Journal of Cultural Heritage 63 (2023) 81–89

Contents lists available at ScienceDirect

Journal of Cultural Heritage

journal homepage: www.elsevier.com/locate/culher

Original article

Subjective Quality Evaluation of Alternative Imaging Techniques for Microfiche Digitization

Dipendra J. Mandal*, Hilda Deborah, Marius Pedersen

Department of Computer Science, Norwegian University of Science and Technology (NTNU), Teknologivegen 22, Gjøvik, 2815, Norway

A R T I C L E I N F O

Article history:
Received 10 March 2023
Revised 4 July 2023
Accepted 24 July 2023
Available online 4 August 2023

Keywords:
Microfiche
Image Quality Assessment
Subjective experiment
Dead Sea Scrolls

A B S T R A C T

Before the advent of digital formats, microfiche, a type of microform, was widely utilized for archiving and preserving historical documents. As a result, numerous historical collections and documents can only be found in microfiche format, transforming them into valuable artifacts and indispensable aspects of our cultural heritage. Although microfiche can last a long time, it is still susceptible to damage and requires digitization for preservation and broader accessibility. In addition, traditional microfiche readers are not always available and are primarily designed for reading rather than digitization. In this study, we evaluated the performance of two alternative imaging devices compared to a traditional microfiche reader and the impact of enhancement on image quality using subjective image quality assessment. The experiments were carried out in a controlled environment with twenty-one participants, including an expert. Our results showed that the reproduction of alternative devices was preferred over that of a traditional microfiche reader. Furthermore, our results demonstrate that image enhancement techniques significantly improved image quality. This study suggests that alternative imaging devices may be a viable option for digitizing microfiche and improving access to historical collections.

1. Introduction

Microform refers to a method of storing information, such as a document or image, in a much smaller form than the original. Information can be stored on a microfilm, microfiche, or other types of microforms. This storage method is useful because it saves space and makes it easier to store large amounts of information in a compact and manageable form. Additionally, the microform can be easily reproduced, allowing for the creation of multiple copies that multiple users can access. Microfiche is a type of microform that consists of a sheet of film containing multiple microimages in a reduced form (commonly referred to as the reduction ratio) and is arranged in a grid-like format [1]. The reduction ratio refers to the extent to which a document is visually scaled down (miniaturized) through photography. It is represented as a ratio between the original linear size and the linear size of the microform image. For instance, if an item is filmed at a reduction ratio of 20 to 1 (20:1), it means it has been reduced in size by a factor of twenty. Reduction ratios are often denoted as 20x, 40x, and so forth. Before the advent of advanced digital technologies, microforms were the exclusive method for archiving and conserving extensive documents, such as newspapers. The cultural heritage domain quickly adopted this technique to capture its collections to ensure preservation, ease of access, and greater distribution. Today, many historical collections can be found only on microforms. For example, Papua New Guinea's colonial-era history is only available in the form of microfiche [2].

Microforms contain an emulsion layer (a thin coating of light-sensitive material) embedded in a base: cellulose (typically nitrate, acetate, or triacetate) or polyester (plastic). Because of its flammability and rapid deterioration, cellulose nitrate was replaced by safer alternatives such as cellulose acetate and cellulose triacetate. The acetate film is prone to rapid deterioration due to fluctuations in temperature and humidity, causing distortion and warping of the emulsion. This deterioration produces acetic acid, resulting in the vinegar syndrome [3]. With better stability, durability, and tear resistance, polyester film became the preferred choice for modern microform production in the 1980s [4,5]. Three main types of emulsion layers used in microforms are silver halide, diazo, and vesicular [4]. Silver halide films are highly light-sensitive, capturing extensive detail and offering a wide tonal range. When properly prepared and stored, they can last up to 500 years [6]. However, diazo films are prone to image fading and loss, with a useful life

* Corresponding author.
E-mail address: dipendra.mandal@ntnu.no (D.J. Mandal).

https://doi.org/10.1016/j.culher.2023.07.014

of approximately 100 years, and are not considered archival quality [7]. Vesicular films offer scratch resistance but can distort under intense use or high heat, lasting between 10 to 100 years when stored appropriately [8].

Microfiche materials, despite their longer life span, are prone to physical degradation caused by various factors [9]. Exposure to light, fluctuating temperatures and humidity levels, improper handling, and poor storage conditions can lead to the deterioration of the microfiche. This degradation often manifests as brittleness in the film base, resulting in compromised image quality or even complete disintegration. Furthermore, microfiche collections are vulnerable to environmental hazards such as water damage, mold growth, pests, and fire. These hazards can lead to the loss of information and render microfiche unusable. Thus, digitizing microfiche can be a valuable solution for addressing such preservation challenges. Nevertheless, preserving the original microfiche could still be necessary for specific purposes. Therefore, a combined approach to digitization and proper storage measures is essential.

Gaining access to numerous historical documents stored in libraries worldwide can prove challenging due to the fragile state of the objects. In some instances, manuscripts or fragments have even been lost [10]. This means that their microfiche copies might be the last surviving records of the objects or the only accessible option for the wider public. However, microfiche cannot be read directly by the human eye and requires a special device for enlarging, printing, and scanning the microforms into a readable format. Only a few specialized archives or libraries still possess microfiche readers to facilitate access to their microform collections.

With the adoption of digital technologies, the digitization of microfiche can offer numerous benefits [11]. With digital formats, they can be accessed and viewed on computers and other digital devices, making it easier to share microfiche collections with a wider audience, regardless of their location. Digitization also makes storing and managing microfiche collections easier, since they no longer require a physical storage space. Additionally, digitization can help to preserve the information in microfiche documents by reducing wear and tear on the physical copies and by making backup copies to guard against loss or damage. Digitization often results in reproductions that differ from the original in various ways. Digitizing microfiche, which requires magnification (the usual reduction factors are from 24, 42 or 48, 96 to 1 [12]), is likely to introduce attributes that can affect the final quality of the reproduction, e.g., noise, distortion, and artifacts. Figure S1, included as a supplementary file, provides a visual illustration of such artifacts showing an original document image and a magnified image obtained from the microfiche. Quality assessment should be included as an essential part of the digitization process to ensure that the digitized object maintains completeness, fidelity, and legibility compared to the original. Both objective and subjective methods [13,14] can be used to assess image quality.

So far, to our knowledge, very few studies have been conducted in the quality assessment of microfiche digitization. Some of the work we found on microfiche quality assessment have concentrated on comparing the quality of the original paper document with its microfiche version. One such study, conducted by Lee et al. [15], used various quantitative measures to evaluate the fidelity of the original paper document and its 35 mm microfilm copies. The study results indicated that the copy was not fully faithful to the original document. In another study by Duff et al. [16] on the early Canadiana material collection, a comparison was made between the original materials and their microfiche counterparts. The findings showed a significant difference between the two formats, with most of the participants preferring the paper format over the digital version. As of today, the original of these are not easily accessible and, in some cases, may not even be available.

Numerous initiatives to convert microfiche into digital format are underway worldwide, utilizing established and innovative digitization methods and technologies. Scanning documents as images and then utilizing Optical Character Recognition (OCR) algorithms to convert the images into text is a widely employed approach for digitizing historical text collections [17,18]. Recently, in the year 2020, the National Security Research Center (NSRC) conducted a trial of an artificial intelligence/ machine learning system aimed at digitizing certain microfilm and microfiche document collection [19]. Deborah and Mandal [20] used objective quality assessment to evaluate the performance of two imaging devices as alternatives to the traditional microfiche reader. They assessed the legibility of the images obtained by these devices when inputted to an OCR and used Levenshtein distance [21] as the text similarity measure. The results showed the superiority of a flatbed scanner at 4k dpi resolution to a traditional microform reader. It is also worth mentioning that to achieve those results, the images obtained by the microform reader had to undergo a post-processing step, meanwhile no enhancement was applied to those from the flatbed scanner. Since the study utilized an OCR, the focus was on microfiche containing typewritten text. However, microfiche materials may also contain handwritten texts that off-the-shelf OCR systems cannot recognize and also photographs where the legibility evaluation may be less relevant. To address this limitation, our study incorporated microfiche materials containing typewritten text, photographs of natural scenes, and photographs of ancient handwritten fragments. Moreover, we conduct a subjective quality assessment with human observers to evaluate the quality of the scanned images.

Enhancing images through image processing is a widely used practice that aims to improve their visual quality. By adjusting image attributes, enhancement can produce a more aesthetically pleasing result for a given scenario [22]. The contrast of an image is widely recognized as a significant quality attribute [23]. Improving contrast is commonly believed to enhance the perceived quality of most natural images [24–26]. Contrast Stretching (CS) is a technique commonly used to enhance low-contrast images, which involves using a piecewise linear curve to expand the dynamic range of gray levels [27]. In our experiment, we applied CS as a post-processing to find out if it enhances the quality of microfiche materials.

2. Research Aim

Many historical collections and documents can only be found in microfiche format, making them valuable artifacts of cultural heritage. Microfiche requires microfiche readers, an imaging device that can enlarge the content of microfiche, allowing it to be read directly by the human eye. Such devices are primarily designed for reading rather than digitization, and only a limited number of specialized archives or libraries may still have them available. In addition, over time, these devices can suffer from various problems that affect their usability. Some common issues include yellow screens (due to aging and prolonged exposure to light), defective reels, degraded image quality, and mechanical failures. These problems can significantly impact the ability to access and read microfiche materials effectively. For instance, a discolored screen can negatively impact the contrast and readability of the microfiche images. Given the importance of preserving and accessing microfiche materials, exploring alternative options that can enhance the reading experience and facilitate the efficient utilization of microfiche becomes crucial. Therefore, this research aims to identify alternative imaging devices compared with traditional microfiche readers as a viable option for digitizing microfiche and improving access to historical collections. To incorporate a wider variability in the types of records, we used microfiche that contains typewrit-

Table 1
Specification and characteristics comparison of three imaging setups. Note that this summary is formulated within the specific context of reading 105 mm × 148 mm microfiches in a monochrome setup.

Factors	Microform reader	Flatbed scanner	In-house film scanner [28]
Model	Zeutschel delta plus	Epson Perfection 4870 Photo	QHY600 16BIT BSI, atx-i 100mm F2.8 FF MACRO
Compatible input types (non-exhaustive)	Microfiche, microcards, 16/35 mm roll microfilm, photographic slides, negatives, 35 mm perforated films	A4 size document, transparencies, photos, 35 mm films, negatives, 4"×5" formats	35 mm photographs and motion picture films, small objects of different kinds
Max. scan area	35×47 mm	216×297 mm	35×40 mm
Max. fiche per scan	$1/n$	$2n$	$6/n$
Effective pixels	10 MP	40,800×56,160 at 4800 dpi	9,576×6,388 (±60 MP)
Illumination	Custom calibrated LED array	Cold cathode fluorescent lamp	Calibrated LEDs
Throughput speed	Medium (±0.3 sec/ image)	High (±0.027 sec/ line)	High (±0.4 sec/ image) *
Operation ease	Low	High	High

*The speed was determined based on the camera's provided specification of 2.5 frames per second (fps) for 16-bit output.

Table 2
Summary of microfiche specification used in this experiment. The assumption is that the accompanying microfiche corresponds to the same time period based on its publication date.

Microfiche	Type	Reduction Ratio	Sheet per Fiche	Time Period
Handbook for Evaluating Microfiche Readers	positive microfiche	1:20	60	1975
Allegro Qumarn Collection	positive silver halide microfiche	1:19	50	1996
Dead Sea Scrolls	positive silver halide microfiche	1:13	50	1992

ten text, photographs of natural scenes, and photographs of ancient handwritten fragments.

3. Materials and Methods

3.1. Imaging Approaches

Access to microfiche readers or other similar microform reader machines is limited today, although they were once a central aspect of archiving. The main challenge with reading microfiche is its significantly reduced size. However, with advances in optical technology in recent years, it is highly probable that alternative imaging solutions are available. The first option to consider is a flatbed scanner, which now offers resolutions up to 6400 dpi. Additionally, by using macro lenses alongside a high-resolution camera, it may be possible to resolve the reduction ratio of microfiche. Based on these considerations, we have opted to utilize a professional grade flatbed scanner at 4800 dpi (FBS) and an in-house film scanning (IFS) system, which includes a monochrome camera and a macro lens, as viable alternatives to a microform reader (FSL). Table 1 summarizes the specifications and characteristics of the imaging devices used in the experiment. For further details, we refer the reader to Deborah and Mandal [20].

3.2. Microfiche Materials

The experiment used microfiche materials from three different sources. They were microfiches provided by 1) a handbook for evaluating microfiche readers [29], 2) the Allegro Qumran Collection on Microfiche [30], and 3) the Dead Sea Scrolls on Microfiche [31]. The microfiche from the first source primarily contained textual content, including typewritten text with various fonts and font sizes. On the other hand, the other two microfiche sources consisted of scroll fragments containing biblical and non-biblical writings. Additionally, the microfiche included photographs of Qumran caves and natural scenes that capture images of natural landscapes and outdoor environments. Fig. 1 illustrates an example of the figure obtained from the IFS device for three distinct microfiches used in an experiment. A summary of the technical specifications for the microfiche is presented in Table 2. It is important to note that the performance of the devices could be impacted by settings.

3.3. Image Acquisition

Each of the three microfiches (containing text, natural scenes, and fragments) was scanned and captured using all three devices. We used a microform reader (Zeutschel Delta Plus), made available through the local library of Gjøvik, Norway. This device required manual lens adjustment for each page within the microfiche. The resulting images were saved in TIF format. We used a professional-grade flatbed scanner manufactured by Epson, which allowed us to scan two complete microfiches in a single capture. This significantly improved operational efficiency by eliminating the need for manual adjustments of equipment and materials prior to each image capture. After scanning, each page of the microfiche was cropped and saved as a TIF file. The in-house multispectral film scanner used for this study is equipped with an LED-based system and a monochrome camera with a macro lens. Grayscale images were captured using a single light source at a wavelength of 415.5 nanometers. The maximum scanning area is determined by its field of view, and the design is optimized to capture images in a transmissive mode. It allowed capturing up to six pages within a single microfiche per scan. Refer to Figure S2, added as a supplementary file, for a detailed setup schematic.

3.4. Psychovisual experiment

A subjective experiment was designed to evaluate the overall image quality of the digitized microfiche using three different imaging devices (FSL, FBS, and IFS). Several psychometric methods exist in the literature for measuring image quality [32–35], such as paired comparison, rank ordering, categorical sort, and magnitude estimation. In this study, a force-choice pair comparison [36] experiment was designed without ties, i.e., observers were forced to choose one of the two preferences randomly if they found a tie between the stimuli. This method was chosen because of its relative simplicity compared to other methods and because it is better

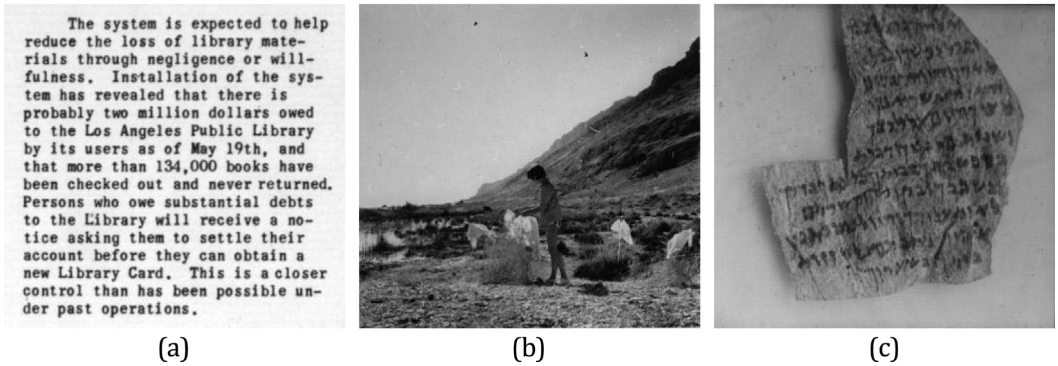

Fig. 1. Illustration of microfiches obtained using the IFS device. (a) Text from a Handbook for Evaluating Microfiche Readers, (b) a natural scene from the Allegro Qumran Collection, and (c) fragments from the Dead Sea Scrolls.

at finding differences between images. In a pair comparison experiment, the task is typically to indicate the preferred option from each pair of stimuli rather than assigning a quality score to each stimulus. A web-based tool called QuickEval [37] was utilized to carry out the experiment. This tool is specifically designed to carry out psychometric scaling experiments related to image quality.

3.4.1. Observers

In this study, there were a total of 21 observers, including one expert. The expert observer had experience with imaging, old/historical manuscripts, and microfiche, all of which were used in the experiment. Among the remaining observers, five had experience with old or historical manuscripts, and one had worked with microfiche. Moreover, most of these observers had experience in imaging. For the study, all observers, except one expert, were combined into a single standard observer group. The selection criteria for the observers followed the guidelines specified in ITU-R BT.500-14 [38], which ensured that none of the observers had any personal involvement in the design of the experiment. A Snellen chart test was also used to ensure that all observers had a normal or corrected-to-normal vision during the selection process.

3.4.2. Treatments

For our study, we selected a set of 20 images, consisting of four text-based images and eight each from fragments and natural scenes. We scanned these images using three devices and utilized CS for image enhancement. For CS, we adjust the minimum and maximum intensity values to encompass the full range of possible values. In the context of an 8-bit grayscale image, this means that the lowest intensity value in the image is expanded to the minimum possible intensity value of 0, while the highest intensity value is extended to reach the maximum value of 255. This stretching process is used to increase the dynamic range of the gray levels in the image. Equation (1) represents the general equation for contrast stretching. Data preprocessing and result analysis were performed using the open-source Python programming language.

$$OutputPixel = \left(\frac{InputPixel - MinInput}{MaxInput - MinInput}\right) \times 255 \quad (1)$$

Here, OutputPixel represents the resulting pixel value after contrast stretching, InputPixel is the original pixel value, MinInput and MaxInput are the minimum and maximum intensity values in the image, respectively. Next, we divided the images into 20 sets, each containing six images, three scanned images from each of the three devices and three enhanced versions of these images. Consequently, each set of images produced 15 pairs, resulting in 300 pairs shown to each observer in a randomized order. Typically, psychophysical experiments should not take more than one hour of total time for the observers. In our experiment, observers were not given a specific time limit, but on average, we found that for 300 pairs, the median time taken by each observer was 30 minutes.

3.4.3. Instructions

At the beginning of the experiment, observers received instructions on the experiment. They were asked to compare two images displayed side by side and choose the one with better image quality. Observers were informed that the subjective test was being conducted to evaluate the performance of an alternative imaging approach for microfiche. Although observers were not provided with a specific definition of 'quality' at the beginning of the experiment, we conducted a survey at the end to gather their feedback. This survey aimed to assess the observer's preferences concerning various predefined image quality attributes that they considered while rating each type of microfiche.

3.4.4. Viewing Condition

The experiment followed the guidelines outlined in ITU-R BT.500-14 [38] regarding viewing conditions on displays (ISO 3664). To ensure consistency, the monitor was calibrated using an Eye-one device prior to the experiment. The chromaticity of the white displayed on the color monitor was set to CIE standard illuminant D65, and the white's luminance level was set to 80 cd/^{m2}. The observers were seated approximately 80 cm away from the monitor, and the lighting in the room was dimmed to approximately 17 lux.

3.4.5. Subjective Data Processing

In a pair comparison experiment, the results can be presented as a winning frequency matrix that illustrates the relative frequencies with which each stimulus is preferred over the others. For example, in this study, comparing reproductions of three different devices (FSL, FBS, and IFS), participants are presented with all possible combinations of devices and asked to select their preferred option. The resulting data is recorded in a 3×3 raw data matrix. By aggregating the responses of all participants, a 3×3 raw frequency matrix is generated, providing a summary of the overall preferences for each device. This matrix visually represents the test outcomes, making it easier to interpret and draw some initial conclusions from the experiment data.

Using the frequency matrix, we computed the Binomial Sign test to show the statistical significance of the result obtained. To account for the possibility of type I errors resulting from multiple condition testing, we applied the Bonferroni correction [39]. Additionally, we computed the Z-score [40] based on Thurstones law of comparative judgment [41]. It is a statistical measure that indicates how far a particular observation is from the mean in

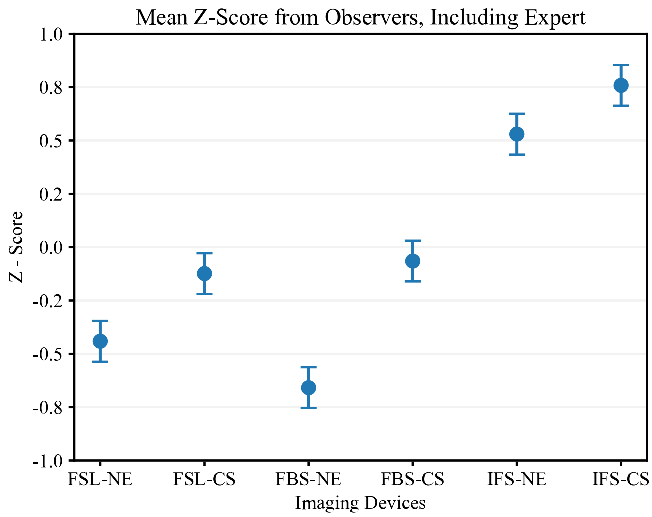

Fig. 2. Comparison of Z-Score for all non-enhanced (NE) and contrast-stretched (CS) enhanced images, evaluated by all observers (including experts and rest of the observers), on all three imaging devices: flatbed scanner (FBS), in-house film scanning (IFS), and microform reader (FSL).

terms of standard deviation and allows us to draw statistical inferences about the differences between the two items being compared.

The most common approach to compute the 95% confidence intervals (CI) for Z-scores involves using the standard deviation (σ) and the number of observations (N) in the analysis and is carried out using Equation (2). When the 95% confidence intervals of the Z-scores do not overlap, we can conclude with 95% confidence that there is a significant difference.

$$CI = 1.96\frac{\sigma}{\sqrt{N}} \tag{2}$$

4. Results and Discussion

The results of this study involving both expert and rest of the observers for all images are shown in Fig. 2. The Z-scores were calculated by comparing non-enhanced and enhanced images (using contrast stretching) from all three devices against each other. From analysis of Fig. 2, it becomes apparent that the observers preferred IFS (both NE and CS) over FSL (NE and CS) and FBS (NE and CS), of which FSL and FBS had comparable responses for NE and CS respectively. In terms of observers' preference for enhanced versus non-enhanced images, the enhanced IFS images are most preferred over the non-enhanced IFS images. Non-enhanced images from IFS are preferred over the enhanced versions from the other two devices, making IFS the preferred device. However, images from FSL and FBS devices have comparable preference levels after enhancement, as shown by overlapping confidence intervals. Non-enhanced FSL images are preferred over non-enhanced FBS. It is important to note that this is for the current settings, devices, and images used for the experiment.

We also performed a separate analysis for the expert and the rest of the observers. The results of the experts were similar to those obtained from all combined observers. When analyzing the results for an expert, we found that IFS remained the preferred option. The Z-scores computed separately for the expert and all other observers are provided as a supplementary file (see Figure S3). Table 3 illustrates the result of hypothesis testing using a sign test for all images between the devices, indicating significant differences between the three devices. We also performed a sign test on the data from the expert, which produced similar results, except for the comparison between FSL and FBS. The formulated null hypothesis that claimed that both devices produce similar results was true. Table S1 in the supplementary file contains a matrix that compares the significant test results for the expert.

Table 3
P-value for each pair of stimuli being compared; here each cell of a lower triangular matrix represents a pair of stimuli being compared; green cells indicate that the corresponding stimulus pair has a statistically significant difference.

	FSL	FBS	IFS
FSL			
FBS	2.78E-04		
IFS	1.36E-126	1.40E-119	

In this experiment, three types of microfiche were used: text, natural scenes, and fragments. We also analyze user preferences for each microfiche category by comparing non-enhanced and enhanced versions of images from the enhanced and non-enhanced images from the three devices against each other. Both non-enhanced and enhanced versions from IFS are preferred for microfiche containing text (Fig. 3 (a)) and natural scenes (Fig. 3 (b)). However, for microfiche with fragments (Fig. 3 (c)), the enhanced versions from FSL and FBS have similar preferences to the non-enhanced version from IFS. However, the enhanced version from IFS remains the most preferred compared to all others. A non-enhanced version of the device IFS is always preferred against the other two devices in both enhanced and non-enhanced cases, indicating that the IFS device exhibits a higher capability for digitizing microfiche than the other two devices. When comparing the performance between the FSL and FBS devices, it becomes difficult to determine which one exhibits better performance. For text-based microfiche, the non-enhanced image from the FSL device is equally preferred as the enhanced version from the FBS device, indicating that the FSL device may be considered better in this aspect. However, the enhanced version of the FBS device is preferred when evaluating natural scenes. Similarly, for microfiche with fragments, the non-enhanced images from the FBS device are less preferred, but both devices demonstrate similar preferences when considering the enhanced versions.

In addition, non-enhanced images from FBS remain the least preferred one for microfiche with text and fragments whereas for natural scene, FBS have similar preference as FSL. When comparing non-enhanced and enhanced versions, there is little or no difference in the case of microfiche with text, as indicated by the overlapping confidence intervals. The same applies to natural scenes, especially for IFS. However, there is a noticeable difference between the non-enhanced and enhanced versions in all other cases. A notable observation from the results is that the observers' ability to differentiate between the devices is becoming less distinct as we compare microfiche with text and natural scene; this distinction is even less for microfiche with the fragment. This can be seen in the range of Z-score values in Fig. 3. Compared to natural scenes and fragments, observers took less time on average to make a decision regarding microfiche with text. It is possible that the devices used in this study were better able to distinguish their performance for simpler and more structured information, while their performance was less distinct for more complex and less structured data.

Table 4 displays the outcomes of hypothesis testing conducted on microfiche with text using various imaging devices, with and without enhancement. It indicates that IFS differs significantly from FSL and FBS, while there is no significant difference between FSL and FBS. It is also evident that when comparing enhanced and non-enhanced versions from the same imaging devices, the re-

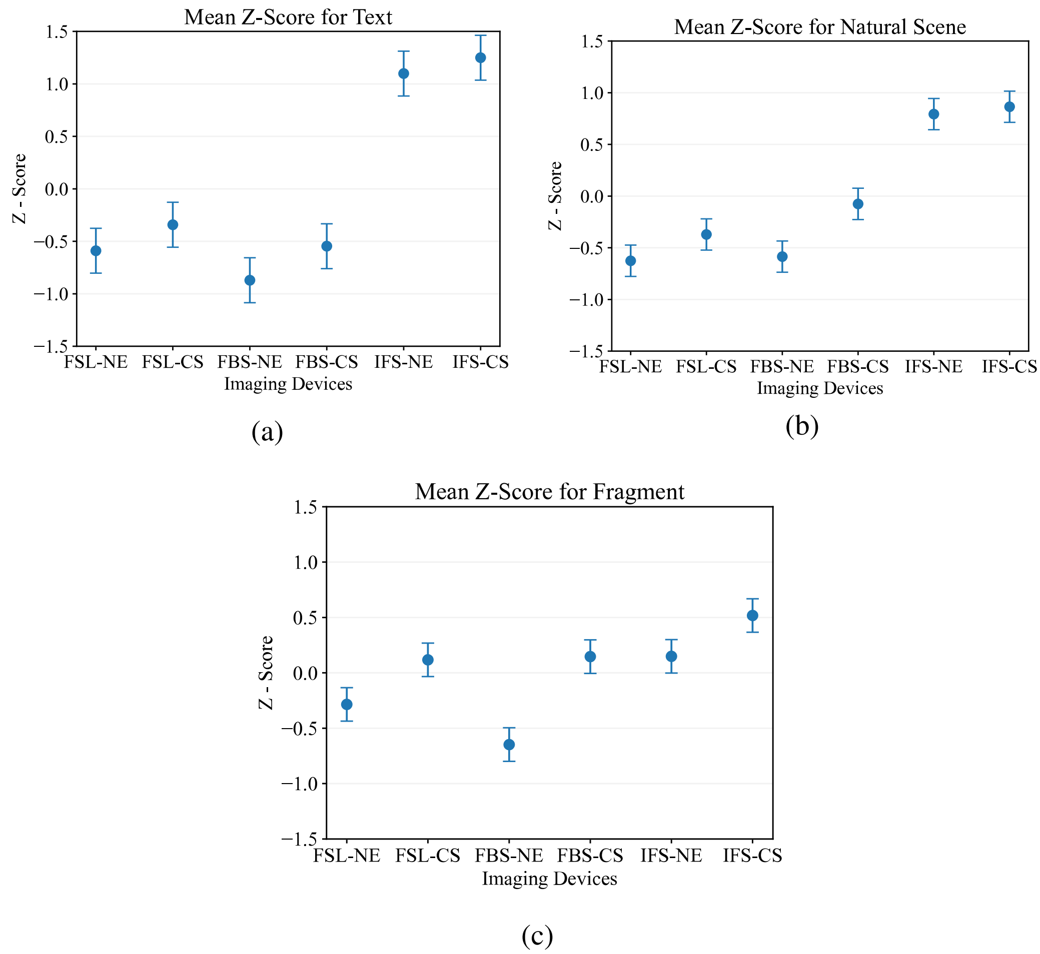

Fig. 3. Z-scores at 95% confidence intervals for each microfiche category (i.e., with (a) Text, (b) Natural Scene, and (c) Fragment) obtained by comparing non-enhanced (NE) and enhanced versions(CS) of images from all three devices, i.e., flatbed scanner (FBS), in-house film scanning (IFS), and microform reader (FSL), against each other.

Table 4
The p-value for each pair of stimuli compared for microfiche with text is shown in a lower triangular matrix where each cell represents a pair of stimuli being compared. Green cells indicate a statistically significant difference between the corresponding stimulus pairs, while red cells indicate no statistically significant difference.

		FSL		FBS		IFS	
		NE	CS	NE	CS	NE	CS
FSL	NE						
	CS	0.59151					
FBS	NE	0.00056	0.00443				
	CS	0.0058	0.03209	0.54429			
IFS	NE	3.2E-09	8.4E-11	1.9E-19	3.1E-17		
	CS	2.3E-09	5.9E-11	1.2E-19	2E-17	1	

sults indicate no statistically significant difference in the images for all three microfiche categories. Similar observations were made for the other two microfiche categories, namely natural scene and fragment, and the results of these categories are included in the supplementary file (Tables S2 and S3).

This study also examined observers' general preferences for enhanced versus non-enhanced images for each device, regardless of the type of microfiche being used. The results shown in Fig. 4 demonstrate that observers consistently preferred enhanced images over non-enhanced images on all imaging devices. This indicates that the application of enhancement techniques improves the overall visual quality of the images, regardless of the specific type of microfiche or the imaging device used.

Table 5 displays the results of hypothesis testing conducted for enhanced versus non-enhanced images for all three imaging devices. Most of the cells are green, indicating a significant difference between the stimuli. However, results were not significantly different when enhanced versus enhanced compared in-between FSL, FBS, and IFS. This is not the case for non-enhanced images except for comparing IFS and FSL. To facilitate visualization, we have included the images in the supplementary file (Fig. S4) that contain enhanced and non-enhanced images, one from each device.

The observers were presented with one image from each type of microfiche, namely text, natural scenes, and fragments, and were provided with five options, including an option for user input. They were then asked to select which attributes they considered when

Table 5
The p-value for each pair of stimuli compared for enhanced versus non-enhanced across three imaging devices; in a lower triangular matrix, green cells indicate a statistically significant difference between the corresponding stimulus pairs, while red cells indicate no statistically significant difference.

		FSL		FBS		IFS	
		NE	CS	NE	CS	NE	CS
FSL	NE						
	CS	7.19E-17					
FBS	NE	0.000497	1.49E-30				
	CS	2.11E-24	0.053128	6.43E-40			
IFS	NE	0.181417	2.04E-12	1.28E-06	4.8E-19		
	CS	1.1E-13	0.356378	2.18E-26	0.0038	1.1E-09	

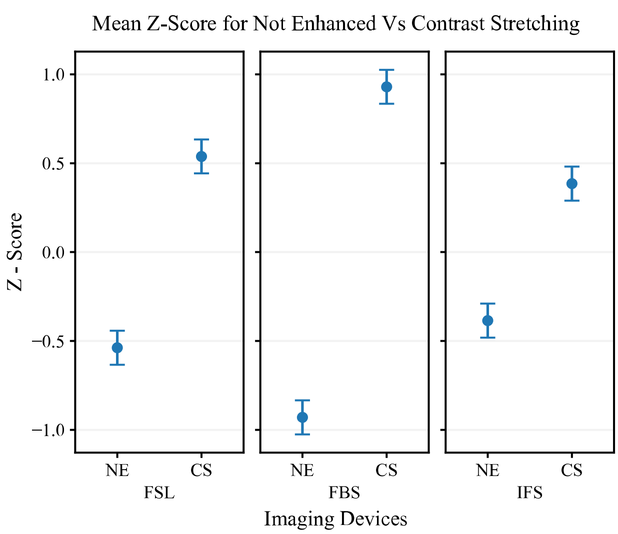

Fig. 4. Z-scores, at 95% confidence intervals, between non-enhanced and contrast-stretched images across three imaging devices.

rating the quality from these options. Fig. 5 illustrates the result of the survey. The result infers that, when assessing the quality of microfiche, legibility is a crucial factor that all observers consider, particularly when examining text. Additionally, sharpness is another key attribute that observers prioritize when evaluating text and fragments on the microfiche. For natural scenes on the microfiche, contrast is the most significant attribute that observers consider when assessing quality. However, the lack of artifacts, such as noise, is regarded as a less important attribute across all three categories of microfiche. This means that the presence of minor imperfections, such as graininess or distortion in the images, would not significantly affect the overall quality assessment of the microfiche.

5. Conclusion

Microfiche has been widely used to preserve historical documents, leading to the formation of valuable artifacts and essential elements of our cultural heritage. Despite its durability, microfiche can still suffer damage and require digitization for preservation and accessibility. However, traditional microfiche readers are not always available and are primarily designed for reading rather than

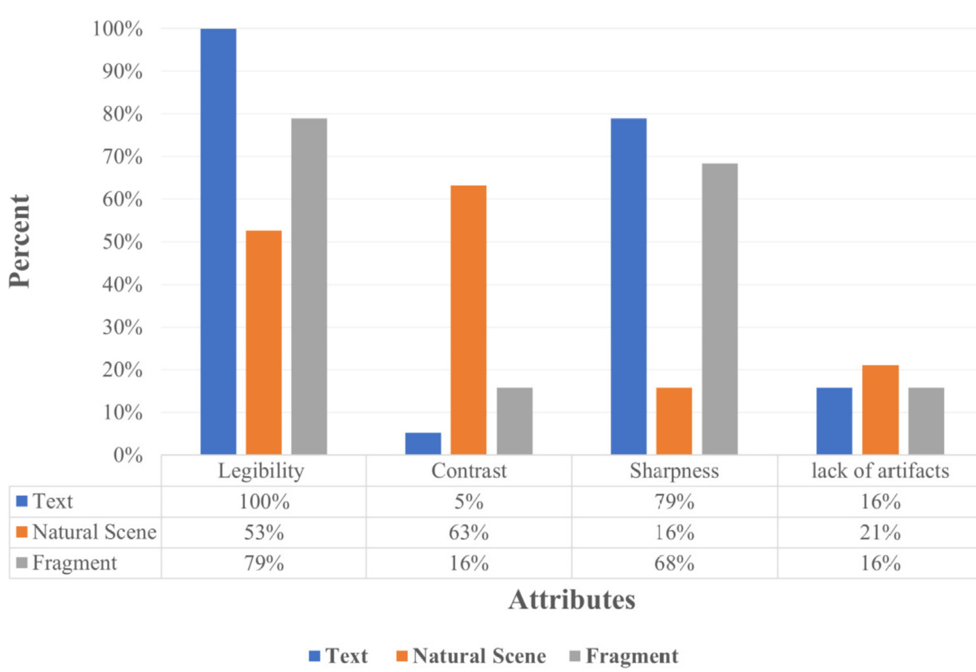

Fig. 5. Results of a survey showing observer's preferences concerning various predefined image quality attributes, i.e., legibility, contrast, sharpness, and lack of artifacts that they considered most significant while rating each type of microfiche, namely text, natural scenes, and fragments.

digitization. This means that a digitization effort using a microfiche reader would also be significantly slower and time-consuming, consult Max. scan area and Max. fiche per scan in Table 1.

In this study, we evaluate the performance of alternative imaging devices compared to a traditional microfiche reader and the impact of image enhancement on image quality using subjective image quality assessment. The experiment was carried out in a controlled environment with twenty-one participants, including an expert. The results of this study show that the reproduction of alternative devices was preferred over a traditional microfiche reader, indicating that alternative imaging devices can be a viable option for digitizing microfiche and improving access to historical collections. Additionally, results also demonstrate that the image enhancement techniques significantly improved image quality for all three categories of microfiche. Therefore, the adoption of alternative imaging devices along with image enhancement techniques could be a feasible approach to digitizing microfiche and facilitate better access to historical collections. This initiative could help conserve precious cultural heritage artifacts that exist only in microfiche format and enable more people to access them.

In future research, it would be beneficial to conduct objective image quality assessments for different quality attributes of microfiche, including contrast, sharpness, etc. This could involve exploring various existing image quality metrics and determining their relevance for assessing microfiche quality. Based on the results, new objective quality metrics could be developed or linked to provide more comprehensive and time-efficient evaluations of microfiche image quality. In addition, when digitizing microfiche technical drawings, it is crucial to ensure that the digitized versions faithfully capture the precise details and accuracy of the original drawings. Thus, further comprehensive studies can be conducted to investigate the geometric quality attributes, such as distortion, scale accuracy, resolution, etc., for microfiche digitization.

Acknowledgment

The authors would like to thank Torleif Elgvin and Gregory High for providing access to the various materials and devices and Gjøvik Biblioteket for allowing us access to the Microform scanner needed for the formulation and execution of this study. This work is supported by The Lying Pen of Scribes–Manuscript Forgeries, Digital Imaging, and Critical Provenance Research funded by the Research Council of Norway (project nr. 275293) and CHANGE-ITN project funded by EU Horizon 2020 (Marie Sklodowska–Curie, grant agreement No. 813789).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.culher.2023.07.014

References

[1] W. De Haas, The microfiche, Journal of the American Society for Information Science 9 (2) (1958) 99, doi:10.1016/j.patcog.2006.04.043.
[2] C. Garcia-Spitz, Patrolling the past: Bringing the papua new guinea colonial-era reports into the digital realm, UC San Diego: Library (2017).
[3] James M. Reilly, The IPI Storage Guide for Acetate Film, 1993(Retrieved from https://s3.cad.rit.edu/ipi-assets/publications/acetate_guide.pdf). Accessed: 27 June 2023.
[4] K. Canepi, B. Ryder, M. Sitko, C. Weng, Microform Terminology, in: Managing microforms in the Digital age, Association for Library Collections & Technical Services (ALCTS), 2013, pp. 267–274. https://www.ala.org/alcts/resources/collect/serials/microforms.
[5] South Carolina Department of Archives and History, Archives and Records Management Division, Microfilm and microforms, 2012, (Retrieved from https://dc.statelibrary.sc.gov/handle/10827/6890). Accessed: 27 June 2023.
[6] ISO 18901:2010, Imaging materials – processed silver-gelatin-type black-and-white films – specifications for stability, 2010, Accessed on June 27, 2023.
[7] ANSI/IS0 8225-1995, ANSI/NAPM IT9.5-1996, Imaging materials – ammonia-processed diazo photographic film – specifications for stability, 1996, Accessed on June 27, 2023.
[8] ISO 18912:2002, Imaging materials – processed vesicular photographic film – specifications for stability, 2002, Accessed on June 27, 2023.
[9] Northeast Document Conservation Center (NEDCC), 6.1 Microfilm and Microfiche, 2017.
[10] E. Tigchelaar, On the unidentified fragments of "DJD" XXXIII and PAM 43.680: A new manuscript of "4QNarrative and poetic composition", and fragments of "4Q13", "4Q269", "4Q525" and "4QSb"(?), Revue de Qumrân 21 (3 (83)) (2004) 477–485.
[11] P.B. Hirtle, The impact of digitization on special collections in libraries, Libraries & Culture (2002) 42–52.
[12] W. Saffady, et al., Computer-output microfilm, American Library Association, 1978.
[13] M. Pedersen, J.Y. Hardeberg, et al., Full-reference image quality metrics: Classification and evaluation, Foundations and Trends® in Computer Graphics and Vision 7 (1) (2012) 1–80.
[14] P. Zhao, M. Pedersen, Extending subjective experiments for image quality assessment with baseline adjustments, in: Image Quality and System Performance XII, volume 9396, SPIE, 2015, pp. 244–256.
[15] T. Lee, M. Ishii, P. Okubo, Assessing the fidelity of seismic records from microfilm and paper media, Seismological Society of America 93 (6) (2022) 3444–3453.
[16] W.M. Duff, J.M. Cherry, Use of historical documents in a digital world: comparisons with original materials and microfiche, Information Research 6 (1) (2000). 6–1
[17] S.S. Bukhari, A. Kadi, M.A. Jouneh, F.M. Mir, A. Dengel, ANYOCR: An open–source OCR system for historical archives, in: 2017 14th IAPR international conference on document analysis and recognition (ICDAR), volume 1, IEEE, 2017, pp. 305–310.
[18] R. Holley, How good can it get? analysing and improving ocr accuracy in large scale historic newspaper digitisation programs, D-Lib Magazine 15 (3/4) (2009).
[19] A.R. Ali, W.J. Spivey, The archives of the future, Technical Report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2021.
[20] H. Deborah, D.J. Mandal, Evaluation of text legibility in alternative imaging approaches to microfiche digitization, in: Proc. IS&T Archiving, 2021, pp. 96–101, doi:10.2352/issn.2168-3204.2021.1.0.22.
[21] Z. Su, B.-R. Ahn, K.-Y. Eom, M.-K. Kang, J.-P. Kim, M.-K. Kim, Plagiarism detection using the levenshtein distance and smith-waterman algorithm, in: 2008 3rd International Conference on Innovative Computing Information and Control, IEEE, 2008. 569–569
[22] G. Singh, A. Mittal, Various image enhancement techniques-a critical review, International Journal of Innovation and Scientific Research 10 (2) (2014) 267–274.
[23] M. Pedersen, N. Bonnier, J.Y. Hardeberg, F. Albregtsen, Attributes of image quality for color prints, Journal of Electronic Imaging 19 (1) (2010). 011016–011016
[24] K. Gu, G. Zhai, X. Yang, W. Zhang, M. Liu, Subjective and objective quality assessment for images with contrast change, in: 2013 IEEE International Conference on Image Processing, 2013, pp. 383–387, doi:10.1109/ICIP.2013.6738079.
[25] K. Gu, G. Zhai, W. Lin, M. Liu, The analysis of image contrast: From quality assessment to automatic enhancement, IEEE Transactions on Cybernetics 46 (1) (2016) 284–297, doi:10.1109/TCYB.2015.2401732.
[26] X. Wu, A linear programming approach for optimal contrast-tone mapping, IEEE transactions on image processing 20 (5) (2010) 1262–1272.
[27] C.-C. Yang, Image enhancement by modified contrast-stretching manipulation, Optics & Laser Technology 38 (3) (2006) 196–201.
[28] G. Trumpy, J.Y. Hardeberg, S. George, B. Flueckiger, B. Cattaneo, M. Picollo, F. Cherubini, V. Marchiafava, Multispectral capture of film colors with led, in: Colour photography and film: Sharing knowledge of analysis, preservation, conservation, migration of analogue and digital materials, Gruppo del Colore–Associazione Italiana Colore, 2021, pp. 195–198.
[29] W.R. Hawken, Evaluating Microfiche Readers: A Handbook for Librarians, Washington: Council on Library Resources, 1975.
[30] G.J. Brooke, H.K. Bond, Allegro Qumran Photograph Collection, Brill, Leiden, The Netherlands, 1999.
[31] E. Tov, S. Pfann, Companion Volume to the Dead Sea Scrolls on Microfiche Edition: Published under the Auspices of the Israel Antiquities Authority, Brill, Leiden, The Netherlands, 1995.
[32] R.K. Mantiuk, A. Tomaszewska, R. Mantiuk, Comparison of four subjective methods for image quality assessment, Computer Graphics Forum 31 (8) (2012) 2478–2491.
[33] M.H. Pinson, S. Wolf, Comparing subjective video quality testing methodologies, in: Visual Communications and Image Processing 2003, volume 5150, SPIE, 2003, pp. 573–582.
[34] B.W. Keelan, Handbook of image quality: characterization and prediction, CRC Press, 2002.
[35] M. Pedersen, Image quality metrics for the evaluation of printing workflows, University of Oslo, 2011 Ph.D. thesis.
[36] H.A. David, The method of paired comparisons, volume 12, London, 1963.
[37] K.V. Ngo, J.J. Storvik, C.A. Dokkeberg, I. Farup, M. Pedersen, QuickEval: A web application for psychometric scaling experiments, in: M.-C. Larabi, S. Triantaphillidou (Eds.), Image Quality and System Performance XII, SPIE, 2015, p. 93960O, doi:10.1117/12.2077548.

[38] BT.500-11 Recomendation ITU-R, Methodology for the subjective assessment of the quality of television images, International Telecommunication Union (2019).
[39] J.M. Bland, D.G. Altman, Multiple significance tests: the bonferroni method, British Medical Journal(BMJ) 310 (6973) (1995) 170.
[40] P.G. Engeldrum, Psychometric scaling: avoiding the pitfalls and hazards, in: Image Processing, Image Quality, Image Capture Systems Conference (PICS), 2001, pp. 101–107.
[41] L.L. Thurstone, A law of comparative judgment, in: Scaling, Routledge, 2017, pp. 81–92.