spectral Angle Mapper (SAM) classification [91]. This measure is extensively used in remote sensing and is derived from the angle formed between a reference spectrum and the image spectrum and is shown in Equation 2.2, where \mathbf{x} and \mathbf{y} are vectors representing the two spectra to be measured.

$$SAM(\mathbf{x}, \mathbf{y}) = \arccos\left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|_2 \|\mathbf{y}\|_2}\right) \quad (2.2)$$

A number of other more recent distance metrics have also been developed, including Spectral Information Divergence [92], Spectral Correlation Mapping [93], Spectral Gradient Angle [94], Spectral Similarity Value [95], and the “(SID, SAM)” metric [96] which combines both Spectral Information Divergence and SAM into a single distance measure and many others. Each of these have their own particular characteristics and properties and an overview of their properties and behavior can be found in **Paper 1**, [97] and in [98].

Spectral distance measures are a key component of many spectral processing workflows, including spectral classification [99] and change detection algorithms [100]. The use of such algorithms typically require the appropriate setting of threshold or other parameters, whose values are often dependent

on the distance metric used. Indeed the choice of distance measure can have a significant impact on the results obtained and on the performance of these algorithms [101, 102].

2.5.3 Round-Robin Tests

Assessing the true performance of an analytical technique or of scientific equipment on real experimental data is important in order to determine both their utility and to understand how the technique or equipment can best be used. Assessments of this sort are usually carried out in a single laboratory, where the technique or equipment is tested against reference standards or against other validated equipment. However, scientific equipment is not always used under the ideal controlled conditions that exist in testing laboratories. Exactly how the equipment or technique is used and the environment in which it is used can significantly influence the results obtained.

A round-robin test attempts to take these influences into account and, instead of carrying out tests within a single laboratory, a number of locations and testing facilities are used to assess performance. The variability found in the obtained results can provide additional insight into the use of a technique and help identify procedures or protocols that require improvement.

Round-robin tests are, therefore, a useful means of comparing equipment or methodologies and have been successfully carried out in a wide range of fields. Within fields related to spectral imaging, round-robin tests have included topics such as the measurement of the BRDF (Bidirectional Reflectance Distribution Function) of diffuse reflectors [103], the radiometric calibration of a satellite multispectral sensor [104] and tests of field spectrometers in laboratory settings [105].

Several papers within this thesis contain or refer to results from a large-scale round-robin test that was carried out to evaluate a range of multispectral and hyperspectral imaging systems. The rest of this section, therefore, provides some supplementary background information on this round-robin test in order to put this work into context.

The experimental data analyzed in **Paper 2** and referred to in **Paper 4** consisted of a subset of the data acquired from a round-robin test carried out for the COSCH project [106]. The round-robin test involved nineteen different institutions, including research laboratories, universities, equipment manufacturers and museums and an overview of the tests is provided in Supporting

Paper **S1**, “*A Study of Spectral Imaging Acquisition and Processing for Cultural Heritage*” [8].

The round-robin test was a coordinated effort to gain a better understanding of the instrumentation, the processes of data acquisition and the effects of the devices and methodology on the reliability of the spectral imaging data. The goal was to evaluate the effective limits in accuracy of the spectral imaging equipment in use within the participating institutions and examine the challenges and issues that arose from bringing these different data sets together. By measuring and understanding the variability seen within them, the goal was to help improve protocols for the acquisition, handling, processing and sharing of spectral data sets.

The data acquisition for the round-robin test was carried out using either multispectral or hyperspectral equipment from different manufacturers and with different experimental acquisition setups, procedures and methodologies. The aim of the comparison was not to compare the hardware specifications or raw performance of the imaging devices themselves, but to measure the resulting effective performances of the systems globally under their standard operating conditions and after the application of the calibration and processing workflows that are usually applied by each institution to their system. The variability due to the different setups, operating procedures and the way data was processed was, therefore, an important factor to take account of and include in the study. In this way, an insight into the practical limits in accuracy of spectral imaging systems within routine operating environments could be gleaned.

The overall aim was to use the insight obtained from these tests to help standardize methodologies and best practices for imaging cultural heritage objects through spectral imaging techniques and the insight gained contributed to the acquisition and calibration workflows defined in **Paper 4**.

A number of test targets were used, each with very different characteristics and each of which aimed to test different aspects of the spectral imaging equipment under evaluation. These test targets consisted of four objects, which can be seen in Figure 2.10 and included: a Macbeth ColorChecker, a wavelength standard, a custom-made pigment panel and a printed polychrome lithography 19th century Russian icon.

The ColorChecker used was a standard Xrite color chart consisting of twenty-four colored patches in a four-by-six grid, which provided a basis for standardized colorimetric comparisons between systems.



Figure 2.10: Targets used in the spectral imaging round-robin test: 24 patch Macbeth ColorChecker (top left), wavelength standard featuring well-defined and narrow features over the UV, visible and near infra-red spectral ranges (top right), custom-made pigment panel with 7 Renaissance-era pigments (bottom left) and a mass-produced 19th century Russian icon made through polychrome lithography (bottom right). (Image source: [8])

The wavelength standard was a Zenith Polymer wavelength standard consisting of a chemically inert diffuse lambertian reflectance standard composed of PTFE (Polytetrafluoroethylene) doped with the oxides of the rare earth elements Holmium, Erbium and Dysprosium. This combination gives the standard a stable spectrum of characteristic, well-defined and narrow features over the ultra-violet, visible and near infra-red spectral ranges, which is suited for use in accurate spectral calibration.

The pigment panel was a test target created especially for the round-robin test by one of the participating institutions. It was made up of seven historical pig-

ments at different concentrations in an egg tempera binder. The paint preparation and application on the panel aimed to authentically reproduce the medieval Tuscan painting technique described in Cennino Cennini's 15th century *Il Libro dell'Arte* [107]. The panel consisted of a wooden support with a gypsum ground, a canvas layer, and a second gypsum ground layer. Before application of the paint layer, five types of drawing materials (watercolor, charcoal, graphite, a lead and tin-based metalpoint and a lead-based metalpoint) were used to create lines and line patterns that were then covered with paints applied with two different thicknesses. Additional details on the panel and on the source and composition of the pigments used for the panel can be found in [108].

The final object was a 19th century, mass-produced Russian icon, printed by polychrome lithography, using eight different inks, onto a tinned steel plate and nailed onto a wooden panel. It has a glossy surface over the colored areas and a high specular reflectance from the golden metallic surface. The icon was used to investigate the spatial imaging characteristics of the spectral imaging systems as well as their behavior with highly reflective surface.

Supporting Paper **S1** provides an overview and preliminary results from the round-robin tests with further results found in several different papers. **Paper 2** provides an in-depth analysis of the hyperspectral data from two of the test targets. Details and results from the Macbeth ColorChecker can be found in both [109] and [110]. Results from the 19th century Russian icon can be found in Supporting Paper **S2**, “*Assessment of Multispectral and Hyperspectral Imaging Systems for Digitisation of a Russian Icon*” [7].

2.6 Spectral Image Calibration

The calibration of spectral imaging systems is one of the main contributions of this thesis and forms the core topic of both **Paper 3** and **Paper 4**. This section, therefore, provides an overview of the topic and provides some supplementary background information.

In order to improve the spectral acquisition workflow, it is first necessary to understand the process of image acquisition. There are essentially two types of calibration that needs to be carried out for spectral imaging: radiometric calibration and geometric calibration.

2.6.1 Radiometric Calibration

A number of different factors contribute to the signal measured by a detector at a particular pixel. This signal will depend on the number of incident photons provided by the illumination, the spectral reflectance of the target, the quantum efficiency of the detector, the transmission characteristics of the filter or dispersion optics and the noise components.

We can represent the raw pixel values obtained from the acquisition of an object by an equation of the form seen in Equation 2.3:

$$D = \int I_{\lambda} R_{\lambda} F_{\lambda} T_{\lambda} Q_{\lambda} d\lambda + n_{\lambda} \quad (2.3)$$

where D , the digital number is the response of the detector at a particular pixel and whose magnitude is the product of various factors: I represents the illuminant spectral power distribution, R the reflectance of the object, F the filter transmission characteristics, T the transmission characteristics of the lens and optics and Q the quantum efficiency of the detector. These all vary with respect to wavelength and so, the detector response is the integral of this product over all wavelengths. The detector response also includes a noise term, n , which represents fixed pattern noise and which is also dependent on wavelength, but independent of the object reflectance or illuminant.

The sensitivity of each individual pixel of a detector varies across the detector and, thus, Equation 2.3 must be used independently for each individual pixel. The reflectance of an object can, therefore, be calculated if we know the values of these various elements. Although factors such as the filter characteristics and the quantum efficiency are often known, it is difficult to know the precise magnitude and spectral content of the illumination at the surface of the object and various optical components within an imaging system also modulate the signal in ways that are difficult to measure. It is practical, there-



Figure 2.11: Multi-step reflectance standard with reflectances at 99%, 50%, 25% and 12.5%

fore, to consider these all as a single combined factor that must be calculated for each acquisition and the most common way to do this is through the use of a reference target or set of targets, such as that shown in Figure 2.11, with a precisely known reflectance spectrum.

The combined response of the illumination, filter, optics and quantum efficiency can therefore be calculated using Equation 2.4, where O_λ is the combined response, D_r is the pixel response for that reference target and L_λ is the spectral reflectance of the reference target at band λ :

$$O_\lambda = \frac{D_r - n_\lambda}{L_\lambda} \quad (2.4)$$

This can then be substituted into Equation 2.3 to provide a simplified equation (Equation 2.5) for the reflectance factor R for a particular band, λ , where L_λ is the spectral reflectance of the reference target at band λ and D_r is the pixel response for that reference target, D is the pixel response for the target object and n is the fixed pattern noise:

$$R_\lambda = L_\lambda \frac{D - n_\lambda}{D_r - n_\lambda} \quad (2.5)$$

2.6.2 Geometric Calibration

Geometric errors and the correction of these errors is addressed in **Paper 2** and **Paper 4**, where the magnitudes of certain geometric errors seen in hyperspectral data are measured and procedures to reduce and correct for these errors discussed.

All imaging systems produce some form of geometric distortion that needs to be accounted for in order to obtain quantitative geometrical data. Any misalignment between spectral bands can result in errors in the spectra at each location [111] and can make registration across different imaging modalities difficult.

Multispectral and snapshot hyperspectral systems have distortions that are similar and which consist of optical distortions and chromatic aberrations that vary with filter and wavelength. Each spectral band may need to be registered together and various models, such as [112], exist for correcting this.

Scanning-based systems such as dispersion-based pushbroom and whiskbroom imagers which combine mechanical movement with complex optics possess dif-

ferent types of geometric distortions that need to be corrected for. These can include spatial distortions across the track of a line scan, known as keystone, and spectral distortions along the spectral axis, known as smile. A discussion of these effects and a method for correcting them can be found in [113]. Camera manufacturers design their systems to minimize these, but they cannot be entirely eliminated.

Other non-parametric optical distortions can also exist due to the complex multi-component optics often used in hyperspectral systems. These require the characterization of a sensor model, which can be produced by using a micrometer and by measuring the angle of view for each pixel. Other kinds of errors can also arise when using mechanical translation stages, such as incorrect synchronization between the movement of the translation stage and the sensor acquisition, which can result in non-square pixels.

2.7 Classification

Classification is one of the most common tasks carried out using spectral imaging data and consists of assigning a label to each pixel in an image. For the study of art, this essentially consists of determining the material or pigment composition of each pixel.

As discussed in Section 2.5.2, distance measures are a key component in many spectral classification workflows and the distances calculated between the spectra at a particular pixel and a set of reference spectra can be used to classify spectral data. The most basic classification workflow essentially consists of defining a set of reference spectra within an image. These spectra are known as *end-members* and define the “purest” spectra within the set. The distance between the pixel spectra and each of the end-members is then calculated and the end-member with the closest spectral distance to the pixel and which is within a user-defined threshold is then assigned as the class label for that pixel.

This basic method can be refined into a multi-step workflow involving first the application of dimensionality reduction techniques in order to reduce noise and reduce computational complexity. Principal Component Analysis (PCA) based algorithms are generally used for dimensionality reduction. The next step involves the determination of the number of end-members that exist within the spectral data. The final step is classification, where each pixel is assigned to an end-member or combination of end-members. Several end-members may be combined within each pixel, so spectral un-mixing can be used to determine the relative proportion (or abundance) of each end-member at each pixel.

The most commonly used pipeline within cultural heritage is the “spectral hourglass wizard” pipeline available in the ENVI remote sensing software suite [114]. This particular pipeline uses Maximum Noise Fraction (MNF) [115] for dimensionality reduction and orders the transformed components by noise content, rather than by variance as in classical PCA. Pixel Purity Index (PPI) [116] is here used for determining the end-members. Variants of this pipeline have included the use of MaxD [117] rather than PPI for end-member extraction and K-Means for classification, which has been applied to the analysis of ancient maps [118]. Another variant has involved the use of the Bidimensional Histogram of Spectral Differences [119] for the classification of paintings.

These pipelines involve a certain amount of manual interaction in order, for example, to select end-members and attempts to automate the classification workflow have been made [120]. More advanced methods based on the use of deep learning have also been attempted [121].

Spectral distance measures can also be combined with other types of metric to improve the results provided by classification algorithms. The spatial content within an image can often provide useful complementary information. There is often a strong spectral correlation between neighboring pixels and this correlation can be exploited through the use of morphology or segmentation to optimize the classification results. For an overview of these methods of combined spectral-spatial classification, see [122].

2.8 Data Structures & Visualization

How to enable access to spectral data, how the data can be shared and how the data should be stored are key issues that are often overlooked by users of spectral imaging equipment within the field of cultural heritage. These are issues that need to be addressed if complex spectral data is to be fully exploited and remain of use in the long term. These questions are at the core of both **Paper 5** and **Paper 6** and this section, therefore, provides an overview of this topic.

2.8.1 Spectral Data Structures

Spectral images can be considered as either a series of images taken at different wavelengths or as a three dimensional image cube where each pixel of the 2D image possesses a third dimension representing the spectral bands at that pixel.

Such image cubes are sometimes referred to as *hypercubes* and an example can be seen in Figure 2.12.

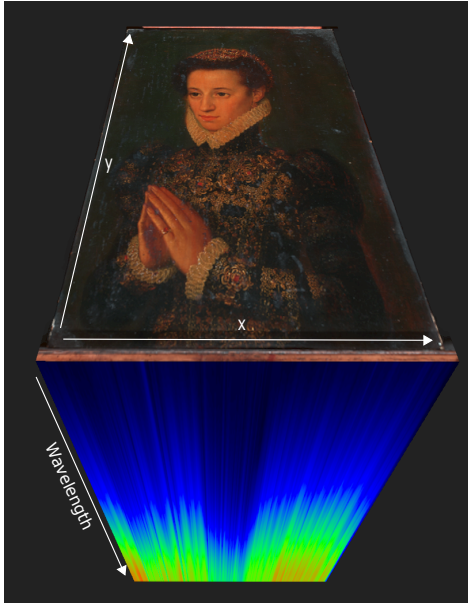


Figure 2.12: 3D representation of a spectral image cube of the painting *Lady Praying* (Anonymous, oil on oak, 1575-1599, Musée du Louvre), showing how the spectral information is projected onto the third axis

most widely used header format is that designed for the ENVI remote sensing software suite. This header provides a flexible metadata format which describes the main elements necessary to decode a spectral data cube. This includes information such as the bit depth of the pixel data, structuring information describing how the pixels, spectra and spatial lines within the cube are interleaved together, the width and height of the image cube, the list of center wavelengths etc.

How the spatial and spectral information is interleaved together has an important impact on the speed at which certain types of data can be read from disk as close or adjacent data locality can result in significantly faster read speeds. Typical interleaving schemes used for spectral data are band sequential (BSQ) where each full 2D image for each spectral band is stored consecutively, band

How this data is structured, encoded and stored can have an important impact on how that data is later used and the performance of any application that uses that data.

Multispectral images have a relatively low number of spectral bands making the storage of such spectral data in separate image files a practical solution. Hyperspectral data, on the other hand, can have hundreds of spectral data values at each pixel, making the use of separate files unwieldy and inefficient. The most common way to store hyperspectral data is, thus, in a single file and in most cases this data format will consist of raw un-structured or semi-structured data cubes with an associated header file that contains the necessary metadata to decode the data.

These data cubes essentially consist of binary pixel-wise data stored with different types of interleaving. The

interleaved line (BIL) where each spectral band for an individual scan line is stored consecutively and band interleaved pixel (BIP) where the spectral points are interleaved together for each pixel in an analogous way to the interleaving that is common in RGB images. The various interleaving schemes involve different trade-offs in terms of performance for different types of processing. BIP interleaving stores the spectral data for each pixel together and is, therefore, optimized for fast access to individual spectra, whereas BSQ is optimized for access to individual image bands. BIL provides a compromise between the two schemes and is, therefore, often the default output format for spectral imaging cameras.

More sophisticated storage formats exist that can be used to store and encode spectral data. Image formats such as TIFF [123] or JPEG2000 [124] are capable of storing multi-band data and are often used for storing multispectral data for remote sensing. Other formats specifically designed for multispectral imaging include Spectral Binary File Format, Natural Vision format and the AIX multispectral image format and an overview of these and their use can be found in the CIE (Commission Internationale de l'Eclairage) TC8-07 Technical Report on multispectral image formats [125]. Work on file formats and standardization for hyperspectral data is also currently being carried by the IEEE Standards Association P4001 Working Group [126].

Paper 6 concerns the field of astronomy, where a high degree of file format standardization has occurred. Astronomy imaging data is usually stored in the FITS format [127], which is a flexible container format that can handle data encoded in up to 64 bits per channel. A more generic data format devised for high performance computing is HDF5 [128], which is a flexible and feature-rich container format designed for arbitrary large data volumes. Despite the potential of HDF5, however, it is not yet widely used for spectral image data.

The number of bytes required to store each spectral value also needs to be taken into account. Consumer cameras, scanners and other image producing equipment usually output data with a bit depth of 8 bits per channel for each pixel (256 possible values for each color channel). Spectral imaging cameras (and scientific imaging systems in general), however, are often capable of acquiring data with up to 12 or even 14 bits of usable information per pixel. Many data formats are only capable of storing byte-aligned data and thus the 12 or 14 bits of data are packed for practical purposes into two bytes per value, thereby using 16 bits of storage. **Paper 6** goes beyond this and tackles floating point data which has 32 bits per channel. Although no image detectors are capable of acquiring raw data with this bit depth, the use of the noise reduction techniques discussed in Section 2.5.1, such as frame averaging, can increase the result-

ing effective bit depth significantly beyond the native capacity of the detector. The file formats used and any encoding / compression algorithms employed, therefore, need to be able to handle such data appropriately.

2.8.2 Visualization of Very Large Image Data

Spectral imaging systems produce complex and quantitative data sets that can have very high spatial resolutions, that can contain tens to hundreds of channels per pixel and that require 16 or even 32 bits of storage per pixel per channel. This combination can result in extremely large volumes of data. The kind of pushbroom hyperspectral imaging equipment shown, for example, in Figure 2.7, is able to capture up to 160 different spectral bands at a resolution of 15 pixels/mm with a bit depth of 16 bits per band for each pixel. The acquisition of a painting of physical size 1m x 1m, therefore, results in over 50 GB of spectral data.

Other high resolution imaging modalities can also result in very large data volumes and when spectral imaging data is combined with these various modalities, data volumes can be extremely large indeed, making the visualization and exploitation of such data difficult.

The storage and processing of such data on local computers is rarely a practical possibility and so an effective solution to making such data accessible is through the use of networked client-server architectures. **Paper 5** and **Paper 6** examine this in more detail and propose efficient solutions together with working open-source implementations. Such architectures avoid the need to transfer large quantities of data and allow both desktop and mobile devices to be used without excessive requirements for processing power. They also have the benefit of allowing data to be made available through standardized network protocols or web APIs and can be coupled with light-weight browser-based viewers.

Such systems can also employ pre-rendering of certain views that may be slow to calculate on-the-fly, such as colorimetric rendering from spectral data or the results of PCA analysis. An example of the client-server system described in **Paper 5** and **Paper 6** is shown in Figure 2.13, where a data-set consisting of 6 GB of high resolution multispectral images of the Renoir painting *Femme Nue dans un Paysage* (1883, oil on canvas, Musée d'Orsay) can be viewed online¹. The viewer interface provides access to each individual spectral band, a D65 colorimetric rendering, the spectral curve at each point, other registered imaging modalities and the ability to blend between any pair of images.

¹Online multi-modal image viewer: <https://tinyurl.com/y5du865l>

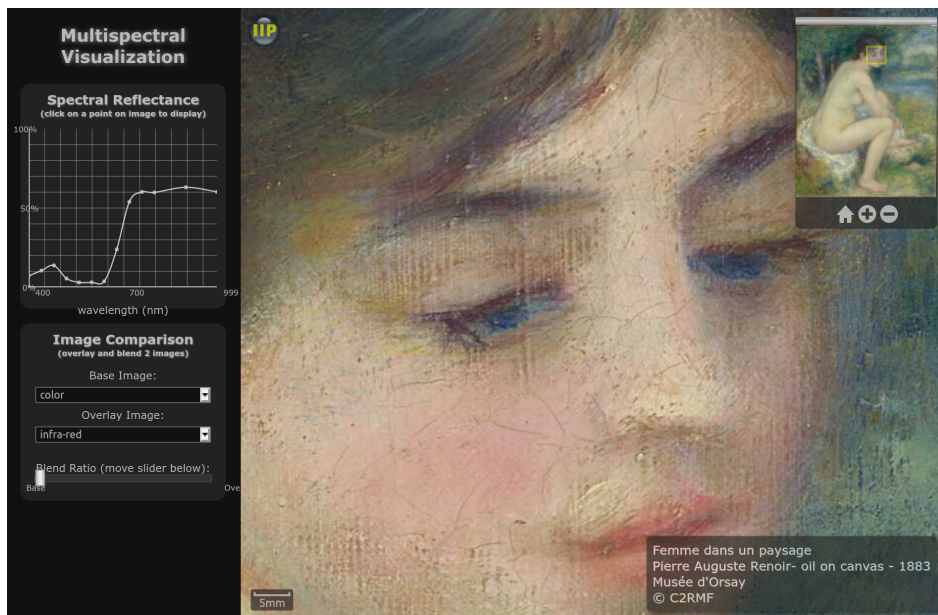


Figure 2.13: Web-based visualization of the high resolution spectral data from Renoir's *Femme Nue dans un Paysage* (1883, oil on canvas, Musée d'Orsay). Functionality includes interactive pan and zoom access to a D65 color rendering, individual spectral bands and other imaging modalities as well as the ability to view the spectral curve at each point

However, efficient networked architectures can only provide a partial solution to the provision of high performance access. In order to fully optimize the performance of such systems, it is necessary to also understand the role played by the underlying storage structures and data formats used.

2.8.3 Optimizing Performance for Very Large Images

Multispectral and, in particular, hyperspectral data can be extremely large, which can result in data that is difficult to manipulate and slow to read from disk.

In order to improve random access to such data, various strategies can be employed. Tiling is a typical means of enabling fast random access on very large data sets and involves splitting the image data into smaller more manageable regions, which are stored in distinct byte ranges within the data format and which can be decoded independently. Several formats are able to handle tiling, including TIFF, JPEG2000, FITS and HDF5.

For very high resolution image data, a further strategy of using multi-resolution storage can be employed in which multiple versions of the data are stored with different spatial resolutions within the file format. This enables fast access to data at different scales, which can be particularly useful for visualization purposes where it is necessary to be able to switch quickly between image data at different levels of resolution. This strategy is typically employed with each successive resolution defined with a size of exactly $\frac{1}{2}$ that of the previous resolution (though other ratios are possible). This creates a pyramid-like structure, which has a certain level of data redundancy, but which in many cases can be worth the extra overhead. When using a ratio of $\frac{1}{2}$, the extra storage space required is approximately 33% as can be seen from Equation 2.6, where M and N are the pixel dimensions of the image.

$$(1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^4} + \dots)(M \times N) \leq \frac{4}{3}(M \times N) \quad (2.6)$$

Multi-resolution pyramids can be combined with tiling to enable very efficient data structures that allow fast random access to any region on the image and at any resolution. A representation of such a tiled multi-resolution structure can be seen in Figure 2.14 where the efficient pyramid structure allows data from any region and resolution to be obtained quickly even for extremely large images.

2.8.4 Compression

Given the very large data volumes that can arise with spectral imaging, compression can be necessary in many situations. The standard compression algorithms supported by image formats such as TIFF or JPEG2000 will work with spectral data in the same way as with standard image data. However, there is a strong correlation between spectral bands, making it possible to potentially obtain significantly improved compression. JPEG2000 has support for compressing image cubes through its JP3D extension and has been shown to be capable of delivering good compression results through this [129]. Within the field of space-born communication, a customized compression standard for spectral data (CCSDS 123.0-B-1) has been also shown to be very effective [130]. Beyond these examples, a wide range of research has been carried out on spectral compression and a review of these can be found in [131].

Scientific data, if compressed, is usually compressed using lossless encoding as this allows data to be reversibly de-compressed without any loss of precision. However, lossy compression can considerably reduce the resultant file

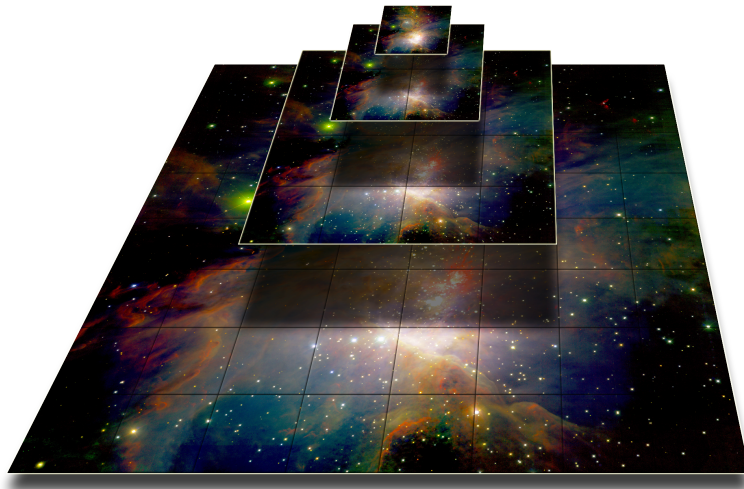


Figure 2.14: Multi-resolution tiled image pyramid structure that allows fast and efficient random access to any region and resolution even for extremely large images. (Image source: [6])

sizes. If perfect data fidelity is not required, lossy compression may provide an acceptable trade-off between storage requirements and data quality. Transform coding such as PCA [132] or related transforms [133] can be used for lossy compression [132] and formats such as JPEG2000 are capable of either lossless or lossy compression. Lossy compression, however, should be used with caution as it can result in differences in the outputs of classification and unmixing algorithms [134].

Chapter 3

Summary of Papers

3.1 Introduction

In this chapter, the papers that form the core of this thesis are introduced with an overview of the relationships between them and how they address each of the questions and areas of research investigated in this thesis. The papers have been grouped by subject area and ordered with respect to the research questions defined in the Introduction (Chapter 1) rather than in chronological order. Each of the papers is then described in detail and the context, results and contributions summarized.

The first of the core research areas described in this thesis is the question of quality: how to quantify quality and how to evaluate the quality of the data produced by multispectral and hyperspectral imaging. To this end, there are two papers that address this issue. Firstly, there is the basic question of how to measure and quantify the quality of spectra. The first of the core papers, **Paper 1**, “*Quality Evaluation in Spectral Imaging - Quality Factors and Metrics*”, provides an overview of the question of spectral quality, establishing methodologies and tools with which to evaluate spectral data. **Paper 2**, “*Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems*”, uses some of these methodologies and examines the quality of real acquired hyperspectral data. The paper provides a detailed quantitative analysis of the results from an extensive round-robin test of hyperspectral systems involving a number of different targets, acquisition systems and acquisition protocols.

Having established what quality is, and having identified, measured and quantified the quality issues found in acquired spectral data, the next logical ques-

tion is, therefore, how to improve the acquisition and calibration workflows in order to maximize the quality of the data acquired from spectral imaging. This question is addressed by two complementary papers, which each focus on different acquisition technologies. **Paper 3**, “*Studying That Smile: A Tutorial on Multispectral Imaging of Paintings*”, examines the calibration workflow for multispectral imaging and provides a tutorial for the imaging of paintings with a discussion of the issues surrounding this. **Paper 4**, “*Hyperspectral Imaging of Art: Acquisition and Calibration Workflows*”, also looks at calibration, but for hyperspectral imaging. The paper looks in detail at the kinds of problems that can occur during the hyperspectral imaging of works of art and proposes techniques to mitigate these problems. The paper goes on to define a comprehensive workflow to address each stage in the acquisition and calibration pipeline in order to optimize the quality of hyperspectral data.

The final research question concerns access to the very large volumes of data produced by high resolution spectral imaging. How this can be achieved and what techniques and technologies can be developed in order to facilitate this. The research focuses, in particular, on the issue of fast online access to and visualization of the kinds of data sets that are not only massive in size, but which also consist of complex scientific quantitative data. Two papers address this question. The first, **Paper 5**, “*Multi-Modal Data Visualization and Analysis of The Bedroom at Arles by Vincent van Gogh*” discusses multi-modal imaging and proposes an open source software platform for providing access to visualization and analytical tools for data combined from different imaging modalities. These modalities encompass high resolution spectral data, 3D surface topologies and other scientific imagery.

The final core paper, **Paper 6**, “*Web-Based Visualization of Very Large Scientific Astronomy Imagery*”, also examines the question of visualization. The paper builds on the results presented in **Paper 5** and explores the issue of extreme data volumes and how to handle imaging data that is quantitative in nature. The paper evaluates the technology required to address this question and provides a case-study from astronomy where the volumes and complexity of imaging data are at their most extreme. The paper proposes an architecture and evaluates the software implementation created to handle spectral and other scientific imaging with data volumes at the terabyte scale.

3.2 Paper 1: Quality Evaluation in Spectral Imaging - Quality Factors and Metrics

Shrestha, R., Pillay, R., George, S. & Hardeberg, J. Y. "Quality Evaluation in Spectral Imaging - Quality Factors and Metrics". *JAIC - Journal of the International Colour Association*, 12, 22–35 (2014).

The issue of the quality of the data produced by spectral imaging is a fundamental one and the ability to quantify differences and errors is an essential element in order to understand the accuracy and limits of the spectral data produced by a system. This paper, therefore, seeks to provide a comprehensive overview of the existing research in the area of spectral image quality. The paper examines a wide range of metrics and classifies them into categories based on how they were developed, their main features and their intended applications. These include metrics based purely on spectral measures, metrics that combine both spatial and spectral measures, task-specific metrics as well as perceptual metrics.

The variety and wide range of spectral image quality metrics highlights the fact that the quality of a spectral imaging system can be evaluated in a number of ways. Indeed each metric is intended for a number of specific domains and/or applications. The paper, therefore, goes on to discuss the notion of a more generalized framework for the evaluation of spectral imaging quality. As a first step towards this, the paper identifies a number of key factors and attributes at each stage of a spectral imaging workflow that influence the quality of the spectral imaging system. The paper then discusses how these could form the basis of a global framework for evaluating spectral imaging quality.

3.3 Paper 2: Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems

Pillay, R., Picollo, M., Hardeberg, J. Y. & George, S. “Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems”, *Sensors*, 20, 381 (2020)

This paper presents the detailed results from the round-robin test of hyperspectral imaging systems described in Section 2.5.3. A total of 14 different hyperspectral system from 8 different institutions were used to acquire spectral cubes from the VNIR and the SWIR spectral regions. Each system was used to acquire a common set of targets under their normal operating conditions with the data calibrated and processed using the standard processing pipeline for each system.

The test targets presented in this paper consist of a spectral wavelength standard and of a custom-made pigment panel featuring Renaissance-era pigments frequently found in paintings from that period. The quality and accuracy of the resulting data was assessed with quantitative analyses of the spectral, spatial and colorimetric accuracy of the data.

The results provide a valuable insight into the accuracy, reproducibility and precision of hyperspectral imaging equipment when used under routine operating conditions. The distribution and type of errors found within the data can provide useful information on the fundamental and practical limits of such equipment when used for applications such as spectral classification, change detection, colorimetry and others.

The comparison of the data showed that the majority of the systems were able to provide data that was usable and accurate enough for general use. There were, however, significant levels of variability in the data, as can be seen for wavelength standard in Figure 3.1. The spectral, geometric and colorimetric accuracies were highly variable. Several of the systems measured had spectral mis-alignment errors and residual errors were common for all systems. The largest amount of variation between the systems was in terms of noise levels with SNR differing by up to a factor of 6 for the VNIR and 12 for the SWIR systems. This variability can have an important impact on the use of such data for tasks such as material classification or for the detection of change over time.

The results, moreover, highlighted the importance of defining better and standardized workflows for hyperspectral imaging. By identifying the nature of these differences and by quantifying their magnitudes and variability, signifi-

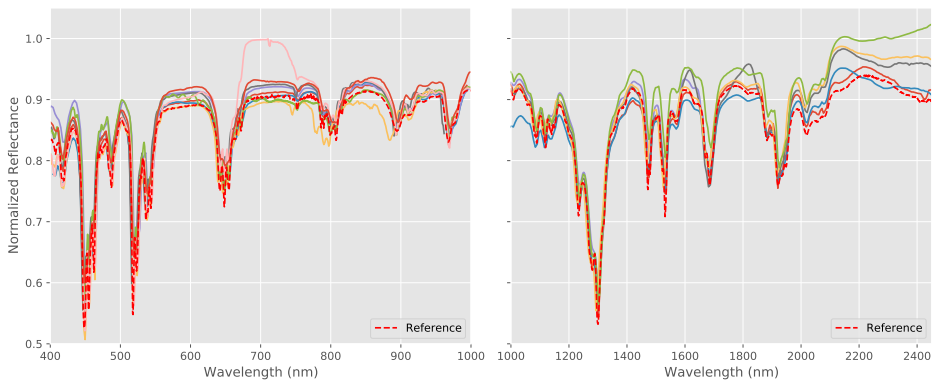


Figure 3.1: Reference and measured reflectances for the wavelength standard for the VNIR systems (left) and SWIR systems (right) used in the round-robin test showing the variability in the spectra

cant improvements can be made to the acquisition and processing pipelines. Furthermore, variability of this sort can have an impact on the behavior of metrics such as spectral distance measures. These metrics form a key component of tasks such as classification, where the choice of threshold parameters can make an important difference to the results obtained. By gaining a better understanding of the characteristics and variability of the underlying data, it will be possible to make better use of hyperspectral data for classification, change detection and other applications.

The paper builds on the results published in Supporting Paper **S1**, “*A Study of Spectral Imaging Acquisition and Processing for Cultural Heritage*”, which introduced and provided an overview of the round-robin test. This supporting paper had included preliminary results from a subset of the acquired data and provided some initial analysis of some of the results. **Paper 2** greatly expands on the results presented in the book chapter and provides a thorough and in-depth analysis of all the available hyperspectral data for the test targets described.

The paper is also related to Supporting Paper **S2**, “*Assessment of Multispectral and Hyperspectral Imaging Systems for Digitisation of a Russian Icon*”, which provides results and analysis from another of the test targets used during the round-robin test.

3.4 Paper 3: Studying That Smile: A Tutorial on Multispectral Imaging of Paintings

Ribés, A., Pillay, R., Schmitt, F & Lahanier, C. “Studying That Smile: A Tutorial on Multispectral Imaging of Paintings”, *IEEE Signal Processing Magazine*, 25, 4, 14-26 (2008)

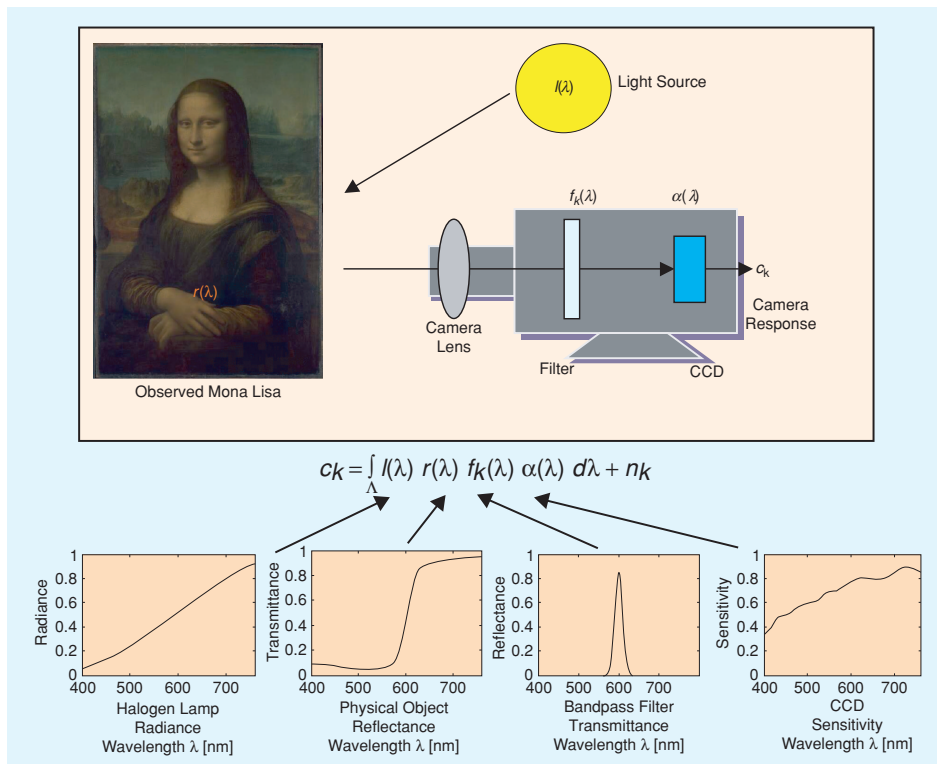


Figure 3.2: Schematic view of the image acquisition process. The camera response depends on the spectral radiance of the light source, the spectral reflectance of the objects in the scene, the spectral transmittance of the color filter, and the spectral sensitivity of the sensor

The paper examines in detail the acquisition and calibration workflow necessary for multispectral imaging of paintings. The paper looks at each of the various key elements that make up a multispectral imaging system and discusses how each of these elements interact and contributes to the final spectral image. These include elements such as sensor sensitivity, lighting conditions, filter design and selection, and the various sources of noise within the system. Models

for each of these elements are proposed, which allow a better understanding of how they behave and what the impact of each component is on the data.

The paper goes on to integrate these elements within a signal-processing framework for multispectral imaging. In order to do this, a global mathematical model is proposed, that can be seen in Figure 3.2. This model includes the key components that make up such an imaging system and which allow the spectral image acquisition process to be modeled quantitatively. The various optimization problems involved in the design of multispectral cameras, their calibration, and the processing of the obtained data are also introduced within the same signal processing model and discussed within the paper.

Finally, the paper takes example data from a multispectral acquisition of Leonardo da Vinci's painting, *Mona Lisa* (1503-1506, oil on poplar), from the Musée du Louvre to illustrate the working of the model and to provide a concrete example of how the model can be applied to correctly acquire and calibrate multispectral image data. The framework is also used to perform virtual de-varnishing of the painting. A de-varnished simulation of the "cleaned" painting is generated which is calculated entirely in the spectral domain.

3.5 Paper 4: Hyperspectral Imaging of Art: Acquisition and Calibration Workflows

Pillay, R., Hardeberg, J. Y. & George, S. “Hyperspectral Imaging of Art: Acquisition and Calibration Workflows”. *Journal of the American Institute for Conservation* 58, 3-15 (2019).

Hyperspectral imaging has become an increasingly used tool in the analysis of works of art. Data that is both spectrally and spatially accurate is an essential step in order to obtain useful and relevant results from hyperspectral imaging. Data that is too noisy or inaccurate will produce sub-optimal results when used for pigment mapping, the detection of hidden features, change detection or for quantitative spectral documentation. However, the quality of the acquired data and the processing of that data to produce accurate and reproducible spectral image cubes can be a challenge to many cultural heritage users. **Paper 2** had indeed highlighted the fact that the experimental acquisition of accurate hyperspectral data is not straightforward and that it is common to observe problems and errors in the data acquired in typical laboratory conditions and this paper makes reference to these results.

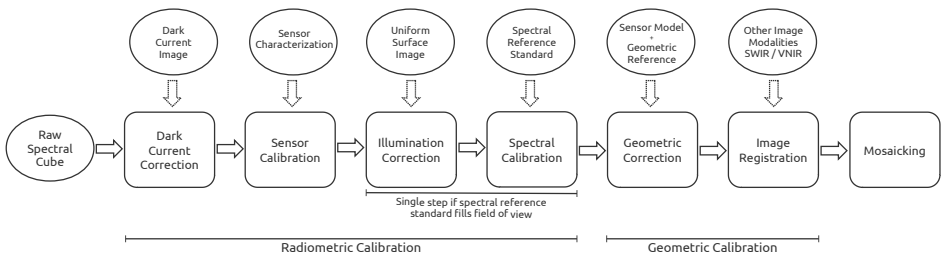


Figure 3.3: Workflow for hyperspectral calibration showing the processing steps and the data necessary to perform them

In order to improve, therefore, on the quality of acquired data, this paper sets out to examine in detail the acquisition and processing workflows necessary to improve data quality and looks, in particular, at the specific issues that surround the hyperspectral imaging of works of art. These workflows include the key parameters that must be addressed during acquisition and the essential steps and issues at each of the stages required during post-processing in order to fully calibrate hyperspectral data. In addition the paper describes in detail the key issues that affect data quality and proposes practical solutions that can make significant differences to overall hyperspectral image quality.

The paper proposes a comprehensive workflow to address each stage in the acquisition and calibration pipeline, which is shown in Figure 3.3. The definition of a workflow suited to the acquisition of works of art will help guide users within the field of cultural heritage and help avoid common problems and errors. In this way, the quality of the resulting hyperspectral data can be greatly improved and data between institutions made more comparable.

The paper complements **Paper 3**, which examines the calibration of multispectral imaging.

3.6 Paper 5: Multi-Modal Data Visualization and Analysis of “The Bedroom at Arles” by Vincent van Gogh

Pillay, R., “Multi-Modal Data Visualization and Analysis of *The Bedroom at Arles* by Vincent van Gogh”, *Proceedings of the 22nd CIPA Symposium*, October 11-15, 2009, Kyoto, Japan.

Scientific imaging of works of art can include a wide range of instrumentation and modalities. These can include traditional scientific imaging techniques such as infra-red imaging, radiography and ultra-violet fluorescence as well as more advanced acquisition techniques such as 3D imaging, multispectral or hyperspectral imaging. The resulting imaging data is often difficult to combine and visualize in a coherent manner, especially if imaging is of high resolution or if remote networked access to the images is required. This paper, therefore, addresses these issues and provides a solution to the question of how to visualize and make accessible online high resolution multi-modal imaging data.



Figure 3.4: Browser-based interface showing image blending functionality with the superposition of the color and X-ray images and the blended smooth transition between them

In this paper, an open source software platform for providing high resolution multi-modal visualization and analysis of advanced heterogeneous scientific imagery is presented. The system is designed around the needs and requirements of cultural heritage and provides a powerful, flexible and easy-to-use means for conservators, researchers or art historians to interact and analyze scientific imagery of works of art. The platform allows cultural heritage institutions to manage and maintain very large image data sets and for multiple simultaneous users to remotely visualize heterogeneous data at high resolution using efficient image streaming techniques. As a case study, the paper presents an example of an analysis carried out of a work by Vincent Van Gogh: *The Bedroom at Arles*. The extensive range of imaging techniques carried out for this painting was unique in its depth and range and showcase the potential of multi-modal visualization techniques, such as the ability to transition between registered sets of images (Figure 3.4).

The visualization methodology is also featured in Supporting Paper **S2**, “*Assessment of Multispectral and Hyperspectral Imaging Systems for Digitisation of a Russian icon*”, which uses the software described here to enable high resolution online access to some of the hyperspectral data presented in that paper.

3.7 Paper 6: Web-Based Visualization of Very Large Scientific Astronomy Imagery

Bertin, E., Pillay, R. & Marmo, C. “Web-Based Visualization of Very Large Scientific Astronomy Imagery”. *Astronomy and Computing* 10, 43–53 (2015).

The final core paper in this thesis extends **Paper 5** and concerns access to the very large volumes of data produced by high resolution quantitative imaging. In particular, the research focuses on the issue of fast instant access to and visualization of the kinds of data sets that are not only massive in size, but which also consist of complex scientific quantitative data. This paper, therefore, looks at the technology required to address this question and provides a case-study from astronomy where the volumes and complexity of imaging data are at their most extreme. The paper describes and evaluates the architecture and software created to handle spectral and other scientific imaging with data volumes at the terabyte scale.

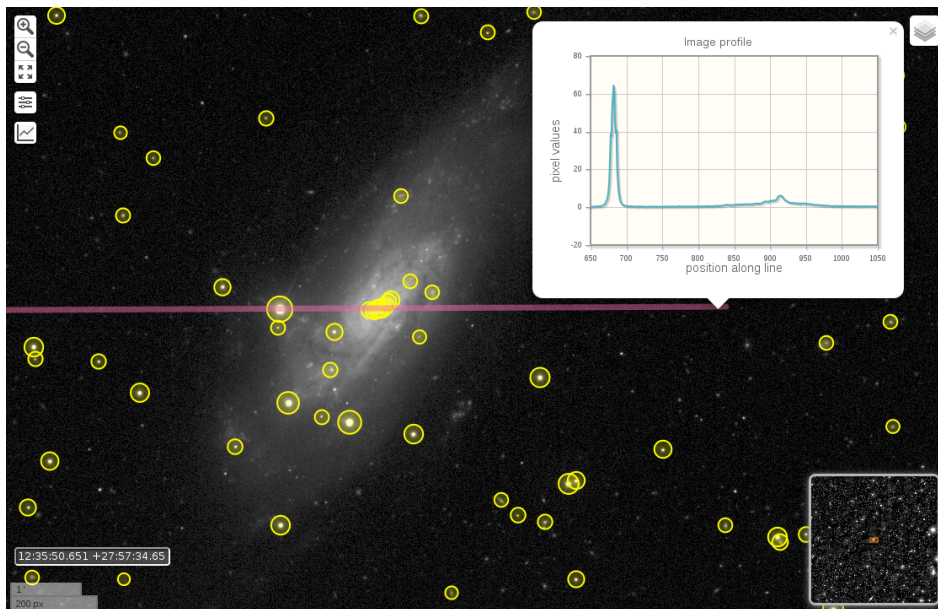


Figure 3.5: Visualizing a 1TB floating point image with vector layers and the ability to extract and graph an image profile

Visualizing and navigating through large astronomy images online can be slow and lacking in ergonomics, especially on mobile devices. The paper, therefore,

examines the issues surrounding this and proposes a high performance, versatile and robust client–server system for remote visualization and analysis of extremely large scientific images. A full working software implementation is developed that can be used for various applications within the field of astronomy such as survey image quality control, interactive data query and exploration, citizen science, as well as public outreach. The proposed software is entirely open source and is designed to be generic and applicable to a variety of data sets. The software is able to provide access to floating point data at terabyte scales, with the ability to precisely adjust image settings in real-time. The clients, such as that shown in Figure 3.5, are light-weight, platform-independent web applications built on standard HTML5 web technologies and are compatible with both touch and mouse-based devices.

As part of the work carried out for the paper, a public demo of a floating point image of one terabyte in size was made available online (the largest single image ever put online) and performance testing was carried out, which demonstrated the scalability of the system. The results show how a single server can comfortably handle more than a hundred simultaneous users accessing full precision 32 bit astronomy data.

Chapter 4

Discussion

In this chapter, we will look more globally at the research presented in this thesis and discuss what this research has contributed to the fields of multispectral and hyperspectral imaging. We will examine how these papers have answered the specific research questions that were put forward in Chapter 1 as motivations for this thesis. We will also discuss the scope and limits of the research and try to put the work into context.

The overarching aim of this research was to address some of the key issues surrounding the multispectral and hyperspectral imaging of art. Although the use of such spectral imaging technologies has advanced within the field of cultural heritage over the last two decades, a number of challenges exist that have prevented spectral imaging from realizing its full potential, from becoming more widely used, and which have prevented it from becoming a routine analytical tool.

Multispectral and hyperspectral imaging systems are complex and precise scientific instruments that require careful acquisition and calibration in order to produce quantitative data that is useful, accurate and repeatable. Understanding and correctly carrying out the full imaging workflow from the instrumentation setup and fixing of parameters, through the acquisition steps through to data processing in order to obtain final fully calibrated spectral images can be challenging for users within the field of cultural heritage and indeed beyond. In addition, multispectral and hyperspectral imaging can produce vast quantities of data that require complex processing and the ability to manage the large volumes of resulting data. Moreover, the final high resolution and multi-dimensional data that is produced can be difficult to store, use and visualize.

The thesis, therefore, covers two main research areas. The first concerns the quality of spectral data and examines how this can be measured, quantified and how the acquisition and calibration of multispectral and hyperspectral images can be improved. The second research area concerns the accessibility and visualization of such data and examines ways to facilitate and make very large and complex image data accessible online.

Within these two research areas, three more specific research questions were defined that the work presented in this thesis was intended to answer:

1. How can the quality of spectral data be measured? What is the quality of acquired experimental spectral data?
2. What are the factors that influence multispectral and hyperspectral data quality and how can acquisition and calibration workflows be improved?
3. How to make high resolution quantitative image data accessible? How can large volumes of spectral and other scientific imaging data be visualized?

4.1 Research Contributions

The work presented in this thesis seeks to address each of the research questions and propose analysis, tools, workflows and methodologies that answer the questions. We will, therefore, look at each of these questions in more detail and discuss what the research has contributed and to what degree the questions have been answered.

4.1.1 How can the quality of spectral data be measured? What is the quality of acquired experimental spectral data?

To be able to understand quality, it is necessary to be able to define what exactly quality is and determine how to measure and quantify it. Assessing the quality of a spectral image is not straightforward and there are many different criteria that depend on exactly how the resulting data will be used and which particular aspects of the spectra are more important. For example, is the shape rather than the absolute value of the spectra more important? Or are we trying to identify the location of specific spectral features? **Paper 1** attempts to provide an answer to the first part of our research question by providing a broad survey of the state of the art of the various spectral distance measures and quality

metrics that exist and that can be applied to spectral images. The paper is essentially a review and its first research contribution is, therefore, to bring these disparate measures and metrics together and provide a typology of metrics that can help users better understand and select appropriate measures. The second important research contribution is the discussion in the paper of the concept of a more generalized global framework for quantifying spectral image quality.

In order to answer the second part of our research question, it is necessary to evaluate the quality of spectral data that is output by real multispectral or hyperspectral imaging systems. This is something that has been widely carried out for individual hardware, as for example in [135] and [136]. However, to fully answer this question, it is also necessary to obtain test data from as wide a range of systems as possible and from as wide a range of test targets and operating conditions as possible. Such a wide range of data is something that is rarely readily available. The spectral imaging round-robin test that is described in Section 2.5.3 provided precisely such an opportunity and included spectral data obtained from different test targets and from a wide range of spectral imaging systems, giving a unique insight into the level of quality that it is possible to obtain in practice.

Paper 2 delivers the most comprehensive reporting and analysis of the results from this round-robin test and provides access to the test results in the form of open data¹. The paper also includes detailed supplementary material on the hardware and acquisition parameters used, which provides valuable reference material for future practitioners. **Paper 2**, therefore, can be seen together with the Supporting Papers, **S1** and **S2**, as answering in large part the second half of our research question.

Limitations

The scope of **Paper 1** is limited to providing an overview and classification of spectral metrics. To be more comprehensive, a more detailed analysis would be required showing the properties of each metric and in which situations each is adapted. In addition, although the issue of a general spectral image quality framework is discussed, no concrete proposition is made for one.

Although **Paper 2** provides a very comprehensive analysis of the results from the round-robin test, it limits itself to the subset of hyperspectral data that was acquired, leaving an important body of multispectral data only published in a fragmentary way. Also the round-robin test protocol limited itself to the collec-

¹<https://github.com/ruven/RRT>

tion of final processed data. It would have been interesting to have also been able to compare the raw data produced.

4.1.2 What are the factors that influence multispectral and hyperspectral data quality and how can acquisition and calibration workflows be improved?

Multispectral and hyperspectral imaging encompass a range of technologies, which although they share common concepts, require different workflows and processing. The subject of how to improve acquisition procedures and how to calibrate data in the best possible way is an important one for fostering the wider use of spectral imaging.

One of the most important contributions of this thesis is to answer this question by examining acquisition and calibration workflows and providing guidelines for both multispectral and hyperspectral systems. As these technologies were initially developed for space and remote sensing, this question has been studied in detail within those fields for multispectral [137] and for hyperspectral systems [138, 35]. For laboratory-based hyperspectral systems a range of research has been carried out [139, 140, 141, 142, 143, 144]. More globally, there is the on-going initiative from the IEEE SA P4001 Working Group on the characterization and calibration of hyperspectral imaging devices [126].

However, the spectral imaging of art can require different methodologies and workflows and the issues surrounding this are addressed in two papers, **Paper 3** and **Paper 4**. These papers cover different technologies: multispectral cameras in **Paper 3** and hyperspectral systems in **Paper 4**. This combination provides a comprehensive overview of the subject, providing both practical insights as well as a conceptual framework for system calibration.

The papers are highly complementary, not only in terms of the different technologies studied, but also in the very different philosophical approaches taken to address the topic. **Paper 3** is presented as a tutorial and establishes a theoretical framework for the multispectral imaging of art based on signal processing approaches and provides a conceptual understanding of the underlying processes. **Paper 4** covers hyperspectral imaging and, in particular, pushbroom-based systems and provides a more practical overview of the topic and looks in detail at the acquisition process, providing advice and highlighting common errors and pitfalls that can occur during acquisition or processing. The paper goes on to define a full workflow for both the acquisition and processing of hy-

perspectral data that aims to enable users to produce the best quality spectral data possible.

The combination of these two papers makes it possible to answer our second research question in a uniquely broad and comprehensive way covering two generations of spectral imaging technology.

The papers also include other research contributions worth highlighting here. **Paper 3** provides an early use-case for the application of spectral imaging and presents an example of the use of virtual restoration applied to a painting. The restoration is performed entirely in the spectral domain rather than using colorimetric information. **Paper 4** also goes beyond simply providing guidelines or proposing abstract workflows and initiates efforts to develop fully open source tools for spectral image calibration.

Limitations

Although these papers provide a solid basis for understanding spectral calibration and the workflow necessary to produce fully calibrated data, several aspects are either not addressed or only addressed superficially. In particular, **Paper 3** concentrates on radiometric calibration and does not address the question of geometric calibration, which is an important issue when using the images for multi-modal analysis or registration. Geometric calibration is addressed more substantially in **Paper 4**, but the topic is a large one and, although research exists for both multispectral [112] and hyperspectral imaging [145], works of art present specific problems, requiring modified or custom solutions.

4.1.3 How to make high resolution quantitative image data accessible? How can large volumes of spectral and other scientific imaging data be visualized?

Our final research question concerns networked access to and the visualization of high resolution spectral and other multi-modal images, and this topic is addressed by the final two papers that make up this thesis: **Paper 5** and **Paper 6**.

Paper 5 presents an architecture and software platform² that aims to provide a solution to our research question. The solution is presented through a case-study involving high resolution multi-modal and heterogeneous scientific imagery. The platform described in the paper shows not only how web-based

²<https://iipimage.sourceforge.io>

interfaces can be used to enable high resolution images to be viewed, but also how more advanced features such as blending between different modalities, on-the-fly processing of CIELAB or elevation data can be integrated into such a platform. The case-study includes an unprecedented range of imaging modalities acquired using a number of different acquisition techniques.

The research contribution of the paper includes the introduction of a number of innovative concepts for the study of paintings. These include the use of standards-based web platforms for carrying out the analysis of art and the introduction of image blending as a means of interactive comparison of different modalities. The platform described in the paper is also an important contribution, which provides a full working implementation of software for high resolution multi-modal web-based visualization.

Paper 6 takes and extends the ideas in **Paper 5** and applies them to, what is on the surface, a very different field. The field of astronomy, however, is also one rich in imaging and where scientific and multi-modal imagery are widely used in ways not too dissimilar to those within cultural heritage. The major difference is in the sizes and complexity of those images. Images acquired for astronomy can be orders of magnitude greater in size than those acquired for paintings. Efficiently handling such image data is a major challenge and **Paper 6** pushes the concepts presented in **Paper 5** to their scientific and technical limits.

The research contributions of **Paper 6** include the adaptation of the software platform to astronomy, its extension to floating-point data and an evaluation of the performance and issues surrounding the visualization of extremely large images. The paper also presents a working demo showcasing this capability with an example image of one terabyte in size, making it the largest single image ever put online.

Thus, the work presented here does indeed provide a powerful solution to our research question, providing a means of making very high resolution quantitative image data accessible and allowing spectral data to be efficiently visualized online.

Limitations

The work and software presented cannot, of course, solve the visualization needs for all types of image data. As discussed in **Paper 6**, the system employs a large degree of server-side processing to maximize efficiency with very large image sizes. However, increases in client-side processing capabilities and

the convergence of web browsers towards web standards now makes a hybrid model a real possibility in which it is now feasible for more sophisticated processing to be carried out client-side. Moreover, although the use of browser-based interfaces makes the model highly portable across systems, dedicated desktop applications will always be capable of better performance and, therefore, of handling more sophisticated processing.

4.2 Scope of the Research

4.2.1 Breadth of the Research

The title of this thesis is *Multispectral and Hyperspectral Imaging of Art: Quality, Calibration and Visualization* and a large part of the work undertaken as well as the research included here is indeed focused on multispectral and hyperspectral imaging within the field of cultural heritage. However, several of the papers present research that is broader in scope than this field or than of the imaging of art. **Paper 1**, for example, provides a broad overview of spectral metrics which is not specific to any single field. **Paper 2** concerns the results from a round-robin test conducted within a project whose subject was, indeed, the documentation of cultural heritage. However, such large-scale round-robin tests are rare and the results and analysis are of interest to a wider range of users, especially to users of laboratory-based spectral imaging equipment or spectral imaging carried out at close range more generally.

The two papers on the topic of calibration, **Paper 3** and **Paper 4**, certainly focus on calibration for art applications. But, the models and workflows they present are sufficiently generic to be applied to other fields. The issues, guidelines and workflows described in **Paper 4**, for example, are applicable to any short-range laboratory-based hyperspectral imaging scenario.

This wider scope is even more clearly apparent for the two papers on remote visualization. Although **Paper 5** presents a case-study involving a work of art, the concepts, architectures and technologies are applicable to any kind of multi-modal imaging. Indeed, **Paper 6**, is explicitly outside of the field of cultural heritage, being a paper on imaging within astronomy. There is, nevertheless, a certain degree of overlap between the two fields in terms of imaging and **Paper 6** has been cited not only from within the field of astronomy but also by papers on cultural heritage as well as from the disciplines of hyperspectral imaging, medical imaging, biology and scientific visualization [146].

Thus, although the research included in this thesis aims to address the spectral imaging of art, much of the work is clearly of wider interest and of a more universal nature.

4.2.2 Chronology and Timeline

The research included within this thesis includes several separate strands that have been conducted in parallel. The papers presented here have, therefore, been grouped and presented by subject rather than in chronological order in order to create a more logical narrative to this thesis.

The research also spans a relatively large period of time with there being over ten years between the publication of the earliest paper and that of the latest. Such a long period creates the risk that parts of the research become obsolete over time. However, a closer examination shows that this is not the case for the body of work presented here.

The earliest paper included in this thesis, **Paper 3**, concerns the calibration of multispectral cameras. The paper is presented as a tutorial and remains relevant. Technology has advanced in that time and many users have migrated to hyperspectral technology, but the fundamental principles described in **Paper 3** remain unchanged and indeed remain applicable to both multispectral and snapshot hyperspectral systems. **Paper 4** was published over a decade after **Paper 3** and covers the related topic of the calibration of hyperspectral systems. Considerable experience has been gained over this intervening period of time and **Paper 4** has benefited from this and has integrated much of this experience into its guidelines.

A related scenario exists for the first of the two papers that cover the subject of image accessibility and visualization. The long timeframe between the publication of **Paper 5** and the publication of this thesis provides a rare opportunity to assess the impact of this research and of the software described. The paper and the research behind it aimed to provide tools and facilitate the availability of scientific imagery online. The notion of data sharing and reuse has had a slow uptake in the field of cultural heritage and was far from accepted at the time of publication. It is only relatively recently that institutions have started to make such image data publicly accessible and many of the ideas presented as well as the software itself have been more widely taken up by major museums [147] as well as by international collaborative archives [148, 149].

Nevertheless, web technologies have advanced since the publication of **Paper 5**, particularly browser technologies such as WebGL [150], which were not gener-

ally available until several years after publication. Such technology potentially makes the kind of sophisticated client-side processing mentioned in Section 4.1.3 possible. However, this kind of processing has yet to be implemented in any comparable platform for scientific or high resolution images. For 3D models, however, WebGL-based tools have been developed and successfully applied to cultural heritage [151].

4.2.3 Research Impact

Making an assessment of the impact of research is often difficult to do. However, the longevity of two of the papers allows a step back to be taken to make a more considered assessment of this impact.

Paper 3 was the culmination of a larger co-coordinated research effort on multispectral imaging, which had produced a series of research [38, 40, 152], which had made a significant contribution to the multispectral imaging of art. **Paper 3** itself has had over 30 citations [153], not only from within cultural heritage, but also in fields such as signal processing [154] and optics [155].

The work presented in **Paper 5** can also be considered to have made an important impact on the use of images within cultural heritage. The work and software has contributed to the development of international initiatives to foster image re-use and sharing over the last few years. This has included, in particular, the work surrounding IIIF (the International Image Interoperability Framework) [156], which has become a de-facto standard within the field and for which the software described here provides a reference implementation. The software has, moreover, become a key part of the infrastructures for image visualization not only within the field of cultural heritage but also beyond, where direct use has been made by a wide range of institutions [157], including within fields such as biology [158], microscopy [159] and, of course, astronomy through **Paper 6**.

Chapter 5

Conclusion

Multispectral and hyperspectral imaging can be powerful tools for analyzing and documenting works of art due to their ability to simultaneously capture both accurate spectral and spatial information. The data can be used for a wide range of diagnostic and analytical purposes, including materials identification, pigment mapping, the detection of hidden features or areas of lost material, for colorimetric analysis or for precise quantitative documentation.

Thanks to this versatility, multispectral and hyperspectral imaging have become increasingly used techniques in the field of cultural heritage. However, a number of technical challenges exist which have prevented them from realizing their full potential and which have prevented the technologies from becoming a more widely used and routine analytical tool.

Both multispectral and hyperspectral imaging systems require careful and precise acquisition workflows in order to produce useful data. In addition, processing and calibration of the acquired data can be a challenge for many cultural heritage users. Hyperspectral imaging, in particular, can produce vast quantities of raw data that require complex processing and the ability to manage the large resulting volumes of data. Moreover, the final high resolution and multidimensional data that is produced can be difficult to use or to visualize. This thesis has, therefore, sought to address some of these issues and has endeavored to analyze and quantify problems and propose tools, workflows and methodologies to resolve and mitigate these issues.

The research has focused on two main areas. The first concerned the quality of spectral data and how to measure, quantify and improve it. The work has sought to determine what spectral quality is and what methods can be used

to quantify it. The research went on to apply these methods to evaluate the levels of quality seen in data acquired under routine operating conditions and, in particular, data acquired from an extensive round-robin test of hyperspectral imaging systems.

In order to improve the quality of spectral data, research was then carried out in order to better understand the elements that affect spectral quality within a system and, thereby, to improve the acquisition and calibration pipelines for spectral imaging systems. This was carried out for both multispectral and hyperspectral imaging with complementary work contributing both a signal processing framework for calibration as well as practical guidelines and workflows.

The second research area concerned the visualization of multispectral and hyperspectral image data and examined ways to facilitate and make large and complex images accessible online. For this, an architecture, visualization techniques and a full software platform were presented for the efficient distribution and visualization of high resolution multi-modal and spectral image data. This work was then extended in order to push the technology to the limits and to apply the techniques to the field of astronomy where image sizes are at their most extreme.

This ensemble of research has helped advance the state of the art along several axes of research and has provided a better understanding of some of the issues surrounding spectral imaging. It has also provided practical solutions and guidelines to help future users better use both multispectral and hyperspectral imaging equipment for the spectral imaging of art.

5.1 Perspectives

The research in this thesis covers some of the key areas within the multispectral and hyperspectral imaging pipelines. The processes of acquisition and calibration form the fundamental basis for any spectral imaging and the work presented here provides a solid foundation for understanding and improving these workflows. Each of the various elements can, of course, be optimized. In particular, there is significant scope for further research in the areas of geometric calibration, mosaicking and the fusion of spectral data from different wavelength ranges.

There is scope also for the further development of the workflows and guidelines developed for hyperspectral acquisition and calibration in association with international standardization efforts, such as the IEEE Standards Association

P4001 Working Group on *Characterization and Calibration of Ultraviolet through Shortwave Infrared Hyperspectral Imaging Devices* [126], which aims to provide recommendations for application-specific tasks.

A particularly important area of research is on the use of spectral data for tasks such as materials analysis or pigment mapping. This topic was outside of the scope of this thesis, but, as described in Section 2.7, is one of the core applications of multispectral and hyperspectral imaging within cultural heritage. Pigment mapping, for example, is based on the use of spectral classification methods, which have been developed around the needs of remote sensing. Strong potential exists for improving the results obtained from these techniques by better employing or adapting them to the field of cultural heritage. Each of the various steps within the classification pipeline can be improved, including the determination of the number of spectral end-members, the calculation of those end-members, the appropriate choice of spectral distance measure and the application of suitable clustering and classification methods. In addition, issues such as the simplification and automation of these processes need to be addressed.

A further perspective concerns the combination of spectral data with other imaging modalities or with other types of chemical mapping techniques, such as X-ray Fluorescence (XRF) mapping. Although, the visualization of such data has been addressed within this thesis through the work on multi-modal visualization, the use of joint analysis for pigment mapping purposes has significant potential and has, thus far, been rarely used within the field.

Other forms of combination that have significant potential include the fusion of spectral images with 3D data or with elevation profiles of the surfaces of paintings.

Bibliography

- [1] Raju Shrestha, **Ruven Pillay**, Sony George, and Jon Y. Hardeberg. Quality evaluation in spectral imaging – Quality factors and metrics. *JAIC - Journal of the International Colour Association*, 12(0):22–35, July 2014.
- [2] **Ruven Pillay**, Marcello Picollo, Jon Yngve Hardeberg, and Sony George. Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems. *Sensors*, 20(14):3812, January 2020.
- [3] Alejandro Ribés, **Ruven Pillay**, Francis Schmitt, and Christian Lahanier. Studying That Smile. *Signal Processing Magazine, IEEE*, 25(4):14–26, July 2008.
- [4] **Ruven Pillay**, Jon Y. Hardeberg, and Sony George. Hyperspectral imaging of art: Acquisition and calibration workflows. *Journal of the American Institute for Conservation*, 58(1-2):3–15, February 2019.
- [5] **Ruven Pillay**. Multi-Modal Data Visualization and Analysis of “The Bedroom at Arles” by Vincent van Gogh. In *Proceedings of the 22nd CIPA Symposium*, Kyoto, October 2008.
- [6] Emmanuel Bertin, **Ruven Pillay**, and Chiara Marmo. Web-based visualization of very large scientific astronomy imagery. *Astronomy and Computing*, 10:43–53, April 2015.
- [7] Lindsay W. MacDonald, Tatiana Vitorino, Marcello Picollo, **Ruven Pillay**, Michał Obarzanowski, Joanna Sobczyk, Sérgio Nascimento, and João Linhares. Assessment of multispectral and hyperspectral imaging systems for digitisation of a Russian icon. *Heritage Science*, 5:41, September 2017.

- [8] Sony George, Jon Ynvge Hardeberg, João Linhares, Lindsay MacDonald, Cristina Montagner, Sérgio Nascimento, Marcello Picollo, **Ruven Pillay**, Tatiana Vitorino, and E. Keats Webb. A Study of Spectral Imaging Acquisition and Processing for Cultural Heritage. In *Digital Techniques for Documenting and Preserving Cultural Heritage*, pages 141–158. Arc Humanities Press, Leeds, January 2018.
- [9] Alexander F.H. Goetz. Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sensing of Environment*, 113(Supplement 1):S5–S16, September 2009.
- [10] Di Wu and Da-Wen Sun. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review — Part I: Fundamentals. *Innovative Food Science & Emerging Technologies*, 19:1–14, July 2013.
- [11] E. Keith Hege, Dan O’Connell, William Johnson, Shridhar Basty, and Eustace L. Dereniak. Hyperspectral imaging for astronomy and space surveillance. In *Imaging Spectrometry IX*, volume 5159, pages 380–391, San Diego, USA, January 2003. International Society for Optics and Photonics.
- [12] F. A. Kruse, J. W. Boardman, and J. F. Huntington. Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping. *IEEE Transactions on Geoscience and Remote Sensing*, 41(6):1388–1400, June 2003.
- [13] Driss Haboudane, John R Miller, Elizabeth Pattey, Pablo J Zarco-Tejada, and Ian B Strachan. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90(3):337–352, April 2004.
- [14] Guolan Lu and Baowei Fei. Medical hyperspectral imaging: A review. *Journal of Biomedical Optics*, 19(1):010901–010901, 2014.
- [15] Athos Agapiou, Diofantos G. Hadjimitsis, Dimitrios Alexakis, and Apostolos Sarris. Observatory validation of Neolithic tells (“Magoules”) in the Thessalian plain, central Greece, using hyperspectral spectroradiometric data. *Journal of Archaeological Science*, 39(5):1499–1512, May 2012.
- [16] David Saunders and John Cupitt. Image Processing at the National Gallery: The VASARI Project. *The National Gallery Technical Bulletin*, 14:72–85, January 1993.

-
- [17] A. Casini, M. Bacci, C. Cucci, F. Lotti, S. Porcinai, M. Picollo, B. Radicati, M. Poggese, and L. Stefani. Fiber optic reflectance spectroscopy and hyper-spectral image spectroscopy: Two integrated techniques for the study of the Madonna dei Fusi. In *Optical Methods for Arts and Archaeology*, volume 5857, pages 58570M–8, Munich, Germany, June 2005. SPIE.
- [18] Kirk Martinez, John Cupitt, and David R Saunders. High-resolution colorimetric imaging of paintings. *Proceedings of SPIE*, 1901(1):25–36, May 1993.
- [19] Henri Maître, Francis Schmitt, Jean-Pierre Crettez, Yifeng Wu, and Jon Yngve Hardeberg. Spectrophotometric Image Analysis of Fine Art Paintings. *Color and Imaging Conference*, 1996(1):50–53, January 1996.
- [20] Roy S. Berns. The Science of Digitizing Paintings for Color-Accurate Image Archives: A Review. *Journal of Imaging Science*, 45(4):305–325, July 2001.
- [21] André Chéron. The Radiography of Pictures. *Journal of the Röntgen Society*, 17(68):120–121, July 1921.
- [22] Marie Farnsworth. Infra-red absorption of paint materials. *Technical Studies in the Field of the Fine Arts*, 7:88–98, 1938.
- [23] J. R. J. van Asperen de Boer. Reflectography of Paintings Using an Infrared Vidicon Television System. *Studies in Conservation*, 14(3):96, August 1969.
- [24] James Joseph Rorimer. Ultra-violet rays and their use in the examination of works of art. *Parnassus*, 4(4), 1931.
- [25] Di-Yuan Tzeng and Roy S. Berns. Spectral-Based Six-Color Separation Minimizing Metamerism. *Color and Imaging Conference*, 2000(1):342–347, January 2000.
- [26] Peter D. Burns and Roy S. Berns. Analysis Multispectral Image Capture. *Color and Imaging Conference*, 1996(1):19–22, January 1996.
- [27] D. Connah, J.Y. Hardeberg, and S. Westland. Comparison of linear spectral reconstruction methods for multispectral imaging. In *Image Processing, 2004. ICIP '04. 2004 International Conference On*, volume 3, pages 1497 – 1500 Vol. 3, October 2004.

- [28] Michael E. Schaepman, Susan L. Ustin, Antonio J. Plaza, Thomas H. Painter, Jochem Verrelst, and Shunlin Liang. Earth system science related imaging spectroscopy—An assessment. *Remote Sensing of Environment*, 113(Supplement 1):S123–S137, September 2009.
- [29] Robert O. Green, Michael L. Eastwood, Charles M. Sarture, Thomas G. Chrien, Mikael Aronsson, Bruce J. Chippendale, Jessica A. Faust, Betina E. Pavri, Christopher J. Chovit, Manuel Solis, Martin R. Olah, and Orlesa Williams. Imaging Spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sensing of Environment*, 65(3):227–248, September 1998.
- [30] C. Durell. IEEE P4001 Hyperspectral Standard: Progress and Cooperation. In *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 4432–4434, July 2019.
- [31] John R. Gilchrist, David W. Allen, and Chris Durell. IEEE P4001: Progress towards a hyperspectral standard. In *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imagery XXV*, volume 10986, page 1098602. International Society for Optics and Photonics, May 2019.
- [32] A. D. Meigs, L. J. Otten, and T. Y. Cherezova. Ultraspectral imaging: A new contribution to global virtual presence. In *1998 IEEE Aerospace Conference Proceedings (Cat. No.98TH8339)*, volume 2, pages 5–12 vol.2, March 1998.
- [33] Thomas S. Pagano. Ultraspectral infrared technology development on the Atmospheric Infrared Sounder (AIRS) and future applications. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X*, volume 5425, pages 304–310. International Society for Optics and Photonics, August 2004.
- [34] Steven A. Macenka and Michael P. Chrisp. Airborne Visible/Infrared Imaging Spectrometer (Aviris) Spectrometer Design And Performance. In *Imaging Spectroscopy II*, volume 0834, pages 32–43. International Society for Optics and Photonics, January 1987.
- [35] Michael Eismann. *Hyperspectral Remote Sensing*, volume PM210. SPIE Press, Bellingham, Washington, April 2012.
- [36] Dimitris G. Manolakis, Ronald B. Lockwood, and Thomas W. Cooley. *Hyperspectral Imaging Remote Sensing: Physics, Sensors, and Algorithms*. Cambridge University Press, October 2016.

-
- [37] Kirk Martinez, John Cupitt, David Saunders, and **Ruven Pillay**. Ten Years of Art Imaging Research. *Proceedings of the IEEE*, 90(1):28–41, August 2002.
- [38] Christian Lahanier, Georges Alquié, Pascal Cotte, Constantinos Christofides, Christophe de Deyne, **Ruven Pillay**, David Saunders, and Francis Schmitt. CRISATEL: High definition spectral digital imaging of paintings with simulation of varnish removal. In *ICOM Committee for Conservation, ICOM-CC : 13th Triennial Meeting, Rio de Janeiro, 22-27 September 2002*, pages 295–300, Rio de Janeiro, Brazil, September 2002. ICOM-CC, James & James.
- [39] Christian Fischer and Ioanna Kakoulli. Multispectral and hyperspectral imaging technologies in conservation: Current research and potential applications. *Studies in Conservation*, 51(Supplement-1):3–16, June 2006.
- [40] Alejandro Ribés, Francis Schmitt, **Ruven Pillay**, and Christian Lahanier. Calibration and Spectral Reconstruction for CRISATEL: An Art Painting Multispectral Acquisition System. *Journal of Imaging Science and Technology*, 49(6):563–573, November 2005.
- [41] Hideaki Haneishi, Takayuki Hasegawa, Asako Hosoi, Yasuaki Yokoyama, Norimichi Tsumura, and Yoichi Miyake. System Design for Accurately Estimating the Spectral Reflectance of Art Paintings. *Applied Optics*, 39(35):6621–6632, December 2000.
- [42] A. Pelagotti, A. Del Mastio, A. De Rosa, and A. Piva. Multispectral imaging of paintings. *IEEE Signal Processing Magazine*, 25(4):27–36, 2008.
- [43] F. Imai, Mitchell R. Rosen, and R. Berns. Multi-spectral Imaging of Van Gogh’s Self-portrait at the National Gallery of Art Washington DC. In *Proceedings of IS&T PICS Conference*, pages 185–189, Rochester, USA, 2001. IS&T.
- [44] Gianluca Novati, Paolo Pellegrini, and Raimondo Schettini. An affordable multispectral imaging system for the digital museum. *International Journal on Digital Libraries*, 5(3):167–178, May 2005.
- [45] Marvin E. Klein, Bernard J. Aalderink, Roberto Padoan, Gerrit De Bruin, and Ted A. G. Steemers. Quantitative Hyperspectral Reflectance Imaging. *Sensors*, 8(9):5576–5618, September 2008.

- [46] Jon Yngve Hardeberg, Francis J. M. Schmitt, and Hans Brettel. Multi-spectral color image capture using a liquid crystal tunable filter. *Optical Engineering*, 41(10):2532–2548, October 2002.
- [47] Gregory H. Bearman and Sheila I. Spiro. Archaeological Applications of Advanced Imaging Techniques. *The Biblical Archaeologist*, 59(1):56, March 1996.
- [48] Haida Liang, David Saunders, and John Cupitt. A New Multispectral Imaging System for Examining Paintings. *Journal of Imaging Science and Technology*, 49(6):551–562, November 2005.
- [49] John Delaney. Multispectral Imaging of Paintings in the Infrared to Detect and Map Blue Pigments. *PNAS Sackler NAS colloquium*, pages 120–136, 2005.
- [50] P. Carcagnì, A. Della Patria, R. Fontana, M. Greco, M. Mastroianni, M. Materazzi, E. Pampaloni, and L. Pezzati. Multispectral imaging of paintings by optical scanning. *Optics and Lasers in Engineering*, 45(3):360–367, March 2007.
- [51] Claudio Bonifazzi, Pierluigi Carcagnì, Andrea Della Patria, Stefano Ferriani, Raffaella Fontana, Marinella Greco, Maria Mastroianni, Marzia Materazzi, Enrico Pampaloni, and Afra Romano. A scanning device for multispectral imaging of paintings. In *Spectral Imaging: Eighth International Symposium on Multispectral Color Science*, volume 6062, page 60620M. International Society for Optics and Photonics, January 2006.
- [52] Claudio Bonifazzi, Pierluigi Carcagnì, Andrea Della Patria, Raffaella Fontana, Marinella Greco, Maria Mastroianni, Marzia Materazzi, Enrico Pampaloni, Luca Pezzati, and Afra Romano. Multispectral imaging of paintings: Instrument and applications. *Proceedings of SPIE*, 6618(1):661816–661816–12, July 2007.
- [53] Pelagotti, L. Pezzati, A. Piva, A. Del Mastio, and Largo E. Fermi. Multi-spectral UV fluorescence analysis of painted surfaces. In *Proceedings of the 2006 14th European Signal Processing Conference*, pages 1–5, Pisa, 2006. IEEE.
- [54] William A. Christens-Barry, Kenneth Boydston, Fenella G. France, Keith T. Knox, Roger L. Easton Jr, and Michael B. Toth. Camera system for multispectral imaging of documents. In *Sensors, Cameras, and Systems for Industrial/Scientific Applications X*, volume 7249, page 724908. International Society for Optics and Photonics, January 2009.

-
- [55] R.L. Easton, K.T. Knox, and W.A. Christens-Barry. Multispectral imaging of the Archimedes palimpsest. In *Applied Imagery Pattern Recognition Workshop, 2003. Proceedings. 32nd*, pages 111–116, October 2003.
- [56] Emanuele Salerno, Anna Tonazzini, and Luigi Bedini. Digital image analysis to enhance underwritten text in the Archimedes palimpsest. *International Journal of Document Analysis and Recognition (IJДАР)*, 9(2-4):79–87, April 2007.
- [57] Lawrence Taplin and Roy Berns. Practical spectral capture systems for museum imaging. *AIC Colour 05*, May 2005.
- [58] Francisco H. Imai, Roy S. Berns, and Di-Y. Tzeng. A Comparative Analysis of Spectral Reflectance Estimated in Various Spaces Using a Trichromatic Camera System. *Journal of Imaging Science and Technology*, 44(4):280–287, July 2000.
- [59] Antonino Cosentino. Identification of pigments by multispectral imaging; a flowchart method. *Heritage Science*, 2(1):8, March 2014.
- [60] G. Antonioli, F. Fermi, C. Oleari, and R. Reverberi. Spectrophotometric Scanner for Imaging of Paintings and Other Works of Art. *Conference on Colour in Graphics, Imaging, and Vision*, 2004(1):219–224, January 2004.
- [61] Mauro Bacci, Andrea Casini, Costanza Cucci, Andrea Muzzi, and Simone Porcinai. A study on a set of drawings by Parmigianino: Integration of art-historical analysis with imaging spectroscopy. *Journal of Cultural Heritage*, 6(4):329–336, December 2005.
- [62] John K. Delaney, Jason G. Zeibel, Mathieu Thoury, Roy Littleton, Kathryn M. Morales, Michael Palmer, and E. René de la Rie. Visible and infrared reflectance imaging spectroscopy of paintings: Pigment mapping and improved infrared reflectography. In *O3A: Optics for Arts, Architecture, and Archaeology II*, volume 7391, page 739103. International Society for Optics and Photonics, July 2009.
- [63] John K. Delaney, Jason G. Zeibel, Mathieu Thoury, Roy Littleton, Michael Palmer, Kathryn M. Morales, E. René de la Rie, and Ann Hoenigswald. Visible and Infrared Imaging Spectroscopy of Picasso’s Harlequin Musician: Mapping and Identification of Artist Materials in Situ. *Applied Spectroscopy*, 64(6):584–594, June 2010.

- [64] Costanza Cucci, John K. Delaney, and Marcello Picollo. Reflectance Hyperspectral Imaging for Investigation of Works of Art: Old Master Paintings and Illuminated Manuscripts. *Accounts of Chemical Research*, September 2016.
- [65] John K. Delaney, Paola Ricciardi, Lisha Deming Glinsman, Michelle Facini, Mathieu Thoury, Michael Palmer, and E. René de la Rie. Use of imaging spectroscopy, fiber optic reflectance spectroscopy, and X-ray fluorescence to map and identify pigments in illuminated manuscripts. *Studies in Conservation*, 59(2):91–101, March 2014.
- [66] Hilda Deborah, Sony George, and Jon Yngve Hardeberg. Pigment Mapping of the Scream (1893) Based on Hyperspectral Imaging. In Abraham Elmoataz, Olivier Lezoray, Fathallah Nouboud, and Driss Mamass, editors, *Image and Signal Processing*, Lecture Notes in Computer Science, pages 247–256, Cherbourg, 2014. Springer International Publishing.
- [67] **Ruven Pillay**. Hyperspectral Imaging, Visualization and Database Integration. In *Proceedings of the ICOM-CC Documentation & Paintings Working Groups Joint Interim Meeting*, pages 15–16, The Hague, September 2013. ICOM-CC.
- [68] Ludo Snijders, Tim Zaman, and David Howell. Using hyperspectral imaging to reveal a hidden precolonial Mesoamerican codex. *Journal of Archaeological Science: Reports*, 9:143–149, October 2016.
- [69] Fabien Pottier. *A Study of the Codex Borbonicus Coloring Materials - Non-Invasive Spectroscopies Applied to Codicology*. PhD Thesis, Université de Cergy Pontoise (UCP), January 2017.
- [70] Haida Liang, Andrei Lucian, Rebecca Lange, Chi Shing Cheung, and Bomin Su. Remote spectral imaging with simultaneous extraction of 3D topography for historical wall paintings. *ISPRS Journal of Photogrammetry and Remote Sensing*, 95:13–22, September 2014.
- [71] John K. Delaney, Paola Ricciardi, Lisha Glinsman, Michael Palmer, and Julia Burke. Use of near infrared reflectance imaging spectroscopy to map wool and silk fibres in historic tapestries. *Analytical Methods*, 8(44):7886–7890, November 2016.
- [72] Sony George and Jon Yngve Hardeberg. Estimation and Correction of Geometric Distortion in Pushbroom Hyperspectral System for Imaging Art Paintings. *Electronic Imaging*, 2016(12):1–4, February 2016.

-
- [73] Roger N. Clark, Trude V. V. King, Matthew Klejwa, Gregg A. Swayze, and Norma Vergo. High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysical Research: Solid Earth*, 95(B8):12653–12680, 1990.
- [74] Kathryn A. Dooley, Suzanne Lomax, Jason G. Zeibel, Costanza Miliani, Paola Ricciardi, Ann Hoenigswald, Murray Loew, and John K. Delaney. Mapping of egg yolk and animal skin glue paint binders in Early Renaissance paintings using near infrared reflectance imaging spectroscopy. *Analyst*, 138(17):4838–4848, July 2013.
- [75] Linda Cséfalvayová, Matija Strlič, and Harri Karjalainen. Quantitative NIR Chemical Imaging in Heritage Science. *Analytical Chemistry*, 83(13):5101–5106, July 2011.
- [76] Taixia Wu, Guanghua Li, Zehua Yang, Hongming Zhang, Yong Lei, Nan Wang, and Lifu Zhang. Shortwave Infrared Imaging Spectroscopy for Analysis of Ancient Paintings. *Applied Spectroscopy*, November 2016.
- [77] Francesca Rosi, Costanza Miliani, René Braun, Roland Harig, Diego Sali, Brunetto G. Brunetti, and Antonio Sgamellotti. Noninvasive Analysis of Paintings by Mid-infrared Hyperspectral Imaging. *Angewandte Chemie International Edition*, 52(20):5258–5261, 2013.
- [78] Mathieu Thoury, John K. Delaney, E. René de la Rie, Michael Palmer, Kathryn Morales, and Jay Krueger. Near-Infrared Luminescence of Cadmium Pigments: In Situ Identification and Mapping in Paintings. *Applied Spectroscopy*, 65(8):939–951, 2011.
- [79] John K. Delaney, Kathryn A. Dooley, Roxanne Radpour, and Ioanna Kakoulli. Macroscale multimodal imaging reveals ancient painting production technology and the vogue in Greco-Roman Egypt. *Scientific Reports*, 7(1):15509, November 2017.
- [80] Aurélie Mounier, Gwénaëlle Le Bourdon, Christian Aupetit, Colette Belin, Laurent Servant, Sylvain Lazare, Yannick Lefrais, and Floréal Daniel. Hyperspectral imaging, spectrofluorimetry, FORS and XRF for the non-invasive study of medieval miniatures materials. *Heritage Science*, 2(1):24, October 2014.
- [81] D. A. Landgrebe and E. Malaret. Noise in Remote-Sensing Systems: The Effect on Classification Error. *IEEE Transactions on Geoscience and Remote Sensing*, GE-24(2):294–300, March 1986.

- [82] E. L. Dereniak and G. D. Boreman. *Infrared Detectors and Systems*. Wiley, April 1996.
- [83] A. Rogalski. HgCdTe infrared detector material: History, status and outlook. *Reports on Progress in Physics*, 68(10):2267, 2005.
- [84] Robert J Kohler and Hutson K Howell. Photographic image enhancement by superimposition of multiple images. *Photographic Science and Engineering*, 7(4):241–245, 1963.
- [85] B. Rasti, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson. Hyperspectral image denoising using 3D wavelets. In *2012 IEEE International Geoscience and Remote Sensing Symposium*, pages 1349–1352, July 2012.
- [86] Q. Yuan, L. Zhang, and H. Shen. Hyperspectral Image Denoising Employing a Spectral-Spatial Adaptive Total Variation Model. *IEEE Transactions on Geoscience and Remote Sensing*, 50(10):3660–3677, October 2012.
- [87] Alamin Mansouri, Ferdinand Deger, Marius Pedersen, Jon Y. Hardeberg, and Yvon Voisin. An adaptive spatial-spectral total variation approach for Poisson noise removal in hyperspectral images. *Signal, Image and Video Processing*, pages 1–8, August 2015.
- [88] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Pearson Education, November 2011.
- [89] M. Bacci, S. Baronti, A. Casini, F. Lotti, M. Picollo, and O. Casazza. Non-Destructive Spectroscopic Investigations on Paintings Using Optical Fibers. *MRS Online Proceedings Library Archive*, 267, 1992.
- [90] Tanvir H. Demetriades-Shah, Michael D. Steven, and Jeremy A. Clark. High resolution derivative spectra in remote sensing. *Remote Sensing of Environment*, 33(1):55–64, July 1990.
- [91] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz. Airbone Imaging Spectrometry The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44(2):145–163, May 1993.
- [92] Chein-I. Chang. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis. *IEEE Transactions on Information Theory*, 46(5):1927–1932, August 2000.

-
- [93] O Abilio De Carvalho and Paulo Roberto Meneses. Spectral correlation mapper (SCM): An improvement on the spectral angle mapper (SAM). In *Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18*, volume 9. JPL Publication Pasadena, CA, 23.
- [94] E. Angelopoulou, S. W. Lee, and R. Bajcsy. Spectral gradient: A material descriptor invariant to geometry and incident illumination. In *The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999*, volume 2, pages 861–867 vol.2, Kerkyra, Greece, September 1999.
- [95] James Sweet, James Granahan, and Mary Sharp. An objective standard for hyperspectral image quality. In *Proceedings of AVIRIS Workshop*, 2000.
- [96] Yingzi Du, Chein-I Chang, Hsuan Ren, Chein-Chi Chang, James O. Jensen, and Francis M. D’Amico. New hyperspectral discrimination measure for spectral characterization. *Optical Engineering*, 43(8):1777–1786, 2004.
- [97] John P. Kerekes, Adam P. Cisz, and Rulon E. Simmons. A comparative evaluation of spectral quality metrics for hyperspectral imagery. In *SPIE 5806, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, volume 5806, pages 469–480, 2005.
- [98] H. Deborah, N. Richard, and J. Y. Hardeberg. A Comprehensive Evaluation of Spectral Distance Functions and Metrics for Hyperspectral Image Processing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6):3224–3234, 2015.
- [99] Antonio Plaza, Gabriel Martín, Javier Plaza, Maciel Zortea, and Sergio Sánchez. Recent Developments in Endmember Extraction and Spectral Unmixing. In Saurabh Prasad, Lori M. Bruce, and Jocelyn Chanussot, editors, *Optical Remote Sensing*, number 3 in Augmented Vision and Reality, pages 235–267. Springer Berlin Heidelberg, Heidelberg, Germany, 2011.
- [100] R.J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam. Image change detection algorithms: A systematic survey. *IEEE Transactions on Image Processing*, 14(3):294–307, March 2005.
- [101] Freek van der Meer. The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation*, 8(1):3–17, January 2006.

- [102] S. A. Robila and A. Gershman. Spectral matching accuracy in processing hyperspectral data. In *International Symposium on Signals, Circuits and Systems, 2005. ISSCS 2005.*, volume 1, pages 163–166 Vol. 1, July 2005.
- [103] E. A. Early, P. Y. Barnes, B. C. Johnson, J. J. Butler, C. J. Bruegge, S. F. Biggar, P. R. Spyak, and M. M. Pavlov. Bidirectional Reflectance Round-Robin in Support of the Earth Observing System Program. *Journal of Atmospheric and Oceanic Technology*, 17(8):1077–1091, August 2000.
- [104] Stanford B. Hooker, Elaine R. Firestone, Scott McLean, Jennifer Sherman, Mark Small, Gordana Lazin, Giuseppe Zibordi, James W. Brown, and Charles R. McClain. The Seventh SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-7), March 1999. Technical report, NASA, February 2002.
- [105] András Jung, Christian Götze, and Cornelia Glässer. Overview of Experimental Setups in Spectroscopic Laboratory Measurements – the SpecTour Project. *Photogrammetrie - Fernerkundung - Geoinformation*, 2012(4):433–442, August 2012.
- [106] Anna Bentkowska-Kafel and Lindsay MacDonald, editors. *Digital Techniques for Documenting and Preserving Cultural Heritage*. Arc Humanities Press, Leeds, January 2018.
- [107] Cennino Cennini and Daniel Varney Thompson. *Il Libro Dell’Arte*. Yale University Press, 1936.
- [108] Tatiana Vitorino. Analysis of the Round Robin Test data acquired at IFAC-CNR. STSM Report, IFAC-CNR, Florence, Italy, September 2014.
- [109] Sony George, Mihaela Ciortan, and Jon Ynvge Hardeberg. Evaluation of Hyperspectral Imaging Systems for Cultural Heritage Applications Based on a Round Robin Test. In *Proceedings, AIC 2015 Tokyo*, Tokyo, 19.
- [110] Tatiana Vitorino, Andrea Casini, Costanza Cucci, Ana Gebejesje, Jouni Hiltunen, Markku Hauta-Kasari, Marcello Picollo, and Lorenzo Stefani. Accuracy in Colour Reproduction: Using a ColorChecker Chart to Assess the Usefulness and Comparability of Data Acquired with Two Hyper-Spectral Systems. In Alain Trémeau, Raimondo Schettini, and Shoji Tominaga, editors, *Computational Color Imaging*, number 9016 in Lecture Notes in Computer Science, pages 225–235. Springer International Publishing, March 2015.

-
- [111] Torbjorn Skauli. Quantifying coregistration errors in spectral imaging. In *Imaging Spectrometry XVI*, volume 8158, page 81580A. International Society for Optics and Photonics, September 2011.
- [112] J. Brauers, N. Schulte, and T. Aach. Multispectral Filter-Wheel Cameras: Geometric Distortion Model and Compensation Algorithms. *IEEE Transactions on Image Processing*, 17(12):2368–2380, 2008.
- [113] Naoto Yokoya, Norihide Miyamura, and Akira Iwasaki. Preprocessing of hyperspectral imagery with consideration of smile and keystone properties. In *Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques, and Applications III*, volume 7857, page 78570B. International Society for Optics and Photonics, November 2010.
- [114] Yukun Xing and Richard B. Gomez. Hyperspectral image analysis using ENVI (environment for visualizing images). In *Geo-Spatial Image and Data Exploitation II*, volume 4383, pages 79–86. International Society for Optics and Photonics, June 2001.
- [115] A.A. Green, M. Berman, P. Switzer, and M.D. Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26(1):65–74, January 1988.
- [116] J.W. Boardman. Geometric mixture analysis of imaging spectrometry data. In *Geoscience and Remote Sensing Symposium, 1994. IGARSS '94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation., International*, volume 4, pages 2369–2371 vol.4, August 1994.
- [117] John R. Schott, Kyungsuk Lee, Rolando Raqueno, Gary Hoffmann, and Glenn Healey. A subpixel Target Detection Technique based on the Invariance Approach. *AVIRIS Airborne Geoscience workshop Proceedings*, February 2003.
- [118] Di Bai, David W. Messinger, and David Howell. A hyperspectral imaging spectral unmixing and classification approach to pigment mapping in the Gough & Selden Maps. *Journal of the American Institute for Conservation*, 58(1-2):69–89, April 2019.
- [119] Hilda Deborah, Sony George, and Jon Yngve Hardeberg. Spectral-divergence based pigment discrimination and mapping: A case study on The Scream (1893) by Edvard Munch. *Journal of the American Institute for Conservation*, 58(1-2):90–107, April 2019.

- [120] Tania Kleynhans, David W. Messinger, and John K. Delaney. Towards automatic classification of diffuse reflectance image cubes from paintings collected with hyperspectral cameras. *Microchemical Journal*, 157:104934, September 2020.
- [121] Tania Kleynhans, Catherine M. Schmidt Patterson, Kathryn A. Dooley, David W. Messinger, and John K. Delaney. An alternative approach to mapping pigments in paintings with hyperspectral reflectance image cubes using artificial intelligence. *Heritage Science*, 8(1):84, August 2020.
- [122] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton. Advances in Spectral-Spatial Classification of Hyperspectral Images. *Proceedings of the IEEE*, 101(3):652–675, March 2013.
- [123] Adobe Systems Inc. TIFF 6.0 Specification, June 1992.
- [124] O. K. Al-Shaykh, I. Moccagatta, and Homer Chen. JPEG-2000: A new still image compression standard. In *Conference Record of Thirty-Second Asilomar Conference on Signals, Systems and Computers (Cat. No.98CH36284)*, volume 1, pages 99–103 vol.1, November 1998.
- [125] CIE. CIE223:2017: Multispectral Image Formats. Technical Report TC8-07, CIE (International Commission on Illumination), 2017.
- [126] S. J. S. Khalsa. Progress on Standards Development for Remote Sensing in the Geosciences [Technical Committees]. *IEEE Geoscience and Remote Sensing Magazine*, 8(3):95–99, September 2020.
- [127] D. C. Wells, E. W. Greisen, and R. H. Harten. FITS - a Flexible Image Transport System. *Astronomy and Astrophysics Supplement Series*, 44:363, June 1981.
- [128] Mike Folk, Gerd Heber, Quincey Koziol, Elena Pourmal, and Dana Robinson. An overview of the HDF5 technology suite and its applications. In *Proceedings of the EDBT/ICDT 2011 Workshop on Array Databases*, AD '11, pages 36–47, New York, NY, USA, March 2011. Association for Computing Machinery.
- [129] Jing Zhang, J.E. Fowler, N.H. Younan, and Guizhong Liu. Evaluation of JP3D for lossy and lossless compression of hyperspectral imagery. In *Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*, volume 4, pages IV-474–IV-477, July 2009.

-
- [130] J. E. Sanchez, E. Auge, J. Santalo, I. Blanes, J. Serra-Sagrasta, and A. Kiely. Review and Implementation of the Emerging CCSDS Recommended Standard for Multispectral and Hyperspectral Lossless Image Coding. In *2011 First International Conference on Data Compression, Communications and Processing*, pages 222–228, June 2011.
- [131] Rui Dusselaar and Manoranjan Paul. Hyperspectral image compression approaches: Opportunities, challenges, and future directions: Discussion. *JOSA A*, 34(12):2170–2180, December 2017.
- [132] Qian Du and J.E. Fowler. Hyperspectral Image Compression Using JPEG2000 and Principal Component Analysis. *IEEE Geoscience and Remote Sensing Letters*, 4(2):201–205, April 2007.
- [133] B. Penna, T. Tillo, E. Magli, and G. Olmo. Transform Coding Techniques for Lossy Hyperspectral Data Compression. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5):1408–1421, May 2007.
- [134] F. García-Vílchez, J. Muñoz-Marí, M. Zortea, I. Blanes, V. González-Ruiz, G. Camps-Valls, A. Plaza, and J. Serra-Sagrastà. On the Impact of Lossy Compression on Hyperspectral Image Classification and Unmixing. *IEEE Geoscience and Remote Sensing Letters*, 8(2):253–257, March 2011.
- [135] K. Lenhard, A. Baumgartner, and T. Schwarzmaier. Independent Laboratory Characterization of NEO HySpex Imaging Spectrometers VNIR-1600 and SWIR-320m-e. *IEEE Transactions on Geoscience and Remote Sensing*, 53(4):1828–1841, April 2015.
- [136] Jeff Secker, Karl Staenz, Robert P Gauthier, and Paul Budkewitsch. Vicarious calibration of airborne hyperspectral sensors in operational environments. *Remote Sensing of Environment*, 76(1):81–92, April 2001.
- [137] Magdeleine Dinguirard and Philip N. Slater. Calibration of Space-Multispectral Imaging Sensors: A Review. *Remote Sensing of Environment*, 68(3):194–205, June 1999.
- [138] J.T. Woodward, S.W. Brown, A.W. Smith, and K.R. Lykke. Hyperspectral imager characterization and calibration. In *Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*, volume 2, pages II–77–II–80, July 2009.
- [139] Gerrit Polder, Gerie van der Heijden, L. Keizer, and Ian Young. Calibration and characterisation of imaging spectrographs. *Journal of Near Infrared Spectroscopy*, 11(1):193, 2003.

- [140] Paul Geladi, Jim Burger, and Torbjörn Lestander. Hyperspectral imaging: Calibration problems and solutions. *Chemometrics and Intelligent Laboratory Systems*, 72(2):209–217, July 2004.
- [141] James Burger and Paul Geladi. Hyperspectral NIR image regression part I: Calibration and correction. *Journal of Chemometrics*, 19(5-7):355–363, May 2005.
- [142] James Burger and Paul Geladi. Hyperspectral NIR image regression part II: Dataset preprocessing diagnostics. *Journal of Chemometrics*, 20(3-4):106–119, March 2006.
- [143] Barbara Boldrini, Waltraud Kessler, Karsten Rebner, and Rudolf Kessler. Hyperspectral imaging: A review of best practice, performance and pitfalls for inline and online applications. *Journal of Near Infrared Spectroscopy*, 20(5):438, 2012.
- [144] Haris Khan, Sofiane Mihoubi, Benjamin Mathon, Jean-Baptiste Thomas, Jon Hardeberg, Haris Ahmad Khan, Sofiane Mihoubi, Benjamin Mathon, Jean-Baptiste Thomas, and Jon Yngve Hardeberg. HyTexiLa: High Resolution Visible and Near Infrared Hyperspectral Texture Images. *Sensors*, 18(7):2045, June 2018.
- [145] Žiga Špiclin, Jaka Katrašnik, Miran Bürmen, Franjo Pernuš, and Boštjan Likar. Geometric calibration of a hyperspectral imaging system. *Applied Optics*, 49(15):2813, May 2010.
- [146] Google Scholar. Citations for "Web-based visualization of very large scientific astronomy imagery". <https://scholar.google.com/scholar?oi=bibs&hl=en&cites=6085262726369496327>, 2020. [Online; accessed 16-December-2020].
- [147] David Beaudet, Alan Newman, and Kenneth Fleisher. IIP-Image Case Study: National Gallery of Art, Washington. <https://iipimage.sourceforge.io/2014/05/iipimage-case-study-national-gallery-of-art-washington/>, May 2014.
- [148] Marco Cardinali. Digital Tools and Technical Views: The Intersection of Digital Art History and Technical Art History in a Digital Archive on the Painting Technique of Caravaggio and His Followers. *Visual Resources*, 35(1-2):52–73, April 2019.

-
- [149] Wietske Donkersloot. The Rembrandt Database: An Inter-Institutional Research Resource for Art-Historical, Technical and Conservation Documentation. In *Proceedings of XVIth Triennial ICOM-CC Conference*, Lisbon, September 2011.
- [150] Catherine Leung and Andor Salga. Enabling WebGL. In *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, pages 1369–1370, New York, NY, USA, April 2010. Association for Computing Machinery.
- [151] Marco Potenziani, Marco Callieri, Matteo Dellepiane, Massimiliano Corsini, Federico Ponchio, and Roberto Scopigno. 3DHOP: 3D Heritage Online Presenter. *Computers & Graphics*, 52:129–141, November 2015.
- [152] Christian Lahanier, Genevieve Aitken, **Ruven Pillay**, Angelo Beraldin, François Blais, Louis Borgeat, Bernd Breuckmann, Philippe Colantoni, Luc Cournoyer, Christophe de Deyne, Michel Picard, Marc Rioux, and John Taylor. The 2D multi-spectral digitisation and the 3D modelling of easel paintings. In *ICOM Committee for Conservation, ICOM-CC: Proceedings of the 14th Triennial Meeting*, pages 322–328. ICOM-CC, James & James, September 2005.
- [153] Google Scholar. Citations for "Studying that Smile: A Tutorial on Multi-spectral Imaging of Paintings". <https://scholar.google.com/scholar?oi=bibs&hl=en&cites=17317665790569856648>, 2020. [Online; accessed 16-December-2020].
- [154] Y. Monno, S. Kikuchi, M. Tanaka, and M. Okutomi. A Practical One-Shot Multispectral Imaging System Using a Single Image Sensor. *IEEE Transactions on Image Processing*, 24(10):3048–3059, October 2015.
- [155] Jennifer D. T. Kruschwitz and Roy S. Berns. SiO₂ Film Thickness Determination on Silicon Using Dual-Angle Imaging. In *Optical Interference Coatings 2016 (2016), Paper ThB.7*, page ThB.7. Optical Society of America, June 2016.
- [156] Stuart Snyderman, Robert Sanderson, and Tom Cramer. The International Image Interoperability Framework (IIIF): A community & technology approach for web-based images. *Archiving Conference*, 2015(1):16–21, May 2015.
- [157] **Ruven Pillay**. Examples of the Use of IIPImage Around the World. <https://iipimage.sourceforge.io/links/>. [Online; accessed 16-December-2020].

- [158] Zsolt L. Husz, Thomas P. Perry, Bill Hill, and Richard A. Baldock. Woolz IIP: A Tiled On-the-Fly Sectioning Server for 3D Volumetric Atlases. In George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Yoshinori Kuno, Junxian Wang, Jun-Xuan Wang, Junxian Wang, Renato Pajarola, Peter Lindstrom, André Hinkenjann, Miguel L. Encarnação, Cláudio T. Silva, and Daniel Coming, editors, *Advances in Visual Computing*, number 5875 in Lecture Notes in Computer Science, pages 924–933. Springer Berlin Heidelberg, January 2009.
- [159] Frank G. A. Faas, M. Cristina Avramut, Bernard M. van den Berg, A. Mieke Mommaas, Abraham J. Koster, and Raimond B. G. Ravelli. Virtual nanoscopy: Generation of ultra-large high resolution electron microscopy maps. *The Journal of Cell Biology*, 198(3):457–469, August 2012.

Part II

Original Papers

Quality Evaluation in Spectral Imaging - Quality Factors and Metrics

Paper 1

This chapter is a reprint of the publication:

Shrestha, R., **Pillay, R.**, George, S. & Hardeberg, J. Y. “Quality Evaluation in Spectral Imaging - Quality Factors and Metrics”. *JAIC - Journal of the International Colour Association*, 12, 22–35 (2014).

Quality Evaluation in Spectral Imaging – Quality Factors and Metrics

Raju Shrestha, Ruven Pillay, Sony George and Jon Y Hardeberg

The Norwegian Colour and Visual Computing Laboratory, Gjøvik University College, Norway
Email: raju.shrestha@hig.no

Spectral imaging has many advantages over conventional three channel colour imaging, and has numerous applications in many domains. Despite many benefits, applications, and different techniques being proposed, little attention has been given to the evaluation of the quality of spectral images and of spectral imaging systems. There has been some research in the area of spectral image quality, mostly targeted at specific application domains. This paper seeks to provide a comprehensive review on existing research in the area of spectral image quality metrics. We classify existing spectral image quality metrics into categories based on how they were developed, their main features, and their intended applications. Spectral quality metrics, in general, aim to measure the quality of spectral images without considering specifically the imaging systems used to acquire the images. Having many different types of spectral imaging systems that could be used to acquire spectral images in an application, it is important to evaluate the performance/quality of these spectral imaging systems too. However, to our knowledge, not much attention has been given in this direction previously. As a first step towards this, we aim to identify different factors that influence the quality of the spectral imaging systems. In almost every stage of a spectral imaging workflow, there may be one or more factors that influence the quality of the final spectral image, and hence the imaging system used for acquiring the image. Identification of these factors, we believe, will be essential in developing a framework, for evaluating the quality of spectral imaging systems.

Received 5 January 2014; revised 26 May 2014; accepted 26 May 2014

Published online: 15 July 2014

Introduction

Spectral imaging has received much attention in recent years for its advantages over conventional three channel based colour imaging (usually RGB), and because of applications in a number of domains such as remote sensing, medical imaging, cultural heritage, biometrics and many more. A spectral imaging system captures image data at specific wavelength intervals (narrow or somewhat wider) across the electromagnetic spectrum. Based on their number of spectral bands, spectral imaging systems can be divided into two major types: multispectral and hyperspectral. There is no fine line separating the two; however, spectral imaging systems with more than 20 bands are generally considered as hyperspectral, and less than 20 as multispectral. We use the term spectral throughout this paper, to refer to both of them in a general sense. Hyperspectral imaging deals with imaging narrow spectral bands over a contiguous spectral range, and produces the spectra of all pixels in the scene. Multispectral imaging systems typically acquire images in a wider and limited number of spectral bands. They do not produce the spectrum of an object directly; but rather use spectral estimation algorithms to obtain spectral reflectances from the sensor responses. Hyperspectral imaging systems produce high measurement accuracy; however, the acquisition time, complexity and cost of these systems are generally high compared to multispectral systems.

A number of different acquisition techniques exist for both multispectral and hyperspectral imaging. Multispectral systems, for example, can be a multi-sensor based [1, 2], filter based using a filter wheel [3, 4] or tunable filter [5, 6], multispectral filter array (MSFA) based [7–9], or light emitting diode (LED) based [10–13]. A filter-less and demosaicking-less colour sensitive device which uses transverse field detectors or tunable sensitivity sensors has also been proposed [14]. Many of these approaches require multiple shots in order to acquire a multispectral image, whereas a filter based one-shot solution which uses a stereo camera has been recently proposed [15–17]. Hyperspectral imaging systems can also be based on tunable filters and multiple shots, but are more often based on gratings, and used using a pushbroom technique [18].

Having many different types of spectral image acquisition systems, an important question that arises is how do we evaluate the quality of the spectral image data captured? In other words, there should be a way to evaluate the quality of a spectral image, and possibly of the spectral imaging system used to acquire the image. Much research has been carried out on image quality for classic three channel colour images and recently, in particular, based on perceptual quality [19, 20]. And there has been research intended for spectral image quality in a number of specific application domains [21–26]. However, to our knowledge, little work has been done on evaluating the quality of spectral imaging systems. This paper provides a comprehensive review of the research carried out so far in the field of spectral image quality metrics, and also identifies important factors involved in describing the quality of spectral imaging systems and spectral image data acquired with them. We believe that this work can further help in the development of general and/or application specific spectral image quality frameworks.

The next section reviews research that has been carried out on spectral image quality metrics. We then describe and classify the metrics into different categories before identifying and discussing the main quality factors and attributes that could form a basis of a global framework for spectral imaging quality.

Spectral image quality metrics

Many studies have been carried out on the quality evaluation of colour images. However, very little has been done on the evaluation of quality of spectral images. Before discussing the quality evaluation of spectral images, it would be useful to first discuss the notion of *quality* in the context of colour imaging and spectral imaging. There is no single universally accepted definition of colour image quality (CIQ), and this is even more true for spectral image quality. A number of tentative definitions can be found in the literature, and most of the definitions depend on particular applications. This is somewhat inevitable as quality always implies some application, nevertheless we think a more general approach will be beneficial.

One such definition adopted by many researchers is that by ISO [27], which defines image quality as the impression of the overall merit or excellence of an image, as perceived by an observer neither associated with the act of photography, nor closely involved with the subject matter depicted. One recent and extensive work that has been done on image quality by Pedersen [20] adopted this definition of image quality. Most of the previous work on image quality considered the visual perception of images (perceptual quality), either on a display or printed images. Unlike colour image quality, the definition of spectral image quality should not be limited to perceptual quality only. Since spectral imaging is used in wider application domains, spectral image quality may be defined differently based on its application domain. For instance spectral imaging can be used for producing

more accurate colour reproduction. In this case, a spectral image and the imaging system, which produces the most accurate colour reproduction, may be considered as having the best quality. In the case of a spectral imaging system which is used to detect and classify the material composition of an object, the system which accurately detects and classifies can be considered as having the best quality. Thus, the notion of the quality of a spectral image should vary depending on the application. Several studies have been made on spectral image quality based on the application, purpose, and type of spectral imaging systems.

There are fundamentally three different types of colour image quality metrics: no-reference, reduced-reference, and full-reference [20]. In the no-reference metric type, only the reproduction is available, and the calculation of quality is based only on the reproduction without use of the reference (i.e. the original). In the reduced-reference metric type, some information of the reproduction and the original is used in the calculation of quality. Both the reference and the reproduction are available in the full-reference metric type. These could be true in the case of spectral image quality metrics also. Unlike for colour images, subjective quality assessments or vision based models are not sufficient to measure the quality of spectral imaging. Moreover, an end user may not be a human as in the case of colour imaging. Quality can be assessed either in the spatial or spectral domain and it is highly driven by the application. A spectral image quality metric, thus, could be calculated pixel-wise in the whole image globally, or it could be calculated based on spatial pixel values in local segments in the image. Based on whether the purpose of a metric is to evaluate the quality of a spectral or perceptual response, or how good a certain task can be carried out, whether it is a full-reference type, and if the calculation is based on the global or local spatial context in a test image, we propose to classify spectral image quality metrics into five categories: global, full reference spectral quality (GFSQ); global, full reference, perceptual quality (GFPQ); spatial, full reference spectral quality (SFSQ); spatial, full reference spectral quality (SFPQ); and task based quality (TBQ) metrics. Before going into the details of the classification, we first define terms and conventions to be used in describing different spectral image quality metrics, in a consistent way. We use I to denote an image. A spectral image of dimension $m \times n \times l$ can be defined as $I(x, y, \lambda_i)$, where $x = 1, \dots, m$, and $y = 1, \dots, n$. λ_i ($i = 1, \dots, l$) denotes a spectral band or wavelength in a l -band spectral image. $I(x, y)$, thus, corresponds to the spectral reflectance at pixel (x, y) in the image. We denote the original reference image as I_r and the test image acquired by an imaging system as I_t .

We now describe the spectral image quality metrics that fall into the five categories below.

1. **Global, Full-reference, Spectral Quality (GFSQ) metrics:** Spectral image quality metrics, which are based on the calculation of the spectral response for every pixel of the test and reference images in a global context, fall into this category. In these types of metrics, metric values are calculated for every pixel of an image globally and the mean value (and possibly additional statistical information for instance, minimum and maximum values, standard deviation etc.) is computed by averaging over the metric values for all the pixels in the image.

A number of quality metrics exist to evaluate the quality of spectral images based on spectral responses. One of the most widely used of these is the root mean square (RMS) error, which provides a statistical estimation of the difference between the spectral responses of test and reference images. RMS calculates the cumulative squared error between the original image and the test image. RMS has been widely used due to its easy calculation and analytical tractability. The mean RMS for the test spectral image I_t with respect to the given reference image I_r is given by:

$$RMS = \frac{1}{m \times n} \sum_{x=1}^m \sum_{y=1}^n RMS(x, y), \quad (1)$$

where $RMS(x, y)$ is the RMS error at pixel (x, y) , and is calculated as:

$$RMS(x, y) = \sqrt{\frac{1}{l} \sum_{i=1}^l [I_t(x, y, \lambda_i) - I_r(x, y, \lambda_i)]^2}. \quad (2)$$

Another metric, the peak signal-to-noise ratio (PSNR) is also widely used, and it can be considered as a GSFQ as it is calculated from the RMS:

$$PSNR = 20 \log_{10} \left(\frac{1}{RMS} \right). \quad (3)$$

An alternative to the RMS is the goodness of fit coefficient (GFC), proposed by Romero et al. [28]. Unlike RMS, GFC is insensitive to the shift in magnitude, and its value is normalised to the range 0 to 1, with 1 indicating the perfect estimation and 0 (zero) indicating the worst estimation. The GFC value at a pixel (x, y) , $GFC(x, y)$, is calculated as:

$$GFC(x, y) = \frac{\sum_{i=1}^l I_t(x, y, \lambda_i) I_r(x, y, \lambda_i)}{\sqrt{\sum_{i=1}^l I_t(x, y, \lambda_i)^2} \sqrt{\sum_{i=1}^l I_r(x, y, \lambda_i)^2}}. \quad (4)$$

Spectral Angle Map (SAM) [29] is another widely used metric, which is usually used for spectral segmentation, but which provides a measure of the difference in terms of spectral angle (α) between two spectra. SAM is nothing but an inverse cosine of the GFC metric, and hence SAM at a pixel (x, y) is calculated as:

$$\alpha(x, y) = \cos^{-1} [GFC(x, y)]. \quad (5)$$

Spectral Information Divergence (SID) is another metric used to compare spectral image data [30]. SID views each pixel spectrum as a random variable, and then measures the discrepancy of probabilistic behaviors between two spectra, thereby determining similarity and variability more effectively than SAM. SID at a pixel (x, y) is calculated using the equation:

$$SID(x, y) = \sum_{i=1}^l \left(\frac{I_t(x, y, \lambda_i)}{\sum_{j=1}^l I_t(x, y, \lambda_j)} - \frac{I_r(x, y, \lambda_i)}{\sum_{j=1}^l I_r(x, y, \lambda_j)} \right) \left(\log \frac{I_t(x, y, \lambda_i)}{\sum_{j=1}^l I_t(x, y, \lambda_j)} - \log \frac{I_r(x, y, \lambda_i)}{\sum_{j=1}^l I_r(x, y, \lambda_j)} \right). \quad (6)$$

Yet another metric that falls into the category GFSQ is the spectral similarity value (SSV), proposed by Sweet et al. [22]. SSV combines magnitude (m) and shape (s) differences between two spectral vectors, giving each equal weighting. SSV at a pixel (x, y) is computed as:

$$SSV(x, y) = \sqrt{m(x, y)^2 + s(x, y)^2}, \quad (7)$$

where $m(x, y)$ is computed as the root mean square value, $RMS(x, y)$:

$$m(x, y) = RMS(x, y) = \sqrt{\frac{1}{l} \sum_{i=1}^l [I_t(x, y, \lambda_i) - I_r(x, y, \lambda_i)]^2} \quad (8)$$

and

$$s(x, y)^2 = 1 - \left(\frac{\frac{1}{l} \sum_{i=1}^l [I_t(x, y, \lambda_i) - \mu[I_t(x, y)]] [I_r(x, y, \lambda_i) - \mu[I_r(x, y)]]}{\sigma[I_t(x, y)] \sigma[I_r(x, y)]} \right)^2. \quad (9)$$

Here $\mu[I_t(x, y)]$ and $\mu[I_r(x, y)]$ are means; and $\sigma[I_t(x, y)]$ and $\sigma[I_r(x, y)]$ are the standard deviation of two spectra at a pixel (x, y) in the test and the reference images, computed across the l wavelengths. The SSV approach is appropriate for use with hyperspectral images.

GFSQ metrics are computed per pixel in an image, and therefore, do not take into account variability in specific regions or spatial information from the whole area of the image. Moreover, these metrics are useful only if reference spectral data/image is available, which would not be the case in many situations.

2. **Global, Full-reference, Perceptual Quality (GFPQ) metrics:** These metrics are similar to the GFSQ metrics, except that in this case the metric calculation is based on visual perception (most commonly, the colour), instead of spectral responses. This type of metric is, therefore, useful in applications where we are interested in how a perceptually accurate image can be reproduced from the acquired spectral images. As in GFSQ, the GFPQ metric values are calculated for each pixels of the image, and then the mean value is calculated by averaging over all the values. Once a spectral image is transformed to a 3-band colour image for visual perception, depending on the need of an application, any colour image quality metric can then be used to evaluate its quality. As an illustration we will discuss a colour difference metric.

Colour difference metrics are used to measure differences in colour between two colour patches. One of the most commonly used colour difference formula is CIE ΔE^*_{ab} [31] which is based on a perceptually uniform colour space, namely the CIELAB colour space. The colour difference is computed as the Euclidean distance between the two colours in the CIELAB space. Extensions of ΔE^*_{ab} have been proposed when it became apparent that it had problems, especially in the blue region with the CIE first proposing ΔE^*_{94} [32] and later ΔE^*_{00} [33]. Since these are increasingly complex compared to ΔE^*_{ab} , ΔE^*_{ab} is still widely used. The mean ΔE^*_{ab} between the test image (I_t) and the original reference image (I_r) is computed by averaging the colour differences in each pixel in the two images. ΔE^*_{ab} at a pixel (x, y) is calculated using the formula:

$$\Delta E^*_{ab}(x, y) = \sqrt{[\Delta L^*(x, y)]^2 + [\Delta a^*(x, y)]^2 + [\Delta b^*(x, y)]^2}, \quad (10)$$

where

$$\Delta L^*(x, y) = L^*_t(x, y) - L^*_r(x, y),$$

$$\Delta a^*(x, y) = a^*_t(x, y) - a^*_r(x, y), \text{ and}$$

$$\Delta b^*(x, y) = b^*_t(x, y) - b^*_r(x, y)$$

are differences in luminance (L^*) and chrominance (a^* and b^*) channels in the CIELAB space, at pixel (x, y) in the test and the reference images.

Metamerism index based metrics which compare the extent to which two spectra have a different colour between a reference condition and a test condition under different illuminants and observers have also been proposed [34, 35]. Viggiano's metamerism index [35], at pixel (x, y) in an image, $M_v(x, y)$ is computed using the equation:

$$M_v(x, y) = \sum_{i=1}^I w(x, y, \lambda_i) \|\Delta\beta(x, y, \lambda_i)\|, \quad (11)$$

where $\Delta\beta(x, y, \lambda_i) = I_t(x, y, \lambda_i) - I_r(x, y, \lambda_i)$, and $w(x, y, \lambda_i)$ are weights computed as follows:

$$w(x, y, \lambda) = \sqrt{\left(\frac{\Delta L^*(x, y)}{\Delta\beta(x, y, \lambda_i)}\right)^2 + \left(\frac{\Delta a^*(x, y)}{\Delta\beta(x, y, \lambda_i)}\right)^2 + \left(\frac{\Delta b^*(x, y)}{\Delta\beta(x, y, \lambda_i)}\right)^2}. \quad (12)$$

Perception based quality metrics including GFPQ work only in the visible part of the spectrum, and thus ignore important information in invisible bands such as the infrared and ultraviolet. This limits the use of such metrics in object detection and classification. Moreover, one single image quality metric is inadequate to indicate the quality of an image [36]. None of the perception based spectral image quality metrics takes into account this fact, and therefore cannot be considered as complete spectral image quality metrics.

3. **Spatial, Full-reference, Spectral Quality (SFSQ) metrics:** Some full-reference metrics aim to calculate spectral quality by taking into account the spatial distribution in images. These metrics can be categorised as SFSQ metrics.

One such metric is the $Q2^n$ index, proposed by Garzelli and Nencini [37]. The $Q2^n$ index extended the universal quality index (UQI) proposed for monochrome images [38], as a generalisation to multispectral and hyperspectral images, through a hypercomplex correlation coefficient (CC) between the reference (I_r) and the test images (I_t). The index jointly measures both spectral and spatial distortions. The $Q2^n$ index is derived from the theory of hypercomplex numbers, particularly of 2^n -ons (two-to-the-any-ons) [39], and made up of different factors to take into account for correlation, mean of each spectral band, intra-band local variance, and the spectral angle. Two hypercomplex image maps corresponding to the test and the reference images, $I_{t,h}$ and $I_{r,h}$ are obtained from a 2^n -on hypercomplex number at each pixel from the 2^n spectral bands. If the number of bands is not a power of two, the image bands are appropriately zero-padded, to analyse the overall data with 2^n spectral bands. The null bands do not influence the image quality measurement. A 2^n -on hypercomplex number for an image, I at a pixel (x, y) is represented as:

$$I_h(x, y) = I(x, y, \lambda_1) + \sum_{i=2}^n I(x, y, \lambda_i) j_{2^i}, \quad (13)$$

where j_2, j_3, \dots, j_{2^n} are hypercomplex unit vectors. Analogously to complex number, the conjugate I_h^* is given by:

$$I_h^*(x, y) = I(x, y, \lambda_1) - \sum_{i=2}^n I(x, y, \lambda_i) j_{2^i}, \quad (14)$$

and the modulus by

$$|I_h(x, y)| = \sqrt{\sum_{i=1}^n I(x, y, \lambda_i)^2} \quad (15)$$

The $Q2^n$ computes the correlation, the mean of each spectral band, and the intra-band local variance, at each pixel (x, y) using a sliding window of size $N \times N$ in the hypercomplex image map. Let \mathbf{t} and \mathbf{r} be the pixel arrays within the sliding window in the test and the reference image maps respectively. The $Q2^n$ at pixel (x, y) is then computed using the equation:

$$Q2^n(x, y) = \frac{\sigma_{\mathbf{tr}}}{\sigma_{\mathbf{t}}\sigma_{\mathbf{r}}} \cdot \frac{2\overline{\mathbf{tr}}}{\overline{\mathbf{t}}^2 + \overline{\mathbf{r}}^2} \cdot \frac{2\sigma_{\mathbf{t}}\sigma_{\mathbf{r}}}{\sigma_{\mathbf{t}}^2 + \sigma_{\mathbf{r}}^2}, \quad (16)$$

where $\overline{\mathbf{t}}$ and $\overline{\mathbf{r}}$ are means, and $\sigma_{\mathbf{t}}$ and $\sigma_{\mathbf{r}}$ are the standard deviations of \mathbf{t} and \mathbf{r} respectively. $\sigma_{\mathbf{tr}}$ is the hypercomplex covariance between \mathbf{t} and \mathbf{r} . These terms are defined as:

$$\begin{aligned} \overline{\mathbf{t}} &= E[\mathbf{t}], \\ \overline{\mathbf{r}} &= E[\mathbf{r}], \\ \sigma_{\mathbf{t}} &= E[|\mathbf{t}|^2] - |\overline{\mathbf{t}}|^2, \\ \sigma_{\mathbf{r}} &= E[|\mathbf{r}|^2] - |\overline{\mathbf{r}}|^2, \\ \sigma_{\mathbf{tr}} &= E[(\mathbf{t} - \overline{\mathbf{t}})(\mathbf{r} - \overline{\mathbf{r}})] = E[\mathbf{tr}^*] - \overline{\mathbf{t}}\overline{\mathbf{r}}^*. \end{aligned}$$

Among the three terms in Equation (16), the first term measures the hypercomplex CC, the second term measures the mean of the spectral band, and the third term computes the intra-band local variance. The mean $Q2^n$ is then obtained by averaging the magnitudes of all $Q2^n$'s over the whole image:

$$Q2^n = E[|Q2^n(x, y)|]. \quad (17)$$

The $Q2^n$ index assumes real values in the interval $[0, 1]$, with 1 being the best value, which can be achieved if and only if the test image is identical to the reference image. This metric is useful to measure the fidelity of a spectral image with respect to a known reference in terms of both spatial and spectral distortions.

4. **Spatial, Full-reference, Perceptual Quality (SFPQ) metrics:** Full-reference metrics which compute the quality of spectral images based on perceptual quality, taking into account local spatial information, can be classified into this category. The SFPQ metrics, therefore, are mainly based on spectral images in the visual range.

LeMoan and Urban [40] recently proposed an evaluation technique of the perceptual quality of spectral images, which is based on pooling (averaging) the image quality indices proposed by Lissner et al. [41], computed under a set of different illuminants, with the images rendered in the perceptually uniform LAB2000HL colour space [42], whose perceptual uniformity is based on the CIE ΔE^*_{uo} colour difference formula [33]. Based on the assumption that given the difference of two images under a certain illuminant, the error added by considering other illuminants can be

summarised solely in terms of chroma and hue difference, they proposed an approximation of the spectral image difference (SpID) between the two images as the average of colour image differences (CID) under different viewing conditions (VC). SpID at a pixel (x, y) , $SpID(x, y)$ is calculated using the equation:

$$SpID(x, y) = \frac{1}{N_{VC}} \sum_{i=1}^{N_{VC}} CID_{VC_i}(x, y), \quad (18)$$

where N_{VC} is the number of viewing conditions considered. CID under a visual condition,

$CID_{vc}(x, y)$ is computed by transforming spectral image to a CIE XYZ image, and then incorporating the chromatic adaptation transform (CAT) employed by CIECAM02. It is calculated using the equation:

$$CID_{VC}(x, y) = 1 - l_L(x, y)l_C(x, y)l_H(x, y)c_L(x, y)s_L(x, y), \quad (19)$$

where l_L , l_C , l_H , c_L , and s_L are image difference features (IDFs): lightness-difference, chroma difference, hue difference, lightness-contrast and lightness-structure respectively. These features are derived from the LAB2000HL images using the structural similarity index (SSIM) [19]. These terms at a pixel (x, y) are computed within a sliding window in the two images. Let \mathbf{t} and \mathbf{r} are the pixel arrays within this window in the test and the reference images. Among the five IDFs, lightness, chroma and hue differences are calculated as follows:

$$l_L(x, y) = \frac{1}{c_1 \cdot \overline{\Delta L(\mathbf{t}, \mathbf{r})}^2 + 1} \quad (20)$$

$$l_C(x, y) = \frac{1}{c_4 \cdot \overline{\Delta C(\mathbf{t}, \mathbf{r})}^2 + 1} \quad (21)$$

$$l_H(x, y) = \frac{1}{c_5 \cdot \overline{\Delta H(\mathbf{t}, \mathbf{r})}^2 + 1} \quad (22)$$

where $\overline{f(\mathbf{t}, \mathbf{r})}$ denotes a Gaussian-weighted mean of $f(x, y)$, computed for the pixel (x, y) , using all the pixel pairs in the two images, within the sliding window. ΔC is the chroma difference, the chroma being defined as $C = \sqrt{a^2 + b^2}$. Hue difference, ΔH is computed using the equation:

$$\Delta H = \sqrt{(a_t - a_r)^2 + (b_t - b_r)^2 + \Delta C^2}. \quad (23)$$

Lightness-contrast and lightness-structure are computed as follows:

$$c_L(x, y) = \frac{2\sigma_t\sigma_r + c_2}{\sigma_t^2 + \sigma_r^2 + c_2} \quad (24)$$

$$s_L(x, y) = \frac{\sigma_{tr} + c_3}{\sigma_t\sigma_r + c_3} \quad (25)$$

where σ_t and σ_r are the standard deviations of the lightness components in the sliding windows. σ_{tr} corresponds to the cosine of the angle between $\mathbf{t} - \bar{\mathbf{t}}$ and $\mathbf{r} - \bar{\mathbf{r}}$ in the lightness component [19]. c_1, \dots, c_5 are parameters that are adjusted for the colour space used, and large colour differences.

Like GFPQ, being perceptual quality based metrics, SFPQ metrics also do not take into account information in the invisible bands.

5. **Task based (Functional) Quality (TBQ) metrics:** There are some spectral image quality metrics which are aimed at an evaluation based on their performance for a certain task or function. These metrics can be categorised as TBQ metrics. One of the most common task is to detect targets/objects in a scene. Several spectral image quality metrics have been proposed in order to evaluate the performance for object detection and/or classification. In most cases, spectral data which cover a wider spectrum including the ultraviolet and/or infrared ranges are used as this allows detection even of targets invisible to the human eye.

Kerekes and Hsu [23] proposed a model-based spectral quality rating scale (SQRS) for target detection in VNIR (Visible Near InfraRed)/SWIR (ShortWave InfraRed) hyperspectral images. The higher the SQRS value, the better the spectral quality of the image. In their latest version of the work, SQRS is computed using the empirically-derived equation:

$$SQRS_{SCR-detection} = 10.6 - 1.6\log_{10}(t) + 3.3\log_{10}(GSD) + 1.6\log_{10}(SCR) \quad (26)$$

where GSD is the ground sample distance in *cm*, SCR is the signal-to-clutter ratio defined for a target and background having spectral mean vectors μ_t and μ_b , and the background having a spectral covariance matrix, Σ_b . t is a threshold on the normalised match filter output test statistic θ that leads to a specified false alarm rate on the image background. SCR , and $\theta(p)$ (θ for a pixel p) are computed using the equations:

$$SCR = \sqrt{(\mu_t - \mu_b)^t \Sigma_b^{-1} (\mu_t - \mu_b)} \quad (27)$$

$$\theta(p) = \frac{(\mu_t - \mu_b)^t \Sigma_b^{-1} (p - \mu_b)}{(\mu_t - \mu_b)^t \Sigma_b^{-1} (\mu_t - \mu_b)} \quad (28)$$

Shen [43] analysed a large number of images with varying spectral image parameters and proposed a target detection probability measure (P_d) based on a regression between some metrics and image parameters:

$$P_d = 6.25 - 0.81 \log_{10}(GSD) + 0.12 \log_{10}(SNR) - 0.20 \log_{10}(\Delta\lambda) - 2.43 \log_{10}(\sigma_{scene}) \quad (29)$$

where GSD is the ground sample distance in cm, SNR is the signal-to-noise ratio, $\Delta\lambda$ is the average spectral resolution of the channels in nm, and σ_{scene} is the average standard deviation in HYDICE scaled radiance units (1 HYDICE = $4/3 \mu\text{W}/\text{cm}^2\text{-sr-}\mu\text{m}$), of the pixels in the scene across all spectral bands.

Martin *et al.* [21] defined spectral quality as the extent to which an image or data set precisely replicates the scene represented by the image or data set. They proposed an approach to subjectively determine the utility through analyst assessments, calculate the quality of an image, and then relate these two metrics to obtain an objective quality metric. The postulated probability of correct material identification, P_{CI} is defined as a function of a number of parameters including the accuracy of signature definition, the sensor performance (spatial, spectral, and radiometric), the analysis of system performance, the sample abundance, and a decision criterion.

Simmons *et al.* [44] tried to combine spectral and spatial information with the aim of a general quality metric based on semantic transformations of the spatial and spectral quality [44]. It calculates spatial and spectral confidences, and a single total confidence value is obtained by combining the two confidences:

$$C_{Total} = 1 - (1 - C_{Spatial}) \cdot (1 - C_{Spectral}), \quad (30)$$

where $C_{Spectral}$ is the spectral confidence, which is obtained through an assessment of the separability of target and background spectral distributions or from results of hyperspectral image analysis techniques. $C_{Spatial}$ is the spatial confidence, which is largely driven by the size of the target relative to the image resolution, and is computed using the equation:

$$C_{Spatial} = \frac{(N / N_{50})^E}{1 + (N / N_{50})^E}, \quad (31)$$

where $E = 2.7 + 0.7(N / N_{50})$, N is the number of resolutions cycles per minimum dimension of the target, and N_{50} is the cycle criteria for 50 percent success and has values of 1.0 ± 0.25 , 4.1 ± 0.35 and 6.4 ± 1.5 for detection, recognition and identification respectively.

Most of the TBQ metrics are based on empirical modeling from a limited set of data, and therefore may not work well in a general sense. Since they use some of the information available from the original scene, they are of reduced-reference type. These metrics, in general, do not take into account spectral and colour accuracies.

Purpose	Full reference	
	Global	Spatial
Spectral	RMS, PSNR, GFC, SAM, SID, SSV	$Q2''$
Perceptual/Colour	ΔE^*_{ab} , M_v , other CIQ metrics	SpID

Table 1: Summary table of the first four categories of the spectral image quality metrics.

Among the five categories of spectral image quality metrics just discussed above, we can summarise the first four as shown in Table 1. From the table we see that all the metrics belonging to these four categories are full-reference type. The TBQ metrics (*SQRS*, P_D , P_{CI} , Simons et al.'s C_{Total}) are mostly reduced-reference type, and they are calculated globally or spatially or a mix of both. We have found no no-reference type spectral image quality metric. From the review, we have seen that all the spectral image quality metrics are aimed at certain application requirements. For instance, GFSQ metrics are aimed at evaluating the accuracy of spectral responses of the test images compared to the original images. None of the metrics so far proposed has been universally accepted as a general spectral quality metric. A comparative study on GFSQ and GFPQ metrics showed that none of the metrics are superior to others for all purposes, and that the choice of metric should be made based on appropriateness to the application [45]. There is, therefore, a need to do further research towards a more effective and possibly a more general spectral image quality metric.

Towards a framework for spectral imaging system quality

We have reviewed a range of spectral image quality metrics, which essentially aim to objectively evaluate the quality of spectral images. As there are many different types of spectral imaging systems that can be used to acquire the spectral images, it is increasingly important to evaluate the quality of these imaging systems themselves. In general one or more appropriate spectral image quality metrics are used to evaluate the quality of the imaging systems also. But from the review of the existing spectral image quality research, we see that the spectral image quality metrics/techniques so far proposed do not take into account all of the parameters that can influence the overall quality, and hence these metrics are not sufficient to fully evaluate the quality of the imaging systems. There is, therefore, a need for a spectral image quality framework which takes into account all of these factors including the characteristics of the scene, acquisition system, algorithms, application requirements etc. Information from the different stages of the spectral imaging workflow such as: spectral acquisition, processing, and resulting spectral data could provide information on these attributes. In this section we try to identify some of the most important of these as a basis for developing a framework for spectral imaging system quality.

- An important attribute that measures quality is the *spectral accuracy*. GFSQ metrics could be used to evaluate this attribute.
- The *perceptual quality* of a colour image rendered from a spectral image can be an important broad level attribute. This could be further detailed using effective image quality metrics from GFPQ, including colour accuracy.
- *Spatial* and *spectral resolution* are two important attributes whose information is available at the very beginning of the acquisition process. There could be a tradeoff between the spatial and spectral resolutions. A good example is the MSFA based spectral imaging system, where there is a need to compromise spatial resolution in order to increase the number of spectral bands.
- Many spectral acquisition systems rely on *image fusion* or *registration*. In such systems, pixel-to-pixel registration is important to obtain a high quality result. In some spectral imaging techniques, for example in satellite imaging, image fusion is one integral part of the spectral

imaging process and some research exists on the evaluation of the spectral images based on the quality of image fusion [25, 26]. Many of the spectral imaging systems use a combination of optical components that can produce both spatial and spectral distortion. These require careful characterisation, including the determination of a sensor model, in order to calibrate fully.

- *Noise* is an inevitable part of digital imaging. Different capturing methods involve different noise characteristics. An effective noise measurement and model should be developed in order to take noise into account more realistically.
- The processing stages may involve *geometrical corrections*, *spectral estimation methods*, and *demosaicking algorithms*, and these algorithms play a vital role in the final quality of a spectral image.
- *Target detection and recognition* capabilities could be other attributes to be considered and TBQ metrics are useful here.
- Repeatability and reproducibility are also very important for a good spectral imaging system. Vilaseca *et al.* [46] studied and analysed the repeatability, reproducibility and accuracy of a pushbroom hyperspectral system, and from their study they concluded that hyperspectral systems have good repeatability, adequate reproducibility and good accuracy. They used the spectral metric *RMS* and the colorimetric metrics ΔE^*_{ab} and ΔE^*_{00} in order to evaluate the accuracy.

We believe that taking into account all of these factors in a general quality framework would lead to a more effective evaluation of the quality of spectral imaging systems.

Conclusions

We have carried out a comprehensive review of previous studies on spectral image quality research. From this review, we found that most of the spectral image quality evaluations are intended for a number of specific domains and/or applications. They do not take into account all the key attributes that influence the quality of the resulting spectral data, and hence are not sufficient to be used to fully evaluate the quality of the spectral imaging systems. We have established a need for a generalised spectral image quality framework, and as a basis for its development, we have identified some of the important factors and attributes that might be involved in one or more of the steps in the workflow of the spectral acquisition process and which will, in turn, influence the overall quality of the spectral image data. Development of a general spectral image quality framework and metric, taking into account those attributes, will therefore be an important and useful area of future research.

References

1. FluxData, *Fluxdata FD-1665 series of 3 CCD cameras*, [www.fluxdata.com] – last accessed 31st January 2014].
2. Murakami Y, Yamaguchi M and Ohya N (2012), Hybrid-resolution multispectral imaging using color filter array, *Optics Express*, **20** (7), 7173-7183.
3. Martinez K, Cupitt J, Saunders D and Pillay R (2002), Ten years of art imaging research, *Proceedings of the IEEE*, **90** (1), 28-41.
4. Ribés A, Schmitt F, Pillay R and Lahanier C (2005), Calibration and spectral reconstruction for CRISATEL: An art painting multispectral acquisition system, *Journal of Imaging Science and Technology*, **49** (6), 463-473.

5. Miller PJ and Hoyt CC (1995), Multispectral imaging with a liquid crystal tunable filter, *Proceedings of SPIE: Optics in Agriculture, Forestry, and Biological*, **2345**, 354-365.
6. Hardeberg JY, Schmitt F and Brettel H (2002), Multispectral color image capture using a liquid crystal tunable filter, *Optical Engineering*, **41** (10), 2532-2548.
7. Miao L and Qi H (2006), The design and evaluation of a generic method for generating mosaicked multispectral filter arrays, *IEEE Transactions on Image Processing*, **15** (9), 2780-2791.
8. Shrestha R, Hardeberg JY and Khan R (2011), Spatial arrangement of color filter array for multispectral image acquisition, *Proceedings of SPIE: Sensors, Cameras, and Systems for Industrial, Scientific, and Consumer Applications XII*, **7875**, 787503-787503-9.
9. Shrestha R and Hardeberg JY (2013), CFA based simultaneous multispectral imaging and illuminant estimation, *Computational Color Imaging Workshop (CCIW)*, **7786**, 158-170.
10. Park JI, Lee MH, Grossberg MDZD and Nayar SK (2007), Multispectral imaging using multiplexed illumination, *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 1-8.
11. Shrestha R, Hardeberg JY and Boust C (2012), LED based multispectral film scanner for accurate color imaging, *Proceedings of the Eighth International Conference on Signal Image Technology and Internet based Systems (SITIS)*, 811-817.
12. Shrestha R and Hardeberg JY (2013), Multispectral imaging using LED illumination and an RGB camera, *Proceedings of the Twenty-first Color and Imaging Conference (CIC)*, 8-13, Albuquerque, USA.
13. Parmar M, Linsel S and Farrell J (2012), An LED-based lighting system for acquiring multispectral scenes, *Proceedings of SPIE: Digital Photography VIII*, **82990**, 82990P-82990P-8.
14. Langfelder G, Longoni AF and Zaraga F (2011), Implementation of a multispectral color imaging device without color filter array, *Proceedings of SPIE: Digital Photography VII*, **7876**, 787608-787608-9.
15. Shrestha R, Mansouri A and Hardeberg JY (2011), Multispectral imaging using a stereo camera: Concept, design and assessment, *EURASIP Journal on Advances in Signal Processing*, **2011** (1), 57-57.
16. Shrestha R and Hardeberg JY (2012), Simultaneous multispectral imaging and illuminant estimation using a stereo camera, in *Image and Signal Processing*, Lecture Notes in Computer Science (LNCS), **7340**, 45-55, Springer Berlin Heidelberg.
17. Shrestha R, Hardeberg JY and Mansouri A (2011), One-shot multispectral color imaging with a stereo camera, *Proceedings of SPIE: Digital Photography VII*, **7876**, 787609-787609-11.
18. Goetz AFH, Vane G, Solomon JE and Rock BN (1985), Imaging spectrometry for earth remote sensing, *Science*, **228** (4704), 1147-1153.
19. Wang Z, Bovik AC, Sheikh HR and Simoncelli EP (2004), Image quality assessment: From error visibility to structural similarity, *IEEE Transactions on Image Processing*, **13** (4), 600-612.
20. Pedersen M (2011), Image quality metrics for the evaluation of printing workflows, *PhD Thesis*, University of Oslo and Gjøvik University College.
21. Martin L, Vrabel J and Leachtenauer J (1999), Metrics for assessment of hyperspectral image quality and utility, *Proceedings of International Symposium on Spectral Sensing Research*, Las Vegas, USA.
22. Sweet J, Granahan J and Sharp M (2000), An objective standard for hyperspectral image quality, *Proceedings of AVIRIS Workshop*, Jet Propulsion Laboratory, Pasadena, California, USA.
23. Kerekes JP and Hsu SM (2004), Spectral quality metrics for VNIR and SWIR hyperspectral imagery, *Proceedings of SPIE: Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X*, **5425**, 549-557.
24. Kalenova D, Toivanen P and Bochko V (2005), Preferential spectral image quality model, in *Image Analysis*, Lecture Notes in Computer Science (LNCS), **3540**, 389-398, Springer Berlin Heidelberg.
25. Babawuro U and Beiji Z (2011), Satellite imagery quality evaluation using image quality metrics for quantitative cadastral analysis, *International Journal of Computer Applications in Engineering Sciences*, **1**, 391-395.

26. Al-Wassai FA, Kalyankar NV and Al-Zaky AA (2012), Spatial and spectral quality evaluation based on edges regions of satellite: Image fusion, *Proceedings of the Second International Conference on Advanced Computing Communication Technologies (ACCT)*, 265-275.
27. ISO (2004), ISO 20462-2 photography - psychophysical experimental methods to estimate image quality - part 2: Triplet comparison method.
28. Romero J, García-Beltrán A and Hernández-Andrés A (1997), Linear bases for representation of natural and artificial illuminants, *Journal of the Optical Society of America A*, **14** (5), 1007-1014.
29. Kruse FA, Lefkoff AB, Boardman JW, Heidebrecht KB, Shapiro AT, Barloon PJ and Goetz AFH (1993), The spectral image processing system (SIPS) - interactive visualization and analysis of imaging spectrometer data, *Remote Sensing of Environment*, **44** (2-3), 145-163.
30. Chang CI (1999), Spectral information divergence for hyperspectral image analysis, *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, **1**, 509-511.
31. CIE (2004), CIE 15:2004: Colorimetry.
32. CIE (1995), Industrial colour-difference evaluation, publication CIE 116-95.
33. Luo MR, Cui G and Rigg B (2001), The development of the CIE 2000 colour-difference formula: CIEDE 2000, *Color Research and Application*, **26** (5), 340-350.
34. Fairman HS (1987), Metameric correction using parametric decomposition, *Color Research and Application*, **12**, 261-265.
35. Viggiano JAS (2002), Perception-referenced method for comparison of radiance ratio spectra and its application as an index of metamerism, *Proceedings of SPIE: The Ninth Congress of the International Colour Association*, **4421**, 701-704, Rochester, USA.
36. Pedersen M and Hardeberg JY (2012), Full-reference image quality metrics: Classification and evaluation, *Foundations and Trends in Computer Graphics and Vision*, **7** (1), 1-80.
37. Garzelli A and Nencini F (2009), Hypercomplex quality assessment of multi/hyperspectral images, *IEEE Geoscience and Remote Sensing Letters*, **6** (4), 662-665.
38. Zhou W and Bovik AC (2002), A universal image quality index, *IEEE Signal Processing Letters*, **9** (3), 81-84.
39. Smith WD (2004), *Quaternions, octonions, and now, 16-ons and 2n-ons; new kinds of numbers*, [<http://www.scorevoting.net/WarrenSmithPages/homepage/nce2.pdf>] – last accessed 31st January 2014].
40. LeMoan S and Urban P (2013), Evaluating the perceived quality of spectral images, *Proceedings of the International Conference on Image Processing (ICIP)*, 2024-2028, Melbourne, Australia.
41. Lissner I, Preiss J, Urban P, Lichtenauer MS and Zolliker P (2013), Image-difference prediction: From grayscale to color, *IEEE Transactions on Image Processing*, **22** (2), 435-446.
42. Lissner I and Urban P (2012), Toward a unified color space for perception-based image processing, *IEEE Transactions on Image Processing*, **21** (3), 1153-1168.
43. Shen SS (2003), Spectral quality equation relating collection parameters to object/anomaly detection performance, *Proceedings of SPIE: Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery IX*, **5093**, 29-36.
44. Simmons RE, Elder TD, Stewart DJ, Cincotta EJ, Kennedy CS and Van Nostrand RC (2005), General spectral utility metric for spectral imagery, *Proceedings of SPIE: Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, **5806**, 457-468.
45. Imai FH, Rosen MR and Berns RS (2002), Comparative study of metrics for spectral match quality, *Proceedings of the First European Conference on Colour in Graphics, Imaging and Vision (CGIV)*, 492-496, Poitiers, France.
46. Vilaseca M, Schael B, Delpueyo X, Chorro E, Perales E, Hirvonen T and Pujol J (2013), Repeatability, reproducibility, and accuracy of a novel pushbroom hyperspectral system, *Color Research and Application*, doi: 10.1002/col.21851.

Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems

Paper 2

This chapter is a reprint of the publication:

Pillay, R., Picollo, M., Hardeberg, J. Y. & George, S. “Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems”, *Sensors*, 20, 381 (2020)



Article

Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems

Ruven Pillay ^{1,2,*} , Marcello Piccolo ³ , Jon Yngve Hardeberg ¹ and Sony George ¹

¹ NTNU-Norwegian University of Science and Technology, Department of Computer Science, Teknologivegen 22, N-2815 Gjøvik, Norway; jon.hardeberg@ntnu.no (J.Y.H.); sony.george@ntnu.no (S.G.)

² C2RMF-Centre de Restauration et Recherche des Musées de France, Porte des Lions-Palais du Louvre, 75001 Paris, France

³ IFAC-CNR-Istituto di Fisica Applicata “Nello Carrara” del Consiglio Nazionale delle Ricerche, Via Madonna del Piano 10, 50019 Firenze, Italy; m.piccolo@ifac.cnr.it

* Correspondence: ruven.pillay@culture.gouv.fr

Received: 26 May 2020; Accepted: 30 June 2020; Published: 8 July 2020



Abstract: In this study, the results from a round-robin test of hyperspectral imaging systems are presented and analyzed. Fourteen different pushbroom hyperspectral systems from eight different institutions were used to acquire spectral cubes from the visible, near infra-red and short-wave infra-red regions. Each system was used to acquire a common set of targets under their normal operating conditions with the data calibrated and processed using the standard processing pipeline for each system. The test targets consisted of a spectral wavelength standard and of a custom-made pigment panel featuring Renaissance-era pigments frequently found in paintings from that period. The quality and accuracy of the resulting data was assessed with quantitative analyses of the spectral, spatial and colorimetric accuracy of the data. The results provide a valuable insight into the accuracy, reproducibility and precision of hyperspectral imaging equipment when used under routine operating conditions. The distribution and type of error found within the data can provide useful information on the fundamental and practical limits of such equipment when used for applications such as spectral classification, change detection, colorimetry and others.

Keywords: round-robin test; hyperspectral imaging; spectral quality; data quality; spectroscopy; colorimetry

1. Introduction

Hyperspectral imaging provides a powerful combination of high spectral resolution and dense spatial mapping and has become a valuable analytical tool in a wide range of fields, including remote sensing [1], food science [2], astronomy [3], mineralogy [4], agriculture [5], medicine [6], the study of art [7] and many others. Hyperspectral data can be used in a number of ways, including, for example, materials mapping and identification, the detection of hidden features, change monitoring and many other applications.

However, the accuracy of real-world hyperspectral imaging systems is not only limited by the technical characteristics of the hyperspectral cameras themselves, but also by the overall setup, how they are used and how the acquired data is processed. These factors include the power and spectral content of the lighting used, the optical geometry, the integration time and acquisition parameters, as well as the kind of processing and calibration of the data that is carried out. The reproducibility of hyperspectral data between systems, institutions and indeed over time is an important consideration in many applications, especially when the comparison of data is important. For example, when monitoring change, when performing classification using 3rd party spectral libraries or when combining a variety of data sets in a multimodal analysis.

In order, therefore, to assess the real-world reproducibility and limits of hyperspectral imaging, we will in this study compare a number of different hyperspectral data sets acquired of a single set of targets, but using different hyperspectral imaging equipment from different institutions.

The experimental data for this study was acquired as part of a round-robin test for the COSCH project [8] involving nineteen different institutions, which included research laboratories, universities, equipment manufacturers and museums. The goal was to evaluate the effective limits in accuracy of the spectral imaging equipment currently in use within these institutions. The data acquisition was carried out using either multispectral or hyperspectral equipment from different manufacturers and with different experimental acquisition setups, procedures and methodologies. The aim of the comparison was *not* to compare the hardware specifications or raw performance of the imaging devices themselves, but to measure the resulting effective performances of the systems globally under their standard operating conditions and after the application of the calibration and processing workflows that are usually applied by each institution to their system. The variability due to the different setups, operating procedures and the way data was processed was, therefore, an important factor to take account of and include in the study. In this way, an insight into the practical limits in accuracy of hyperspectral systems within routine operating environments could be gleaned.

Round-robin tests are a useful means of comparing equipment or methodologies and have been successfully carried out in various related fields including the measurement of the BRDF (Bidirectional Reflectance Distribution Function) of diffuse reflectors [9], for radiometric calibration of a satellite multispectral sensor [10] or in laboratory settings for field spectrometers [11]. Although the round-robin test was mainly concerned with the study of cultural heritage, the methodology and results from the test presented here are generic and relevant to a wide range of fields.

The round-robin test included both multispectral as well as hyperspectral imaging equipment. However, for this study, we have limited ourselves to a subset of the participants who were able to provide data from pushbroom hyperspectral imaging devices. This subset of data was obtained from eight participating institutions including cultural heritage conservation centers, specialist research laboratories, research libraries, universities as well as data obtained directly from two different hyperspectral camera manufacturers.

These participants all used broadly similar pushbroom hyperspectral imaging devices, configured to use either horizontal or vertical linear translation stages for scanning. For all systems, the light sources moved together with the camera, providing constant, though not necessarily uniform, illumination. Each hyperspectral camera system was used to acquire the set of targets under their usual operating conditions. Calibration and processing of the data was carried out according to the procedures that were routine for that particular system and institution. The data presented here includes data from the visible and near infra-red (VNIR) spectral region (approximately 400–1000 nm) as well as from the short-wave infra-red (SWIR) spectral region (approximately 1000–2500 nm). Hyperspectral cameras with different detectors were used for the VNIR and SWIR regions, with Silicon-based CCD (Charged Couple Device) or CMOS (Complementary Metal Oxide Semiconductor) detectors for the VNIR and for the SWIR either InGaAs (Indium Gallium Arsenide) based detectors with spectral ranges of approximately 1000–1700 nm or MCT (Mercury Cadmium Telluride) detectors with spectral ranges of approximately 1000–2500 nm.

All eight institutions were able to provide usable data in the VNIR region with 6 of these also providing data in the SWIR region. The systems included 8 different VNIR cameras from 3 different manufacturers and 6 different SWIR cameras from 2 different manufacturers. The essential spectral and spatial characteristics of the acquired data are given in Table 1, which show the hardware used, the spectral ranges, the number of bands captured and the spatial resolution at which the data was acquired. More detailed hardware characteristics and acquisition parameters used can be found in the Appendix B in Table A2.

Each hyperspectral system was a production-ready system which had been characterized and “factory-calibrated” for spectral alignment either directly by the manufacturer or by in-house specialists

within the institution. As we can see from Table 1, although the hardware used was often very similar or even identical, each system was configured to operate differently with different spectral ranges, different numbers of spectral bands, different average bandwidths and different spatial resolutions.

Table 1. Hyperspectral cameras and the spectral and spatial characteristics of the acquired data from the round-robin test.

System	Wavelengths (nm)	Number of Bands	Average Bandwidth (nm)	Resolution (Pixels/mm)
VNIR				
Hypex VNIR1600	414.70–994.01	160	3.62	29.2
Hypex VNIR1800	404.16–994.55	186	3.17	10.7
Specim V10E #1	400.11–900.77	406	1.23	3.84
Specim V10E #2	395.58–1006.05	210	2.91	10.7
Specim V10E #3	394.80–1009.13	776	0.79	10.26
Specim V10E #4	400.65–949.95	393	1.40	4.66
Specim sCMOS-50-V10E	400.73–999.98	472	1.27	9.46
Headwall Hyperspec	382.27–1001.14	346	1.79	11.6
SWIR				
Specim V17E #1 (InGaAs)	954.45–1661.65	337	2.10	7.92
Specim V17E #2 (MCT)	965.18–2563.78	256	6.24	2.34
Specim V17E #3 (MCT)	919.66–2521.58	256	6.26	2.34
Specim V17E #4 (MCT)	932.20–2530.50	255	6.27	3.26
Specim SWIR (MCT)	907.16–2523.43	288	5.61	3.10
Hypex SWIR384 (MCT)	954.98–2511.03	288	5.40	2.26

In addition, each system possessed slightly different acquisition optical geometries; different illuminant light sources; different acquisition times and differed as to whether techniques such as binning and averaging were used etc. There was also variability in the way raw data was processed by each institution. In all cases, radiometric calibration was carried out, which entailed the use of a known traceable reflectance standard (typically Spectralon[®], manufactured by Labsphere Inc.). Although there were variations in the sophistication of the calibration procedures used by each institution, the basic principles used by all groups were identical, consisting essentially of subtraction of the dark noise current and then signal normalization to “absolute” reflectance. In all cases, this involved the acquisition of a dark current signal using an identical integration time to the scan, which was averaged over 10–100’s of acquisitions. And for normalization, this involved scanning a reflectance standard and a uniform neutral target large enough to fill the entire field of view in order to correct for inhomogeneity in the illumination and to scale the pixel responses to absolute reflectance factor. Several systems performed this in a single step with a reflectance standard large enough to fill the field of view. Several systems additionally had accurate pixel-wise relative responsivity values generated using an integrating sphere carried out in-house or from a manufacturer-supplied characterization. The calibration routines were, therefore, similar, but not identical and followed to varying degrees the hyperspectral acquisition, processing and calibration workflows described in Reference [12]. Although differences in the resulting spectra in such circumstances are to be expected, the scale of this variability provides important information on the ability to compare data across systems and institutions.

The results presented here are from two of the test targets used for the round-robin test: a diffuse lambertian Zenith Polymer[®] wavelength standard and a custom-made pigment panel made up of 7 different historical pigments in a tempera binder. For an overview of the round-robin test, see Reference [13]. Details and results from the other test targets used in the round-robin test (a Macbeth ColorChecker chart and a 19th Century Russian icon) can be found in References [14] and [15] respectively.

To evaluate the data, a number of different analyses were carried out on the data to ascertain the accuracy of the data, the types of errors within the data and the variability between the different sets of data. These analyses included quantitative evaluations of the accuracy of the data in terms of spectral accuracy, geometric accuracy and colorimetric accuracy. In addition, the quality of the data in terms of noise was also quantified and compared between data sets.

2. Test Targets and Methodology

The two test targets used for this analysis were a wavelength standard and a custom-made pigment panel made up of historical pigments in an egg tempera binder.

2.1. Wavelength Standard

Wavelengths standards are reflectance targets designed for precise wavelength calibration of spectrophotometers, reflectometers and other spectral instruments. The standard used in this study was a Zenith Polymer[®] wavelength standard (manufactured by SphereOptics GmbH) (Figure 1), which is a chemically inert diffuse lambertian reflectance standard composed of PTFE (Polytetrafluoroethylene) doped with the oxides of the rare earth elements Holmium, Erbium and Dysprosium. This combination gives the standard a stable spectrum of characteristic, well-defined and narrow features over the UV, visible and near infra-red spectral ranges, which is suited for use in accurate spectral calibration. The standard is supplied with traceable, laboratory-certified reference reflectance measurements covering the entire spectral range.

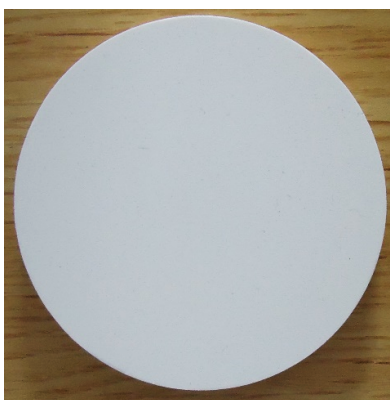


Figure 1. Rare-Earth Doped Wavelength Standard.

2.2. Pigment Panel

The other test target evaluated in this study was a painted panel of approximately 22 × 29 cm that was made up of historical pigments at different concentrations in an egg tempera binder. The panel, shown in Figure 2, was created especially for the round-robin test by IFAC-CNR, one of the participating institutions. The paint preparation and application aimed to authentically reproduce the medieval Tuscan painting technique described in Cennino Cennini's 15th century *Il Libro dell'Arte* [16].

The panel consists of a wooden support with a gypsum ground, a canvas layer, and a second gypsum ground layer. Before application of the paint layer, five types of drawing materials (watercolor, charcoal, graphite, a lead and tin-based metalpoint and a lead-based metalpoint) were used to create lines and line patterns that were then covered with paints applied with two different thicknesses.

Seven pigments (obtained from Zecchi - Colori, Belle Arti, Florence) were chosen which were widely used during the Renaissance and were commonly applied using an egg tempera technique. The pigments were Burnt Umber, Carmine, Vermilion, Malachite, Azurite, Lead White, and Ivory Black (the compositions of the pigments can be found in the Appendix B in Table A1). These pigments were

mixed in an egg tempera binder and applied to the panel creating the vertical strips seen in Figure 2. Additional details on the panel and on the source and composition of the pigments used for the panel can be found in Reference [17].

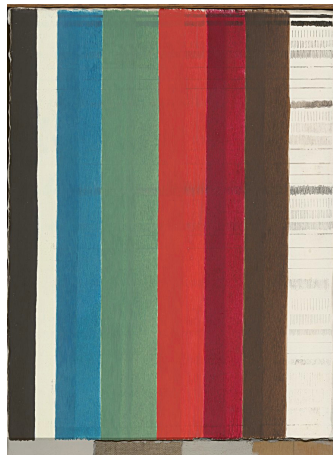


Figure 2. Pigment panel with vertical strips of historical pigments mixed in an egg tempera binder. From left to right: ivory black, lead white, azurite, malachite, vermilion, carmine and burnt umber. To the right of the pigments is an unpainted gypsum ground layer with various types of metalpoint drawing.

The pigments were chosen as they all have distinct spectra in both the VNIR and SWIR spectral ranges, allowing useful spectral analysis to be carried out in either spectral region. In particular, each pigment has strong absorption bands in the VNIR range and all behave differently with respect to each other in the SWIR region. Carmine, for example, is very transparent in the SWIR whereas Burnt Umber is opaque. In addition, Lead White, Azurite and Gypsum possess sharp and medium-strong absorption bands in the SWIR region which can be used to test the ability of the different hyperspectral systems to detect such spectral features.

The choice of a hand-made painted panel using historical pigments was made because one of the goals of the spectral imaging carried out for the project was the documentation and study of works of art and, in particular, of paintings. Natural pigments typically do not contain the kinds of abrupt absorption features seen in the wavelength standard and possess smooth continuous spectra [18]. Such a target, therefore, provides very different characteristics to more industrial materials and provides, therefore, an important test case.

As with the wavelength standard, a set of reference spectra was required with which to compare the acquired hyperspectral data to. These were obtained using Fiber Optic Reflectance Spectroscopy (FORS) at IFAC-CNR. FORS measurements in the UV-Visible-NIR-SWIR range were performed using two single-beam spectroanalyzers (Zeiss MCS601 Si UV-NIR & MCS611 InGaAs NIR 2.2WR) using a 0/45° geometry. Data was calibrated and processed using the manufacturer supplied software, giving spectra with 0.8 nm and 6 nm resolutions in the VNIR and SWIR respectively resulting in a total spectral range from 350–2200 nm. The reflectance spectra of the seven pigments and ground layer are shown in Figure 3. Of course, as the hand-made target is not perfectly homogeneous, the FORS data does not necessarily provide objective spectral data. However, for the purposes of this study, it does provide a sufficiently accurate base-line from which to compare the hyperspectral data cubes.

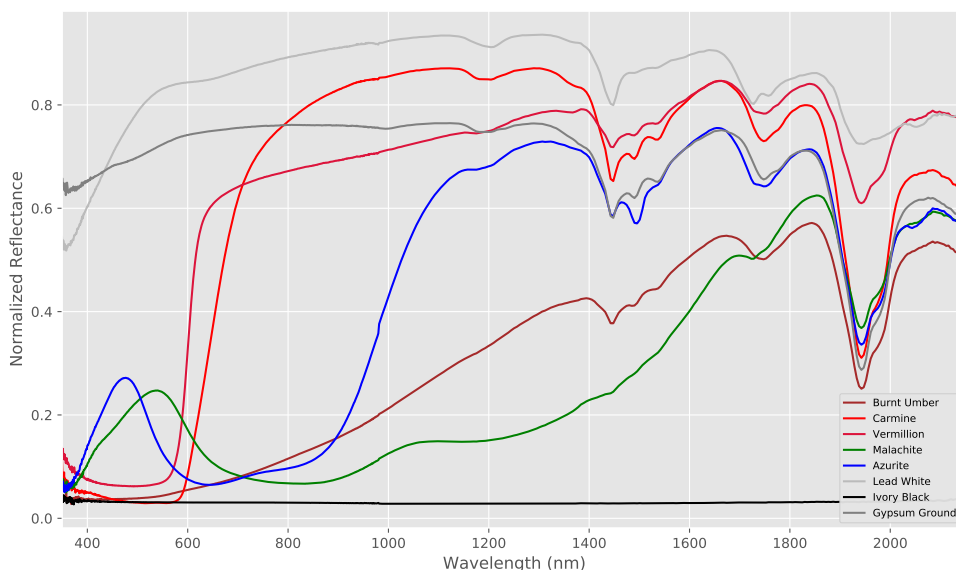


Figure 3. FORS reference reflectance spectra for the panel pigments.

2.3. Acquisition and Measurement Methodology

Each hyperspectral system acquired data with each of the two test targets. However, one of the participants (using a VNIR Headwall Hyperspec) did not provide data for the pigment panel, resulting in 8 data sets in the VNIR spectral range for the wavelength standard and only 7 data sets for the pigment panel. All 6 SWIR systems provided data for both targets.

The hyperspectral data itself was acquired and processed by each institution under the operating conditions that were standard for the institution and equipment. The acquisition parameters were all set to values considered appropriate for obtaining high quality data from the targets. The list of available parameters is provided in the Appendix B in Table A2 and the characteristics of the final data can be found in Table 1. Other parameters such as detector gain, were kept at the default level of 1 by all participants. For the illumination, all systems used similar 100 or 150 W variable Tungsten-Halogen light sources with smooth spectra that covered the whole of the spectral ranges of the cameras. However, many of the systems allow variable output and the precise illumination levels are, therefore, not known. Additionally, for the majority of systems, the acquisition integration times were not recorded by the manufacturer-supplied acquisition software and are, therefore, not available. Nor is the closely related translation stage scan speed, which was not recorded by any of the acquisition software.

After acquisition and processing carried out by the participating institutions, the resulting final calibrated spectral images were provided for analysis. The spectral data used in this paper (available as Supplementary Material) was extracted directly from these spectral images using identical protocols to ensure any variability in the spectra came from the data and not from the analytical methodology.

For both the wavelength standard and the pigment panels, the representative reflectance spectra from the acquired hyperspectral data was obtained by averaging the spectra over several hundred or thousand pixels. In this way, noise was significantly reduced and inhomogeneities averaged out.

In the case of the wavelength standard, the used pixel spectra were taken from a central circular region of size $\frac{2}{3}$ of the diameter of the target. For the pigment panel a square region within each pigment within the hyperspectral image was extracted, which was chosen to be approximately coincident to that used for physically obtaining the FORS reference spectra. The region was chosen to be, once again, $\frac{2}{3}$ of the size of the width of the pigment strip in order to allow the data to be averaged.

3. Results and Analysis

In this section, we will look in detail at the results obtained from the different VNIR and SWIR hyperspectral systems for both the wavelength standard and pigment panel. There are a number of ways to evaluate and quantify the hyperspectral data obtained and we will examine a range of these. These include various quantitative measures for how accurately the spectra have been reproduced, the level of noise within each data set, the geometric accuracy of the images and the fidelity of the colorimetry of each system.

3.1. Spectral Accuracy

The most important criteria for a hyperspectral imaging system is how accurately the spectra are reproduced. However, quantifying this is not always straightforward as there are various ways to quantify spectral accuracy and the appropriate measure of accuracy can depend on how the results will be used. For example, when using hyperspectral data for classification tasks, measures of “spectral distance” are generally used, whereas when using data for spectroscopy, factors such as the accuracy of the localization of spectral features will be more important. In this section, therefore, we will use a number of different ways to evaluate the spectral accuracy of the systems, including residual errors, distance metrics, spectral mis-alignment, spectral feature detection and how accurately the derivatives are reproduced.

3.1.1. Spectral Errors

The spectral responses of the different systems for the wavelength standard are shown in Figure 4 together with the reference values supplied by the manufacturer. All systems were able to produce spectra that are similar to the reference values for both the VNIR and SWIR cameras and which correctly detect the main spectral features present in the wavelength standard. However, a more detailed examination shows that there are small but clear differences in the form of the spectra and, in addition, small spectral mis-alignments in the location of the spectral features between the reference spectra and the data from the various hyperspectral imaging systems.

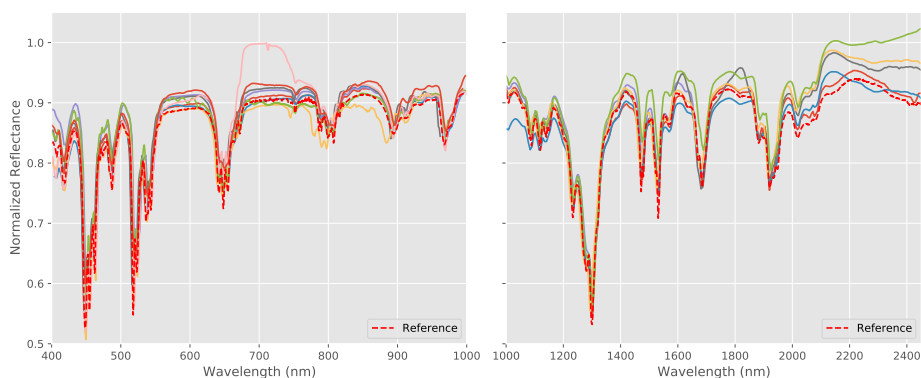


Figure 4. Reference and measured spectral reflectances for the wavelength standard for the visible and near infra-red (VNIR) systems (**left**) and short-wave infra-red (SWIR) systems (**right**), where the colored lines represent the individual spectra measured by each system. The systems all produced similar spectra and were able to reproduce the main spectral features in the standard. However, there are small but clear differences in the measured spectra and small spectral mis-alignments for certain systems.

If we consider the responses with respect to wavelength, we see that the spectral responses are close to the reference for the majority of the spectral ranges of each system, but a number of the

SWIR systems have problems beyond 2000 nm, where differences in both the shape and amplitude are evident.

The reference target contains sharp narrow peaks and troughs, which are in several cases, beyond the spectral resolution of the hyperspectral imaging systems. Each hyperspectral system has slightly different wavelength ranges, different numbers of bands with different central wavelengths and different bandwidths for each band. Therefore, in order to make meaningful quantitative comparisons, it was first necessary to resample each of our data sets to a common discretization. As the traceable laboratory reference values have a higher sampling resolution of 0.5 nm, this reference was resampled using the given center wavelengths and the given FWHM (full width half maximum) for each band for each camera. The reference spectra were essentially convolved at each center wavelength for each band with a Gaussian with the given FWHM in order to mimic the spectral and bandwidth characteristics of each of the different hyperspectral systems.

From these resampled reference spectra, the residual error was calculated for each system by subtracting the measured spectra from the reference spectra for each band. These residual errors can be seen in Figure 5. As we can see, there are large errors at the sharp peaks and troughs, but otherwise taking the average errors and trends, we see that the error ranges are relatively small, ranging from 0.05–0.18 for the VNIR systems and 0.01–0.15 for the SWIR.

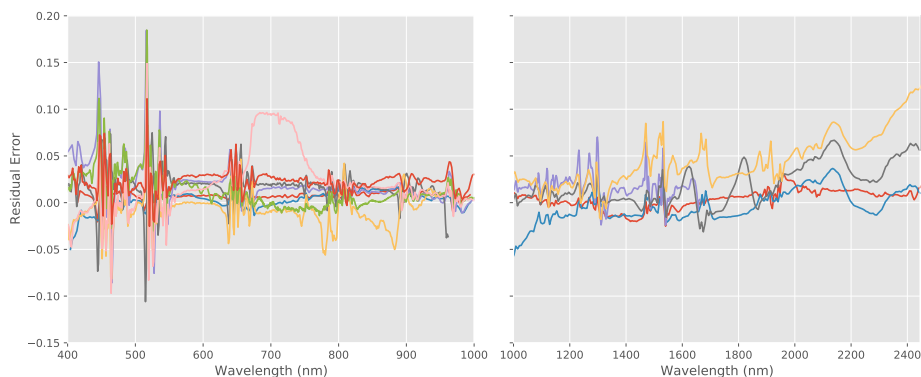


Figure 5. Residual errors between measured reflectance and reference spectra for wavelength standard for the VNIR (**left**) and SWIR systems (**right**) where the colored lines represent the individual errors calculated for each system. Errors are relatively small ranging from 0.05–0.18 for the VNIR systems and 0.01–0.15 for the SWIR systems.

The results for the pigment panel are shown in Figure 6 which shows the processed spectra measured by each hyperspectral system together with the FORS reference for the ensemble of pigments used on the pigment panel plus the Gypsum ground layer. As with the wavelength standard, both the VNIR and SWIR systems were able to broadly reproduce the overall shape and essential features of the spectra of each pigment. Although the spectra match very closely in the VNIR in most cases, there is again more variation in the SWIR spectra both in terms of spectral shape and in terms of the amplitudes of the spectra.

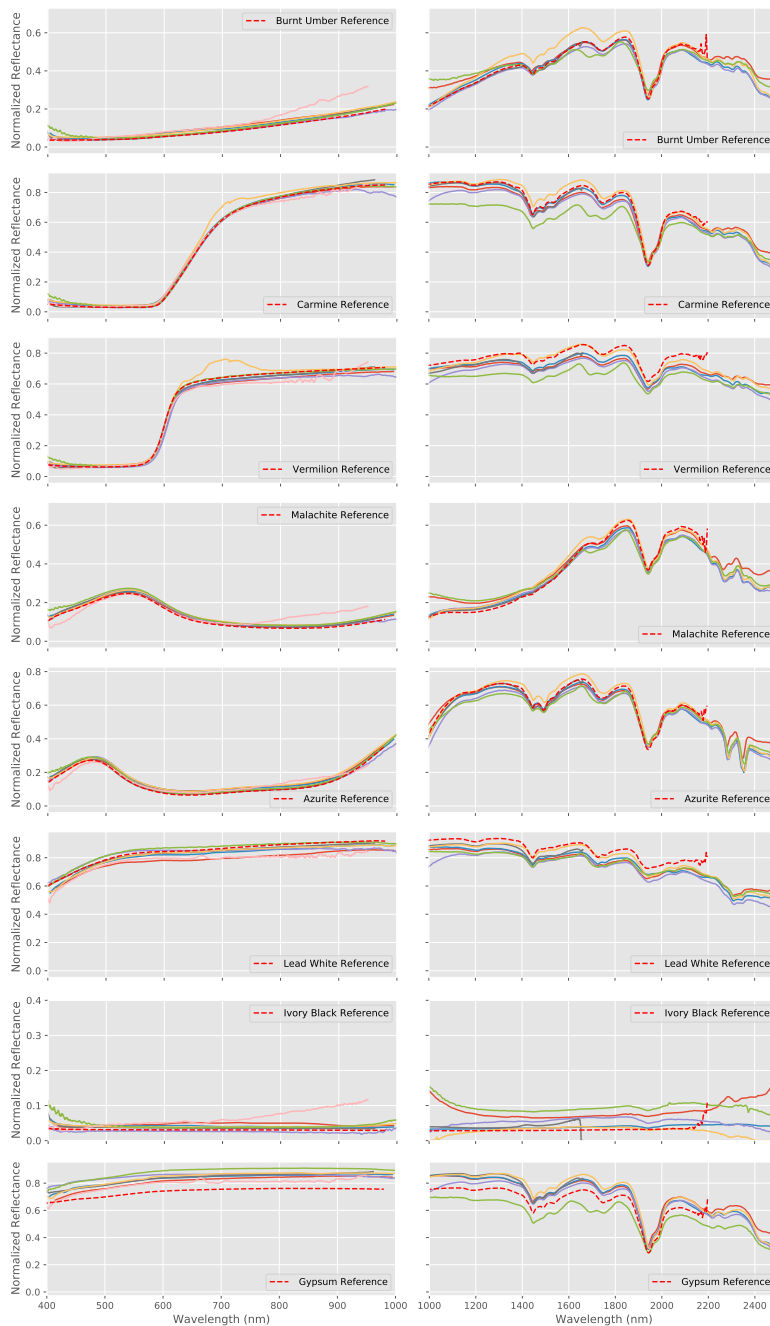


Figure 6. Spectral reflectances measured by each hyperspectral system (VNIR left and SWIR right) together with the FORS reference for the ensemble of pigments used on the pigment panel plus the Gypsum ground layer, where the colored lines represent the individual spectra measured by each system (note that the FORS wavelength range has a maximum of only 2200 nm in the SWIR region). The spectral shapes are all accurately reproduced, especially in the VNIR region.

3.1.2. Spectral Alignment

Closer inspection of the data from Figure 4 shows that there are slight mis-alignments between the spectra acquired by the hyperspectral systems with respect to the reference. This can be seen more clearly if we take the results from a single system and Figure 7 shows the full VNIR spectra for the reference and for a single data set. The spectral mis-alignment is clearly visible in the zoomed view on the right, showing a narrower range of wavelengths.

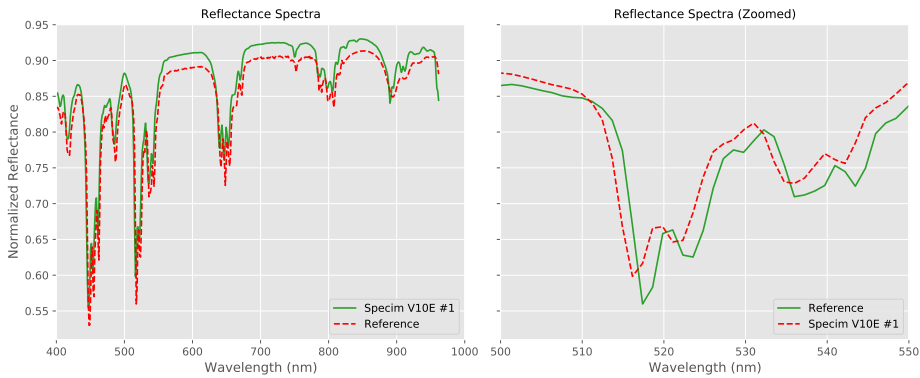


Figure 7. Comparison of the reflectance spectra of wavelength standard for one of the hyperspectral systems with the reference values. The zoomed view (right) shows more clearly the spectral mis-alignment.

In order to quantify this spectral mis-alignment seen in several of the hyperspectral cameras, the derivative of the spectra was used, as the target contains a large number of narrow peaks and troughs. The zero-crossings in the derivative correspond to the top and bottoms of the peaks and troughs of the reflectance spectra and phase correlation [19] was applied to the derivative of the spectra to calculate a global offset between the resampled reference signal and the measured spectra. In Figure 8, we can see the alignment error for each system (the values are provided in the Appendix B in Table A3). The mis-alignment ranges from 0.0–2.6 nm for our different VNIR hyperspectral systems with an average mis-alignment magnitude of 0.93 ± 0.83 nm, which is less than the average distance between center wavelengths. The SWIR systems behave very similarly with error magnitudes ranging from 0.1–5.5 nm and an average mis-alignment magnitude of 1.75 ± 1.05 nm, which is again less than the average system bandwidth.

Mis-alignment can also vary with respect to wavelength and to evaluate this, the mis-alignment was calculated at different wavelengths by using small sub-sections of the spectra. However, no measurable difference was found with respect to wavelength for any of the systems.

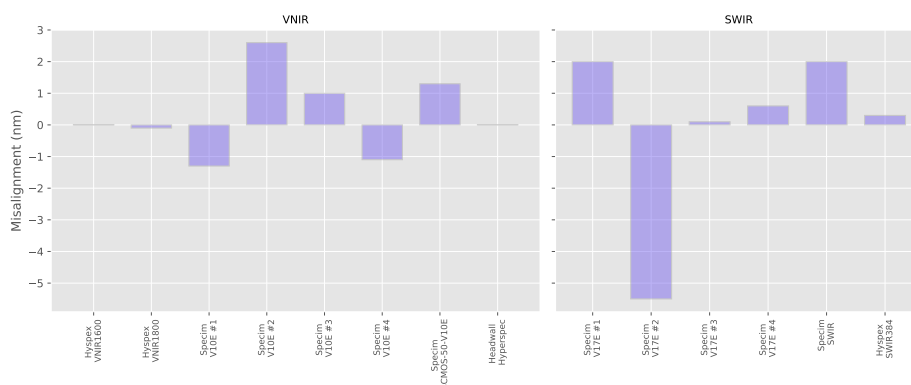


Figure 8. Errors in the spectral alignment for each hyperspectral system (colored lines) calculated from the data from the wavelength standard.

3.1.3. Distance Metrics

Hyperspectral imaging is often used for the purposes of material identification, which is usually performed through some form of classification. Classification algorithms [20] require a quantitative measure of the difference (or distance) between different spectra. In addition, they typically require the appropriate setting of threshold parameters, whose values are often dependent on the distance metric used and which can make an important difference to the classification results obtained. Given the central role played by distance metrics in classification workflows, it is useful to be able to quantify the range and variability of the spectral distances observed in real data sets.

A large number of ways exist for calculating this distance, each of which have their own particular characteristics and properties [21]. The most widely used of these distance metrics is the “Spectral Angle” used in Spectral Angle Mapper (SAM) classification. This measure is extensively used in remote sensing and is derived from the angle formed between a reference spectrum and the image spectrum of each pixel [22]. A number of other more recent distance metrics exist, including Spectral Information Divergence [23], Spectral Correlation Mapping [24], Spectral Gradient Angle [25] amongst many others. Although these metrics improve on SAM, SAM remains the most widely used distance measure thanks to its relative simplicity and integration in many hyperspectral processing software packages.

If we calculate the spectral distance between the measured reflectance spectra and the reference spectra, we can obtain a value for the accuracy of each system. Knowing how the measurement errors found in typical hyperspectral data affect distance metrics provides useful information for the setting of thresholds in classification workflows and for understanding the results and limits from classification.

These errors can be seen in Figure 9, which shows the Spectral Angle distance (Equation (A1) in the Appendix A) between each measured spectra and the appropriately resampled reference for the wavelength standard. The distances are listed in the Appendix B in Table A3. The average Spectral Angle for the VNIR systems is 0.0211 for the VNIR and 0.567 for the SWIR systems.

The results for each individual pigment in the pigment panel can be seen in Figure 10. The median of the result over all pigments (both thickly and thinly applied as well as for the Gypsum ground) for each individual system can be seen in Figure 11 with the distances listed in Table A3. As the FORS reference values in the SWIR region cover a smaller range of wavelengths (up to around 2200 nm) than the MCT-based systems, only the subset of the hyperspectral data that was coincident with the reference was used for calculating the Spectral Angle. The average over all pigments in the VNIR is 0.0608 ± 0.0261 and 0.0890 ± 0.0481 in the SWIR, so although the average magnitude of the errors are similar in both wavelength ranges, there is more variability in the results from the SWIR systems. If we look in more detail at the individual pigments, we note that in the VNIR region, the Spectral Angle error looks to be correlated to the average level of reflectance with the most reflective materials, Lead

White and Gypsum. having lower errors than the colored pigments and the very dark pigment, Ivory Black which has a reflectivity of only around 3%, having by far the largest errors.

As we saw in Section 3.1.1, the average error is again higher in the SWIR than the VNIR systems for both targets.

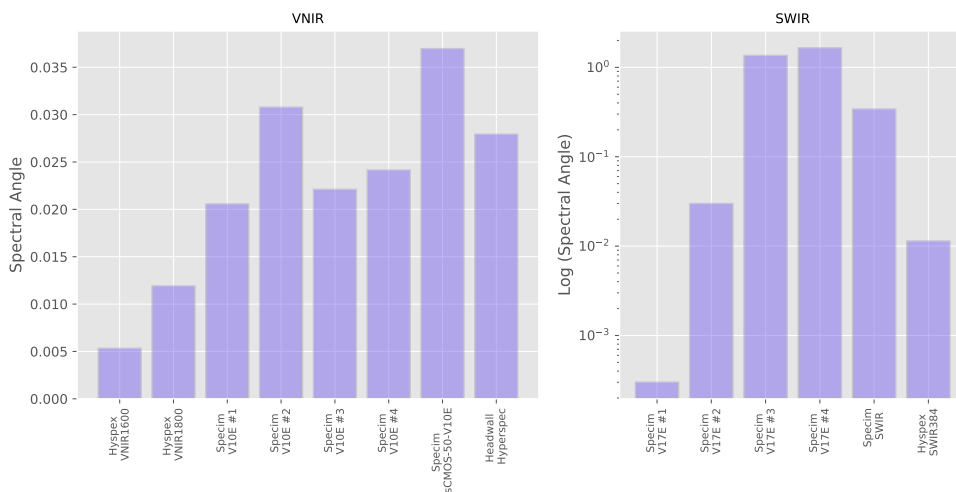


Figure 9. “Spectral Angle” calculated between the spectra measured by each hyperspectral system and the reference for the wavelength standard. (Note that the vertical axis of the SWIR data is a logarithmic scale due to the wide range of scores).

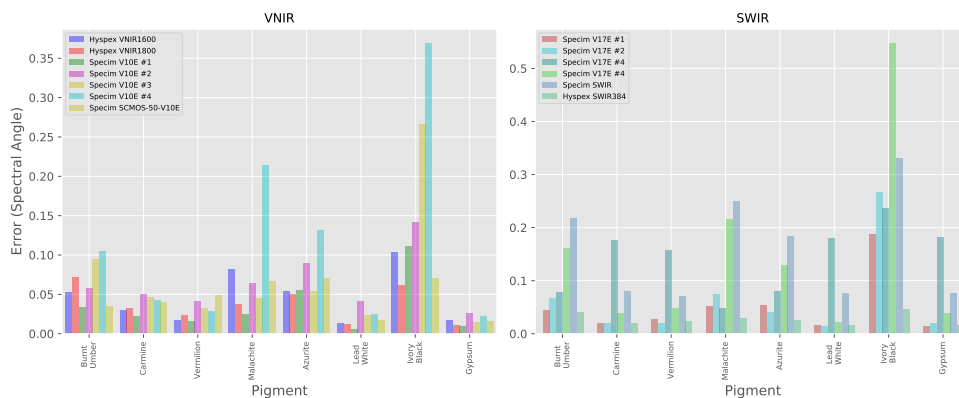


Figure 10. Spectral Angle error between the spectra measured by each hyperspectral system and the FORS reference for each (thickly applied) pigment.

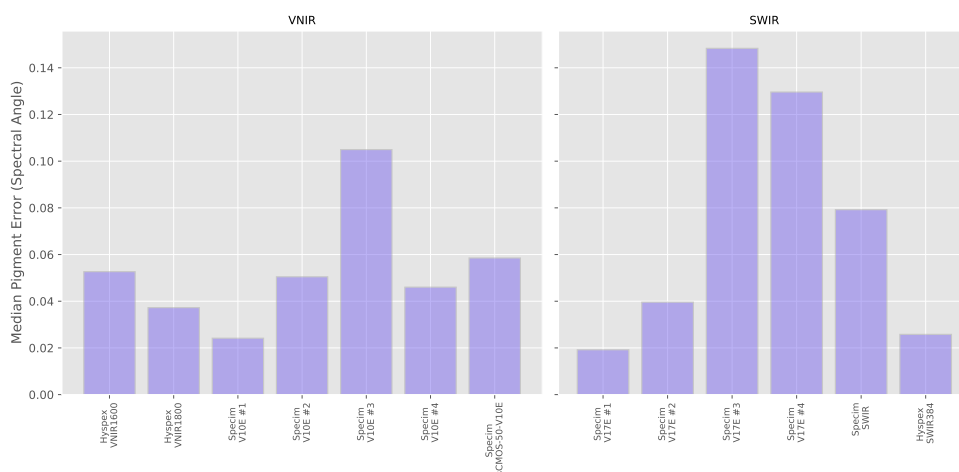


Figure 11. Median of the Spectral Angle errors over the ensemble of pigments and Gypsum ground layer for each hyperspectral system.

3.1.4. Spectral Feature Detection

The top row of Figure 12 shows the spectra from the different VNIR systems for two pigments, carmine and vermilion, which have a similar red color. The samples used on the panel are in their pure form and can be easily differentiated visually in Figure 2. However, when used in a painting, such pigments will often be mixed with a white pigment, such as lead white, in order to obtain the desired shade of red. It is often, therefore, difficult to identify a pigment from its color alone and the use of distinguishing spectral features can be an important tool in pigment identification. As can be seen from their spectra, these two pigments are indeed distinguishable from the shapes of their spectra with vermilion characterized by a very sharp upward slope between around 575–625 nm, whereas carmine possesses a less steep upward slope between 600–725 nm.

A common way of studying the kind of spectral shapes seen in carmine and vermilion is through the use of derivative analysis, which can be a useful tool for the identification of pigments [26] as well as for material identification more generally [27].

The derivatives for the thickly applied patches of carmine and vermilion from each of the hyperspectral systems in the VNIR region are shown beneath their respective reflectance spectra in the bottom row of Figure 12. As certain data sets had significant levels of noise, the derivatives have been smoothed using a moving average and the derivatives normalized in order to make them comparable between systems. As we can see, the location and shape of the derivatives are faithfully reproduced by all systems.

Azurite also possesses a very characteristic spectral shape within the VNIR region (Figure 13) with a strong peak in the blue region of the spectrum at around 480 nm which provides the pigment's dominant color, a minima at around 645 nm and a rapid increase in reflectance after around 850 nm. Between this and the minima, there is a subtle inflection in the spectra, which gives a characteristic derivative signature. As we can see at the bottom of Figure 13, the derivative (again after application of a moving average and normalization) is accurately reproduced by all the systems, though considerable levels of noise is visible in one of the results.

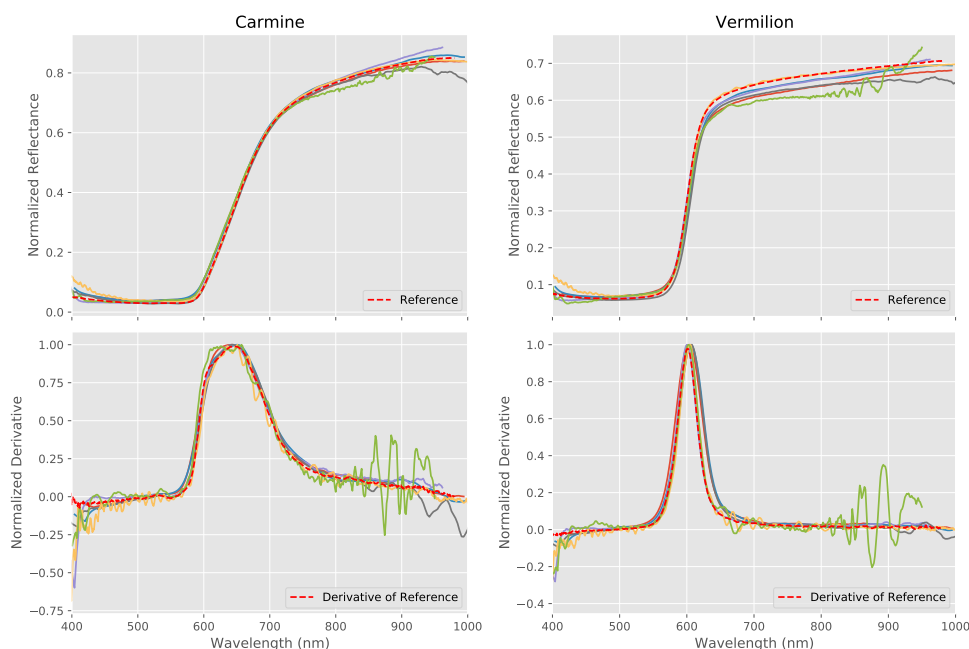


Figure 12. Reflectance spectra measured by each VNIR hyperspectral system along with the FORS reference for the two red pigments, Carmine (top left) and Vermilion (top right). The calculated derivatives of the respective pigment spectra from each system are shown in the bottom row. All systems were able to faithfully reproduce the reflectance spectra as well as reproduce distinct and accurate derivatives.

To evaluate this quantitatively, the wavelength location of the peak in the blue region of the spectra and the local minima at around 645 nm were calculated for each system as well as for the FORS reference data by identifying the wavelength at which the derivative crosses zero on the Y-axis. Piece-wise linear interpolation was used to find an accurate location for the zero crossing. The results for Azurite are given in Table 2 and show that the majority of the systems were able to accurately identify the wavelength of the dominant blue peak with 4 of the systems accurate to within just 2 nm. The local minimum in the reflectance spectra proved more difficult to find accurately due to it being much broader and less well-defined. The sharp minimum in the derivative was harder to identify precisely due to noise in the derivative of the spectra, but was nonetheless correctly identified within an average of 5.2 nm.

Within the SWIR region, Azurite also has a characteristic but subtle spectral feature at around 1500 nm [26]. This feature is not easy to distinguish in Figure 6 as it is partially transparent in the infra-red and its spectra is, therefore, modulated by the Gypsum ground layer. This makes it an interesting test-case for the hyperspectral imaging systems and the “Azurite” column in Table 3 shows the error in the position of the calculated location of this feature for each system with respect to that found in the FORS reference. As we can see, despite the small size of this feature, the systems were all able to accurately locate the minima to within 3–4 nm, which is within the spectral resolution of the cameras.

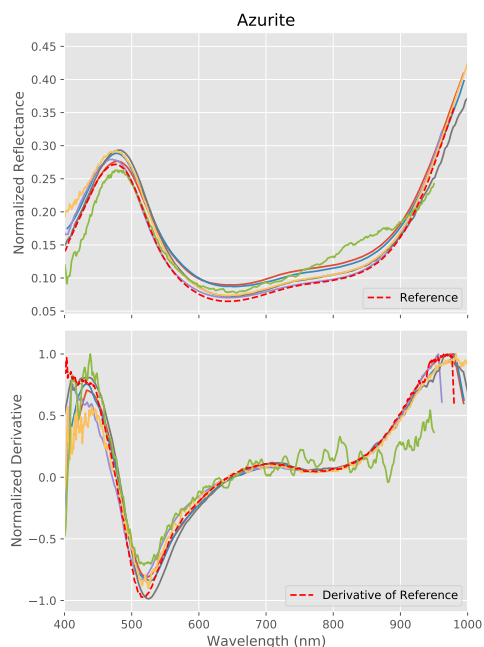


Figure 13. Reflectance spectra measured by each VNIR hyperspectral system (colored lines) for Azurite together with the FORS reference (**top**) and their derivatives (**bottom**). The characteristic form of the derivative is accurately reproduced, though significant levels of noise is visible in one of the results.

Table 2. Errors with respect to the FORS reference in the location of the wavelengths identified by the VNIR cameras for the blue peak, the minima in the spectral reflectance and the minima of the derivative of the pigment Azurite. (The FORS reference values for these wavelengths were 476.3, 644.1 and 518.7, respectively). The blue peak was accurately identified by the majority of the systems, while the reflectance minimum proved more difficult for the systems to accurately locate. The minimum in the derivative shows the most error, but was nonetheless correctly identified with an average error of 5.2 nm.

System	Wavelength Location Error (nm)		
	Blue Peak	Reflectance Minimum	Derivative Minimum
Hypex VNIR1600	04	38	−42
Hypex VNIR1800	13	90	−13
Specim V10E#1	−79	−09	−51
Specim V10E#2	60	12	104
Specim V10E#3	14	−11	63
Specim V10E#4	14	116	−72
SpecimSCMOS – 50 – V10E	51	30	22
<i>Average Error Magnitude</i>	3.4 ± 2.7	4.4 ± 3.9	5.2 ± 2.9

A much more prominent feature can be found in the Gypsum ground layer, which possesses a strong absorption band at around 1950 nm. This can be clearly seen in Figure 6, but is beyond the spectral range of the InGaAs-based SWIR camera. Table 3 shows the errors in the wavelength location calculated for this feature for each MCT-based system and again, there is a good level of accuracy within the results with the location of the spectral feature identified with errors that are within the spectral resolution of the cameras.

Table 3. Error in the wavelength location with respect to the FORS reference determined by each system for two characteristic spectral features in the SWIR region for the pigments Azurite and Gypsum. For Azurite, a local minimum at 1495.5 nm and for Gypsum a strong absorption feature at 1942.5 nm (Note that the wavelength range of the InGaAs-based camera is limited to a maximum of 1661 nm and so was not able to detect the Gypsum feature).

System	Error (nm)	
	Azurite	Gypsum
Specim V17E #1	0.5	—
Specim V17E #2	−7.3	−2.7
Specim V17E #3	−1.2	4.1
Specim V17E #4	−3.3	2.8
Specim SWIR	2.7	5.5
Hypspec SWIR384	−3.7	−0.7
<i>Average</i>	3.1 ± 2.2	3.2 ± 1.6

3.2. Noise

The level of noise is an important criteria for the quality of the acquired data. The targets used in this paper have homogeneous areas which allow spectral measurements to be calculated from data averaged over several hundreds or thousands of pixels, resulting in relatively noise-free and “clean” data. However, when acquiring hyperspectral data of inhomogeneous real-world objects rather than test targets, such averaging is rarely possible and data needs to be as noise-free as possible in order to maximize its utility. Quantifying this noise is, therefore, useful and to do this we will use the most commonly used metric for measuring noise, namely the Signal to Noise Ratio (SNR).

The spectral cubes acquired from the wavelength standard were used to calculate the SNR at each spectral band with pixels were taken from a circular region with a fixed diameter of 100 pixels located at the center of the wavelength standard. For the SNR statistics to be comparable, it was necessary to use the same number of pixels from each data set. However, the resolutions of the acquired data are not identical and, thus, the collected pixels do not cover precisely the same area on the wavelength standard. Nevertheless, although not perfect, the target is sufficiently homogeneous to be able to use the data in this way to make indicative comparisons between the data.

The graph shown in Figure 14 shows how the SNR varies as a function of wavelength for each system.

Each of the spectral cubes was acquired with slightly different ranges of wavelength, as already seen in Table 1. The sensitivity of a sensor can drop dramatically at the extremes of its nominal operating range and in Figure 14 we indeed see much reduced SNR at the edges of the wavelength ranges. In order to make a meaningful quantitative comparison between the spectral cubes, a subset of the data was used which corresponds to the wavelength range common to all systems. From this subset, the SNR was calculated for each band and the max, min and average SNR over all bands were calculated. The results can be seen in Figure 15 and detailed in the Appendix B in Table A3 and we can see a wide degree of difference between the minima and maxima for each system as well as a wide spread across the systems. The minimum SNR ranges from 10.88 to 64.01 and the maximum from 70.26 to 277.88 with the average ranging from 42.77 to 203.87.

It should be noted that the SNR measures represent noise within the final processed data and not noise that is intrinsic to the system, such as detector noise etc. The differences in SNR between the systems are, therefore, not due to the hardware used, but mainly to how the acquisition was performed and the parameters used. These include factors such as the level and type of illumination, the acquisition integration time and whether noise reduction techniques such as binning or averaging have been used.

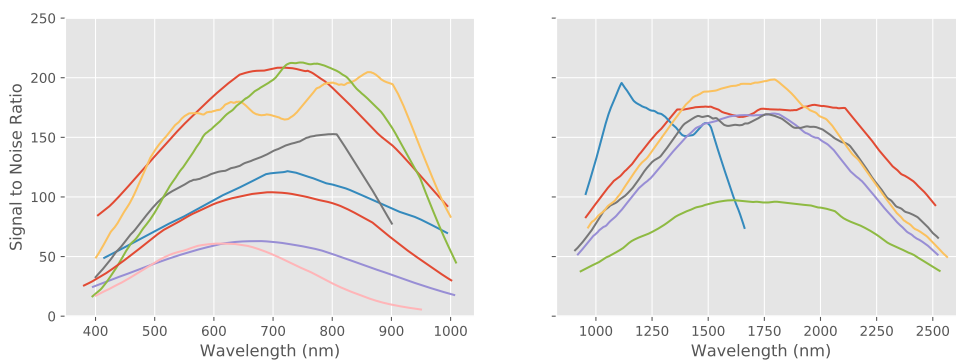


Figure 14. Signal to Noise Ratio (SNR) with respect to wavelength (smoothed) for each system (where the colored lines represent the SNR calculated for each system) showing how the SNR varies as a function of wavelength for each camera. As expected, SNR drops at the extremes of the wavelength range for each system. However, large differences in the amplitudes of the SNR can be seen between the acquired hyperspectral data.

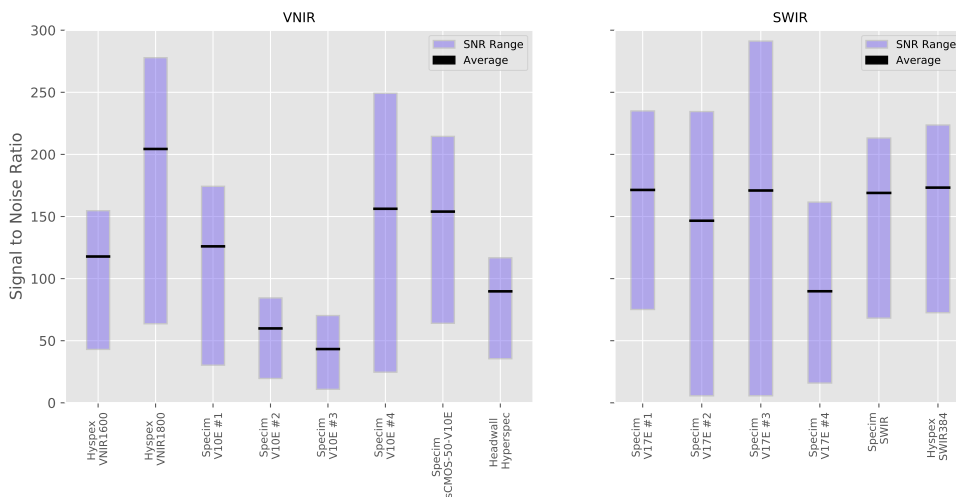


Figure 15. Minimum, maximum and average signal to noise ratios for each acquisition of the wavelength standard. All systems exhibit a wide range of SNR between their spectral bands and the average SNR across all wavelengths also varies considerably between systems.

3.3. Spatial Accuracy

3.3.1. Spatial Resolution

The hyperspectral systems tested are capable of varying levels of spatial resolution. The scans were not necessarily acquired at the maximum resolution possible for each system, but at the resolution deemed appropriate by each user given the target size, optical arrangement and the capabilities of the systems themselves. The field of view and working distances of the systems can be seen in the Appendix B Table A2 and the final spatial resolutions of the acquired data in Table 1. As we can see, there is a wide range of resolutions spanning from 3.84 to 29.2 pixels/mm with an average of 11.2 pixels/mm for the VNIR data and a narrower range of 2.26 to 3.26 pixels/mm for the MCT-based SWIR data, whose average is much lower at 3.54 pixels/mm. The single InGaAs-based SWIR system

provided data at a higher resolution of 7.92 pixels/mm, though of course with less spectral range than the MCT-based systems.

3.3.2. Geometric Distortion

In many applications, the spatial accuracy of the data can often be as important as the quality of the spectral data. Geometric distortion can be a common problem and there are many types of such distortion that can occur in hyperspectral imaging systems, which can make registration between VNIR and SWIR data sets or with other imaging modalities more difficult. Various optical distortions can exist that require correction and the complex multi-component optics often used in hyperspectral systems can result in complex non-linear distortions in the data that require a sensor model to correct. Data from pushbroom scan-based systems can also have distortions due to the movement and inaccuracy of the translation stages. Indeed a common problem for such scanning systems is for synchronization between the horizontal pixel resolution and the scan speed along the axis of movement. Incorrect calculation of the acquisition speed in the direction of scan movement can result in errors in the resulting aspect ratio of the data.

Many of these distortions require specialized equipment or targets to measure. The aspect ratio, however, is something that can be accurately and reliably calculated from the data acquired during the round-robin test and this was, moreover, the type of geometric distortion that was most visible and obvious in several of the data sets.

To assess aspect ratio accuracy, the effective resolution was measured in both the X and Y axes and the error calculated with respect to a perfect square pixel aspect ratio of 1.0. To do this, the images from the wavelength standard were used. The wavelength standard is a perfect circle and so a Hough transform [28] was used to automatically extract the coordinates of the circle with sub-pixel accuracy and the eccentricity (if any) of the shape calculated. The results are shown in Figure 16 and the detailed results can be found in the Appendix B in Table A3. As we can see, the errors for most systems were relatively small.

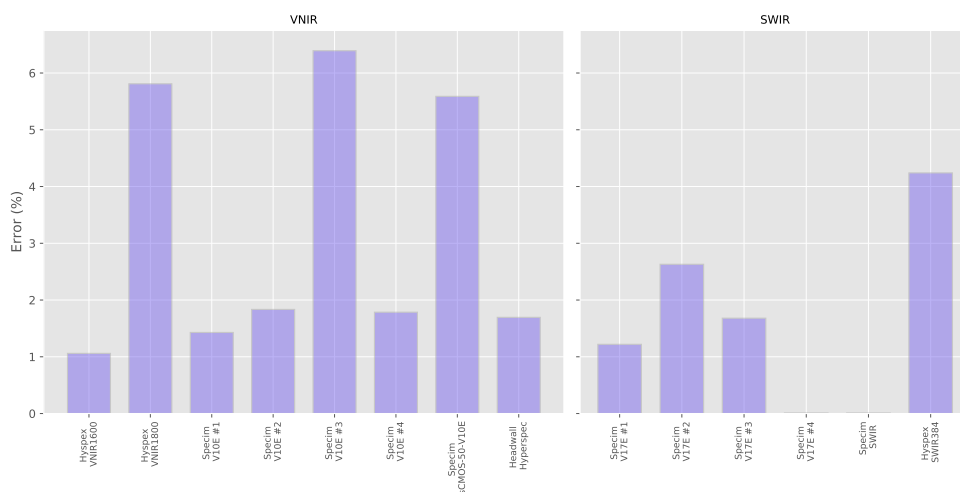


Figure 16. Aspect ratio errors expressed as a percentage for each system. The errors vary from zero for two of the systems to 6.39%.

Five of the eight VNIR systems had errors of less than 2% with the three others having errors of between 5% and 6%. For the SWIR systems, two systems had perfect aspect ratios, while the other four had errors ranging from 1.22% to 4.24%. These lower levels of error can be explained by the fact that the SWIR data was acquired at much lower resolution.

3.4. Colorimetric Analysis

Accurate colorimetry that is illuminant-independent and free of metamerism has always been an important goal for spectral imaging in fields such as cultural heritage [29,30]. It is, therefore, useful to also look at the colorimetric performance of each system. The VNIR system data was, therefore, used to calculate color coordinates in the CIELAB color space under a D65 standard daylight illuminant for each of our targets.

This was done by first determining the CIE XYZ tristimulus values through multiplying the spectral reflectance curves by the CIE standard 2° observer color matching functions and by the spectral energy of the illuminant (D65 standard daylight). The values were summed and normalized for each of X, Y and Z (Equation (A2) in the Appendix A where E is the spectral energy and R is the reflectance). CIELAB coordinates were then calculated from the CIE XYZ coordinates.

The colorimetric coordinates were calculated for each target for each hyperspectral system as well as for the reference spectra. The color differences between the measured spectra and the reference were determined using CIE ΔE_{2000} [31].

Although the primary goal of the use of a wavelength standard is the assessment of spectral performance, an evaluation of the accuracy of color rendering can provide complementary information. The CIELAB color coordinates for the wavelength standard are shown in Table 4 along with the error with respect to the reference.

Table 4. CIELAB Color Coordinates and Error for Wavelength Standard Ordered by Error.

System	L	a	b	ΔE_{2000}
Reference	92.79	2.93	6.09	0.00
Hypspec VNIR1600	93.04	0.20	11.43	5.72
Hypspec VNIR1800	93.01	1.77	8.18	2.45
Specim V10E #1	93.64	2.66	6.55	0.36
Specim V10E #2	93.79	3.16	5.04	1.15
Specim V10E #3	93.78	2.10	5.03	1.41
Specim V10E #4	92.61	2.70	6.98	5.72
Specim sCMOS-50-V10E	93.13	3.55	7.34	1.15
Headwall Hyperspec	93.82	3.45	6.00	0.93

The colorimetric errors range from 0.36 to 5.72 ΔE_{2000} (with an average of 2.41) and the L^* lightness values all closely correspond to the reference, being all within 1% of the reference. However, the a^* and b^* chromacity values vary far more between the data sets with there being sometimes significant errors in the b^* values.

CIELAB color coordinates were also calculated for each of the different pigments from the hyperspectral data as well as for the reference spectra. The colorimetric error for each of the thickly applied pigments can be seen in Figure 17 and the average errors for each system over the ensemble of pigments are given in Table 5.

Table 5. CIE ΔE_{2000} errors averaged over all six pigments (both thin and thick layers) plus the Gypsum ground for each hyperspectral system.

System	Average ΔE_{2000}	Std Deviation
Hypspec VNIR1600	3.09	1.42
Hypspec VNIR1800	2.48	0.93
Specim V10E #1	1.68	0.77
Specim V10E #2	2.37	0.92
Specim V10E #3	3.90	2.11
Specim V10E #4	4.46	2.34
Specim sCMOS-50-V10E	2.74	1.04

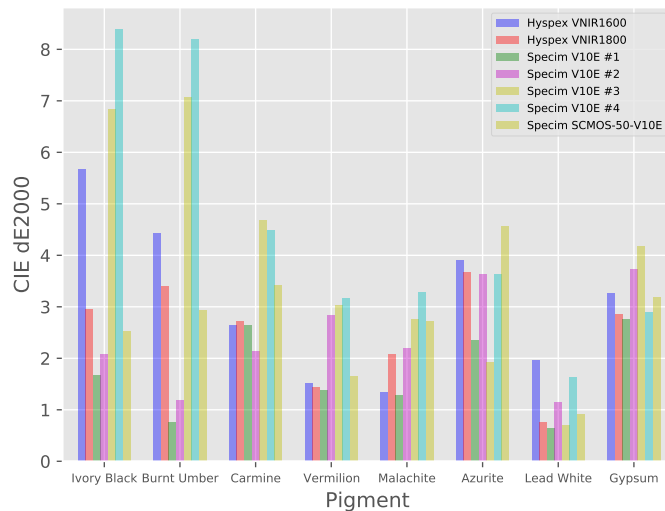


Figure 17. CIE ΔE_{2000} errors for panel pigments (only the thickly applied pigments are shown for clarity).

4. Discussion

The spectral data obtained in this study for the two test targets from the various systems and institutions show a good degree of coherence and accuracy. As can be seen in Figures 4 and 6, they were all capable of distinguishing the various spectral features in the wavelength standard and were able to accurately reproduce the shapes and essential features of the pigment spectra. Using the acquired data, it was also possible to carry out reliable spectroscopic analysis and all systems were able to accurately identify sometimes subtle spectral features in both the reflectance spectra and their derivatives.

There was, nevertheless, a certain amount of variation between the results from the various hyperspectral imaging systems that were used in the study and it is useful to look in more detail at where these occur and how these errors can be reduced.

As we saw in several of the analyses, it is interesting to note that the SWIR systems had higher average levels of error than the VNIR system in the majority of the analyses and this is noticeable in Figures 4 and 6, particularly at the higher end of their wavelength ranges. Particular care, therefore, clearly needs to be taken when acquiring and calibrating SWIR data.

Looking at the systems individually, no single data set was perfectly accurate, even data that had been acquired directly by the camera manufacturers. Indeed it is interesting to note that there was no consistent ranking between the various data sets when using the various measures of accuracy. In other words, systems that had the best accuracy in one category did not necessarily have the best in another.

If we consider the various measures of accuracy used, it is difficult to compare or rank the absolute magnitudes as it is the intended end-use of the data which will determine which types of accuracy are more important for a particular task. For example, for classification, accurate spectral distance measures will be critical, whereas for spectroscopy applications, spectral alignment and the ability to detect spectral features will be more important.

However, the level of variation between each of the systems for a particular measure can provide useful information and we see that the levels of SNR found within the final processed data have by far the most variability across the systems with SNR differing by up to a factor of 6 for the VNIR and 12 for the SWIR systems. The SNR measured here is the SNR found in the final processed data and is something that is affected far more by the acquisition parameters, setup and calibration than by

the hardware used. This can be clearly seen in our results where there are large differences in the SNR found in the processed data even when using identical hardware. A number of straightforward improvements can be made to improve the SNR within a system [12], including tuning the acquisition parameters to the spectral characteristics of the target, the use of binning, the use of averaging or the use of equalization filters.

Other errors, such as spectral alignment errors, are difficult to correct without specialist equipment and accurate measurement, but errors in, for example, basic spatial accuracy can be improved relatively straightforwardly in many cases. For example, the aspect ratio measure used here to assess geometry only provides a very partial view of the spatial accuracy of a system, but this error is often the most visible and can encompass several types of distortion. These include optical distortions as well as errors in translation stage speed or accuracy, all of which can be corrected for.

5. Conclusions

A round-robin test was conducted using 14 different pushbroom hyperspectral imaging systems from 8 different institutions encompassing both the VNIR and SWIR spectral ranges. A single set of targets was used and hyperspectral data acquired using each of the imaging systems under the operating procedures, conditions and acquisition parameters that were standard for each institution and for each system. The data from each system was calibrated using the post-processing workflow that was also standard for that institution and system in order to obtain final calibrated spectral data. The aim was to compare these final data sets in order to better ascertain the reliability and comparability of hyperspectral imaging under “real-world” operating conditions where the equipment is often, to some extent, considered to be a black box. Although the study was focused on the acquisition of hyperspectral data for cultural heritage, the results will be of interest to the wider community performing laboratory-based hyperspectral imaging.

The comparison of the data shows that the majority of the systems were able to provide data that was usable and accurate enough for general use. There were, however, significant levels of variability in the data. The spectral, geometric and colorimetric accuracies were highly variable. Several of the systems measured had spectral mis-alignment errors and residual errors were common for all systems. The largest amount of variation between the systems was in terms of the levels of noise found in the processed data with SNR differing by up to a factor of 6 for the VNIR and 12 for the SWIR systems.

This variability can have an important impact on the use of such data for tasks such as material classification or for the detection of change over time. The results, moreover, highlight the importance of defining better and standardized workflows for hyperspectral imaging. By identifying the nature of these differences and by quantifying their magnitudes and variability, significant improvements can be made to the acquisition and processing pipelines. Measures, such as SNR, are highly dependent on the acquisition parameters and significant improvements are possible by the setting of more optimal parameters.

Furthermore, variability of this sort can have an impact on the behavior of metrics such as spectral distance measures. These metrics form a key component of tasks such as classification, where the choice of threshold parameters can make an important difference to the results obtained. By gaining a better understanding of the characteristics and variability of the underlying data, it will be possible to make better use of hyperspectral data for classification, change detection and other applications.

Supplementary Materials: The spectral data used in this paper is available from <https://github.com/ruven/RRT>. The site includes processed spectral data of the references and the acquired system spectra from each system for both test targets. The site also features interactive graphs allowing exploration of the references and acquired spectra in more detail.

Author Contributions: Scientific methodology, data analysis and writing: R.P.; Round-robin test data curation: R.P., S.G. and M.P.; review and editing: J.Y.H., M.P. and S.G.; Round-robin test administration: M.P. and S.G. All authors have read and agreed to the published version of the manuscript.

Funding: The round-robin test was carried out as part of the European Union COST Action TD1201 *Colour and Space in Cultural Heritage* (COSCH). The APC for this article was funded by NTNU.

Acknowledgments: The authors would like to thank the participants in the spectral imaging round-robin test and especially those who contributed the hyperspectral data used in the paper. In particular, we gratefully acknowledge the data supplied by the CRCC (Centre de Recherche sur la Conservation des Collections), University College London, the Bodleian Library, the National Museum in Krakow, Norsk Elektro Optikk AS and Specim Ltd. The authors would also like to thank Elena Prandi and Marina Ginanni for the preparation of the Pigment Panel.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

VNIR	Visible Near Infra-Red
SWIR	Short Wave Infra-Red
BRDF	Bidirectional Reflectance Distribution Function
MCT	Mercury Cadmium Telluride
InGaAs	Indium Gallium Arsenide
CCD	Charged Couple Device
CMOS	Complementary Metal Oxide Semiconductor
FORS	Fiber Optic Reflectance Spectroscopy
UV	Ultra Violet
SAM	Spectral Angle Mapper
SNR	Signal to Noise Ratio

Appendix A. Equations

$$SAM(\mathbf{x}, \mathbf{y}) = \arccos \left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|_2 \|\mathbf{y}\|_2} \right) \quad (\text{A1})$$

$$[X, Y, Z] = 100 * \frac{\int E_\lambda R_\lambda [\bar{x}, \bar{y}, \bar{z}]_\lambda d\lambda}{\int E_\lambda \bar{y}_\lambda d\lambda} \quad (\text{A2})$$

Appendix B. Tables

Table A1. Chemical composition of the pigments used in the Pigment Panel target.

Pigment	Composition
Burnt Umber	Manganese Hydroxide and Oxide with Iron Hydroxide
Carmine	Carminic acid lake
Vermilion	Mercury(II) Sulphide: HgS
Malachite	Basic copper carbonate: $\text{CuCO}_3 \cdot \text{Cu}(\text{OH})_2$
Azurite	Hydrated copper carbonate: $2\text{CuCO}_3 \cdot \text{Cu}(\text{OH})_2$
Lead White	Basic lead carbonate: $2\text{PbCO}_3 \cdot \text{Pb}(\text{OH})_2$
Ivory Black	Phosphorus, calcium and magnesium with carbonate

Table A2. Hyperspectral cameras and the hardware and acquisition parameters used for data acquisition during the round-robin test. Note that information on the entry slit width is not available for the Hypspec systems. Also acquisition integration times are not recorded by the majority of the acquisition software and are, therefore, not provided here.

System	Detector Type	Aperture	Field of View (°)	Working Distance (cm)	Entrance Slit (μm)	Spectral Binning	Frame Averaging	Equalization Filter
VNIR								
Hypspec VNIR1600	CCD	<i>f</i> /2.5	17	30	–	2	8	True
Hypspec VNIR1800	CCD	<i>f</i> /2.5	17	100	–	2	20	True
Specim V10E #1	CCD	<i>f</i> /1.7	38	50	18	2	1	True
Specim V10E #2	CCD	<i>f</i> /1.7	38	50	30	4	1	False
Specim V10E #3	CCD	<i>f</i> /1.7	38	50	30	1	1	False
Specim V10E #4	CCD	<i>f</i> /1.7	38	50	30	2	1	False
Specim sCMOS	sCMOS	<i>f</i> /2.5	34	50	30	2	1	False
Headwall Hyperspec	sCMOS	<i>f</i> /2.0	34	40	25	1	1	False
SWIR								
Specim V17E #1	InGaAs	<i>f</i> /2.0	16	50	18	2	1	False
Specim V17E #2	MCT	<i>f</i> /2.0	16	50	30	1	1	False
Specim V17E #3	MCT	<i>f</i> /2.0	16	50	30	1	1	False
Specim V17E #4	MCT	<i>f</i> /2.0	16	50	30	1	1	False
Specim SWIR	MCT	<i>f</i> /2.0	16	50	30	1	1	False
Hypspec SWIR384	MCT	<i>f</i> /2.0	17	30	–	1	20	False

Table A3. Analysis data for both the VNIR and SWIR systems for the spectral mis-alignment, signal to noise statistics (mean, minimum and maximum), aspect ratio errors and Spectral Angle errors for the wavelength standard and the pigment panel (the median over the ensemble of the pigments) respectively. (Note that one of the VNIR systems did not acquire pigment panel data and, thus, the Spectral Angle error for that system and target is not available).

System	Spectral	SNR			Aspect Ratio	SAM Error	
	Mis-Alignment nm	Mean	Min	Max	Error %	Wavelength Standard	Pigment Panel
VNIR							
Hyspex VNIR1600	0.0	117.29	42.99	154.7	1.06	0.00534	0.0527
Hyspex VNIR1800	−0.1	203.87	63.62	277.88	5.81	0.0119	0.0373
Specim V10E #1	−1.3	125.46	30.31	174.32	1.43	0.0206	0.0242
Specim V10E #2	2.6	59.39	19.59	84.44	1.83	0.0307	0.0505
Specim V10E #3	1.1	155.72	24.67	249.26	1.78	0.0242	0.0460
Specim V10E #4	1.0	42.77	10.88	70.26	6.39	0.0221	0.1050
Specim sCMOS	1.3	151.46	64.01	214.5	5.59	0.0370	0.0585
Headwall Hyperspec	0.0	89.24	35.47	116.75	1.69	0.0279	—
Average Magnitudes	0.93 ± 0.83	118.40 ± 50.04	36.44 ± 18.22	167.76 ± 70.90	3.20 ± 2.14	0.0211 ± 0.0117	0.0534 ± 0.0234
SWIR							
Specim V17E #1	0.1	170.89	75.12	235.01	1.22	0.000301	0.0192
Specim V17E #2	0.3	170.44	5.63	291.26	2.63	0.0300	0.0396
Specim V17E #3	0.6	146.11	5.62	234.5	1.68	1.36	0.1483
Specim V17E #4	−5.5	89.32	15.91	161.61	0.0	1.66	0.1296
Specim SWIR	2.0	168.48	68.19	213.21	0.0	0.342	0.0792
Hyspex SWIR384	2.0	172.77	72.52	223.68	4.24	0.0114	0.0258
Average Magnitudes	1.75 ± 1.84	153.00 ± 29.88	40.50 ± 31.70	226.54 ± 38.13	1.63 ± 1.49	0.567 ± 0.682	0.0736 ± 0.0503

References

- Goetz, A.F. Three Decades of Hyperspectral Remote Sensing of the Earth: A Personal View. *Remote Sens. Environ.* **2009**, *113*, S5–S16. [\[CrossRef\]](#)
- Wu, D.; Sun, D.W. Advanced Applications of Hyperspectral Imaging Technology for Food Quality and Safety Analysis and Assessment: A Review—Part I: Fundamentals. *Innov. Food Sci. Emerg. Technol.* **2013**, *19*, 1–14. [\[CrossRef\]](#)
- Hege, E.K.; O’Connell, D.; Johnson, W.; Basti, S.; Dereniak, E.L. Hyperspectral Imaging for Astronomy and Space Surveillance. In Proceedings of the Imaging Spectrometry IX. International Society for Optics and Photonics, San Diego, CA, USA, 6–7 July 2003; pp. 380–391. [\[CrossRef\]](#)
- Kruse, F.A.; Boardman, J.W.; Huntington, J.F. Comparison of Airborne Hyperspectral Data and EO-1 Hyperion for Mineral Mapping. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1388–1400. [\[CrossRef\]](#)
- Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral Vegetation Indices and Novel Algorithms for Predicting Green LAI of Crop Canopies: Modeling and Validation in the Context of Precision Agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [\[CrossRef\]](#)
- Lu, G.; Fei, B. Medical Hyperspectral Imaging: A Review. *J. Biomed. Opt.* **2014**, *19*, 010901. [\[CrossRef\]](#) [\[PubMed\]](#)
- Casini, A.; Bacci, M.; Cucci, C.; Lotti, F.; Porcinai, S.; Piccolo, M.; Radicati, B.; Poggesi, M.; Stefani, L. Fiber Optic Reflectance Spectroscopy and Hyper-Spectral Image Spectroscopy: Two Integrated Techniques for the Study of the Madonna Dei Fusi. In Proceedings of the Optical Methods for Arts and Archaeology, Munich, Germany, 13–14 June 2005; p. 58570M-8.
- Bentkowska-Kafel, A.; MacDonald, L. *Digital Techniques for Documenting and Preserving Cultural Heritage*; Arc Humanities Press: Leeds, UK, 2018.
- Early, E.A.; Barnes, P.Y.; Johnson, B.C.; Butler, J.J.; Bruegge, C.J.; Biggar, S.F.; Spyak, P.R.; Pavlov, M.M. Bidirectional Reflectance Round-Robin in Support of the Earth Observing System Program. *J. Atmos. Ocean. Technol.* **2000**, *17*, 1077–1091. [\[CrossRef\]](#)
- Hooker, S.B.; Firestone, E.R.; McLean, S.; Sherman, J.; Small, M.; Lazin, G.; Zibordi, G.; Brown, J.W.; McClain, C.R. *The Seventh SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-7)*; Technical report for NASA; NASA: Pasadena, CA, USA, 2002.
- Jung, A.; Götze, C.; Glässer, C. Overview of Experimental Setups in Spectroscopic Laboratory Measurements—the SpecTour Project. *Photogramm. Fernerkund. Geoinf.* **2012**, *2012*, 433–442. [\[CrossRef\]](#)
- Pillay, R.; Hardeberg, J.Y.; George, S. Hyperspectral Imaging of Art: Acquisition and Calibration Workflows. *J. Am. Inst. Conserv.* **2019**, *58*, 3–15. [\[CrossRef\]](#)
- George, S.; Hardeberg, J.; Linhares, J.; MacDonald, L.; Montagner, C.; Nascimento, S.; Piccolo, M.; Pillay, R.; Vitorino, T.; Webb, E.K. A Study of Spectral Imaging Acquisition and Processing for Cultural Heritage. In *Digital Techniques for Documenting and Preserving Cultural Heritage*; Arc Humanities Press: Leeds, UK, 2018; pp. 141–158.
- George, S.; Ciortan, I.; Hardeberg, J.Y. Evaluation of Hyperspectral Imaging Systems for Cultural Heritage Applications Based on a Round Robin Test. In Proceedings of the International Colour Association (AIC), Tokyo, Japan, 19–22 May 2015.
- MacDonald, L.W.; Vitorino, T.; Piccolo, M.; Pillay, R.; Obarzanowski, M.; Sobczyk, J.; Nascimento, S.; Linhares, J. Assessment of Multispectral and Hyperspectral Imaging Systems for Digitisation of a Russian Icon. *Herit. Sci.* **2017**, *5*, 41. [\[CrossRef\]](#)
- Cennini, C.; Thompson, D.V. *Il Libro Dell’Arte*; Yale University Press: New Haven, CT, USA, 1936.
- Vitorino, T. *Analysis of the Round Robin Test Data Acquired at IFAC-CNR*; STSM Report for IFAC-CNR; IFAC-CNR: Florence, Italy, 2014.
- Maitre, H.; Schmitt, F.; Cretiez, J.P.; Wu, Y.; Hardeberg, J.Y. Spectrophotometric Image Analysis of Fine Art Paintings. *Color Imaging Conf.* **1996**, *1996*, 50–53.
- Rabiner, L.R.; Gold, B. *Theory and Application of Digital Signal Processing*; Prentice-Hall: New York, NY, USA, 1975.
- Plaza, A.; Martín, G.; Plaza, J.; Zortea, M.; Sánchez, S. Recent Developments in Endmember Extraction and Spectral Unmixing. In *Optical Remote Sensing*; Prasad, S., Bruce, L.M., Chanussot, J., Eds.; Springer: Heidelberg, Germany, 2011; pp. 235–267.

21. Deborah, H.; Richard, N.; Hardeberg, J. A Comprehensive Evaluation of Spectral Distance Functions and Metrics for Hyperspectral Image Processing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, 1–11. [[CrossRef](#)]
22. Kruse, F.A.; Lefkoff, A.B.; Boardman, J.W.; Heidebrecht, K.B.; Shapiro, A.T.; Barloon, P.J.; Goetz, A.F.H. Airborne Imaging Spectrometry The Spectral Image Processing System (SIPS)—Interactive Visualization and Analysis of Imaging Spectrometer Data. *Remote Sens. Environ.* **1993**, *44*, 145–163. [[CrossRef](#)]
23. Chang, C.I. An Information-Theoretic Approach to Spectral Variability, Similarity, and Discrimination for Hyperspectral Image Analysis. *IEEE Trans. Inf. Theory* **2000**, *46*, 1927–1932. [[CrossRef](#)]
24. De Carvalho, O.A.; Meneses, P.R. Spectral Correlation Mapper (SCM): An Improvement on the Spectral Angle Mapper (SAM). In Proceedings of the Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18, Pasadena, CA, USA, 23–25 February 2000.
25. Angelopoulou, E.; Lee, S.W.; Bajcsy, R. Spectral Gradient: A Material Descriptor Invariant to Geometry and Incident Illumination. In Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, 20–27 September 1999; pp. 861–867. [[CrossRef](#)]
26. Bacci, M.; Baronti, S.; Casini, A.; Lotti, F.; Picollo, M.; Casazza, O. Non-Destructive Spectroscopic Investigations on Paintings Using Optical Fibers. *Mrs Online Proc. Libr. Arch.* **1992**, 267. [[CrossRef](#)]
27. Demetriades-Shah, T.H.; Steven, M.D.; Clark, J.A. High Resolution Derivative Spectra in Remote Sensing. *Remote. Sens. Environ.* **1990**, *33*, 55–64. [[CrossRef](#)]
28. Duda, R.O.; Hart, P.E. *Use of the Hough Transformation to Detect Lines and Curves in Pictures*; Technical Report SRI-TN-36: Menlo Park, CA, USA, 1971.
29. Martinez, K.; Cupitt, J.; Saunders, D.R. High-Resolution Colorimetric Imaging of Paintings. *Proc. Spie* **1993**, *1901*, 25–36. [[CrossRef](#)]
30. Berns, R.S. The Science of Digitizing Paintings for Color-Accurate Image Archives: A Review. *J. Imaging Sci.* **2001**, *45*, 305–325.
31. Luo, M.R.; Cui, G.; Rigg, B. The Development of the CIE 2000 Colour-Difference Formula: CIEDE2000. *Color Res. Appl.* **2001**, *26*, 340–350. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Studying That Smile: A Tutorial on Multispectral Imaging of Paintings

Paper 3

This chapter is a reprint of the publication:

Ribés, A., **Pillay, R.**, Schmitt, F. & Lahanier, C. “Studying That Smile: A Tutorial on Multispectral Imaging of Paintings”, *IEEE Signal Processing Magazine*, 25, 4, 14-26 (2008)

©2009 IEEE. Reprinted with permission.

This paper is not included in NTNU Open due to copyright

Hyperspectral Imaging of Art: Acquisition and Calibration Workflows

Paper 4

This chapter is a reprint of the publication:

Pillay, R., Hardeberg, J. Y. & George, S. "Hyperspectral Imaging of Art: Acquisition and Calibration Workflows". *Journal of the American Institute for Conservation* 58, 3-15 (2019).

Hyperspectral Calibration of Art: Acquisition and Calibration Workflows

Ruven Pillay^{1,*}, Jon Y Hardeberg¹ and Sony George¹

¹NTNU – Norwegian University of Science and Technology, Department of Computer Science, Gjøvik, Norway and ²C2RMF – Centre de Restauration et Recherche des Musées de France, Palais du Louvre, Paris, France

*Corresponding author: ruven.pillay@culture.gouv.fr

Abstract

Hyperspectral imaging has become an increasingly used tool in the analysis of works of art. However, the quality of the acquired data and the processing of that data to produce accurate and reproducible spectral image cubes can be a challenge to many cultural heritage users. The calibration of data that is both spectrally and spatially accurate is an essential step in order to obtain useful and relevant results from hyperspectral imaging. Data that is too noisy or inaccurate will produce sub-optimal results when used for pigment mapping, the detection of hidden features, change detection or for quantitative spectral documentation. To help address this, therefore, we will examine the specific acquisition and calibration workflows necessary for works of art. These workflows includes the key parameters that must be addressed during acquisition and the essential steps and issues at each of the stages required during post-processing in order to fully calibrate hyperspectral data. In addition we will look in detail at the key issues that affect data quality and propose practical solutions that can make significant differences to overall hyperspectral image quality.

Key words: hyperspectral imaging; radiometric calibration; geometric calibration; registration; spectralon; works of art; equalization filter; open-source

1 Introduction

1.1 Hyperspectral Imaging for Cultural Heritage

Spectral reflectance imaging can be a valuable tool for analyzing and documenting works of art due to its ability to simultaneously capture both accurate spectral and spatial data. The introduction of spectral reflectance imaging technologies to the cultural heritage field occurred initially through multispectral imaging systems able to capture a handful of spectral bands (Saunders and Cupitt 1993; Martinez et al. 2002; Lahanier et al. 2002). However, the development and commercial availability of push-broom hyperspectral cameras able to capture hundreds of narrow contiguous wavelengths soon became a practical possibility and a small number of institutions were able to develop bespoke equipment (Bacci et al. 2005; Delaney et al. 2010).

A number of system designs have been implemented for push-broom hyperspectral imaging of art. Standard laboratory systems for industrial use are usually horizontally mounted with the target lying flat on a horizontal surface. This set-up works well for manuscripts (figure 1(a)), books or drawings. However, for easel paintings, a vertical alignment is better suited, which therefore requires a custom-made set-up, such as that in figure 1(b). Various push-broom based hyperspectral imaging systems have been put in place for cultural heritage to analyze paintings (Bacci et al. 2005; Delaney et al. 2009), codices (Snijders et al. 2016; Pottier 2017), wall paintings (Liang et al. 2014) and many others. In almost all of these systems, the push-broom hyperspectral camera moves linearly along one or more axes, while the painting remains mounted on a static easel. It is also possible, however, to mount the hyperspectral camera on a motorized rotating tripod (George and Hardeberg 2016) or use a rotating scan mirror in front of



(a) Horizontally-mounted push-broom hyperspectral system with motorized Y-axis scanning a manuscript.



(b) Vertically-mounted push-broom hyperspectral system with motorized XY axes scanning a painting.

Figure 1. Hyperspectral imaging equipment used in the cultural heritage field.

the hyperspectral camera to enable snapshot imaging of the work of art (Delaney et al. 2010).

Although the use of such equipment within cultural heritage institutions remains relatively rare, an increasing number of studies have made use of such equipment and hyperspectral imaging has successfully been used for a variety of applications including the revealing of otherwise hidden features and areas of lost material (Bacci et al. 2005), for materials identification and pigment mapping (Delaney et al. 2014; Deborah et al. 2014) where methods and algorithms from the field of remote sensing have been applied as well as for documentation and visualization (Pillay 2013).

Nevertheless, a number of important challenges exist, which will need to be addressed before hyperspectral imaging can become more widely used within the cultural heritage sector. The sheer volumes of hyperspectral data that can be produced from an acquisition requires high-end processing, data storage capacity and the expertise to go with it. Such facilities are rarely available in museum environments, making this a major challenge. In addition, existing methodologies, algorithms and software for calibration, spectral mapping and unmixing methods etc. have been developed and optimized mainly for remote sensing applications. These will require careful customization and tuning to the needs of cultural heritage if the full potential of hyperspectral imaging is to be realized.

1.2 Calibration Workflow

The accurate calibration of spectral data is an essential step in order to obtain useful and relevant results from hyperspectral imaging. Colorimetric or spectral analysis of the data is highly sensitive to incorrect or sub-optimal calibration. The issue of spectral calibration has been well studied in the field of remote sensing (Woodward et al. 2009; Eismann 2012), environmental imaging (Aspinall et al. 2002), plant biology (Behmann et al. 2015) and for astronomy (Fiacchione et al. 2009). For laboratory-based systems, there have been several studies carried out, in particular (Polder et al. 2003; Geladi et al. 2004; Burger and Geladi 2005,

2006; Boldrini et al. 2012) and more recently (Khan et al. 2018). However, none of these have addressed the specific requirements for the spectral imaging of works of art such as paintings or manuscripts.

In the following sections, we will, therefore, look at the issues arising from the imaging of works of art where the goal is to obtain high-resolution spectral as well as spatial data. We will consider the overall acquisition and calibration pipeline and examine ways to improve the accuracy and precision of the resulting data.

We will draw, in particular, on the results obtained from a round-robin test (RRT) of spectral imaging equipment (George et al. 2018), which involved a number of institutions and a variety of different hyperspectral systems and set-ups. The RRT provided an important insight into how such equipment is used in real-world settings, the kind of data that is typically obtained and the specific problems and errors that can arise when applying hyperspectral imaging to works of art.

The results revealed a number of issues related to system calibration, reproducibility and the lack of standardized workflows that have acted as a barrier to the greater take-up and use of hyperspectral imaging within the cultural heritage community. We will, therefore, attempt to address some of the key issues by examining in detail the acquisition and processing steps, the factors that affect performance and common errors that can arise when using hyperspectral imaging and, thereby, help to provide a more standardized workflow and guidance for users seeking to make better use of their equipment.

A full workflow for the set-up and processing of hyperspectral acquisition for 2D artworks is described in the following sections. The typical pipeline in spectral calibration involves first the removal of fixed pattern noise (dark current), pixel sensitivity normalization, denoising, illumination correction and finally spectral calibration to reflectance factor. Spatial geometric correction then needs to be carried out and finally mosaic assembly is required if the data has been acquired in multiple sections.

The resultant output from this calibration workflow is

a spectral image cube that has been fully calibrated both radiometrically and spatially and which contains quantitative reflectance data suitable for use in applications such as materials mapping.

2 Hyperspectral Acquisition Parameters

Hyperspectral imaging systems are very sensitive to acquisition parameters and operating conditions. Ensuring the best possible acquisition conditions will not only help improve the quality of the resulting data but can also greatly simplify post-processing and data handling.

2.1 Reference Targets

Reflectance Targets

In order to be able to obtain quantitative spectral data, a known reference must be used to normalize the acquired spectra to absolute reflectance factor. To do this, an acquisition must be made of a known reflectance standard or set of reflectance standards under the same operating conditions and acquisition parameters used for the acquisition of the work of art itself. These conditions include lighting levels, integration time, scan speed, acquisition geometry, camera gain settings etc.

Ideal reference targets are uniform and diffuse lambertian planar surfaces that have constant reflectance across a wide spectral range. These are typically durable and chemically inert PTFE (Polytetrafluoroethylene) targets, the most widely used of which is Spectralon®¹. Typically, a single reflectance standard is used and, indeed, the majority of the participants of the RRT acquired a single 99% reference standard image.

However, the quality of the radiometric calibration can often be improved by acquiring a set of different reference targets with a range of reflectivities. An example of such a target is shown in figure 2. Again, these must be acquired using the same parameters as those for the painting acquisition either during the same scan sequence or separately.



Figure 2. Multi-step reflectance standard with reflectances at 99%, 50%, 25% and 12.5%

For many works of art, tuning the acquisition parameters in order to acquire a good image of the 99% reference standard will result in sub-optimal results for the work of art itself. The maximum reflectance found in even a bright painting is rarely close to that of a 99% reference target and dark paintings can have maximum reflectances of 50% or less, meaning that half of the dynamic range of the hyperspectral sensor is lost. Indeed as detectors become non-linear when operated at close to saturation, users are recommended by the manufacturers to set their acquisition parameters in order to capture data below the saturation level and often at around 80% or less. Thus, in practice, a dark painting would only use 40% of the sensor's full dynamic range, significantly degrading the potential performance of the system.

By using several reference targets together, however, the integration time of the system can be set according to the requirements of the artwork itself rather than be limited by any particular reference target. The integration time of the acquisition can, therefore, be increased for darker works of art and set at an optimal level for them. In practice, this means performing a test scan over the brightest part of the work of art and increasing the integration time until the detector is at, for example, 80% saturation. In this way, the acquisition uses as much of the available dynamic range as possible, significantly improving signal-to-noise and the quality of the resulting data.

An additional advantage to using multiple reference targets is the ability to combine the measurements from all of the (unsaturated) reference targets and calculate a more accurate linear fit, especially when using low light levels or with less sensitive or noisy equipment (Burger and Geladi 2005).

If the integration time has been tuned to the work of art and only a single reference target is to be used, it should have a reflectance that is similar or less than the brightest part of the work of art and it is essential that the data from this target is not saturated.

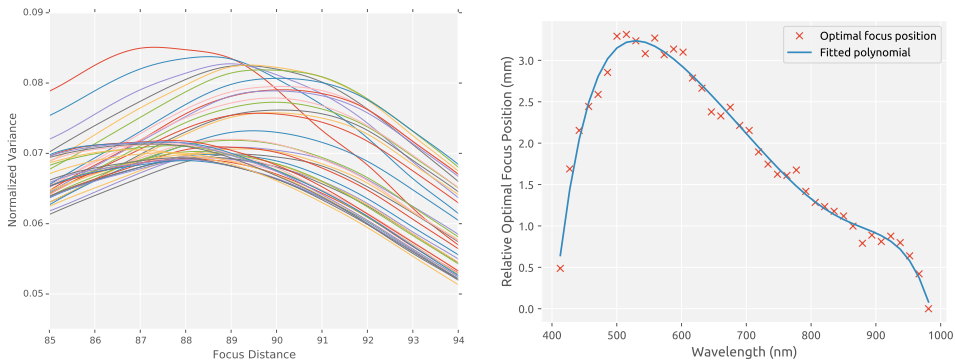
A spectral image cube of one or more reference standards is critical for calibrating to reflectance factor and it is important that this data is of high quality and, in particular, that only those not saturated in any spectral band are used for radiometric calibration.

Geometric Targets

Push-broom hyperspectral systems are essentially line-scanners that acquire one single spatial line at a time. In order to acquire a full 2D image, it is necessary to physically move the camera perpendicular to this scan line along the Y-axis. In order to avoid errors in the aspect ratio of the resulting image, it is important to correctly calculate the scan speed along this direction of motion. Some hyperspectral camera manufacturers supply acquisition software that automatically calculates this scan speed, but many do not, thereby requiring the user to input a scan speed manually. This can often be tricky as tiny errors in the calculation of the scan speed can result in large accumulated errors over extended scan areas. Indeed, incorrect aspect ratios due to incorrectly calculated scan speeds were a common error seen in the RRT.

In order to better calculate the scan speed and to help correct for any residual errors during post-processing, a geometric calibration target can be placed in the scene. Such a geometric calibration target can be something as simple

¹ manufactured by Labsphere Inc.



(a) Image sharpness with respect to focus position for each spectral band.

(b) Optimal sharpness location as a function of wavelength.

Figure 3. Focus metrics for the HySpex VNIR1600 showing how the optimal focus distance varies with respect to wavelength.

as a ruler or printed grid, which can be used to calculate the appropriate scan speed or can be used to re-scale the data during post-processing.

2.2 Focusing and Image Sharpness

The alignment of the work of art is an important initial step in most hyperspectral imaging workflows. The work of art must be at an appropriate distance to the hyperspectral camera and as perpendicular as possible to both the camera and to the direction of scanning. This alignment is important for obtaining a good sharp focus across the entire field of view and to avoid introducing geometric errors.

Hyperspectral imaging systems need to capture as much light as possible and so typically use lenses with low F -numbers and, therefore, very open apertures. This, however, results in relatively shallow depths of field. Hyperspectral imaging was originally devised for use in remote sensing where focus is obtained at infinity and depth of field is less of an issue. For close-range imaging, however, obtaining a spectral image cube that is sharp and in-focus across the whole spectral range can be difficult. This is especially true for targets such as panel paintings which can often be warped or have uneven surfaces and sub-optimal focusing was indeed an issue for several participants in the RRT, where data were often blurred.

Hyperspectral acquisition systems can be either fixed-distance focusing or can be focused through the lens. In both cases, the need for appropriate focusing is essential. The ideal focus setting generally varies with respect to wavelength. Although spectral acquisition systems are designed to minimize spectral aberration, this is difficult to eliminate entirely. Finding the ideal focus is, therefore, a compromise over the entire spectral range. In figure 3(a), we see the results from measuring the level of sharpness of an image for each captured wavelength at different distances for a visible – near infrared (VNIR) hyperspectral system with a fixed-distance lens. In this example, the measure of “sharpness” is the simple, but widely used variance (Groen et al. 1985) of the image for each individual band.

We see from figure 3(a) that the peak for each wavelength occurs at a different distance of the camera to the target. If we plot the locations of the peaks with respect to

wavelength, figure 3(b), we see that the ideal focus distance is different at each wavelength. Knowing this, it is possible to choose a focus setting either suited to the spectral region most of interest or which is a suitable compromise that provides good focus for the data as a whole.

An improvement to using a single focus distance would be to extend the depth of field through image processing techniques such as focus-stacking (Pieper and Korpel 1983) whereby multiple images are acquired at different focus distances and combined into a single image in focus at every wavelength.

2.3 Filters

Equalization Filters

Hyperspectral camera detectors use various semiconductors that have different optimal operating ranges and sensitivities. This sensitivity (or quantum efficiency) is dependent on the wavelength of the light being measured and is different for each semiconductor material. The factory-measured quantum efficiency curve from the Silicon-based CCD used in several VNIR hyperspectral cameras, such as the HySpex VNIR1600² is given in figure 4(a). The signal at the low and high extremes of the spectral range of the detector has a significantly lower sensitivity and is, therefore, prone to significant noise.

However, the sensitivity of the system as a whole is dependent not only on the camera detector, but also on the illumination used. The spectral power distribution of a typical extended spectral range illuminant for VNIR systems is given as a function of wavelength in figure 4(b), where we can see that although power at the upper limit of the detector’s range towards 1000nm is excellent, the power towards the lower end at 400nm drops to very low levels. If we multiply the quantum efficiency with the output of the illuminant, we obtain the combined efficiency of the system (the second plot in figure 4(d)) which shows that our peak sensitivity occurs at just under 600nm and that our sensitivity at 1000nm, despite the strong output of the light source at that wavelength, is very weak indeed with

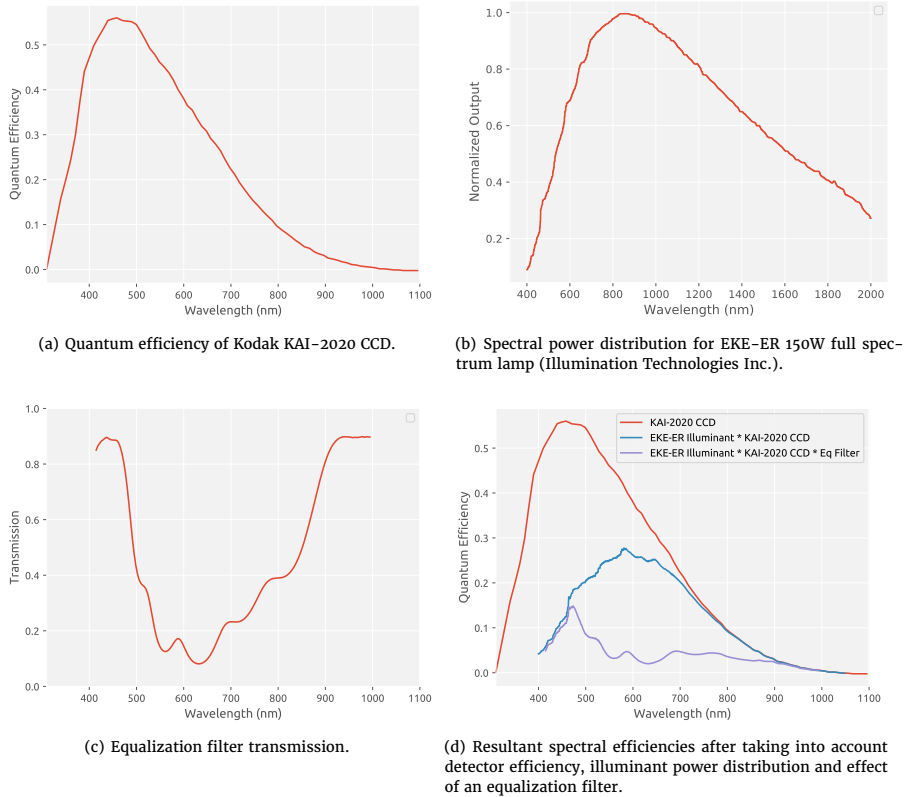


Figure 4. The quantum efficiency of the detector and the impact of the illuminant and equalization filter on it.

an efficiency that is over 50 times less than that at 600nm.

Unlike in a snapshot system (as typically used for multi-spectral imaging), where each spectral band is acquired independently, the integration time for a push-broom hyperspectral camera must be set globally for the entire spectral range, making data quality very variable between regions of high and low sensitivity. There are, nevertheless, several ways to improve the signal-to-noise at the extremes of the spectral range.

The first is to increase the intensities of the spectral content of the light sources in those wavelengths where our illuminant is weak and the detector has low sensitivity. Thus, in our example, this means in the blue region towards 400nm, where secondary narrow-wavelength lamps may be employed.

Alternatively, or in addition to the previous method, it is possible to use an equalization filter. Equalization filters are designed to reduce the sensitivity in the central most sensitive region of the detector's spectral range, thereby improving the *relative* signal-to-noise at the extremes. This can greatly improve the overall signal-to-noise of the data, but necessarily requires longer integration times and reduces signal-to-noise for the central wavelengths.

The ideal equalization filter would be the exact inverse of the quantum efficiency. However, obtaining an exact inverse is difficult to obtain in practice and figure 4(c) shows the measured transmission of a commercially avail-

able equalization filter. The combination of the illuminant, camera quantum efficiency and equalization filter is shown by the third plot in figure 4(d), which has a relatively flatter response over the spectral range, thereby improving the relative efficiency at 1000nm with respect to the average and maximum efficiencies. As we can see from the plot, however, the overall efficiency is much lower, therefore, requiring an integration time in this case around three times as long as it would be without a filter.

Polarizing Filters

For works of art that have areas that are very reflective and which have large amounts of specular reflection, it is possible to use polarizing filters in order to reduce this effect. Both the light source and hyperspectral camera optics require polarized filters with the filter in front of the optics rotated with respect to the those on the illumination in order to block light directly reflected from the surface.

The use of polarized filters in this way, however, necessarily reduces the light incident on the camera detector by half and will, therefore, require longer integration times.

2.4 Frame Averaging

A complementary technique to the use of equalization filters is frame averaging, whereby a number of co-incident

scans of the same line on the work of art are averaged together to improve signal-to-noise (Kohler and Howell 1963). This technique improves signal-to-noise by a factor of \sqrt{N} , where N is the number of scans used in the averaging. In general, push-broom hyperspectral systems achieve this by slowing down or momentarily halting the acquisition scan movement and scanning a single line multiple times. This, however, increases the overall acquisition time, so may not be practical in many cases.

By combining the use of both an equalization filter and averaging, significant improvements in data quality can be achieved at the expense of increased scan time.

3 Data Processing - Calibration Workflow

The acquired raw hyperspectral data will need to be processed to produce accurate and calibrated reflectance data. The processing can be divided into two domains: radiometric calibration and geometric calibration and an overview of the steps can be seen in figure 5. As was seen in the RRT, all participants applied varying levels of radiometric calibration to their data. Geometric calibration, was, however, rarely carried out.

3.1 Radiometric Calibration

Radiometric calibration takes the raw pixel data and produces accurate quantitative reflectance values for each pixel. This step needs to compensate for the variations in the spectral sensitivity for each pixel and both the spatial inhomogeneity and the spectral content of the illumination. In addition, it must also take into account various sources of noise, which can often be significant in hyperspectral imaging.

The first step in the calibration workflow is to remove dark current (also sometimes referred to as fixed pattern) noise. Dark current is essentially electronic noise within the detector which increases with respect to integration time and the operating temperature of the camera. To correct for this, a dark current scan is acquired using the same acquisition settings as for the main scan with no light incident on the detector. In practice this involves blocking the camera optics with either a lens cap or by using an electronic shutter. This should be carried out several times in order to average the data, thereby minimizing random noise and revealing the underlying fixed pattern offset. For push-broom hyperspectral systems, this involves performing a

scan over typically a hundred lines with the optics blocked. The dark current is then calculated by simply averaging the individual pixel values from each scan line to provide a dark current for each pixel on both the spatial and spectral axes. This dark current offset should then be subtracted pixel-wise from all subsequent scan data. Certain manufacturers include this functionality within their acquisition software, thereby automating and simplifying this step.

The following step involves correcting for the variability in the sensitivities of each pixel. This is a particular problem for MCT (Mercury Cadmium Telluride) SWIR detectors, which often exhibit very marked striping in their raw data (Rogalski 2005). To correct this, an image must be acquired under constant and extremely uniform lighting conditions, such as in an integrating sphere. This characterization data is usually acquired by the manufacturer and often automatically corrected for by the acquisition software.

Spatial variability in the illumination and from the optics then also needs to be corrected for. This can be performed by acquiring a diffuse uniform grey or white target that fills the entire field of view and the data must be scaled accordingly.

Once, the pixel sensitivities and spatial illumination has been accounted for, a final spectral correction must be performed by acquiring an image of a known spectral reflectance standard. The spectral image cube can then be scaled to true reflectance factor.

For push-broom hyperspectral systems, it is often practical to combine these last 2 steps into a single step by using a spectral reflectance standard that is large enough to fill the entire field of view.

Reflectance Standard Reliability

The use of one or more reflectance standards is essential in order to calibrate spectral reflectance data. Often, however, this is carried out in a very rudimentary form where the spectral image cube is simply scaled to the nominal reflectance factor given for the reference target – in practice this means dividing the spectral image cube by the measured pixel values of an image of a reflectance standard. Indeed the vast majority of the participants of the RRT used a reference target in this way.

These spectral reflectance standards are widely considered to have a constant reflectance at their nominal reflectance value. However, although they are highly uniform compared to other reflective materials, their reflectances are never perfectly constant and can vary with respect to wavelength. Reflectance standards are usually supplied with individual traceable laboratory-certified absolute refer-

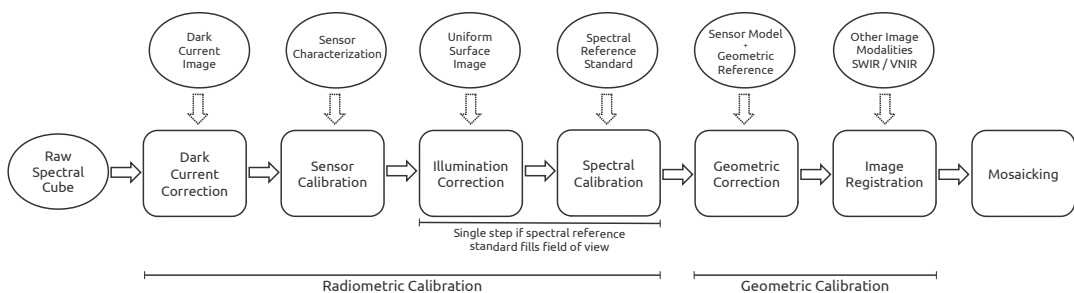


Figure 5. Workflow for hyperspectral calibration showing the processing steps and the data necessary to perform them.

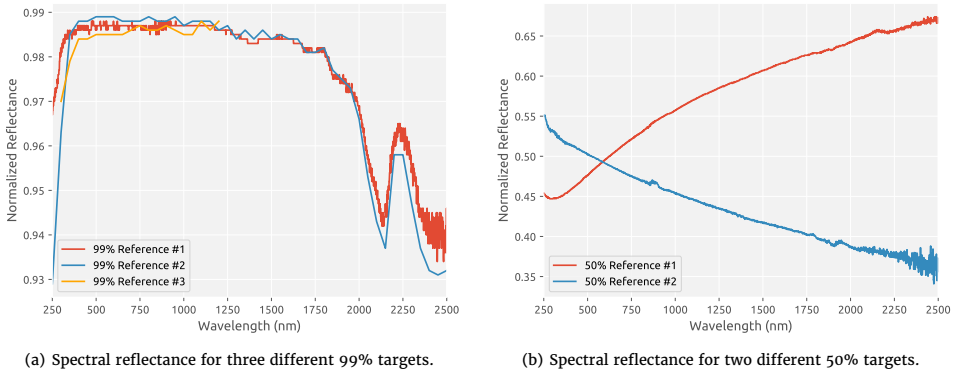


Figure 6. Variability in spectral reflectance for various reflectance standards at 99% and 50% nominal reflectance.

ence reflectance values across the spectral range for which these targets are intended to be used, typically covering the UV, visible and SWIR (short-wave infrared) spectral regions. Figure 6 shows the variability that exists for the certified reflectances of five different reflectance standards with nominal reflectances of 99% and 50%.

Although the 99% reflectance targets are close to uniform over the visible to near infrared region, this is no longer the case from 1500nm to 2500nm. Indeed there are inflections in the spectral curves at around 2130 and 2230nm for all three 99% targets with the reflectance in reality ranging from 90% to 99%. For the 50% reflectance targets examined, the differences are greater still with our two examples exhibiting diverging gradients and real reflectances which vary from 37% to 67%.

Thus, in order to accurately calibrate hyperspectral data to absolute reflectance factor, especially in the short-wave infrared region, the supplied certified reference reflectance values should be used to normalize the hyperspectral data *individually* for each spectral band rather than just assuming a constant fixed reflectance across the entire spectral range.

Reflectance Standard Statistics

There are also a number of subtle statistical errors that can be introduced when processing reflectance standards, as even the best maintained reference standards can get dirty, causing incorrect values to be obtained. One way to avoid this is to exploit the fact that we have a line-scan set-up and can average the data points for each column of the reference target. In this way, the effect of dirt can be reduced. However, such a method, although an improvement on a spot measurement is still subject to bias if done naively. In order to better understand our data, it is useful to first analyze the results obtained from the reference targets in terms of homogeneity and spatial behavior.

A typical measured spatial profile at a particular wavelength from a line along the X (horizontal) direction of a reference target can be seen in figure 7(a). The individual pixel values within the spatial profile are a function of both the reflectance at each point, the illumination and camera noise. Ideally this distribution should be as flat as possible to ensure that the signal-to-noise is constant across the field of view. However, a perfectly even distribution is difficult to achieve in practice and as we can see from

the example, the measured values can vary quite markedly across the field of view despite careful adjustment of the light sources and of their alignment. The signal is also quite noisy with several outliers due to dirt or imperfections on the target. In order to reduce the effect of this noise, it is usual with push-broom hyperspectral systems to scan along the Y direction of the reference standard and average along this axis in order to obtain a relatively noise-free measure of the reflectance for each pixel.

However, although the average is the most widely used statistic, dirt and other imperfections have an asymmetric effect on the distribution resulting in a biased result. For a more accurate statistic, a skewed Gaussian distribution (Azzalini and Valle 1996) provides a better fit and we see from figure 7(b) the skewed form of the distribution of pixel values and the close fit provided by the skewed Gaussian model. Both the mean and median values understate the true value due to the asymmetric nature of the outliers and, in this example, there is almost a 5% difference in the calculated result, which will impact any spectral values calculated from this.

3.2 Noise Reduction

Hyperspectral systems are particularly prone to noise, given the narrow wavelengths used and the wide variability in quantum efficiency (as already seen in figure 4(a)) across the spectral range of the sensor. Although an optimal use of illumination and integration time combined with the use of equalization filters and averaging can significantly improve the quality of the data, noise will invariably exist, particularly at the extremes of the spectral range of the sensor. And this noise can have a significant negative effect on the use of hyperspectral data for tasks such as classification, anomaly detection or materials identification (Landgrebe and Malaret 1986).

Denoising through post-processing is often, therefore, a necessary step in maximizing the utility of the data obtained. A wide range of noise reduction methods exist in image processing. Classic methods developed for greyscale or color images include linear filtering methods such as gaussian blurring, as well as a number of non-linear methods, such as median filtering.

However, spectral data consists of a sequence of highly correlated spectral bands, and it is advantageous to take this

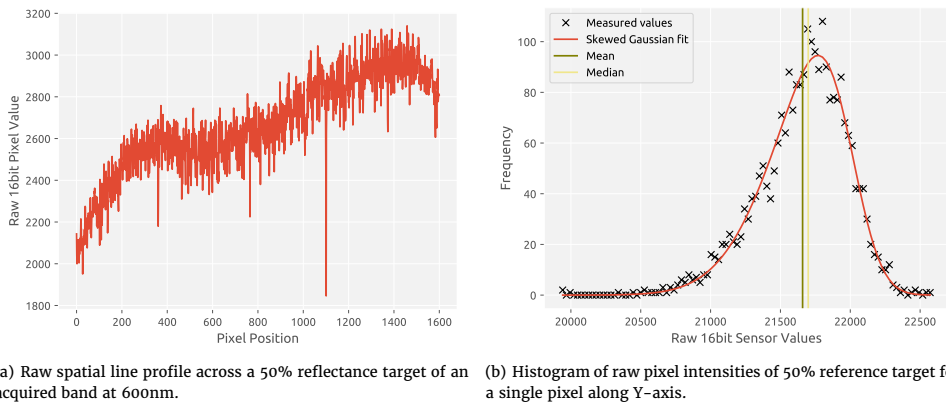


Figure 7. The raw spatial line profile from a hyperspectral camera of a single target (left) and the histogram of pixel values for a single point on the profile when scanned along the Y-axis (right).

extra dimensionality into account. A considerable amount of research exists covering denoising methods adapted to hyperspectral data. These generally consist of adapting existing 2D image denoising methods to take into account the spectral dimension. Examples of these adapted methods include the use of wavelets (Rasti et al. 2012), an extension to non-local means (Qian et al. 2012), variational methods (Mansouri et al. 2015) and combined spectral-spatial approaches (Yuan et al. 2012).

3.3 Geometric Calibration

Geometric calibration is rarely carried out for laboratory-based hyperspectral imaging as, in general, only the spectral values are of interest or the data is to be used in isolation. Such calibration is, therefore, seldom available in the acquisition or calibration software provided by the manufacturer. However, if registration to other imaging modalities is intended or if mosaicking will be carried out, geometric calibration (Špiclin et al. 2010; Khan et al. 2018) will need to be performed on the acquired data.

Hyperspectral systems produce geometric distortions due to the scan motion, to the arrangement of dispersion optics, from the lens itself and also from alignment errors with respect to the work of art. The RRT results showed that not only were the spectral results highly variable, but the resulting image geometries were often equally so (MacDonald et al. 2017).

The complex multi-component optics often used in hyperspectral systems can produce non-parametric distortions that require the characterization of a sensor model in order to correct. A precise sensor model can be produced by using a micrometer and measuring the angle of view for each pixel and several system manufacturers provide such characterization data. A typical sensor model is given in figure 8 and we see that there are differences of up to 20% in the effective pixel sizes across the field of view that must be corrected if spatially accurate data is to be obtained.

In addition, if the scan speed has not been correctly calculated, the aspect ratio of the resulting spectral image cube may be incorrect. This can be the case even if the correct scan speed has been calculated due to tiny residual errors in

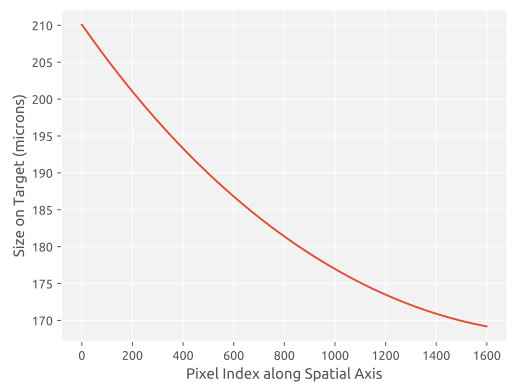


Figure 8. Variation in effective pixel size for each pixel along X-axis for the HySpex VNIR1600

the calculation or in the precision of the movement of Y-axis translation stage, which can accumulate over extended scan areas. In order to correct for this, the geometric calibration target should be used to re-scale the data appropriately along one of the axes.

3.4 Spatial Co-registration

It is often the case that data from spectral cameras with different spectral ranges are available, such as VNIR and SWIR. Data with an extended spectral range can be important for segmentation and materials identification. In order to make full use of this, it is necessary to co-register the data cubes into a single spectral image cube. SWIR detectors, such as those based on MCT or InGaS (Indium Gallium Arsenide) have much lower numbers of pixels (typically 256–320 pixels for MCT and up to 640 pixels for InGaAs), than VNIR cameras and so this step necessarily involves the resampling of data.

Although several participants of the RRT acquired both VNIR and SWIR spectral data, none were able to provide combined co-registered data cubes. Indeed, these comple-

mentary data sets were treated as if they were independent of each other.

If geometric correction has been applied, co-registration should only require alignment and resampling in order to match the spatial pixel sampling. However, even after geometric calibration, small errors can still exist due to minor optical aberrations etc. Co-registration can be optimized by several methods, such as, for example, feature-based matching over an overlapping spectral range (Schwind et al. 2014), which can provide an automated and robust result with sub-pixel level accuracy.

3.5 Mosaicking

Hyperspectral detectors have made great progress in terms of both spectral and spatial resolution. However, spatial resolutions remain limited to at best 2200 pixels for current generation VNIR cameras. Thus, in order to obtain sufficient spatial resolution of a work of art, it is necessary to perform mosaicking of the data cube as the final step in the processing pipeline.

Such mosaicking is now performed routinely for remote sensing data (Palubinskas et al. 2003), which often has the benefit of both accurate GPS positional and height data to help align and ortho-rectify hyperspectral scans. Laboratory push-broom hyperspectral systems that possess both X- and Y-axes are able to scan very regularly spaced strips. However, this level of precision is often insufficient and any variability in the planarity of the artwork can result in distortions due to the close-range nature of the acquisition. Many panel paintings, for example, display warping and even canvas paintings are rarely perfectly flat due to the chassis or from the topographical relief produced by uneven or thick paint layers.

Although the majority of participants in the RRT carried out mosaicking manually using software such as ENVI³, more advanced methods for both mosaicking and registration have been used in the cultural heritage field (Conover et al. 2015).

4 Software

A range of both commercial and open-source software exists for handling data from hyperspectral imaging systems, which can be used to perform a calibration workflow. These include commercial remote sensing software such as ENVI, open source software such as SpectralPython⁴ and Orfeo Toolbox⁵ as well as the manufacturer supplied acquisition software which can occasionally include basic calibration functionality. However, this software is often ill-suited and limited for use on works of art, lacking in key functionality or difficult to use. Indeed, commercial spectral imaging software can be prohibitively expensive for most cultural heritage institutions.

The participants of the RRT used a mixture of commercial remote sensing software such as ENVI, custom-made tools and generic data processing packages such as Matlab⁶. Many participants had difficulties in using the software

and cited the lack of specific functionality adapted to the imaging of works of art.

Therefore, in order to address this issue and allow cultural heritage users to more easily process their data, a suite of custom open-source software tools has been developed and made available⁷ that provides efficient implementations of the various steps of the calibration workflow described in this paper. By developing software collaboratively within the cultural heritage community, it will be possible to extend the software to handle the variety of equipment that is used and provide software adapted to the various needs found within cultural heritage. The use of a common set of software specifically designed for the cultural heritage sector will, furthermore, help improve the quality of hyperspectral data and foster the use of hyperspectral imaging more generally.

5 Conclusion

Hyperspectral imaging has become an increasingly used tool in the analysis of works of art. However, the quality of the acquired data and the processing of that data to produce accurate and reproducible spectral image cubes can be a challenge to many cultural heritage users. As seen in the COSCH RRT (round-robin test), a significant number of users had difficulties in acquiring high fidelity data

The calibration of data that is both spectrally and spatially accurate is an essential step in order to obtain useful and relevant results from hyperspectral imaging. Data that is too noisy or inaccurate will produce sub-optimal results when used for pigment mapping, the detection of hidden features, change detection or for quantitative spectral documentation. To help address this, therefore, we have examined the specific acquisition and calibration workflows required for the hyperspectral imaging of works of art. These workflows include the key parameters that must be addressed during acquisition and the essential steps and issues at each of the steps required during post-processing in order to fully calibrate hyperspectral data.

To accompany this, we have also released a suite of open-source software tools that will enable users to more easily apply the steps described within the workflow. In addition, we have described a number of areas where image quality can be compromised and have proposed techniques to mitigate these problems and help improve overall data quality.

6 Acknowledgments

The authors would like to thank the participants in the spectral imaging round-robin test, which was carried out as part of the European Union COST Action TD1201 *Colour and Space in Cultural Heritage* (COSCH)⁸.

References

Saunders D, Cupitt J. Image Processing at the National Gallery: The VASARI Project. The National Gallery Technical Bulletin 1993 Jan;14:72–85.

3 Harris Geospatial Solutions, Inc.

4 <https://www.spectralpython.net/>

5 <https://www.orfeo-toolbox.org/>

6 Mathworks Inc.

7 <https://hyperspectral-calibration.github.io>

8 <https://cosch.info>

- Martinez K, Cupitt J, Saunders D, Pillay R. Ten Years of Art Imaging Research. *Proceedings of the IEEE* 2002 Aug;90(1):28–41.
- Lahaniar C, Alquié G, Cotte P, Christofides C, de Deyne C, Pillay R, et al. CRISATEL: High Definition Spectral Digital Imaging of Paintings with Simulation of Varnish Removal. In: ICOM Committee for Conservation, ICOM-CC : 13th Triennial Meeting, Rio de Janeiro, 22–27 September 2002 Rio de Janeiro, Brazil: ICOM-CC, James & James; 2002. p. 295–300.
- Bacci M, Casini A, Cucci C, Muzzi A, Porcinai S. A Study on a Set of Drawings by Parmigianino: Integration of Art–Historical Analysis with Imaging Spectroscopy. *Journal of Cultural Heritage* 2005 Dec;6(4):329–336.
- Delaney JK, Zeibel JG, Thoury M, Littleton R, Palmer M, Morales KM, et al. Visible and Infrared Imaging Spectroscopy of Picasso's Harlequin Musician: Mapping and Identification of Artist Materials in Situ. *Applied Spectroscopy* 2010 Jun;64(6):584–594.
- Delaney JK, Zeibel JG, Thoury M, Littleton R, Morales KM, Palmer M, et al. Visible and Infrared Reflectance Imaging Spectroscopy of Paintings: Pigment Mapping and Improved Infrared Reflectography. *SPIE*; 2009. p. 739103–739103–8.
- Snijders L, Zaman T, Howell D. Using Hyperspectral Imaging to Reveal a Hidden Precolonial Mesoamerican Codex. *Journal of Archaeological Science: Reports* 2016 Oct;9:143–149.
- Pottier F. A Study of the Codex Borbonicus Coloring Materials – Non-Invasive Spectroscopies Applied to Codicology. PhD Thesis, Université de Cergy Pontoise (UCP); 2017.
- Liang H, Lucian A, Lange R, Cheung CS, Su B. Remote Spectral Imaging with Simultaneous Extraction of 3D Topography for Historical Wall Paintings. *ISPRS Journal of Photogrammetry and Remote Sensing* 2014 Sep;95:13–22.
- George S, Hardeberg JY. Estimation and Correction of Geometric Distortion in Pushbroom Hyperspectral System for Imaging Art Paintings; 2016.
- Delaney JK, Ricciardi P, Glinsman LD, Facini M, Thoury M, Palmer M, et al. Use of Imaging Spectroscopy, Fiber Optic Reflectance Spectroscopy, and X-Ray Fluorescence to Map and Identify Pigments in Illuminated Manuscripts. *Studies in Conservation* 2014 Mar;59(2):91–101.
- Deborah H, George S, Hardeberg JY. Pigment Mapping of the Scream (1893) Based on Hyperspectral Imaging. In: Elmoataz A, Lezoray O, Nouboud F, Mammas D, editors. *Image and Signal Processing Lecture Notes in Computer Science*, Cherbourg: Springer International Publishing; 2014. p. 247–256.
- Pillay R. Hyperspectral Imaging, Visualization and Database Integration. In: *Proceedings of the ICOM-CC Documentation & Paintings Working Groups Joint Interim Meeting The Hague: ICOM-CC*; 2013. p. 15–16.
- Woodward JT, Brown SW, Smith AW, Lykke KR. Hyperspectral Imager Characterization and Calibration. In: *Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*, vol. 2; 2009. p. II–77–II–80.
- Eismann M. *Hyperspectral Remote Sensing*, vol. PM210. Bellingham, Washington: SPIE Press; 2012.
- Aspinall RJ, Marcus WA, Boardman JW. Considerations in Collecting, Processing, and Analysing High Spatial Resolution Hyperspectral Data for Environmental Investigations. *Journal of Geographical Systems* 2002 Mar;4(1):15–29.
- Behmann J, Mahlein AK, Paulus S, Kuhlmann H, Oerke EC, Plümer L. Calibration of Hyperspectral Close-Range Pushbroom Cameras for Plant Phenotyping. *ISPRS Journal of Photogrammetry and Remote Sensing* 2015 Aug;106:172–182.
- Filacchione G, Capaccioni F, Ammannito E, Coradini A, De Sanctis MC, Piccioni G. Calibration Pipeline of VIS–NIR Imaging Spectrometers for Planetary Exploration: The Rosetta VIRTIS–M Case. In: *First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2009. WHISPERS '09; 2009. p. 1–4.
- Polder G, van der Heijden G, Keizer L, Young I. Calibration and Characterisation of Imaging Spectrographs. *Journal of Near Infrared Spectroscopy* 2003;11(1):193.
- Geladi P, Burger J, Lestander T. Hyperspectral Imaging: Calibration Problems and Solutions. *Chemometrics and Intelligent Laboratory Systems* 2004 Jul;72(2):209–217.
- Burger J, Geladi P. Hyperspectral NIR Image Regression Part I: Calibration and Correction. *Journal of Chemometrics* 2005 May;19(5–7):355–363.
- Burger J, Geladi P. Hyperspectral NIR Image Regression Part II: Dataset Preprocessing Diagnostics. *Journal of Chemometrics* 2006 Mar;20(3–4):106–119.
- Boldrini B, Kessler W, Rebner K, Kessler R. Hyperspectral Imaging: A Review of Best Practice, Performance and Pitfalls for Inline and Online Applications. *Journal of Near Infrared Spectroscopy* 2012;20(5):438.
- Khan H, Mihoubi S, Mathon B, Thomas JB, Hardeberg J, Khan HA, et al. HyTexiLa: High Resolution Visible and Near Infrared Hyperspectral Texture Images. *Sensors* 2018 Jun;18(7):2045.
- George S, Hardeberg J, Linhares Ja, MacDonald L, Montagner C, Nascimento S, et al. A Study of Spectral Imaging Acquisition and Processing for Cultural Heritage. In: *Digital Techniques for Documenting and Preserving Cultural Heritage Leeds: Arc Humanities Press*; 2018. p. 141–158.
- Groen FCA, Young IT, Lighthart G. A Comparison of Different Focus Functions for Use in Autofocus Algorithms. *Cytometry* 1985 Mar;6(2):81–91.
- Pieper RJ, Korpel A. Image Processing for Extended Depth of Field. *Applied Optics* 1983 May;22(10):1449.
- Kohler RJ, Howell HK. Photographic Image Enhancement by Superimposition of Multiple Images. *Photographic Science and Engineering* 1963;7(4):241–245.
- Rogalski A. HgCdTe Infrared Detector Material: History, Status and Outlook. *Reports on Progress in Physics* 2005;68(10):2267.
- Azzalini A, Valle AD. The Multivariate Skew–Normal Distribution. *Biometrika* 1996 Jan;83(4):715–726.
- Landgrebe DA, Malaret E. Noise in Remote-Sensing Systems: The Effect on Classification Error. *IEEE Transactions on Geoscience and Remote Sensing* 1986 Mar;GE-24(2):294–300.
- Rasti B, Sveinsson JR, Ulfarsson MO, Benediktsson JA. Hyperspectral Image Denoising Using 3D Wavelets. In: *2012 IEEE International Geoscience and Remote Sensing Symposium*; 2012. p. 1349–1352.
- Qian Y, Shen Y, Ye M, Wang Q. 3-D Nonlocal Means Filter with Noise Estimation for Hyperspectral Imagery Denoising. In: *2012 IEEE International Geoscience and Remote Sensing Symposium*; 2012. p. 1345–1348.
- Mansouri A, Deger F, Pedersen M, Hardeberg JY, Voisin Y. An Adaptive Spatial–Spectral Total Variation Approach for Poisson Noise Removal in Hyperspectral Images. *Signal, Image and Video Processing* 2015 Aug;p. 1–8.
- Yuan Q, Zhang L, Shen H. Hyperspectral Image Denoising Employing a Spectral–Spatial Adaptive Total Variation Model. *IEEE Transactions on Geoscience and Remote Sensing* 2012 Oct;50(10):3660–3677.
- Špiclin Ž, Katrašnik J, Bürmen M, Pernuš F, Likar B. Geometric Calibration of a Hyperspectral Imaging System. *Applied Optics* 2010 May;49(15):2813.
- MacDonald LW, Vitorino T, Picollo M, Pillay R, Obarzanowski M, Sobczyk J, et al. Assessment of Multispectral and Hyperspectral Imaging Systems for Digitisation of a Russian Icon. *Heritage Science* 2017 Sep;5:41.
- Schwind P, Schneider M, Müller R. Improving HySpex Sensor Co-Registration Accuracy Using BRISK and Sensor–Model Based RANSAC. *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 2014 Nov;XL-1:371–376.
- Palubinskas G, Muller R, Reinartz P. Mosaicking of Optical Remote Sensing Imagery. In: *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477)*, vol. 6; 2003. p. 3955–3957 vol.6.
- Conover DM, Delaney JK, Loew MH. Automatic Registration and Mosaicking of Technical Images of Old Master Paintings. *Applied Physics A* 2015 Apr;119(4):1567–1575.

Multi-Modal Data Visualization and Analysis of “The Bedroom at Arles” by Vincent van Gogh

Paper 5

This chapter is a reprint of the publication:

Pillay, R., “Multi-Modal Data Visualization and Analysis of *The Bedroom at Arles* by Vincent van Gogh”, *Proceedings of the 22nd CIPA Symposium*, October 11-15, 2009, Kyoto, Japan.

Multi-modal Data Visualization and Analysis of “The Bedroom at Arles” by Vincent van Gogh

Ruven Pillay*

Abstract—In this paper we present an open source software platform for providing high resolution multi-modal visualization and analysis of advanced heterogeneous scientific imagery. The system has been designed around the needs and requirements of cultural heritage and provides a powerful, flexible and easy-to-use means for conservators, researchers or art historians to interact and analyze scientific imagery of works of art. The platform allows cultural heritage institutions to manage and maintain very large image data sets and for multiple simultaneous users to remotely visualize heterogeneous data at high resolution using efficient image streaming techniques. As a case study, we present an example of a recent laboratory analysis carried out by the C2RMF of a work by Vincent Van Gogh: The Bedroom at Arles. The extensive range of imaging techniques carried out for this painting was unique in its depth and range and showcase the potential of multi-modal visualization techniques.

Index Terms—multispectral, 3D, multi-modal, registration, visualization

I. INTRODUCTION

A. Motivation

IN order to accurately and more fully document works of art for conservation or scientific study, multiple orthogonal digital acquisition techniques are often required. High resolution scientific imaging techniques, such as multispectral imaging, 3D structured light acquisition, high dynamic range (HDR) imaging, light transmission imaging, X-ray radiography, UV fluorescence, raking light and infrared reflectography all provide different views of the holistic whole. Although each technique can be used in isolation, in order to fully understand the nature of the work of art, the ensemble of data needs to be brought together in a meaningful way. Advanced software tools adapted to the needs of cultural heritage are necessary for restorers, art

historians and the general public to make sense of such data. The fusion and visualization of these very large and diverse data sets not only allow existing images to be better exploited, but also allow new kinds of analysis to be carried out.

B. The Painting

The subject of this study was a painting by the Dutch post-impressionist painter Vincent Van Gogh (30 March 1853 – 29 July 1890). *La Chambre à Arles* or the *Bedroom in Arles* is a composition showing the interior of the bedroom of his house at 2, Place Lamartine in Arles, France. Three versions of the composition exist: at the Van Gogh museum in Amsterdam, the Art Institute of Chicago and the Musée d’Orsay in Paris.

The first version, believed to be that currently at the Van Gogh museum, was painted in 1888 with the other two painted in 1889. The different versions are similar, but are not exact copies having several details differing. In addition the third version, that of the Musée d’Orsay, is significantly smaller than the others, measuring 57.5 x 74 cm compared to 72 x 90 cm for the other two. It is this third version that was the subject of this particular study.

C. Multi-national Study

An extensive range of scientific analyses was carried out at the Centre de Restauration et de Recherche des Musées de France (C2RMF), the national restoration and research centre for French museums, housed within the Louvre Museum. The analyses carried out formed part of a multi-national study involving all three of the museums holding versions of the painting. Good timing and the availability of equipment allowed us to apply an unprecedented range and combination of advanced imaging techniques to this particular version of the painting.

*C2RMF, Palais du Louvre - Porte des Lions, Paris 75001, France
Email: ruven.pillay@culture.gouv.fr

D. Ultra High Resolution Colorimetric Imaging

A high resolution and high fidelity color rendering is an essential part of any image acquisition of a painting. Color digital cameras, however, are limited in their resolution and color accuracy. In order to fully study this painting, it was necessary to go well beyond the limitations of our camera hardware. Our high-end Imacon/Hasselblad digital SLR has a good native resolution of 20 mega pixels. However, we were aiming for a resolution of around 20 pixels per mm on the painting, which equates to around 180 mega-pixels of data. In order to achieve such a resolution with our equipment an image mosaic approach was required. Numerous images of details were taken under fixed lighting conditions and camera geometry and stitched together. Corrections were made for lens distortion, vignetting and key-stoning etc. Color accuracy was crucial, so a full colorimetric characterization of the camera and lighting was made. This involved dark current correction, pixel sensitivity normalization, measurement of the lighting inhomogeneity, spectral measurement of the light source and target calibration using a MacBeth ColorChecker SG reference chart. The resulting high resolution image was profiled to CIEL*a*b* color-space. Verification of the final result (Figure 1) was carried out both visually under daylight D65 lamps and using photo-spectrometer spot tests on several points on the painting.



Fig. 1. Colorimetric 180 mega-pixel mosaic of The Bedroom at Arles by Vincent Van Gogh, oil on canvas, 1889. Copyright © C2RMF 2009

E. Multispectral Infra-Red Imaging

Two forms of multispectral imaging were carried out in the near (NIR) and short wavelength (SWIR) infra-red spectral regions. These infra-red regions are particularly useful for pigment identification as many pigments with similar colors are only distinguishable by their spectral signatures in this region (Casadio, 2001).

1) *Narrow-band imaging*: Narrow-band multispectral imaging was carried out by the Cultural Heritage group of the National Institute of Applied Optics (INOA-CNR - Gruppo Beni Culturali) of Italy. The system used was an experimental multispectral imaging system based on an InGaAs photodetector linked to interferential filters through fiber optics. These fiber optics relay light from a single source lens mounted on an automated mechanical X-Y translation stage (Bonnifazzi, 2008). In effect, the apparatus acts as a high speed spectrometer that automatically takes millions of spot readings along the entire surface of the painting. The advantage of such an approach is that it avoids errors due to the lens and fore-optics, and due to geometrical mis-alignment caused by the different refractive indices of the filters. In all, 14 narrow-band filters with an approximate band-width of 50nm were used for this scan with an overall spectral range of 800-2300 nm and a spatial resolution of 2.5 pixels per mm. The resulting data was corrected for noise, gain, normalized using reference targets and compiled into image files per filter.

2) *Wide-band imaging*: In addition to the multispectral infra-red imaging, wide-band infra-red imaging was carried out using the high resolution InGaAs based Osiris camera (Falco, 2009). This camera was used with bandpass filters of 900-1200nm and 1200-1700nm to obtain several infra-red acquisitions with higher spatial resolution than was possible from the narrow-band multi-spectral apparatus. Affordable InGaAs sensors are often of relatively low resolution. Thus, in order to obtain sufficiently high spatial resolution, the camera is able to automatically employ a mosaicing technique.

Figure 2 shows a selection of the wide and narrow-band infra-red images.

F. Structured Light 3D Imaging

The paintings of Van Gogh often possess a strong coarse brushwork. This is certainly the case for



Fig. 2. Infra-red images: wide-band 900-2300nm (top), narrow-band 1600nm (bottom). Copyright © C2RMF 2009

the Bedroom at Arles, making a 3D acquisition of the surface particularly interesting. The 3D scanning technique used was that of structured light (Frankowski, 2000). In this method, a sequence of pattern fringes of different sizes is projected onto the surface of the painting and scanned by a digital camera. The distortion in the fringes resulting from the surface roughness makes it possible to calculate the height at each point (Figure 3a). This is done by a combined Gray Code/Phase shift technique (Park, 2001). The scan itself was carried out by the Breuckmann company (Breuckmann, 2000) using their OptoScan system (Figure 3b). In order to obtain a sufficiently high spatial resolution, once again a mosaic technique was used. The resulting dense point cloud of three dimensional coordinates for the surface of the painting had a spatial resolution of

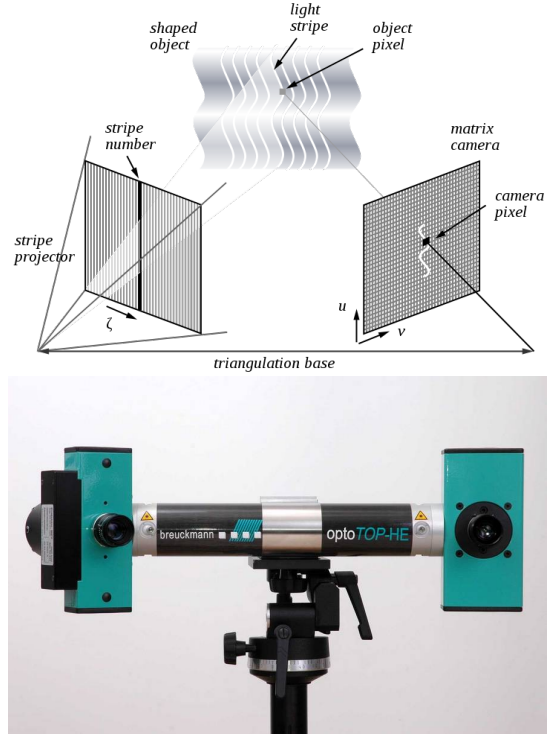


Fig. 3. (a) Principal of Structured Light Projection, (b) Breuckmann OptoTop equipment

around 20 pixels per mm in X-Y and a resolution in Z (depth) of approximately 20 μ m.

G. Infra-red, Ultra-violet Florescence and X-ray

Together with the advanced techniques used above, more mainstream scientific imaging were carried out at the C2RMF. To cover other important parts of the electro-magnetic spectrum, UV florescence and X-rays were used. In addition, techniques such as raking light photography (from multiple directions) and light transmission imaging (where the light source is placed behind the painting) was carried out in both the visible and infra-red regions. Indeed the range of imaging techniques applied to this painting was quite unprecedented. Never before had a painting been so comprehensively photographed at the C2RMF with such a wide range of equipment and techniques.

II. MULTI-MODAL REGISTRATION

Because of the differing imaging techniques, optics, acquisition geometries and sensor characteristics, registration of the different images was not straightforward. Geometric distortion, key-stoning, chromatic aberration, and errors in the image mosaics all make a simple linear transformation impossible. In addition, the content of the images themselves can not always be exactly correlated. For example, raking light images show the shadow cast by a feature on the surface of the painting and not the feature itself. X-rays and Infra-red images penetrate into the painting and are not, therefore, in the same geometric plane as, for example, a UV fluorescence image, which originates at the upper surface layer of the painting. In order to resolve this, a combination of different registration techniques adapted to each imaging type was employed (Zitová, 2003). This consisted of a combination of semi-automatic registration based on cross-correlation, phase correlation as well as manual alignment in certain cases.

III. VISUALIZATION

A. Remote Visualization

In order to be useful for conservators, researchers or member of the public, the large quantity of heterogeneous data produced by scientific imaging needs to be synthesized and made available to the user in an efficient manner. In order to achieve this, remote visualization software was employed.

B. IIPImage

The remote visualization platform is based on IIP-Image¹, an open source image streaming system for ultra-high resolution and scientific imagery (Pitzalis, 2006; Saunders, 1999). It works by streaming tiles corresponding to the area currently under view at the requested resolution to the client. In this way, the user can efficiently view, navigate and zoom around at the maximum resolution possible. In addition to sRGB imagery, scientific image types such as 16 bit, colorimetric CIEL*a*b*, multispectral and 3D height/normal data are supported. The software is based on a client-server architecture with an imaging server providing tiles, image processing and metadata resources via a RESTful protocol. In

order to be as efficient as possible and reduce the load of the server, image data is stored in a multi-resolution tiled TIFF format.

The client software, embedded within a web interface, is platform independent, giving users access to multi gigabyte size scientific images without having to install any 3rd party software or requiring any particular hardware / software configuration or operating system. Excellent performance is available even on old or slow hardware. The client incorporates numerous functions allowing the user to easily navigate and zoom within the data. In addition several more advanced features exist.

C. Image Blending

The ability to easily compare and contrast any of the images and do so interactively is a potentially extremely powerful tool in order to fully understand the material composition, layering or technique of a painting. An extension to the IIPImage client has been especially designed to give the user the ability to select any pair of registered images and then simultaneously navigate and zoom within both. At any time, the user can perform a superposition and smooth blending back and forth between the pair of images. Subtle differences that may not be apparent on viewing them statically and separately can be far more easily discerned from such an interactive manipulation. An example of the blending between the colour and X-ray is shown in Figure 4.

D. Spectral Visualization

The ability to determine the spectral reflectance at any point on the painting gives crucial information on the material composition of the paint layer. Spectral curves can be compared to reference spectra for pigment or material identification. Within IIPImage, a graphing area is available within the interface, allowing the user to click on any point on the painting and instantly see the spectral curve for that point. This feature was applied to the narrow-band multispectral infra-red data acquired (Figure 5).

E. 3D Hillshade Rendering

High resolution 3D data is not easy to stream or make multi-resolution. Furthermore as our software is essentially image-centric, a 2D rendering approach to visualization was favoured. Inspiration

¹<http://iipimage.sourceforge.net>

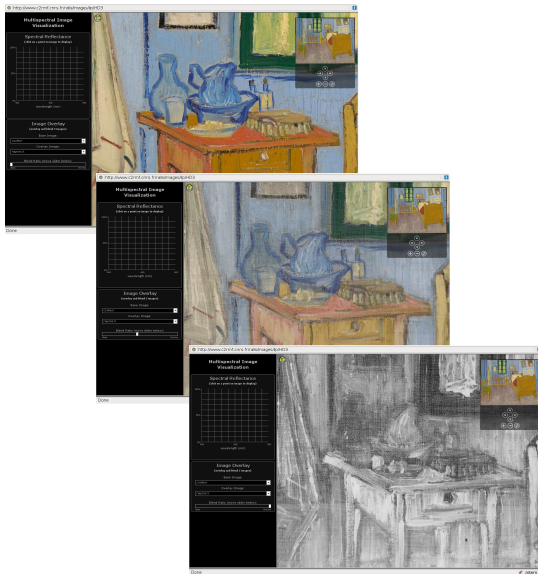


Fig. 4. Image blending showing the superposition of the color and X-ray images and the blended smooth transition between them

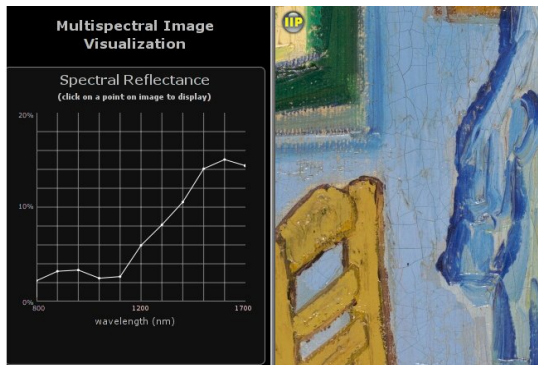


Fig. 5. Spectral reflectance curve for any point on painting

was taken from the domain of cartography and the surface of the painting was considered as a digital relief map (DEM). In order to facilitate the integration of such data into IIPImage, the 3D model data was converted to point data representing the vector normal at each point and the height at each point. In this form, they are able to be stored in standard TIFF format, and are thus able to be tiled and used directly for multi-resolution streaming with IIPImage. A basic rendering technique is that of hill-shading (Horn, 1981) where a virtual

directional illumination is used to create shading on virtual "hills". A fast hill-shading algorithm has been implemented in IIPImage allowing the user to interactively set the angle of incidence of the lighting source and view a dynamically rendered hill-shaded relief map. By moving the light source, a better understanding of the surface relief is possible (Figure 6).

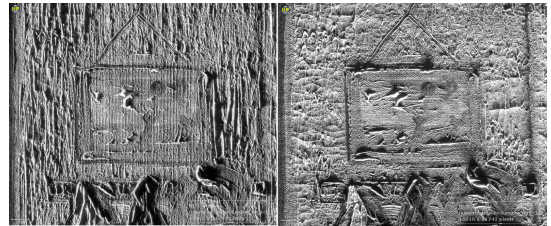


Fig. 6. Hill-shading applied in two different directions to 3D model data

IV. CONCLUSION

Advanced imaging techniques give conservators, restorers and researchers unprecedented access to highly detailed information from numerous kinds of analytical technique. However, this heterogeneous mass of data can often be unwieldy and difficult to view let alone undertake basic analyses with. We have presented here an open source streaming-based solution designed around the needs and requirements of cultural heritage that addresses these issues. The example of The Bedroom at Arles provided us with an exceptional and unprecedented range of imaging data to handle. The integration of the ensemble into a single visualization system demonstrates the potential of such a platform for museums and researchers. In addition, ability to stream the data online make it especially useful when research is undertaken by different research teams or in an international context.

In addition to the already implemented functionality, many potentially interesting applications are conceivable. For example automatic pigment spectra classification or identification. Or the use of the height data from a 3D acquisition to provide renderings of cross sections etc.

ACKNOWLEDGMENTS

We would like to thank the Breuckmann company of Germany, the Gruppo Beni Culturali of the

INOA-CNR in Italy, the staff of the C2RMF who contributed to the imaging analyses (Elsa Lambert for infra-red/raking light/transmission and Elisabeth Ravaud for the X-rays), the Van Gogh museum and the musée d'Orsay for their help in carrying out this study.

REFERENCES

- [1] Casadio, F., Toniolo, L., 2001. "The analysis of polychrome works of art: 40 years of infrared spectroscopic investigations". *Journal of Cultural Heritage* Vol 2, Issue 1, pp. 71-78
- [2] C Bonifazzi, C., Carcagni, P., Fontana, F., Greco, M., Mastroianni, M., Materazzi, M., Pampaloni, E., Pezzati, L., Bencini, D., 2008. "A scanning device for VIS-NIR multispectral imaging of paintings". *J. Opt. A: Pure Appl. Opt.* 10 064011 pp. 9
- [3] Frankowski, G., Chen, M., Huth, T., 2000. "Real-time 3D Shape Measurement with Digital Stripe Projection by Texas Instruments Micromirror Devices (DMD)". *Proceedings. Of SPIE-Vol. 3958*, pp. 90 – 106
- [4] Park, J., G. N. DeSouza, G. N., Kak, A. C., 2001. "Dual-beam structured-light scanning for 3-D object modeling. 3D Digital Imaging and Modeling", *IEEE Computer Society*, pp. 65-72
- [5] Breuckmann, B., Klass, E., 2000. "Topometric Extends Image Processing Into The Third Dimension". *Euromicro*, 46 vol 12, pp. 32-33
- [6] Falco, C., 2009. "High resolution digital camera for infrared reflectography". *Review of Scientific Instruments*, 80, 071301
- [7] Pitzalis, D., Pillay, R., Lahanier, C., 2006. "A New Concept in High Resolution Internet Image Browsing". *Proceedings of 10th International Conference on Electronic Publishing - June 2006*
- [8] Saunders D., Cupitt J., Pillay R., Martinez K., 1999. "Maintaining colour accuracy in images transferred across the Internet". *Colour Imaging : Vision and Technology*, pp. 215-230
- [9] Zitová, Z., Flusser, J., 2003. "Image registration methods: a survey". *Image Vision Computing* 21(11), pp. 977-1000
- [10] Horn, B.K.P., 1981. "Hill shading and the reflectance map". *Proceedings of the Institute of Electrical and Electronics Engineers IEEE*, Vol. 69, Issue 1, pp. 14-47

Web-Based Visualization of Very Large Scientific Astronomy Imagery

Paper 6

This chapter is a reprint of the publication:

Bertin, E., **Pillay, R.** & Marmo, C. “Web-Based Visualization of Very Large Scientific Astronomy Imagery”. *Astronomy and Computing* 10, 43–53 (2015).

Web-Based Visualization of Very Large Scientific Astronomy Imagery

Emmanuel Bertin^{a,d}, Ruven Pillay^{b,*}, Chiara Marmo^{c,d}

^aUniv Pierre et Marie Curie, Institut d'Astrophysique de Paris, UMR7095, Paris, F-75014, France

^bC2RMF, Palais du Louvre - Porte des Lions, Paris 75001, France

^cUniv Paris-Sud, Laboratoire GEOPS, UMR8148, Orsay, F-91405, France

^dCNRS, France

Abstract

Visualizing and navigating through large astronomy images from a remote location with current astronomy display tools can be a frustrating experience in terms of speed and ergonomics, especially on mobile devices. In this paper, we present a high performance, versatile and robust client-server system for remote visualization and analysis of extremely large scientific images. Applications of this work include survey image quality control, interactive data query and exploration, citizen science, as well as public outreach. The proposed software is entirely open source and is designed to be generic and applicable to a variety of datasets. It provides access to floating point data at terabyte scales, with the ability to precisely adjust image settings in real-time. The proposed clients are light-weight, platform-independent web applications built on standard HTML5 web technologies and compatible with both touch and mouse-based devices. We put the system to the test and assess the performance of the system and show that a single server can comfortably handle more than a hundred simultaneous users accessing full precision 32 bit astronomy data.

Keywords: visualization, scientific data, web application, high resolution, HTML5

1. Introduction

Although much of the extraction of information from astronomy science images is now performed “blindly” using computer programs, astronomers still rely on visual examination for a number of tasks. Such tasks include image quality control, assessment of morphological features, and debugging of measurement algorithms.

The generalization of standardized file formats in the astronomy community, such as FITS (Wells et al., 1981), has facilitated the development of universal visualization tools. In particular, SAOIMAGE (Vanhilst, 1990), ALADIN (Bonnarel et al., 1994), SKYCAT (Albrecht et al., 1997), GAIA (Draper, 2000) and ds9 (Joye and Mandel, 2003). These packages are designed to operate on locally stored data and provide efficient access to remote image databases by downloading sections of FITS data which are subsequently read and processed locally for display; all the workload, including image scaling, dynamic range compression, color compositing and gamma correction, is carried out client-side.

However, the increasing gap between storage capacities and data access bandwidth (Budman, 2011) makes it increasingly

efficient to offload part of the image processing and data manipulations to the server, and to transmit some form of pre-processed data to clients over the network.

Thanks to the development of wireless networks and light mobile computing (tablet computers, smartphones), more and more scientific activities are now being carried out on-the-go outside an office environment. These possibilities are exploited by an increasing number of scientists, especially experimentalists involved in large international collaborations and who must interact remotely, often in real-time, with colleagues and data located in different parts of the world and in different time zones. Mobile devices have increasingly improved display and interfacing capabilities, however, they offer limited I/O performance and storage capacity, as well as poor battery life when under load. Web-based clients, or simply *Web Apps*, are the applications of choice for these devices, and their popularity has exploded over the past few years.

Thanks to the ubiquity of web browsers on both desktop and mobile platforms, *Web Apps* have become an attractive solution for implementing visual interfaces. Modern web browsers feature ever faster and more efficient JavaScript engines, support for advanced standards such as HTML5 (W3C, 2012) and CSS3 (W3C, 2011), not to mention interactive 3D-graphics with the recent WebGL API (Khronos Group, 2013). As far as data visualization is concerned, web applications can now be made sufficiently feature-rich so as to be able to match many of the functions of standalone desktop applications, with the additional

*Corresponding author

Email addresses: bertin@iap.fr (Emmanuel Bertin),
ruven.pillay@culture.gouv.fr (Ruven Pillay),
chiara.marmo@u-psud.fr (Chiara Marmo)

benefit of having instant access to the latest data and being embeddable within web sites or data portals.

One of the difficulties in having the browser deal with science data is that browser engines are designed to display gamma-encoded images in the GIF, JPEG or PNG format, with 8-bits per Red/Green/Blue component, whereas scientific images typically require linearly quantized 16-bit or floating point values. One possibility is to convert the original science data within the browser using JavaScript, either directly from FITS (Lowe, 2011; Kapadia, 2013), or from a more “browser-friendly” format, such as e.g., a special PNG “representation file” (Mandel, 2014), or compressed JSON (Federl et al., 2011). In practice this is currently limited to small rasters, as managing millions of such pixels in JavaScript is still too burdensome for less powerful devices. Moreover, lossless compression of scientific images is generally not very efficient, especially for noisy floating-point data (e.g. Pence et al., 2009). Hence, currently, server-side compression and encoding of the original data to a browser-friendly format remains necessary in order to achieve a satisfactory user experience on the web client, especially with high resolution screens.

Displaying images larger than a few megapixels on monitors or device screens requires panning and/or pixel rebinning, such as in “slippy map” implementations (Google Maps™, OpenStreetMap¹ etc.). On the server, the images are first decomposed into many small tiles (typically 256×256 pixels) and saved as PNG or JPEG files at various levels of rebinning, to form a “tiled pyramid”. Each of these small files corresponds to a URL and can be loaded on demand by the web client. Notable examples of professional astronomy web apps based on this concept include the Aladin Lite API (Schaaff et al., 2012), and the Mizar plugin² in SITools2 (Malapert and Marseille, 2012).

However, having the data stored as static 8-bit compressed images means that interaction with the pixels is essentially limited to passive visualization, with little latitude for image adjustment or interactive analysis. Server-side dynamic processing/conversion of science-data on the server and streaming of the processed data to the web client are necessary to alleviate these limitations. Visualization projects featuring dynamic image conversion/streaming in Astronomy or Planetary Science have mostly relied on browser plugins implementing proprietary technologies (Federl et al., 2012) or Java clients/applets (Muller et al., 2009; Kitaëff et al., 2012). Notable exceptions include Helioviewer (Hughitt et al., 2008), which queries compressed PNG tiles directly from the browser with the tiles generated on-the-fly server-side from JPEG2000 encoded data.

In this paper we describe an open source and multi-platform high performance client-server system for the processing, streaming and visualization of full bit depth scientific imagery at the terabyte scale. The system consists of a light-weight C++ server and W3C standards-based JavaScript clients capable of running on stock browsers. In section 2, we present our approach, the protocols and the implementation of both the

server and the client. Sections 3 and 4 showcase several applications in Astronomy and Planetary Science. In Section 5, we assess the performance of the system with various configurations and load patterns. Finally in Section 6, we discuss future directions in the light of current technological trends.

2. Material and Methods

The proposed system consists of a (or several) central image server(s) capable of processing 32 bit floating point data on-demand and of transcoding the result into an efficient form usable by both light-weight mobile devices or desktop computers.

2.1. Image Server

At the heart of the system is the open source IIPImage³ image server (Pitzalis et al., 2006). IIPImage is a scalable client-server system for web-based streamed viewing and zooming of ultra high-resolution raster images. It is designed to be fast, scalable and bandwidth-efficient with low processor and memory requirements.

IIPImage has a long history and finds its roots in the mid 1990s in the cultural heritage field where it was originally created to enable the visualization of high resolution colorimetric images of paintings (Martinez et al., 1998). The original system was designed to be capable of handling gigapixel size, scientific-grade imaging of up to 16 bits per channel, colorimetric images encoded in the CIEL*a*b* color space and high resolution multispectral images (Martinez et al., 2002) (Fig. 1). It had hitherto been very difficult to simply even view such image data locally, let alone access it remotely, share or collaborate between institutions. The client-server solution also enabled integration of full resolution scientific imaging such as infra-red reflectography, Xray, multispectral and hyperspectral imagery (Fig. 2) into museum research databases, providing for unprecedented levels of interactivity and access to these resources (Lahanier et al., 2002).

Beyond cultural heritage, the system has also been adapted for use in the field of biomedical imaging. For example, to visualize ultra-large high resolution electron microscopy maps created by ultra-structural mapping or *virtual nanoscopy* (Faas et al., 2012), or to explore high resolution volumetric 3D cross-sectional atlases (Husz et al., 2012).

In practice the IIPImage platform consists of a light-weight C++ Fast-CGI (Brown, 1996) server, `iipsrv`, and an AJAX-based web interface. Image data stored on disk are structured in order to enable efficient and rapid multi-resolution random access to parts of the image, allowing terapixel scale data to be efficiently streamed to the client. As only the region being viewed needs to be decoded and sent, large and complex images can be managed without onerous hardware, memory or network requirements by the client. The IIPImage server performs on-the-fly JPEG compression for final visualization, but as the underlying data is full bit depth uncompressed data, it

¹<http://www.openstreetmap.org>

²<https://github.com/TPZF/RTWeb3D>

³<http://iipimage.sourceforge.net>

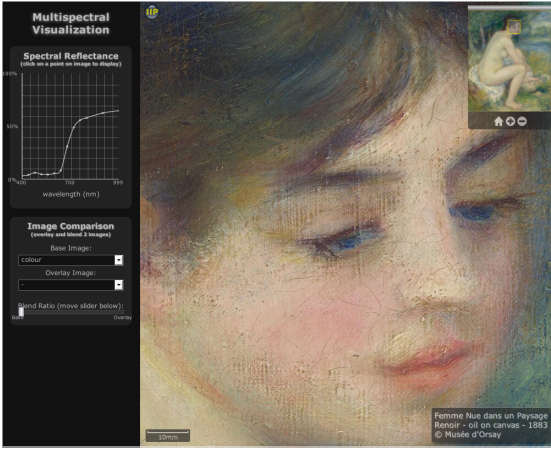


Figure 1: Spectral visualization of Renoir’s *Femme Nue dans un Paysage*, Musée de l’Orangerie, showing spectral reflectance curve for any location and controls for comparing different imaging modalities

can operate directly on scientific images, and perform operations such as rescaling or filtering before sending out the results to the client.

IIPImage, therefore, possessed many of the attributes necessary for astronomy data visualization and rather than develop something from scratch, it was decided to leverage this existing system and extend it. A further benefit to this approach would be access to a larger scientific community beyond that of astronomy with certain similar data needs. Moreover, as IIPImage forms part of the standard Debian, Ubuntu and Fedora Linux distributions, access to the software, installation and maintenance would be greatly simplified and sustainable in the longer term.

Hence a number of modifications were made to the core of IIPImage in order to handle astronomy data. In particular, to extend the system to handle 32 bit data (both integer and IEEE floating point), FITS metadata, functionality such as dynamic gamma correction, colormaps, intensity cuts, and to be capable of extracting both horizontal and vertical data profiles. The resulting code has been integrated into the main IIPImage software development repositories and is available from the project website⁴, where it will form part of the 1.0 release of iipsrv.

2.2. Data Structures and Format

Extracting random image tiles from a very large image requires an efficient storage mechanism. In addition to tile-based access, the possibility to rapidly zoom in and out imposes some sort of multi-resolution structure on the data provided to the server. The solution adopted for “slippy map” applications is often simply to store individual tiles rebinned at the various resolution levels as individual image files. For a very large image this can translate into hundreds of thousands of small files being



Figure 2: Hyperspectral imaging of paintings

created. This approach is not convenient from a data management point of view, and for IIPImage a single file approach has always been preferred.

The current version of IIPImage supports both TIFF and JPEG2000 formats. Multi-resolution encoding is one of the major features of JPEG2000, but the lack of a robust, high performance open source library has been a serious issue until recently. Nevertheless, the encoding of floating point values spanning a large dynamic range remains a concern with current open-source libraries, as in practice input data is managed with only fixed point precision (Woodring et al., 2011; Kitaef et al., 2014).

The combining of tiling and multi-resolution mechanisms is also possible with TIFF. TIFF is able to store not only 8 bit and 16 bit data, but also 32 bit integers, and single or double precision floating point numbers in IEEE format. As a well supported and mature standard with robust and widely used open source development libraries readily available, TIFF was adopted as the main server-side storage format, rather than creating a completely new format or adapting existing science formats in some way.

2.3. Image Transcoding

Astronomy imaging data are usually stored in the FITS format (Wells et al., 1981). FITS is a flexible container format that can handle data encoded in up to 64 bits per value. FITS supports image tiling, whereby the original raster is split into separate rectangular tiles, which can be retrieved quickly and read and decoded independently (Pence, W. et al., 2000). Versions of the same image could be stored at multiple resolution levels in different extensions, at the price of an increased file

⁴<http://iipimage.sourceforge.net>

size. However currently neither tiling nor multi-resolution is present in archived FITS science images. Hence, regardless of the adopted storage format (TIFF in our case), a considerable amount of pixel shuffling and rebinning must be carried out in order to convert FITS data before they can be handled by the server.

Transcoding from basic FITS to multi-resolution tiled TIFF is carried out via the STIFF conversion package (Bertin, 2012). The multi-resolution structure consists of an image “pyramid” whereby pixels in each image are successively rebinned 2×2 and stored in separate TIFF virtual “directories” in tiled format. Tile size remains constant across all resolution levels (Fig. 3). The total number of pixels stored in the pyramid is increased by approximately one third compared to the original raster, but TIFF’s widespread support for various lossless compression algorithms (e.g., LZW, Deflate) mitigates some of this extra structural overhead. Note that using pixel rebinning instead of decimation (as in traditional astronomy image display tools) averages out background noise as one zooms out: this makes faint background features such as low surface brightness objects or sky subtraction residuals much easier to spot.

The default orientation for TIFF images (and most image formats) is such that the first pixel resides in the upper left corner of the viewport, whereas FITS images are usually displayed with the first pixel in the lower left corner. To comply with these conventions, STIFF flips the original image content along the y direction by proceeding through the FITS file backwards, line-by-line.

STIFF takes advantage of the TIFF header “tag” mechanism to include metadata that are relevant to the IIPIMAGE server and/or web clients. For instance, the ImageDescription tag is used to carry a verbatim copy of the original FITS header. Another set of information of particular importance, especially with floating point data, is stored in the SMinSampleValue and SMaxSampleValue tags: these are the minimum and maximum pixel values (S_{\min} and S_{\max}) that define the display scale. These values do not necessarily represent the full range of pixel values in the image, but rather a range that provides the best visual experience given the type of data. STIFF sets S_{\max} to the 999th permil of the image histogram by default. S_{\min} is computed in a way that the sky background S_{sky} should appear on screen as a dark grey $\rho_{\text{sky}} \approx 0.001$ (expressed as a fraction of the maximum display radiant emittance: $1 \equiv$ full white):

$$S_{\min} = \frac{S_{\text{sky}} - \rho_{\text{sky}} S_{\max}}{1 - \rho_{\text{sky}}}. \quad (1)$$

STIFF currently takes simply the median of all pixel values in the FITS file to compute S_{sky} , although better estimates could be computed almost as fast (Bertin and Arnouts, 1996).

Transcoding speed can be a critical issue, for instance in the context of real-time image monitoring of astronomy observations. On modern hardware, the current STIFF conversion rate for transcoding a FITS file to an IIPImage-ready tiled pyramidal TIFF ranges from about 5Mpixel/s to 25Mpixel/s (20-100MB/s) depending on the chosen TIFF compression scheme and system I/O performance. This means that FITS frames with



Figure 3: Illustration of tiled multi-resolution pyramid with 4 levels of resolution.

dimensions of up to $16k \times 16k$ pixels can be converted in a matter of seconds, and just-in-time conversion is a viable option for such images. Note that although STIFF is multithreaded, all calls to libtiff for writing tiles are done sequentially in the current implementation and there may, therefore, be some room for significant performance improvements.

2.4. Protocol and Server-Side Features

IIPIMAGE is based on the Internet Imaging Protocol (IIP), a simple HTTP protocol for requesting images or regions within an image, which allows the user to define resolution level, contrast, rotation and other parameters. The protocol was originally defined in the mid-1990s by the *International Imaging Industry Association* (Hewlett Packard, Live Picture, Eastman Kodak, 1997), but has since been extended for IIPIMAGE. The use of such a protocol provides a rich RESTful-like interface to the data, enabling flexible and consistent access to imaging data. IIPIMAGE is also capable of communicating using the simpler tile request protocols used by Zoomify or Deepzoom and the more recent IIIF image access API (International Image Interoperability Framework, Sanderson et al., 2013).

Table 1 lists the main commands already available in the original, cultural heritage-oriented version of IIPIMAGE. For a complete description of the protocol, see the full IIP protocol specification (Hewlett Packard, Live Picture, Eastman Kodak, 1997).

For this project the entire IIPIMAGE codebase was updated and generalized to handle up to 32 bits per pixel, with support for single precision floating point data. Support for double precision (at the expense of performance) would require a relatively simple code update. In addition, several extensions were implemented that allow the application of predefined colormaps to grayscale images, adjust the gamma correction, change the minimum and maximum cut-offs of the pixel value range, and that enable the export of image data profiles. A list of the new available commands is given in table 2.

Command	Description
FIF	Image path p . [FIF= p]
OBJ	Property/ies $text$ to be retrieved from image and server metadata. [OBJ= $text$]
QLT	JPEG quality factor q between 0 (worst) and 100 (best). [QLT= q]
SDS	Specify a particular image within a set of sequences or set of multi-band images. [SDS= s_1, s_2]
CNT	Contrast factor c . [CNT= c]
CVT	Return the full image or a region, in JPEG format. [CVT= $jpeg$]
WID	Width w in pixels of the full sized JPEG image returned by the CVT command (interpolated from the nearest resolution). [WID= w]
HEI	Height h in pixels of the full sized JPEG image returned by the CVT command (interpolated from the nearest resolution). [HEI= h]
RGN	Define a region of interest starting at relative coordinates x, y with width w and height h . [RGN= x, y, w, h]
ROT	Rotate the image by r (90, 180 or 270 degrees). [ROT= r]
JTL	Return a tile with index n at resolution level r , in JPEG format. [JTL= r, n]
SHD	Apply hillshading simulation with azimuth and altitude angles a, b . [SHD= a, b]
SPECTRA	Return pixel values in all image channels for a particular point x, y on tile t at resolution r in XML format. [SHD= r, t, x, y]

Table 1: Main commands available in IIPImage

2.4.1. Examples

In order to better understand how these commands can be used, here are several examples showing the typical syntax and usage for applying colormaps, setting a gamma correction and for obtaining a full bit-depth profile.

All requests take the general form:

```
<protocol>://<server address>/<iipsrv>?<IIP Commands>
```

The first IIP command must specify the image path and several IIP command–value pairs can be chained together using the separator & in the following way:

```
FIF=<image path>&<command>=<value>&<command>=<value>
```

Thus, a typical request for the tile that fits into the smallest available resolution (tile 0 at resolution 0) of a TIFF image named `image.tif` is:

```
http://server/iipsrv.fcgi?FIF=image.tif&JTL=0,0
```

Let us now look at some more detailed examples using the new functionality created for IIPImage. For example, in order

Command	Description
CMP	Set the colormap for grayscale images. Valid colormaps include GREY, JET, COLD, HOT, RED, GREEN and BLUE. [CMP=JET]
INV	Invert image or colormap. [Does not require an argument]
GAM	Set gamma correction to g . [GAM= g]
MINMAX	Set minimum min and maximum max for channel c . [MINMAX= $c:min,max$]
PFL	Request full bit-depth data profile for resolution r along the line joining pixel x_1, y_1 to x_2, y_1 . [PFL= $r:x_1, y_1-x_2, y_2$] <i>Note: Only horizontal ($y_1 = y_2$) and vertical profiles ($x_1 = x_2$) currently supported</i>

Table 2: List of new commands implemented in IIPImage.

to export a profile in JSON format from pixel location x_1, y_1 horizontally to pixel location x_2, y_2 at resolution r on image `image.tif`, the request would take the form:

```
FIF=image.tif&PFL=r:x1,y1-x2,y2
```

In order to request tile t at resolution r and apply a standard *jet* colormap to image `image.tif`, the request would take the form:

```
FIF=image.tif&CMP=JET&JTL=r,t
```

and the equivalent inverted colormap request:

```
FIF=image.tif&CMP=JET&INV&JTL=r,t
```

In order to obtain metadata containing the minimum and maximum values per channel:

```
FIF=image.tif&OBJ=min-max-sample-values
```

In order to request tile t at resolution r and apply a gamma correction of g and specify a minimum and maximum of m_1 and m_2 respectively for image band b :

```
FIF=image.tif&MINMAX=b:m1,m2&GAM=g&JTL=r,t
```

Commands are not order sensitive excepting JTL and CVT that must always be specified last.

2.5. Security

A client-server architecture also has the advantage in terms of control and security of the data as the raw data at full bit depth does not necessarily need to be made fully available to the end user. Indeed, the raw data need never be directly accessible by the public and can be stored on firewalled internal storage and only accessible via the IIPImage server. Thus only 8 bit processed data is ever sent out to the client and restrictions and limits can be applied if fully open access is not desired. The IIPImage server also contains several features for added security, such as a path prefix, which limits access to a particular subdirectory on the storage server. Any requests to images higher up or outside of this subdirectory tree are blocked.

If an even greater level of security is required on the transmitted data, the IIPIMAGE server can also dynamically apply a watermark to each image tile with a configurable level of opacity. Watermarking can be randomized both in terms of which tiles they are applied to as well as their position within the tile itself, making removal of watermarks extremely difficult.

2.6. Web Clients

Two web clients, developed using different approaches and different goals in mind, are presented in this paper as examples to illustrate the capabilities of the system.

The first one, known as VisiOMATIC, is built on top of the LEAFLET JavaScript mini-framework, and is designed to display large celestial images through a classic image tile-based view.

The second client builds on the existing IIPMooVIEWER client to demonstrate two experimental features more specifically relevant to planetary surface studies: hillshading and advanced compositing / filtering performed at the pixel level within the browser.

3. Astronomy Applications

3.1. Celestial Images

Two essential features of astronomy image browsers are missing in the IIPMooViewer client originally developed for cultural heritage applications: the handling of celestial coordinates and a comprehensive management system for vector layers (overlays). It soon became clear that developing such a system from scratch with limited human resources would raise severe maintenance issues and portability concerns across browsers and platforms. We investigated several JavaScript libraries that would provide such functionality and decided to build a new client, VisiOMATIC, based on the LEAFLET library (Agafonkin, 2010). LEAFLET is open-source and provides all the necessary functions to build a web client for browsing interactive maps. It is, in fact, not simply a client, but a small framework, offering features not directly available in standard JavaScript such as class creation and inheritance. It has a well-documented, user-friendly API and a rich collection of plug-ins that significantly boost its potential, while providing many advanced programming examples. Indeed, VisiOMATIC operates as a LEAFLET plug-in and as such comes bundled as a NodeJS package. Documentation for the VisiOMATIC API is available on the VisiOMATIC GitHub page⁵.

Once the iipsrv server has been installed, embedding a zoomable astronomy image in a web page with the VisiOMATIC and LEAFLET JavaScript libraries is very simple and can be done with the following code:

```
<div id="map"></div>
<script>
  var map = L.map('map'),
      layer = L.tileLayer.iip('/fcgi-bin/iipsrv.fcgi?
        FIF=/path/to/image.ptif').addTo(map);
</script>
```

LEAFLET was built from the ground up with mobile device support in mind. VisiOMATIC capitalizes on this approach by defining the current map coordinates at the center of the viewport instead of the mouse position. This also makes the coordinate widget display area usable for input, copy or paste as coordinates do not change while moving the mouse. Celestial coordinates are handled through a custom JavaScript library that emulates a small subset of the WCS standard (Calabretta and Greisen, 2002), based on the FITS header content transmitted by the IIPIMAGE server. Our simplified WCS library fits into LEAFLET's native latitude-longitude coordinate management system, giving access to all layer contents directly in celestial coordinates. This makes it particularly easy to synchronize maps that do not use the same projection for e.g., orientation maps, "smart" magnifiers, or multi-band monitoring.

Changing image settings is done by appending the relevant IIP commands (see e.g., Table 2) to the http GET tile requests. Metadata and specific data queries, such as profile extractions, are carried out through AJAX requests. VisiOMATIC also uses AJAX requests for querying catalogs from other domains, with the restriction that the *same origin security policy*⁶ present in current browsers requires that all requests transit through the image server domain, which must, therefore, be configured as a web proxy.

The VisiOMATIC website⁷ showcases several examples of applications built with the VisiOMATIC client. They involve large images of the deep sky stored in floating point format, including a one terabyte (500,000 × 500,000 pixels) combination of 250,000 exposures from the 9th Sloan Digital Sky Survey data release (Ahn et al., 2012), representing about 3TB worth of raw image data (Fig. 4).

Display performance with the VisiOMATIC client varies from browser to browser. Browsers based on the WEBKIT rendering engine (e.g., CHROME, SAFARI) generally offer the smoothest experience on all platforms, especially with complex overlays. User experience may also vary because of the different ways browsers are able to deal with data. For example, examining images at exceedingly high zoom levels and scrutinizing groups of pixels displayed as blocks is common practice among astronomers. LEAFLET takes advantage of the built-in resampling engines in browsers to allow image tiles to be zoomed in smoothly through CSS3 animations. VisiOMATIC uses the image-rendering CSS property to activate nearest-neighbor interpolation and have the pixels displayed as blocks at higher zoom levels. Although this works in, for example, FIREFOX and INTERNET EXPLORER 11, other browsers, such as CHROME, do not offer the possibility to turn off bilinear interpolation at the present time, and zoomed images will not appear pixelated in those browsers. Hopefully, it is expected that such residual differences will eventually disappear as browser technology converges over standards.

⁶The Cross-Origin Resource Sharing (CORS) mechanism implemented in modern browsers could in principle prevent that, but it is not supported by the main astronomy data providers at this time.

⁷<http://visiomatic.org>

⁵<https://github.com/astromatic/visiomatic>

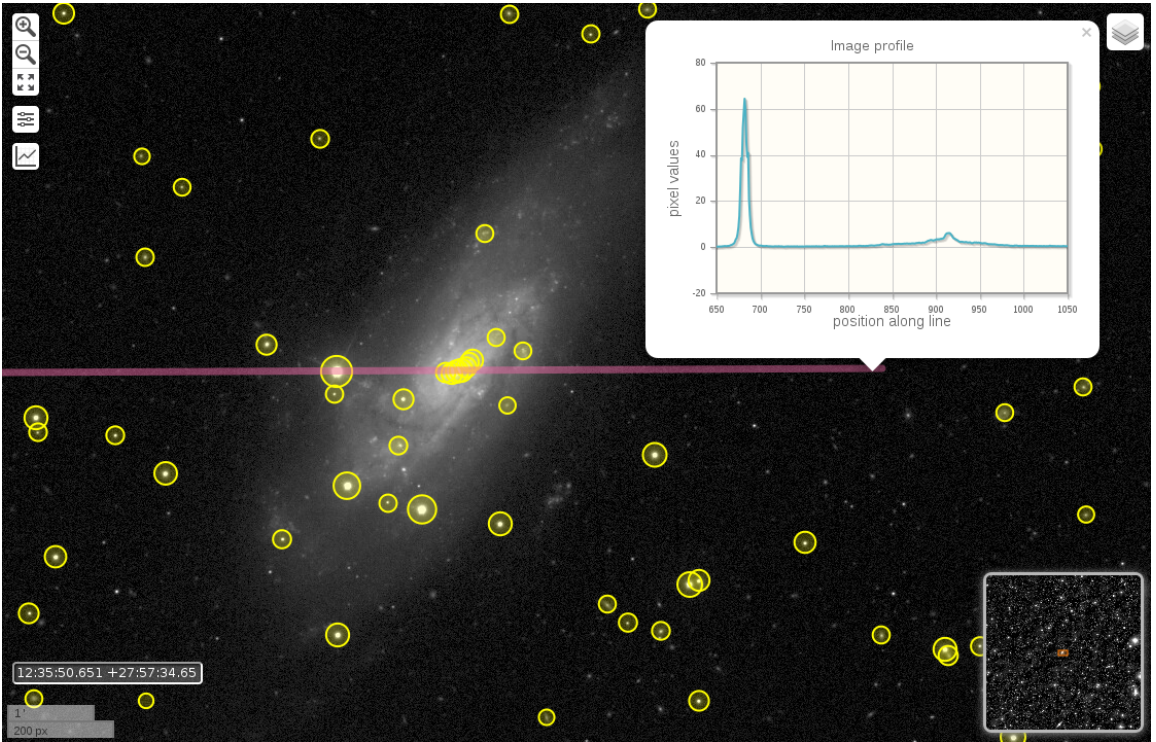


Figure 4: The VisiOmatic web client showing a part of an SDSS release nine image stack (Ahn et al., 2012) provided by the IIPImage server in the main layer, plus two vector layers superimposed. Yellow: local detections from the photometric SDSS catalog provided by the Vizier service (Ochsenbein et al., 2000). Purple: horizontal profile through the image extracted by the IIPImage server.

4. Planetary Science

Planetary Science data are largely heterogeneous with respect to the physical quantities they describe (chemical abundances, atmospheric composition, magnetic and gravitational fields of Earth-like planets and satellites, reflectance, surface composition) and with respect to the formats they are encoded in (raster, vector, time-series, in ASCII or various binary formats). Two scientific communities are essentially involved in Planetary Science research: astronomers and geologists / geophysicists.

Geographical Information Systems (GIS) are the basis for planetary surface studies but they often suffer from a lack of systematic and controlled access to pixel values for quantitative physical analyses on raster data (Marmo, 2012).

In Earth Sciences, distributors of GIS commercial software have been ready to exploit the potential of the Web. This is the case, for example, of the online ArcGIS WebMap Viewer⁸.

However, the difficulty in sending 16 or 32 bit precision scientific data using current web technologies has, hitherto, limited web visualization to public outreach applications such as

the Microsoft World Wide Telescope available for images of Mars (Scharff et al., 2011) or Google Earth for Mars⁹.

Nevertheless, remote scientific visualization has been achieved with tools such as JMars¹⁰ (Christensen et al., 2009) and HiView¹¹, which are both Java clients that aim to visualize both remote and local data. The first is GIS based (layer superposition oriented) while HiView is more a remote sensing software (raster manipulation oriented) that is an ad-hoc product for HiRISE¹² and which uses the JPIP protocol for remote access to JPEG2000 imagery.

The visualization system we propose already supports basic manipulations on raster layers, raster layer superposition and could easily manage vector layer creation and superposition. It, moreover, enables access to full precision pixel values and provides a simple and generic solution for planetary applications, efficiently and elegantly blending both GIS and remote sensing approaches.

⁹<http://www.google.com/earth/explore/showcase/mars.html>

¹⁰<http://jmars.mars.asu.edu>

¹¹<http://hirise.lpl.arizona.edu/hiview/>

¹²<http://hirise.lpl.arizona.edu>

⁸<http://www.arcgis.com/home/webmap/viewer.html>



Figure 5: Example of a planetary application using HRSC Mars images (ESA/DLR/FU Berlin/G. Neukum). The resulting color image is a linear combination of input channels. The mixing matrix is defined by a user-adjustable combination of Red, Green, Blue and a contrast factor for each input channel.

4.1. Color Compositing

Color compositing is an essential feature in both GIS and remote sensing applications and is used to point out differences in surface composition by performing on-demand composition of specific color bands. Interactive color composition on the Web can be achieved using the HTML5 canvas element which allows us to directly access and manipulate image pixel values. This, therefore, enables more complex real-time image processing directly within the client and we have developed a version of our client making extensive use of HTML5 canvas properties¹³ in order to implement on-demand color composition with multiple input channels (Fig. 5).

Color compositing performance depends essentially on canvas size (the overall image size is irrelevant as only the displayed part of the image needs to be processed). For the example cited above the processing time is about 1 to 3 ms per tile (256×256 pixels), depending on browser and client hardware.

4.2. Terrain Maps - Hillshading

High resolution 3D data is not easy to stream or to make multi-resolution. Furthermore as IIPImage is essentially image-centric, a 2D rendering approach to visualization was favored. In order to facilitate the use of DEM (digital elevation map) data, two approaches have been developed in our IIPImage framework.

The first approach is the dynamic application of custom colormaps to grayscale images. A new command CMP has been

added to the IIP protocol, which can also be useful for the visualization of other physical map data such as gas density, temperature or chemical abundances in real or simulated data.

In the second approach, elevation point data is converted to vector normal and height data at each pixel. In this form, they are also able to be stored within the standard TIFF format. The XYZ normal vectors can be packed into a 3 channel “color” TIFF, whereas the height data can be packed into a separate 1 channel monochrome TIFF. They are both, therefore, able to be tiled, compressed and structured into a multi-resolution pyramid for streaming with IIPImage. A basic rendering technique for DEM data is that of hill-shading (Horn, 1981) where a virtual directional illumination is used to create shading on virtual “hills”. A fast hill-shading algorithm has been implemented server-side in IIPImage and extended to 32 bit data allowing the user to interactively set the angle of incidence of the light source and view a dynamically rendered hill-shaded relief map. An example showing a Mars terrain map from Western Arabia Terra can be seen online¹⁴ and in Fig. 6.

5. Performance Analysis

Although the use of JPEG compression as a delivery format significantly reduces the bandwidth required, data access, dynamic processing and compression of 32 bit data can impose significant server-side overhead, that will ultimately dictate the maximum number of users that a server will be able to handle. Timings and memory usage depend on image type, server settings and commands in the query; we chose to focus on the typical case of browsing a large, single channel, single precision floating-point image stored with tiles of 256×256 pixels in size. In order to fully test this, we created a large $131,072 \times 131,072$ pixel FITS image by combining contiguous SDSS i-band images using the SWARP package (Bertin et al., 2002). This large image was then converted to a 92GB multi-resolution TIFF comprising 9 resolution levels using STIFF.

Our tests were performed on two Dell PowerEdge servers running GNU/Linux (Fedora distribution with kernel 3.11) and equipped with 2.6GHz processors, 32 and 48 GB of RAM and a Perc5i internal RAID controller. In order to check the influence of the I/O subsystem on server performance, we installed the TIFF file on two different types of RAID:

- a RAID 6 array of 12×3 TBytes SAS (6Gb/s) hard drives formatted with the XFS filesystem.
- a RAID 5 array of 6×1 Tbytes SATA3 (6Gb/s) solid-state drives (SSDs) formatted with the Ext4 filesystem.

On both systems we obtain a typical sequential read speed of 1.2 GB/s for large blocks; but obviously access times are much lower on the RAID of SSDs ($< 1\text{ms}$ vs 15ms for the one with regular hard drives).

The client consists of a third machine sending requests to any of the two servers through a dedicated 10GbE network. We

¹³<http://image.iap.fr/iipcanvas/hrsc.html>

¹⁴<http://image.iap.fr/iipdiv/hirise.html>

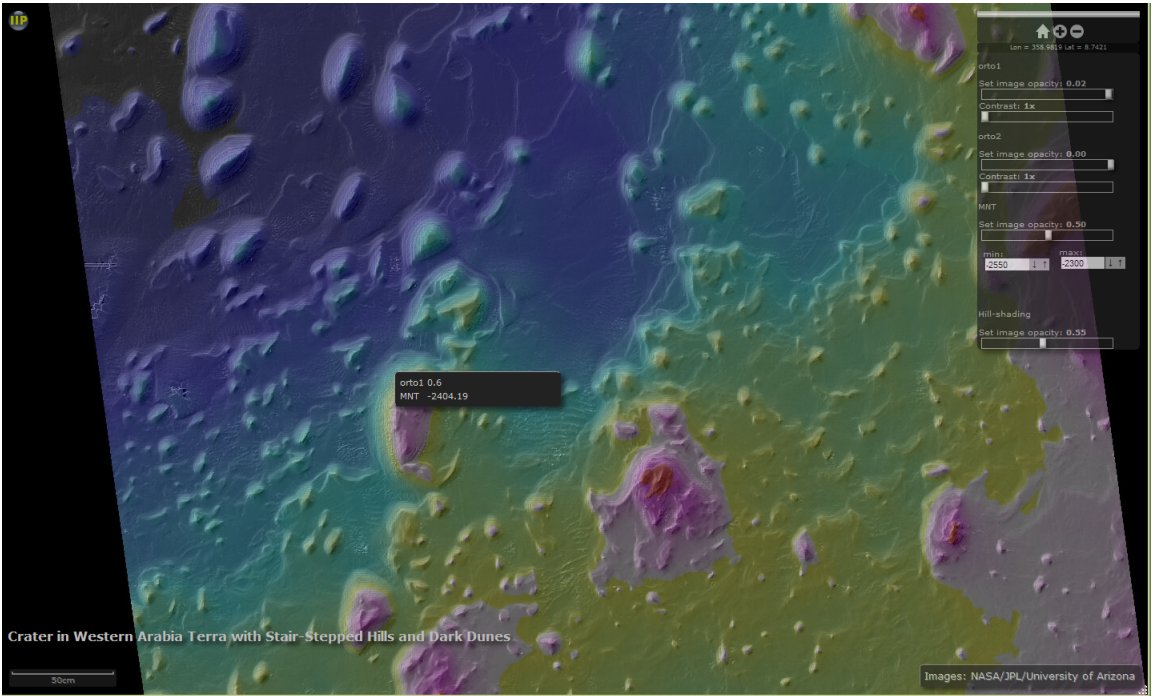


Figure 6: Example of a planetary web application using HiRISE Mars images (NASA/JPL/University of Arizona). The digital elevation model (a floating point raster) is displayed using the JET colormap with cuts set by the user from the control panel. Superimposed is the hill-shading layer computed by the IIPImage server from the DEM; the azimuth incidence angle can be adjusted from the control panel.

used a modified version of AB, the `APACHEBENCH` HTTP server benchmarking package, to send sequences of requests to random tiles among the 262,144 that compose the highest image zoom level. Appropriate system settings, as prescribed in [Veal and Foong \(2007\)](#), were applied server-side and client-side to ensure that both ends would stand the highest possible concurrency levels with minimum latencies and maximum throughput. We conducted preliminary tests through Apache’s `httpd`¹⁵, Lighty Labs’ `lighttpd`¹⁶, a combination of Nginx¹⁷ and Lighty Labs’ `spawn-fcgi` and finally LiteSpeed Technologies’ `OpenLiteSpeed`¹⁸. We found the latter to offer the best combination of performance and robustness, especially at high concurrency levels; hence all the requests to `iipsrv` in the tests reported below were served through `OpenLiteSpeed` (one single `lshttpd` instance).

5.1. Timings

Figure 7 shows the distribution of timings of the main tasks involved in the server-side processing of a tile, for several system and `iipsrv` cache settings. In order to probe the impact of

I/O latencies, we set up an experiment where the server system page cache is flushed and the `iipsrv` internal cache is deactivated prior to running the test (upper row of Fig. 7). With such settings most accesses to TIFF raw tiles do not benefit from caching. As a consequence, `iipsrv` timings are dominated by random file access times when the data are stored on spinning hard disks, with access latencies reaching up to ≈ 25 ms in unfavorable situations. As expected, switching to SSDs reduces the uncached file access latencies to less than 1ms.

However, in practice much better timings will be obtained with regular hard drives, as tiles are generally not accessed randomly. Moreover, leaving the system page cache un-flushed between test sessions when using spinning hard drives reduces access latencies to a few milliseconds (lower row in Fig. 7). Activating `iipsrv`’s internal cache system will further reduce latencies close to zero for tiles that were recently visited.

Further testing with TIFF images of different size was carried out in order to ensure that the system would also scale in terms of file size and the timings reported above remain roughly identical as file size increases up to at least 1.8TB.

5.2. Concurrency and Data Throughput

Each single-threaded FastCGI process takes about 5-10ms to complete, and is therefore capable of serving up to 100-200 256×256 “new” tiles per second. Higher tile serving rates

¹⁵<http://httpd.apache.org>

¹⁶<http://www.lighttpd.net>

¹⁷<http://nginx.org>

¹⁸<http://open.litespeedtech.com>

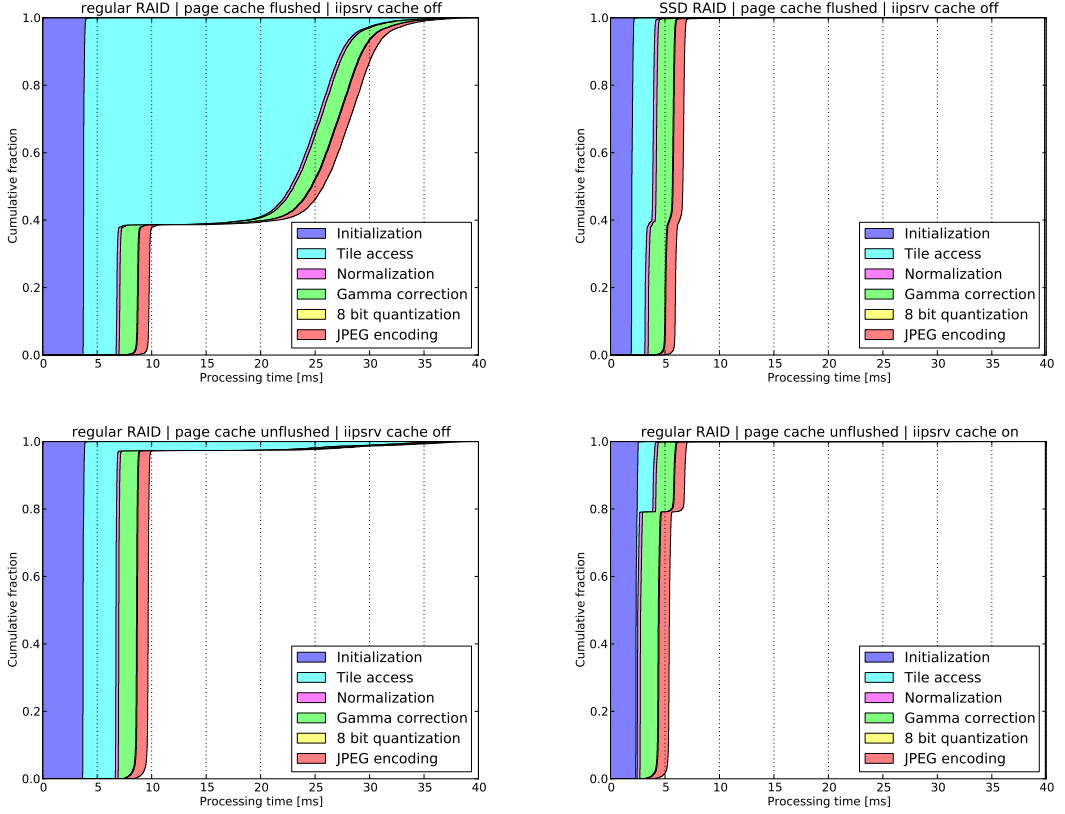


Figure 7: Cumulative distributions for the timings of the main tasks involved in the processing of random 256×256 pixel tiles in four different contexts (see text for details). “Initialization” is the time taken to initialize various objects and (re-)open the TIFF file that contains the requested raw tile (call to `TIFFOpen()`). “Tile access” is the time spent accessing and reading the content of a raw tile. “Normalization” and “Gamma correction” respectively measure the time it takes to apply intensity cuts and to compress the dynamic range of pixel values for the whole tile. “8 bit quantization” is the time spent converting the tile to 8 bit format, while “JPEG encoding” is the time taken to encode the tile in JPEG format.

are obtained by running several instances of `iipsrv` on servers with multiple CPU cores. But how is the system able to keep up with a large number of concurrent requests?

As Fig. 8 shows, the tile serving rate remains remarkably flat, and latency scales linearly with the concurrency level when the number of concurrent requests exceeds that of CPU cores. Setting a limit for average latency to ~ 500 ms for comfortable image browsing, we see that a single 12-core web server can handle ~ 700 concurrent 256×256 tile requests, which corresponds to about 100 users frantically browsing large, uncompressed, single-channel floating-point images. This estimation is well verified in practice, although it obviously depends on tile size and on the amount of processing carried out by `iipsrv`. Note that the average tile serving rate obtained with a single 12-core web server corresponds to a sustained data rate of 60MB/s for 256×256 tiles encoded at a JPEG quality factor of 90; higher JPEG quality factors bring the data rate close to the saturation limit of a 1GbE connection.

6. Conclusion and Future Work

A high performance web-based system for remote visualization of full resolution scientific grade astronomy images and data has been developed. The system is entirely open-source and capable of efficiently handling full resolution 32 bit floating point image and elevation map data.

We have studied the performance and scalability of the system and have shown that it is capable of handling terabyte-size scientific-grade images that can be browsed comfortably by at least a hundred simultaneous users, on a single server.

By using and extending an existing open source project, a system for astronomy has been put together that is fully mature, that will benefit from the synergies of the wider scientific imaging community and that is ready for use in a busy production environment. In addition the `IIPImage` server, is distributed as part of the default Debian, Ubuntu and Fedora package repositories, making installation and configuration of the system very

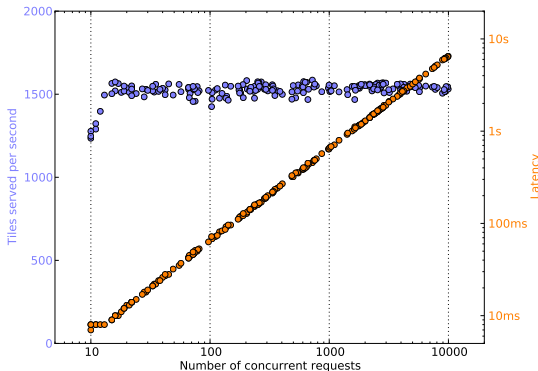


Figure 8: Tile serving rate (in blue) and latency (in orange) as a function of the number of concurrent tile requests using 12 instances of `iipsrv` on a 12-core server equipped with an SSD RAID.

straightforward. All the code developed within this project for `iipsrv` has been integrated into the main code base and will form an integral part of the 1.0 release. However, there are still many potential directions for improvements, both server-side and client-side. Most importantly:

- The TIFF storage format used on the server currently restricts pixel bit depth, the number of image channels, and I/O performance (through `libtiff`). A valid alternative to TIFF could be the Hierarchical Data Format Version 5 (HDF5) (HDF Group, 2000), which provides a generic, abstract data model that enables POSIX-style access to data objects organized hierarchically within a single file; some radio-astronomers have been trying to promote the use of HDF5 for storing massive and complex astronomy datasets (Masters et al., 2012). A more radical approach would be to adopt JPEG2000 as the archival storage format for astronomy imaging archives (Kitaef et al., 2014), which could also remove the need for transcoding images for visualization purposes.
- Additional image operations could be implemented within `iipsrv`, including real-time hyperspectral image processing and compositing.
- Although the `IIPImage` image tile server already supports simple standard tile query protocols and interfaces easily with most image panning clients, a welcome addition would be to offer support for the more GIS-oriented WTMS (Web Map Tile Service) protocol (Open Geospatial Consortium, 2010).
- The International Virtual Observatory Alliance (IVOA) has agreed on a standard set of specifications for discovering and accessing remote astronomical image datasets: the Simple Image Access Protocol (SIAP) (Tody et al., 2011). The response to an SIAP query consists of meta-data and download URLs for matching image products.

Current SIAP specifications¹⁹ do not provide specific ways to access pyramids of tiled images. Still, support for SIAP could be implemented within or outside of `iipsrv` for generating, for example, JPEG cutouts or lists of tiles that match a given set of coordinates/field of view/pixel scale.

- Both `IIPMooVIEWER` and `LEAFLET` clients require all layers displayed on a map at the same moment to share the same “native” pixel grid (projection). Although this limitation does not prevent “blinking” images with different pixel grids, it precludes overlapping different observations/exposures on screen. For instance it makes it impossible to display accurately the entire focal plane of a mosaic camera on a common viewport, without prior resampling. Having different images with different native pixel grids sharing the same map would require the web-client to perform real-time reprojection. Client-side reprojection should be possible e.g., with version 3 of the `OPENLAYERS` library²⁰.

7. Acknowledgments

The authors would like to thank the anonymous referees whose comments helped not only in improving the clarity of this paper, but also the performance of the code.

CM wishes to acknowledge Prof. Joe Mohr for hospitality at USM, Munich, and the SkyMapper team, in particular Prof. Brian Schmidt, Dr. Patrick Tisserand and Dr. Richard Scalzo for support during her stay at MSO-ANU, Canberra where part of this work was completed. EB thanks Raphael Gavazzi, Valérie de Lapparent, and the Origin and Evolution of Galaxies group at IAP, Paris for financial support with the `VisiOMATIC` hardware, and Dr. Hervé Bouy at CAB, Madrid for providing content for the `VisiOMATIC` demos.

The `VisiOMATIC` client implements services provided by the Sesame Name Resolver and the `VizieR` catalog access tool developed at CDS, Strasbourg, France.

Some of our demonstration data are based on SDSS-III²¹ images. Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New

¹⁹<http://www.ivoa.net/documents/SIA/>

²⁰<http://ol3js.org/>

²¹<http://www.sdss3.org>

York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

References

- Agafonkin, V., 2010. Leaflet - a JavaScript library for mobile-friendly maps. <http://www.leafletjs.com>.
- Ahn, C.P., Alexandroff, R., Allende Prieto, C., Anderson, S.F., Anderton, T., Andrews, B.H., Aubourg, É., Bailey, S., Balbinot, E., Barnes, R., et al., 2012. The Ninth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey. *Astrophysical Journal Supplement Series* 203, 21. [1207.7137](https://doi.org/10.1088/0007-1385/203/1/21).
- Albrecht, M.A., Brighton, A., Herlin, T., Bieriichel, P., Durand, D., 1997. Access to Data Sources and the ESO SkyCat Tool, in: Hunt, G., Payne, H. (Eds.), *Astronomical Data Analysis Software and Systems VI*, p. 333.
- Bertin, E., 2012. Displaying Digital Deep Sky Images, in: Ballester, P., Egret, D., Lorente, N.P.F. (Eds.), *Astronomical Data Analysis Software and Systems XXI*, p. 263.
- Bertin, E., Arnouts, S., 1996. SExtractor: Software for source extraction. *Astronomy and Astrophysics Supplement Series* 117, 393–404.
- Bertin, E., Mellier, Y., Radovich, M., Missonnier, G., Didelon, P., Morin, B., 2002. The TERAPIX Pipeline, in: Bohlender, D.A., Durand, D., Handley, T.H. (Eds.), *Astronomical Data Analysis Software and Systems XI*, p. 228.
- Bonnarel, F., Paillou, P., Ochsenein, F., Creze, M., Egret, D., 1994. Aladin: Towards an Interactive Atlas of the Digitized Sky, in: D. R. Crabtree, R. J. Hanisch, & J. Barnes (Ed.), *Astronomical Data Analysis Software and Systems III*, pp. 215–218.
- Brown, M., 1996. FastCGI Specification. <http://www.fastcgi.com>.
- Budman, G., 2011. Price gap: Storage vs Bandwidth. <http://blog.backblaze.com/2011/06/22/price-gap-storage-vs-bandwidth>. Last accessed 07/03/2012.
- Calabretta, M.R., Greisen, E.W., 2002. Representations of celestial coordinates in FITS. *Astronomy and Astrophysics* 395, 1077–1122. [astro-ph/0207413](https://doi.org/10.1051/astro-ph/0207413).
- Christensen, P.R., Engle, E., Anwar, S., Dickenshied, S., Noss, D., Gorelick, N., Weiss-Malik, M., 2009. JMARS - A Planetary GIS. AGU Fall Meeting Abstracts, A6.
- Draper, P.W., 2000. GAIA: Recent Developments, in: Manset, N., Veillet, C., Crabtree, D. (Eds.), *Astronomical Data Analysis Software and Systems IX*, p. 615.
- Faas, F.G.A., Avramut, M.C., Berg, B.M.v.d., Mommaas, A.M., Koster, A.J., Ravelli, R.B.G., 2012. Virtual nanoscopy: Generation of ultra-large high resolution electron microscopy maps. *The Journal of Cell Biology* 198, 457–469.
- Federl, P., Grimstrup, A., Kiddle, C., Taylor, A.R., Robinson, K., Stephure, M., Yee, G., 2012. Remote Visualization of Large Multi-dimensional Radio Astronomy Data Sets, in: Ballester, P., Egret, D., Lorente, N.P.F. (Eds.), *Astronomical Data Analysis Software and Systems XXI*, p. 95.
- Federl, P., Grimstrup, C., A.K., Taylor, A.R., 2011. On-Line Access and Visualization of Multi-Dimensional FITS Data, in: I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots (Eds.), *Astronomical Society of the Pacific Conference Series*, pp. 467–470.
- HDF Group, 2000. Hierarchical data format version 5.
- Hewlett Packard, Live Picture, Eastman Kodak, 1997. Internet Imaging Protocol specification. <http://iipimage.sourceforge.net/IIPv105.pdf>.
- Horn, B., 1981. Hill shading and the reflectance map. *Proceedings of the IEEE* 69, 14–47.
- Hughitt, V.K., Ireland, J., Lynch, M.J., Schmeidel, P., Dimitoglou, G., Müller, D., Fleck, B., 2008. Helioviewer: A web 2.0 tool for visualizing heterogeneous heliophysics data. AGU Fall Meeting Abstracts -1, 1617.
- Husz, Z., Burton, N., Hill, B., Milyaev, N., Baldock, R., 2012. Web tools for large-scale 3D biological images and atlases. *BMC Bioinformatics* 13, 122.
- Joye, W.A., Mandel, E., 2003. New Features of SAOImage DS9, in: Payne, H.E., Jedrzejewski, R.L., Hook, R.N. (Eds.), *Astronomical Data Analysis Software and Systems XII*, p. 489.
- Kapadia, A., 2013. AstroJS library. <http://www.astrojs.org>. Last accessed 12 August 2013.
- Khronos Group, 2013. WebGL Specification Version 1.0.2. <http://www.khronos.org/registry/webgl/specs/1.0/>. Last accessed 29/07/2014.
- Kitaeff, V.V., Cannon, A., Wicenc, A., Taubman, D., 2014. Astronomical Imagery: Considerations for a Contemporary Approach with JPEG2000. ArXiv e-prints [1403.2801](https://arxiv.org/abs/1403.2801).
- Kitaeff, V.V., Wu, C., Wicenc, A., Cannon, A.D., Vinsen, K., 2012. Skuare-View: client-server framework for accessing extremely large radio astronomy image data, in: *Proceedings of the 2012 workshop on High-Performance Computing for Astronomy Data*, ACM, New York, NY, USA. p. 25–32.
- Lahner, C., Aitken, G., Shindo, J., Pillay, R., Martinez, K., Lewis, P., 2002. EROS: an open source multilingual research system for image content retrieval dedicated to conservation-restoration exchange between cultural institutions, in: *ICOM Committee for Conservation, ICOM-CC : 13th Triennial Meeting, Rio de Janeiro, 22-27 September 2002, ICOM-CC; James & James, Rio de Janeiro, Brazil*, p. 287–294.
- Lowe, S., 2011. jsfits repository. <https://github.com/slowe/jsFITS>. Last accessed 12 August 2013.
- Malapert, J.C., Marseille, M., 2012. SITools2: A Framework for Archival Systems, in: Ballester, P., Egret, D., Lorente, N.P.F. (Eds.), *Astronomical Data Analysis Software and Systems XXI*, p. 821.
- Mandel, E., 2014. JS9 website. <http://js9.si.edu/>. Last accessed 17 July 2014.
- Marmo, C., 2012. Development of a Planetary Web GIS at the “Photothèque Planétaire” in Orsay, in: Ballester, P., Egret, D., Lorente, N.P.F. (Eds.), *Astronomical Data Analysis Software and Systems XXI*, p. 411.
- Martinez, K., Cupitt, J., Perry, S., 1998. High resolution colorimetric image browsing on the web. *Computer Networks and ISDN Systems* 33, 399–405.
- Martinez, K., Cupitt, J., Saunders, D., Pillay, R., 2002. Ten years of art imaging research. *Proceedings of the IEEE*, 28–41.
- Masters, J., Alexov, A., Folk, M., Hanisch, R., Heber, G., Wise, M., 2012. The AstroHDF Effort, in: Ballester, P., Egret, D., Lorente, N.P.F. (Eds.), *Astronomical Data Analysis Software and Systems XXI*, p. 871.
- Muller, D., Fleck, B., Dimitoglou, G., Caplins, B.W., Amadigwe, D.E., Ortiz, J.P.G., Wamsler, B., Alexanderian, A., Hughitt, V.K., Ireland, J., 2009. Jhelioviewer: Visualizing large sets of solar images using jpeg 2000. *Computing in Science and Engineering* 11, 38–47.
- Ochsenein, F., Bauer, P., Marcout, J., 2000. The VizieR database of astronomical catalogues. *Astronomy and Astrophysics Supplement Series* 143, 23–32. [astro-ph/0002122](https://doi.org/10.1051/astro-ph/0002122).
- Open Geospatial Consortium, 2010. OpenGIS Web Map Tile Service Implementation Standard. <http://www.opengeospatial.org/standards/wmts>. Last accessed 07/28/2014.
- Pence, W.D., Seaman, R., White, R.L., 2009. Lossless Astronomical Image Compression and the Effects of Noise. *PASP* 121, 414–427. [0903.2140](https://doi.org/10.1086/314003).
- Pence, W., White, R. L., Greenfield, P., Tody, D., 2000. A FITS image compression proposal, in: *Astronomical Data Analysis Software and Systems IX*, N. Manset, C. Veillet, D. Crabtree, p. 551.
- Pitzalis, D., Pillay, R., Lahner, C., 2006. A new concept in high resolution internet image browsing, in: 10th International Conference on Electronic Publishing, Bulgarian Academy of Sciences, Bulgaria, Bansko, Bulgaria. pp. 75–85. 10th International Conference on Electronic Publishing, organised by the Bulgarian Academy of Sciences, Bulgaria, 14-16 June 2006.
- Sanderson, S., Albritton, B., Enders, M., 2013. International Image Interoperability Framework. <http://lib.stanford.edu/iiif>.
- Schaaff, A., Boch, T., Fernique, P., Kaestlé, V., 2012. The CDS at the Age of Multitouch Interfaces and Mobility, in: Ballester, P., Egret, D., Lorente, N.P.F. (Eds.), *Astronomical Data Analysis Software and Systems XXI*, p. 443.
- Scharff, E.B., Beyer, R.A., Broxton, M., Lundy, M., Fay, J., Turcan, P., Fay, D., Messeri, L., 2011. WorldWide Telescope Mars, in: *Lunar and Planetary Institute Science Conference Abstracts*, pp. 2337–2338.
- Tody, D., Plante, R., Harrison, P., 2011. IVOA Recommendation: Simple Image Access Specification Version 1.0. ArXiv e-prints [1110.0499](https://arxiv.org/abs/1110.0499).
- Vanilist, M., 1990. SAOImage. Smithsonian Astrophysical Observatory. Report for the period through Jan 1990., in: *Bulletin of the American Astronomical Society*, p. 935.
- Veal, B., Foong, A., 2007. Performance scalability of a multi-core web server, in: *Proceedings of the 3rd ACM/IEEE Symposium on Architecture for Networking and Communications Systems*, ACM, New York, NY, USA. pp.

57–66.

- W3C, 2011. W3C Working Group Note 12 May 2011. <http://www.w3.org/TR/css-2010/>. Last accessed 29/07/2014.
- W3C, 2012. HTML5 A vocabulary and associated APIs for HTML and XHTML. <http://www.w3.org/TR/html5/>. Last accessed 07/03/2012.
- Wells, D.C., Greisen, E.W., Harten, R.H., 1981. FITS - a flexible image transport system. *Astronomy and Astrophysics Supplement Series* 44, 363.
- Woodring, J., Mniszewski, S., Brislawn, C., DeMarle, D., Ahrens, J., 2011. Revisiting wavelet compression for large-scale climate data using jpeg 2000 and ensuring data precision, in: *Large Data Analysis and Visualization (LDAV)*, 2011 IEEE Symposium on, pp. 31–38.

ISBN 978-82-326-6787-1 (printed ver.)
ISBN 978-82-326-6204-3 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)



NTNU

Norwegian University of
Science and Technology