

## Article

# Finding the Best Lighting Mode for Daguerreotype, Ambrotype, and Tintype Photographs and Their Deterioration on the Cruse Scanner Based on Selected Methods

Zuzana Berger Haladová <sup>1,\*</sup>, Róbert Bohdal <sup>1</sup>, Zuzana Černeková <sup>1</sup>, Paula Štancelová <sup>1</sup>, Andrej Ferko <sup>1</sup>, Janka Blaško Križanová <sup>2</sup>, Jana Hojstričová <sup>2</sup> and Kitti Baráthová <sup>2</sup>

<sup>1</sup> Faculty of Mathematics Physics and Informatics, Comenius University, Mlynska Dolina, 842 48 Bratislava, Slovakia

<sup>2</sup> Conservation and Restoration Department, Academy of Fine Arts and Design in Bratislava, Hviezdoslavovo Námestie 18, 814 37 Bratislava, Slovakia

\* Correspondence: zuzana.bergerhaladova@fmph.uniba.sk; Tel.: +421-903-850-930

**Abstract:** In the digital heritage field, the accurate reproduction of hard-to-photograph items, such as daguerreotypes, ambrotypes, and tintypes, is an ongoing challenge. Industrial contactless sensors offer the potential to improve the quality of scanned images, but their capabilities and limitations have not been fully explored. In our research, a dataset of 48 scans was created using the hi-tech industrial contactless sensor CRUSE. Moreover, 3 rare original photographs were scanned in 16 different modes, the most suitable images were determined by specialists in the restoration, and validated through experiments involving eye-tracking, multiple computer vision, and image processing methods. Our study identified the Cruse scanning modes, which can be utilized to achieve the most accurate digital representation of scanned originals. Secondly, we proposed several methods for highlighting the degradation and minor scratches on photographs that otherwise might not be detected by the restorer's naked eye. Our findings belong to four overlapping areas, i.e., image understanding, digital heritage, information visualization, and industrial sensors research. We publish the entire dataset under the CC BY-NC 4.0 license. The CRUSE sensor shows promise as a tool for improving the quality of scanned images of difficult-to-photograph items. Further research is necessary to fully understand its capabilities and limitations in this context.

**Keywords:** CRUSE; ambrotype; daguerreotype; tintype; scanning modes; eye-tracking; concurrent color channels; sharpness function; image differences; image similarity; UIQI; SSIM; ESSIM



**Citation:** Berger Haladová, Z.; Bohdal, R.; Černeková, Z.; Štancelová, P.; Ferko, A.; Blaško Križanová, J.; Hojstričová, J.; Baráthová, K. Finding the Best Lighting Mode for Daguerreotype, Ambrotype, and Tintype Photographs and Their Deterioration on the Cruse Scanner Based on Selected Methods. *Sensors* **2023**, *23*, 2303. <https://doi.org/10.3390/s23042303>

Academic Editor: Gwanggil Jeon

Received: 31 December 2022

Revised: 7 February 2023

Accepted: 15 February 2023

Published: 18 February 2023



**Copyright:** © 2023 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 (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The digitization of historical ambrotypes, daguerreotypes, and other photographic materials is incredibly important. First, it allows us to preserve these fragile and delicate items for future generations. Without digitization, many of these items would be at risk of being lost or damaged due to age and environmental factors.

Second, digitization makes these items more accessible to researchers and the general public. Instead of having to visit a physical archive or museum, people can now view these photographs online, making it easier for them to study and appreciate them.

Third, digitization allows us to enhance and improve these photographs. With the use of specialized software, we can restore damaged or faded images, bringing them back to life and making them more legible. We can also make high-resolution copies of these photographs, which can be used for research and educational purposes.

Overall, the digitization of historical ambrotypes, daguerreotypes, and other photographic materials is essential for preserving our cultural heritage and making it more accessible to the public. It allows us to protect these items for future generations and unlock the valuable information and stories they contain. Digitization facilitates the preservation

of original heritage objects and provides a digital equivalent available for further study. This is especially important in information-bearing artifacts and paper-based artifacts. The emphasis on digitization may lead to the misconception that information matters, whereas the physical preservation of the artifact does not [1]. Digitization is not preservation. However, digital copies of paper documents may be used for further material analysis protecting valuable originals from handling.

Virtualization of cultural or natural heritage has become an important part of the knowledge economy. We create the digital twins using the hi-tech CRUSE CS 220ST1100 contactless scanner, providing high precision both in geometry and radiometry. We report on our “multiple ground truths” findings in the context of the digitization of hundreds of thousands of museum originals (cultural objects) in Slovakia.

We experimented with sharpening, scanning mode selection, and properties of scans. In this paper, we offer a detailed description of the in-depth processing of a triplet of intentionally selected hard-to-photograph originals. They are named by the creation procedure, i.e., daguerreotype, ambrotype, and tintype (DAG, AMB, and TIN for short).

Globally, there are a few major initiatives, aggregators, or harvesters of the daily growing volume of online heritage items, including scanned paintings or photographs from more than 7000 heritage institutions. Google Arts and Culture (formerly Google Art Project) [2] is an online platform of high-resolution images and videos of artworks and cultural artifacts from 2000+ partner cultural organizations throughout the world. Europeana [3] is a virtual repository of artworks, literature, cultural objects, relics, and musical recordings/writings from over 2000 European institutions. The Digital Museum Canada (formerly Virtual Museum of Canada) [4] is a virtual collection presenting content from more than 3000 Canadian heritage institutions. Wikipedia GLAM [5] (“galleries, libraries, archives, and museums”, also including botanic and zoological gardens) helps cultural institutions share their resources with the world through collaborative projects with experienced Wikipedia editors, as report the authors of Wikipedia article on Google Arts and Culture. Rare originals from a local collection at the Academy of Fine Arts and Design (AFAD) (Figure A1) will be available through both Europeana and Google Arts and Culture. This will happen after being included in Slovákiana [6], which is guided only in the Slovak language. All three are used for the education of restoration experts, supported by, e.g., a textbook [7]. In the area of digitized old photographing techniques, there is major computer research being done within generative adversarial network (GAN) methods for the digital restoration of damaged photographs, e.g., [8]; in our work, we will not reflect on digital restoration of the photographs.

## 2. Materials and Methods

### 2.1. Materials, a Triplet of Cultural Objects

In total, we studied 3 material originals (cultural objects) and utilized 2 sensors (a scanner and an eye tracker). The primary data in the real-world documented three discoveries in the evolution of photography, daguerreotype, ambrotype, and tintype. Using 16 scanning modes with the given parameters, we obtained secondary data, which are described in the next section and were deposited in a publicly available website under the Creative Commons license [9]. Moreover, the samples of processed images were provided in the same way. Unlike tintypes, daguerreotypes and ambrotypes [10,11] belong to the group of cased photographic objects [12–14]. The examined photographs are from the Photographic reference collection at AFAD. The Laboratory of Restoration of Photography deals with research, identification, and the actual restoration of various types of photographic techniques [15]. Emphasis is placed on the gradual acquisition of current non-destructive research techniques and the subsequent choice of restoration intervention. The advantages that we see in using the CRUSE scanning modes versus high-definition photographic images on photographic objects are mainly in the ability to scan objects in very high magnification.

### 2.1.1. DAG

All DAG scans are shown in Appendix A Figure A2. The daguerreotype was the first commercially patented photographic process in 1839 by Louis Daguerre in France. From the material point of view, a daguerreotype is a very complex object consisting of a thin sheet of copper, coated with highly polished silver, and sensitized by iodine fumes exposed in the camera. Afterward, the exposure image is developed in the fumes of mercury and fixed with sodium thiosulfate. As a finish, the object was housed in a protective case, due to its very sensitive nature.

The daguerreotype in Figure A2 was chosen for the examination because it is a good example of this technique. It is a bare plate (missing the protective case), degraded physically and chemically. In general, daguerreotypes have very reflective surfaces, such as a mirror; therefore, it is complicated to photograph them. The scanner has the advantage of minimizing the mirroring effect and the ability to emphasize the actual image.

The following DAG metadata will serve for image documentation, cataloging, and/or retrieval.

Category: historic photo.

Subcategory: reference object for study purposes.

Administered by: University of Fine Arts in Bratislava, Department of Restoration, and Department of Photography and New Media.

Description of the object: The daguerreotype shows a portrait of an elderly woman with glasses and a scarf. The original decorative protective packaging of the daguerreotype is absent.

Technique: daguerreotype

Physical description: number 1, width 70.0 mm, height 80.0 mm.

Degree of authenticity: original.

Legal Status: free.

Inventory number: F-00046-DAG.

Identifier: this field will be supplied according to the future virtual exhibition (e.g., in Europeana).

The original daguerreotype was scanned in *LRFB* (scanning mode first) at a 600 ppi resolution, 48-bit RGB, 0.25 cm was the original height, with 1.10 gain settings. The resulting TIFF file size is about 30 MB and, for further processing, the master scan was converted using GIMP into the PNG file of 21.5 MB. Scan number 1831 refers to the number of scans after the previous scanner calibration. All scans were of the same size, as given by the original width and height, 70 × 80 mm. The naming convention stores the settings, e.g., *Dagero 600.00 ppi LRFB Gain 1.10 0.250 cm 1 831.tif*, see Figure A2 in Appendix A.

### 2.1.2. AMB

The ambrotype process was invented by Frederick Scott Archer in the 1840s in England. Collodion is poured on a clean sheet of glass, and while still wet, it is dipped into the solution of silver nitrate, exposed in the camera, and fixed. The result is a negative image on the glass. In 1850, Louise Désire Blanquart-Evrard discovered that this image could appear positive when placed on a dark backing. This was possibly due to very little or no silver in the shadows allowing the black background to show through the glass. Due to the very fragile nature of glass, the ambrotypes were housed in protective cases and frames same as daguerreotypes.

The ambrotype in Appendix A Figure A1 is a representative example of the ambrotype image. Ambrotypes are hard to be photographed because they are 3D objects; there is a glass with an image on its verso and the black backing is underneath it. Thus, one looks through the thick glass to see the image underlaid with a black backing (which is not always in direct contact with the image; sometimes it is heavily degraded or is missing completely). All AMB scans are shown in Appendix A Figure A1.

The following AMB metadata will serve for image documentation, cataloging, and/or retrieval.

Category: historic photo.

Subcategory: reference object for study purposes.

Administered by: University of Fine Arts in Bratislava, Department of Restoration, and Department of Photography and New Media.

Description of the object: The ambrotype depicts a portrait of a seated man and a standing woman. The work is heavily hand-colored in the area of the face, hands, and jewelry of the figures. The ambrotype is inserted into a decorative protective cover/cassette, the front closing part of which is absent.

Technique: ambrotype.

Physical description: number 1, ambrotype width with cassette 83.0 mm, ambrotype height with cassette 96.0 mm.

Degree of authenticity: original.

Legal Status: free.

Inventory number: F-00047-AMB.

Identifier: this field will be supplied according to the future virtual exhibition (e.g., in Europeana).

### 2.1.3. TIN

The ferrotype (tintype) technique belongs to the group of melanotypes and it is inspired by the ambrotype–collodion negative with a black background. Instead of glass, which is overlaid with black paper, velvet, varnish, or paint on a collodion negative background, was designed in 1853 by the French teacher Adolphe Alexandre Martin (1824–1896) to apply the emulsion to a thin black lacquered sheet of iron plate. Black paint was applied to the plate colored paint, varnish (japaning), or enamel. The commonly used term tintype is garbled by transference of meaning because in fact tin was not used, but an iron plate (lat. ferrum). A wet collodion process was used on black iron sheets. Ferrotypes have little contrast, whites are a soft gray, and detail is lost in the shadows. Ferrotype (tintype) was used from 1853 until 1950 when it was still used by street photographers in Turkey or India. As a customer, you could buy a photograph right on the street, made in two minutes.

The following tin metadata will serve for image documentation, cataloging, and/or retrieval.

Category: historic photo.

Subcategory: reference object for study purposes.

Administered by: University of Fine Arts in Bratislava, Department of Restoration, and Department of Photography and New Media.

Description of the object: The ferrotype shows a portrait of a seated man and a standing woman. The work is in the area the faces and hands of the characters are gently colored by hand.

Technique: ferrotype, tintype.

Physical description: number 1, width 61.0 mm, height 103.0 mm.

Degree of authenticity: original.

Legal Status: free.

Inventory number: F-00048-TIN.

Identifier: this field will be supplied according to the future virtual exhibition (e.g., in Europeana).

All TIN scans are shown in Appendix A Figure A3.

## 2.2. Methods, Modeling the Scanner Focusing, InfoVis, Eye Tracking

We divide the originally designed workflow into naturally given groups of methods.

1. We explain the standard scanner settings of a given push-broom camera [16] and the

mathematical model of the manual sharpening process, based on the Fourier transform. 2. Within a data state reference model in information visualization [17], we visualize and distinguish properties of scanning modes. This was done by transforming the image differencing problem to a well-studied assignment problem, using Harris corners [18] or Co-SIFT feature points [19]. 3. We analyze the differences between various scans by multiple image quality metrics. 4. The qualitative evaluation by restoration experts from the Laboratory of Restoration of Photography [15] is supported by eye-tracking measurements of their perception.

### 2.2.1. Contactless Scanning

The CRUSE scanner CS 220ST1100 (CCD “14.400 Pixel Tri Linear RGB Line-Sensor” with Schneider optics) offers 4 light sources *LRFB* (left, right, front, back) and *LRFB* scanning mode is the default one for heating the lamps, which requires 30 min. This starting mode is the fastest one. Another mode, recommended for scanning artwork, is *LTx*, which better documents the fine relief of the given original. It is possible to scan the given scan-line of real artwork under constant illumination accurately up to 1200 ppi (pixels per inch) in the TIFF format. The scanned original weight may be up to 300 kg, measuring 120 × 180 × 30 cm [20]. The current software CSx does not create a 3D model, but allows for precise scanning with variable illumination, 15° sensor rotation, and variable depth of field, thus offering dozens of scan versions of a single original. The ultimate and unique truth does not exist here. Consequently, there are many best scans, each one for a given mode and settings. The wooden decors illustrate a sample from natural heritage [21]. From cultural heritage we selected hard-to-distinguish cases, such as ambrotypes, daguerreotypes [19,21], synagogue plans, or handmade papers [22]. For educational purposes, Comenius University students created virtual herbarium, 3D models using photogrammetry, depth of field, or shadows [23].

The CRUSE scanner offers different light modes: *LRFB*, *LR*, *LFB*, *LF*, *LTx*, *LTx 5cm*, *LTx 10cm*, *RTx*, *RTx 5cm*, *RTx 10cm*, *LLa*, *LRFB 15deg*, *LR 15deg*, *LFB 15deg*, *LF 15deg*, *L 15deg*. They can be subdivided into several groups using several criteria, i.e., symmetric, rotated, and shifted. Their effect is captured in scans. Given pairs of scans are hard to distinguish by the naked eye. Using computer vision methods we can either segment the objects to support the semantics or mimic the human vision by low-level operations, adding points and edges first or image evaluation then. The computed and perceived versions of interpretation can be consequently compared to gain more insight. We elaborate on this approach in more detail in the following parts.

### 2.2.2. Perceptually Meaningful Dot Patterns and Density

The motivation for adding edges into the image data after Toussaint in 1980 [24]:

In... computational approaches to perception... it is desired to find some structure in the form of edges connecting the subset of the pairs of points... such that the final graph obtained is perceptually meaningful in some sense.

The proposed solution by Toussaint was a relative neighborhood graph (RNG), a superset of the minimum spanning tree (MST), and the subset of Delaunay triangulation (DT). Our idea was to relax from the connected visual descriptor or spanner (MST, RNG) by connecting the dots with the “Hungarian” edges [25,26]. For adding dots, we experimented with searching the gray level Harris corners [18] and concurrent color channels coSIFT [27]. The latter performed better on Cruse scans and already the density of Harris or coSIFT corners indicated “interesting parts” in scanned wooden decors. For more details on the Hungarian edges approach, refer to [19,21]. Methodologically, adding corners and edges can serve for image (or scanning mode) differentiation in terms of the data state reference model in information visualization [17]. The primary scan is analytically abstracted with added feature points, which are consequently visualized by adding edges and displayed.

Note that DT edges can be too long for certain dot configurations but the Hungarian edges' total length is warranted to be minimal.

Technically, three sequential processes occur with visual data: acquisition, representation, and processing/simulation of human vision. These processes support understanding and ultimately lead to a decision. Metaphorically, an inexperienced cultural object owner might view the resulting versions in a similar way to how people used to observe the night sky, i.e., by identifying the brightest stars in a given area, connecting them with edges, and then searching for suitable symbolism. Our approach can be helpful in the early stages by providing different corners and edges.

In this study, we move away from focusing on edges because we can directly associate the density of dots with appealing image components, such as salient regions, or semantic primitives, such as human faces. These elements can ultimately contribute to the overall meaning of a portrait.

Image quality measuring is based on an assumption that there exists one ideal, ground truth image. If such an image is not available, we have to give up the algorithmic problem-solving, inevitably opt for the heuristic approach, and select one of the available options as a relatively quality prototype. For multiple sensors, the quality evaluation of nearly identical images was proposed, e.g., in [28]. This scenario can be observed in CRUSE scanning multiple modes offered by the CSx software tool. The authors in [21] compared *LRFB* scans (in the role of ground truth) against other ones, using the Harris corner detector, Hungarian edges, and SSIM (structural similarity index measure) comparisons. This detector suffered from a low count of Harris corners for certain hard-to-scan images. Our goal in coSIFT [27] research was to improve the gray level method [21] by a more color-sensitive corner detector, which we describe in detail below. These data will be used to compute the coSITF heatmaps, which will be confronted with the measured heatmaps from the evaluation of eye-tracking sessions.

### 2.2.3. Eye Tracking Session

To find out which parts of the scanned images are important or interesting for the experts from the Academy of Fine Arts and Design, we recorded their eye movements during an eye-tracking session. The data were collected during a free-viewing of three different digitized photographs, scanned with different scanning modes.

Two restoration professionals with at least five years of experience in the field participated in the study. The participants were seated in front of a computer monitor and asked to view scans of Ambrotype, Daguerreotype, and Tintype photographs. The eye-tracking equipment recorded the participants' gaze patterns as they viewed the scans. The recording took 5 minutes for each participant. The eye-tracking data were analyzed to determine the visual attention patterns of the participants. The setup of the session can be seen in Figure 1.

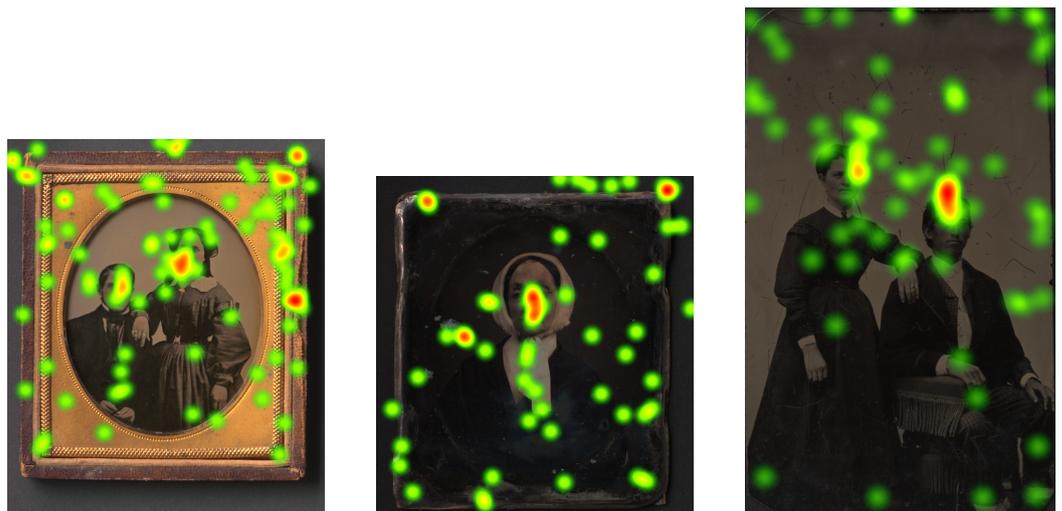
At the beginning of each recording, the eye tracker was calibrated. The calibration precision for the first participant was 92% for the second one 95%. The focus of the visual attention of participants can be visualized using heatmaps. They show how looking is distributed over the images. heatmaps obtained from our eye tracking session can be seen in Figure 2. The heatmaps were generated using the software Tobii Pro Lab [29].

### 2.2.4. Eye Tracker Tobii

We used Tobii Pro Glasses 3 (Figure 3), the third-generation wearable eye-tracking glasses, to capture areas of interest in the images. Tobii Pro 3 glasses provide robust eye tracking and accurate gaze data while giving users freedom of movement. The glasses contain 16 illuminators and 4 eye cameras integrated into scratch-resistant lenses, to provide optimal eye-tracking performance and create an unobstructed view for the user. The scene camera offers a wide field of view (106° H: 95°, V: 63°).



**Figure 1.** Image of the eye-tracking session at Comenius University. The participant, Zuzana Berger Haladová agrees with the publishing of her photo here. Photo by Zuzana Černeková, 2022.



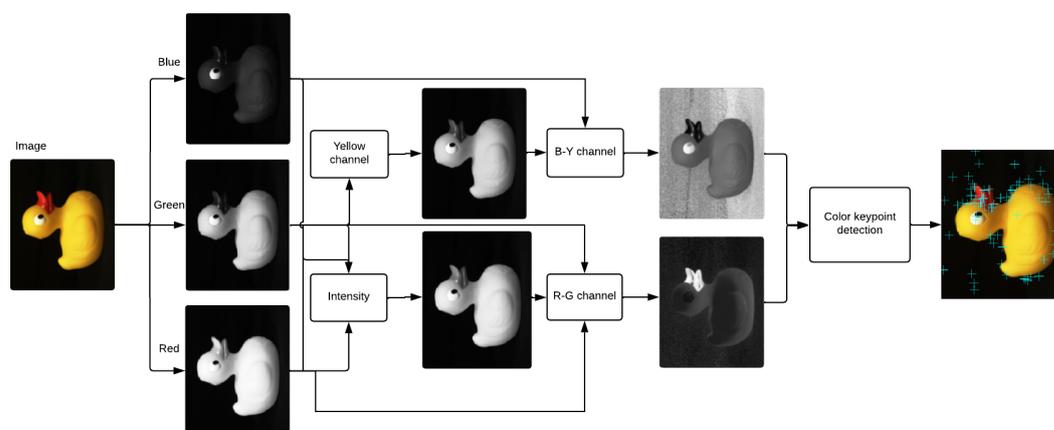
**Figure 2.** These three heatmaps show looking behavior of two restoration expert participants to the Ambrotype, Daguerreotype and Tintype scanned with the *LLa* scanning mode.



**Figure 3.** Tobii Pro Glasses 3.

### 2.3. Co-SIFT

Although the color is perceived as an irreplaceable element describing the world around us, the techniques for extracting the local features are mostly based on the description of the shape, while the color information is entirely ignored. The pipeline of the proposed Co-SIFT [27] algorithm is illustrated in Figure 4. The key idea of our solution is to incorporate color information from the image into the SIFT method by replacing grayscale information so that key points are detected on two separate chromatic channels separately (red–green (RG) and a yellow–blue (YB)) and the achromatic brightness channel. Individual steps of the algorithm are described in detail in the following subsections.



**Figure 4.** The pipeline of the Co-SIFT [27].

#### 2.3.1. Pre-Processing

Stephen Engel et al. in [30] examined the human visual cortex and its color matching using magnetic resonance. The human visual cortex processes signals from photoreceptors on the retina of the eye and interprets the color information. The authors [31] experimentally found that the strongest response is to red–green stimuli. In yellow–blue stimuli, the reaction is also strong, but compared to red–green stimuli, it decreases rapidly. The combination of the trichromatic process and the opponent color-coding process was until recently considered impossible. The trichromatic process, however, speaks of composing colors from several cones, the process of the opponent’s colors, on the other hand, involves finding the colors from their differences. However, the eye works on a much more complicated level, and these two processes perfectly combine. Therefore, as a basic model for our method, we chose, an approach based on experiments in [30], using chromatic opponent channels proposed in [32] and trichromatic color theory, for color image processing and the SIFT method.

The intensity channel (1) is computed as the weighted average of the  $R$ ,  $G$ , and  $B$  values, where the weights were acquired by measuring the intensity perceived by the people with undisturbed trichromatic vision. The weights are the same as used in the standard sRGB space by [33]

$$I = 0.2126 \cdot R + 0.7152 \cdot G + 0.0722 \cdot B. \quad (1)$$

In this color space, the two chromatic channels ( $RG$  and  $YB$ ) proposed in [34] are normalized by the intensity channel, which removes the effect of intensity variations. The channels are defined as follows

$$RG = \frac{(R - G)}{I}$$

and

$$YB = \frac{(B - Y)}{I},$$

where

$$Y = \frac{(R + G)}{2}.$$

Now we proceed with the detection of the interesting points in the chromatic channels.

### 2.3.2. Interesting Point Detection

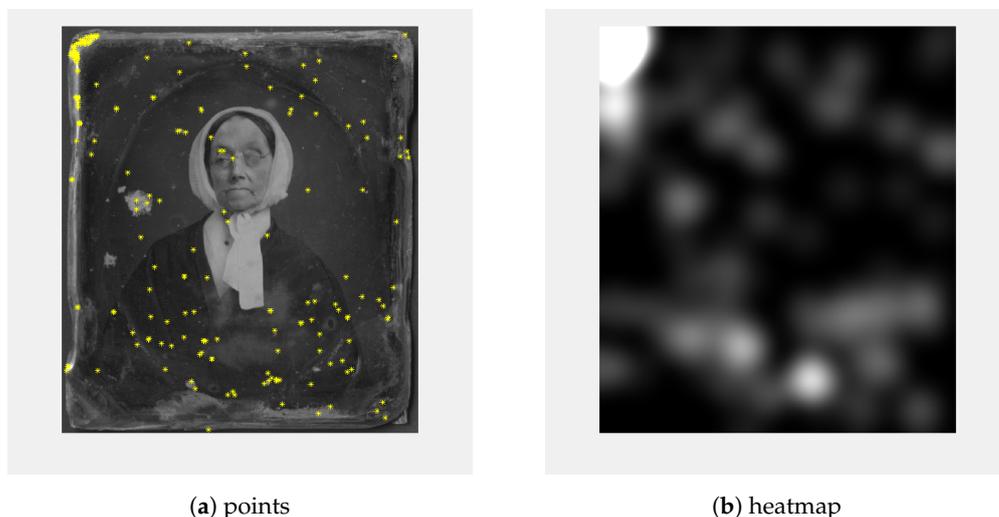
SIFT algorithm introduced by [35] consists of a scale and rotation invariant detector and a HOG (histogram of oriented gradients) descriptor. SIFT detector uses a Gaussian scale pyramid. The image is scaled to  $K$  sizes—octaves. Each octave is then recurrently filtered by a 2D Gaussian.

Two consecutive images in each octave are then subtracted and the resulting  $N - 1$  DoG (difference of Gaussian) images are approximations of LoG (Laplacian of Gaussian) images. The points of interest are identified in the  $3 \times 3 \times 3$  neighborhood in the DoG space. The octave in which the interesting point was found represents the “scale” of the IP and determines the size of the neighborhood for descriptor extraction of that point.

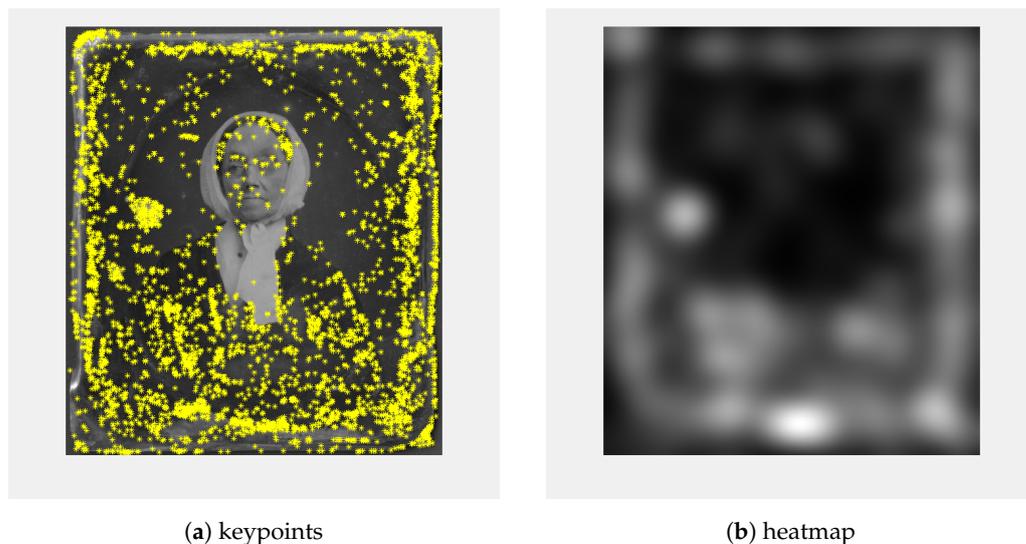
In our method, the Co-SIFT interesting point detection is applied directly to each opponent’s chromatic channel.

As feature detectors, we used co-SIFT and Harris corner detectors [18]; 200 Harris corners were calculated (Figure 5a) and a heatmap (Figure 5b) was created so that for every point a circle blurred with Gaussian blur was added. Similar to the keypoints calculated using Co-SIFT (Figure 6a), the heatmap (Figure 6b) was created.

The number of co-SIFT features, resp. keypoints for input images varied from 4 to 5143.



**Figure 5.** A total 200 points (a) and heatmap (b) calculated using the Harris corner detector.



**Figure 6.** Keypoints (a) calculated using Co-SIFT detector, heatmap (b) calculated from the Co-SIFT keypoints.

### 3. Manual Focus of the Push Broom-like Scanners

Scanners with lenses that capture images in a similar way to cameras do not always have autofocus capability. If we want the resulting scanned image to be sharp enough, we need to know the height of the object to be scanned and make sure that the optical system is well-adjusted and calibrated. Since it is not always easy to meet both conditions, scanners often provide a manual focusing option that uses one of the image sharpness functions mentioned later in this paper.

The sharpness function calculates the sharpness measure or degree of sharpness using the pixel values of the captured image. The value of the sharpness function varies depending on the distance of the lens from the image sensor. The distance of the lens from the sensor is provided by a precision stepper motor, which searches for the position at which the image captured by the sensor is the sharpest. A suitable sharpness function should meet at least the following criteria [36]:

- Unimodality—the sharpness function contains one significant extreme;
- Accuracy—the maximal sharpness is at the extreme of the function;
- Monotonicity—the sharpness function is monotonic at a sufficient distance from the extreme.

There are many sharpness functions. According to [37,38], they can be divided into functions using differences, image convolution, Fourier transform, image entropy, and others. All sharpness functions exploit the fact that details in an image are more visible when the image is sharper.

When looking for an extreme, due to oscillations (rapid changes) of the sharpness function, it is possible to use smoothing of the function, e.g., using a moving average [39]. Since a color image usually consists of three color components  $R$ ,  $G$ ,  $B$ , it is necessary to either convert the image to grayscale or calculate the sharpness function for only one color component [40].

In this section, we will consider the image of size  $M \times N$  as a 2D function  $f(x, y)$ , for  $x = 0, 1, \dots, M - 1$  and  $y = 0, 1, \dots, N - 1$ . The  $x, y$  values are the spatial coordinates and the value of  $f(x, y)$  at any point  $(x, y)$  denotes the intensity at that point (pixel).

### 3.1. Sharpness Functions Based on Differences

These functions compute a measure of sharpness using the sum of the differences of the values of adjacent pixels in either the horizontal or vertical direction, or in both directions simultaneously [36–38,40]

$$S_{df,1}(x,y) = \sum_{x=0}^{M-k-1} \sum_{y=0}^{N-1} |I(x+k,y) - I(x,y)| \quad \text{or}$$

$$S_{df,1}(x,y) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-k-1} |I(x,y+k) - I(x,y)|,$$

$$S_{df,2}(x,y) = \sum_{x=0}^{M-k-1} \sum_{y=0}^{N-k-1} (|I(x+k,y) - I(x,y)| + |I(x,y+k) - I(x,y)|).$$

In general, the difference between more than two rows or columns can also be used, which partially eliminates the sensitivity of these measures to noise in the image [36]. The value of  $k$  is usually chosen as 1 or 2, but [36] states that a higher value can be used.

### 3.2. Sharpness Functions Based on Image Convolution

These sharpness functions use image filters that highlight edges in the image at an appropriately chosen threshold  $T$  [36–40].

The resulting sharpness function is computed using the convolution operator “\*” of the image function  $I(x,y)$  with the given convolution function  $\omega(x,y)$

$$I'(x,y) = (\omega * I)(x,y) = \sum_{r=-k}^k \sum_{s=-k}^k \omega(r,s)I(x-r,y-s), \text{ for } k = 1, 2, \dots \quad (2)$$

Of the relatively large number of sharpness functions using convolution, the following are the most used:

$$S_{ic,1}(x,y) = \sum_{x=1}^{M-2} \sum_{y=1}^{N-2} |L(x,y)|, \text{ for } |L(x,y)| > T,$$

$$S_{ic,2}(x,y) = \sum_{x=1}^{M-2} \sum_{y=1}^{N-2} \left( G_x^2(x,y) + G_y^2(x,y) \right), \text{ for } G_x^2(x,y) + G_y^2(x,y) > T^2,$$

$$\text{where } L(x,y) = (L * I)(x,y), \quad L = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix},$$

$$G_x(x,y) = (S_x * I)(x,y), \quad S_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \text{ and}$$

$$G_y(x,y) = (S_y * I)(x,y), \quad S_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}.$$

### 3.3. Sharpness Functions Based on Fourier Transform

Sharpness functions using the Fourier transform are computationally intensive but are sufficiently robust to the moderate noise in the image coming from image sensors. These functions require that the image be first transformed into the frequency domain using a discrete Fourier transform (DFT) [36,38,40,41] or a discrete cosine transform (DCT) [42]. In a sharp image, the higher frequency values are larger, in an unfocused image they are smaller. Similarly, the values at lower frequencies are larger for an unfocused image and

smaller for a focused image [41]. To define sharpness functions, it is often convenient to use only a certain frequency interval  $[T_{low}, T_{high}]$  [36].

The discrete Fourier transform  $F(u, v)$  of the image function  $I(x, y)$  is expressed using the formula

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) e^{-i2\pi(u\frac{x}{M} + v\frac{y}{N})},$$

but to calculate the sharpness function we use a formula in which the DFT is written using the real and imaginary components

$$\begin{aligned} F(u, v) &= \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) \cos\left(2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right) \\ &\quad - i \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) \sin\left(2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right) \\ &= \operatorname{Re}(F(u, v)) - i\operatorname{Im}(F(u, v)). \end{aligned}$$

Based on the previous formula, we calculate the sharpness function using one of the following formulas

$$\begin{aligned} S_{ft,1}(x, y) &= \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} |F(u, v)|, \text{ while } T_{low} < u + v < T_{high}, \\ S_{ft,2}(x, y) &= \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} ||F(u, v)| \cdot \arg(F(u, v))|, \end{aligned}$$

where  $|F(u, v)| = \sqrt{\operatorname{Re}^2(F(u, v)) + \operatorname{Im}^2(F(u, v))}$ ,  $\arg(F(u, v)) = \arctan\left(\frac{\operatorname{Im}(F(u, v))}{\operatorname{Re}(F(u, v))}\right)$ .

### 3.4. Experiments with Manual Focusing

For use in photography, sharpness measures based on differences and measures based on image convolution are the most suitable, according to [37,38]. Sharpness functions based on the Fourier transform are sufficiently robust not only to the surface texture of the captured object and different lighting conditions but also to noise in the image [41,42]. However, their accuracy is slightly worse than that of the previously mentioned methods.

The Cruse scanner uses one of the above-mentioned sharpness functions for manual focusing. It has an advanced interface for manual focusing, which can be seen in Figures 7 and 8. In this interface, it is possible to change the distance of the lens from the sensor (current position) by adjusting the speed of the stepper motor (lens speed) that moves the lens or the entire sensor head (head speed). We selected the green color channel to calculate the sharpness function. We set the minimum threshold (black threshold) to 0.25 and the maximum (white threshold) to 0.75. The scanner control algorithm then calculated the sharpness measure (installation items—current) for the current moving lens position and continuously displayed the calculated maximum value (installation items—max). After verifying the maximum value using the graph of sharpness function (see Figure 9), we set the scanning head and lens to the position (Optimum At:) where the sharpness function had the largest value.

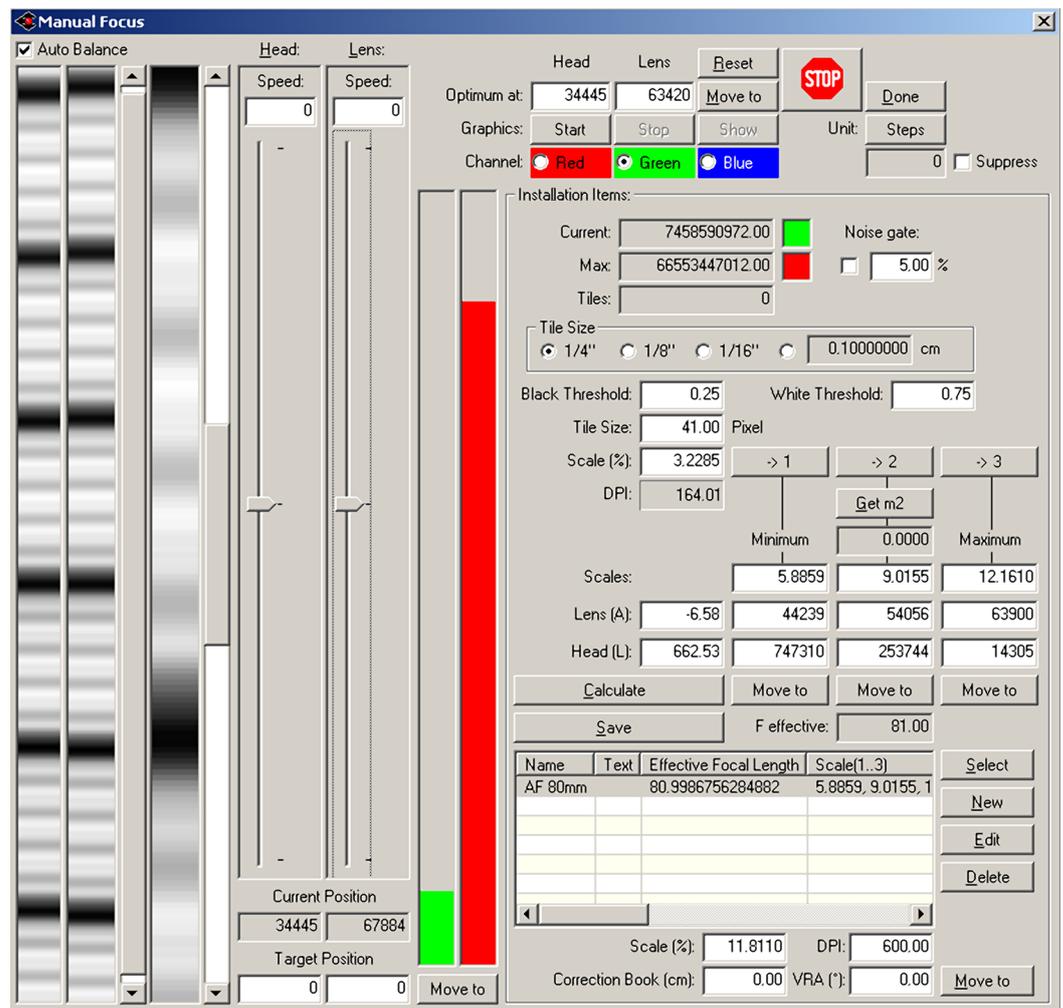


Figure 7. Cruise manual focus dialog for the unfocused picture.

A magnified picture of one column of the unfocused scanned image can be seen in Figure 7, and a picture of the focused image can be seen in Figure 8. The sharpened image has larger sharpness values, which are manifested by a distinct and clearly delineated alternation of dark and light bands.

We verified the manual focusing procedure by scanning on an inaccurately calibrated scanner on multiple objects and at different resolutions. One of them was a black and white ruler (see Figure 10) with height  $h = 0.0$  cm, for which we used a scanner resolution of 600 ppi. The  $4\times$  magnified cutout (see Figure 11 left) shows a poorly focused image. The scanner makes it possible to obtain a sharper image by setting a different height than the actual height of the object being scanned, but at the cost of increased chromatic aberration (see Figure 11 in the middle). After manual sharpening using the sharpness function  $S_{ft,1}(x, y)$ , we obtained a significantly sharper image without chromatic aberration (see Figure 11 right).

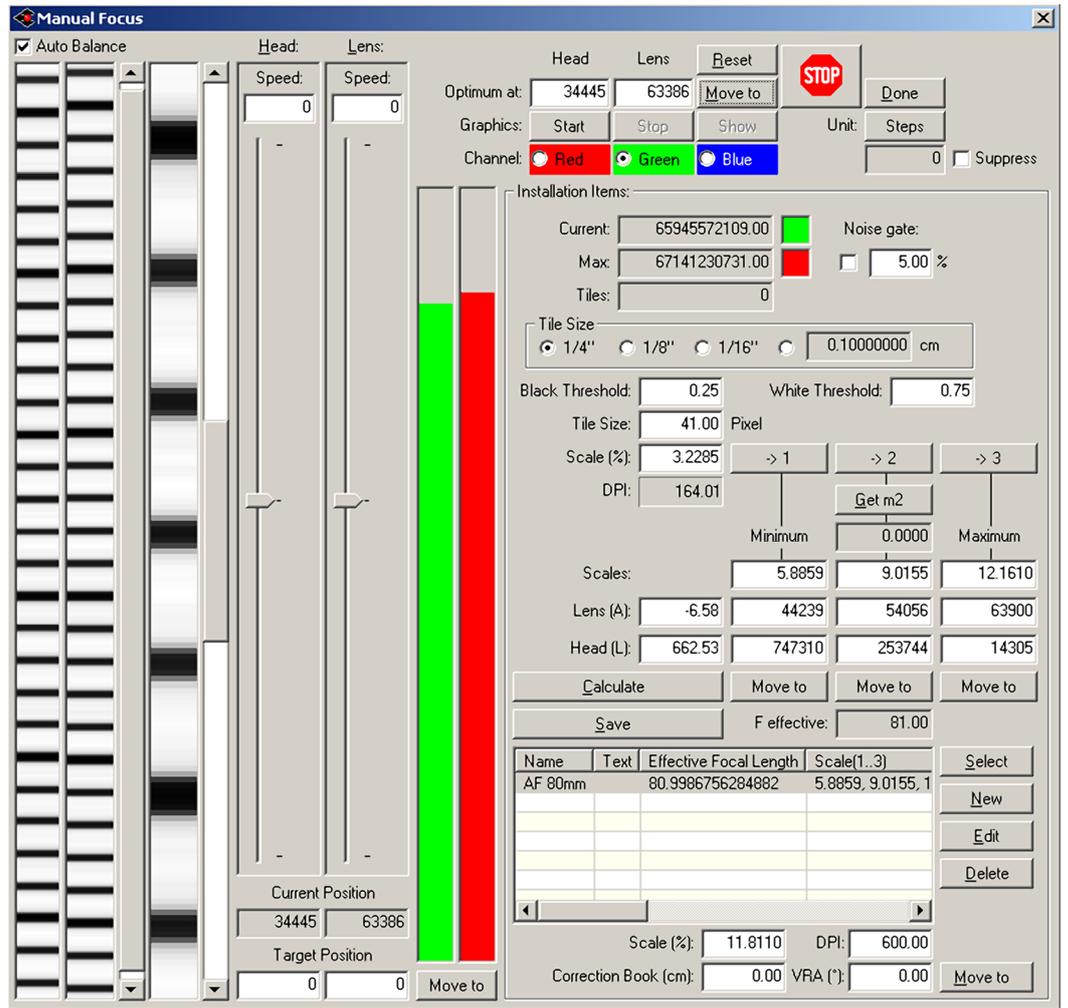


Figure 8. Cruse manual focus dialog for the focused picture.

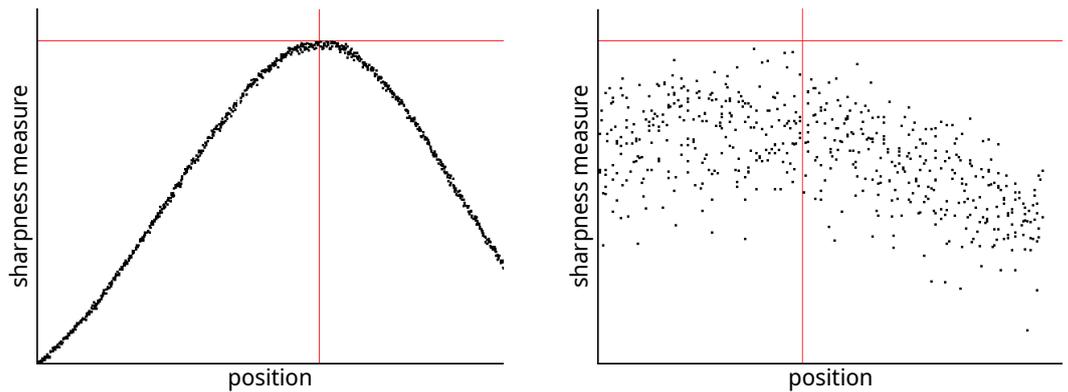
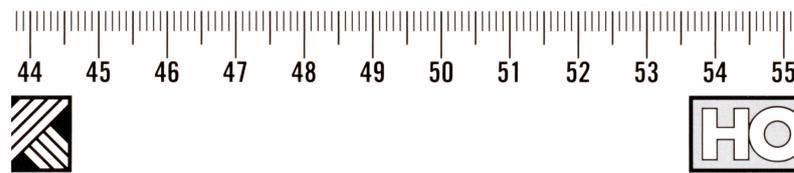
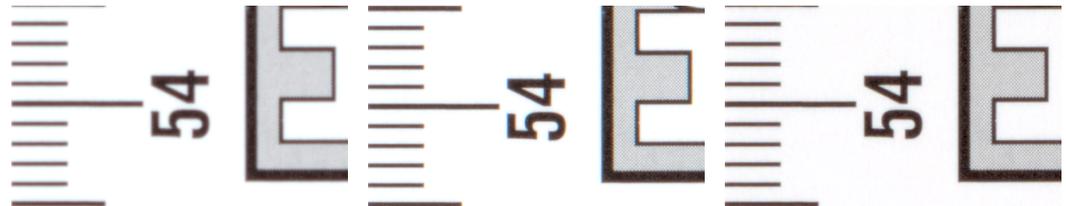


Figure 9. Cruse Focus Value Diagram: graph of sharpness function (left), detail of the graph with noise (right).



**Figure 10.** Picture of scanned black-and-white ruler.



**Figure 11.** Unfocused with  $h = 0.0$  cm (left), unfocused with  $h = 1.4$  cm (center), focused with  $h = 0.0$  cm (right).

The CRUSE Synchron Table scanner makes it possible to manually focus the image even in the case of insufficient calibration or incorrectly estimated height of the scanned object with complex relief. For this purpose, it uses a fairly robust sharpness function based on the aforementioned list of sharpness functions.

#### 4. Comparison of Images Captured in Different Scan Modes of the CRUSE Synchron Table Scanner

As we mentioned earlier, the CRUSE Synchron Table scanner has many lighting modes to choose from. Our goal was to compare different lighting modes and select one that captures the small surface damage of the scanned objects sufficiently well. Thanks to its design, the scanner allows us to scan objects that are not completely flat with sufficient sharpness. This is made possible because when scanning with a resolution of 600 ppi, the depth of field is around 1.0 cm, and at a lower resolution, it is even slightly larger.

Several methods can be used to compare the similarity of the two images. The simplest of them, such as MSE, RMSE, and similar ones, use well-known statistical measures such as the mean squared error, mean absolute error, mean relative error, and the standard deviation or correlation. We could not use simple statistical measures when comparing images, since different lighting modes produce images not only with different average pixel intensities but sometimes with different colors as well. This is most evident when scanning glossy surfaces. Useful measures could be UIQI (universal image quality index), SSIM (structural similarity index measure) and its variations, image entropy value, or histogram comparison, as shown by other authors when comparing two images [43]. It is also possible to use other, less well-known, measures based on spectral distance or the human visual system (HVS) [44]. In our comparison, we did not have to take into account measures that compare images of different sizes or images that are scanned with a different rotation.

Before comparison, we converted all evaluated images into 8-bit grayscale images with sRGB color profile. To convert from a color image to a grayscale image, we used the Equation (1) that takes into account the human perception of color intensity [45].

In the following text, we compare two equal-sized grayscale images  $I(x, y)$  and  $J(x, y)$  of size  $M \times N$ , with pixel intensities from 0 to 255, for  $x = 0, 1, 2, \dots, M - 1$  and  $y = 0, 1, 2, \dots, N - 1$ .

The simple differences in the image pixels generally do not give good results, so we used other measures to compare the two images.

#### 4.1. Universal Image Quality Index

The first measure we used to compare two images was the Wang and Bovik UIQI measure [46], which uses common image measures, such as mean and standard deviation. This measure simultaneously compares the structural distortion (correlation loss)

$$s(I, J) = \frac{\sigma_{IJ}}{\sigma_I \sigma_J},$$

luminance distortion

$$l(I, J) = \frac{2\mu_I \mu_J}{\mu_I^2 + \mu_J^2},$$

and contrast distortion

$$c(I, J) = \frac{2\sigma_I \sigma_J}{\sigma_I^2 + \sigma_J^2}$$

in the resulting formula for calculating the similarity index of the two compared images

$$\text{UIQI}(I, J) = s(I, J) \cdot l(I, J) \cdot c(I, J) = \frac{4\sigma_{IJ}\mu_I\mu_J}{(\mu_I^2 + \mu_J^2)(\sigma_I^2 + \sigma_J^2)}, \quad (3)$$

where  $\mu_I, \mu_J$  are the average pixel values of the first and second image,  $\sigma_I$  and  $\sigma_J$  are standard deviations for the first and second image, and finally  $\sigma_{IJ}$  is the covariance of both images.

#### 4.2. Structural Similarity Index Measure

The second measure we used was Wang et al. [47,48] SSIM measure, which is derived from UIQI by adding stabilization constants  $C_1, C_2, C_3$  to the structural distortion

$$s(I, J) = \frac{\sigma_{IJ} + C_1}{\sigma_I \sigma_J + C_1},$$

luminance distortion

$$l(I, J) = \frac{2\mu_I \mu_J + C_2}{\mu_I^2 + \mu_J^2 + C_2},$$

and contrast distortion

$$c(I, J) = \frac{2\sigma_I \sigma_J + C_3}{\sigma_I^2 + \sigma_J^2 + C_3},$$

which also serves as a tolerance factor for the comparison.

In the formula for SSIM, the powers of alpha, beta, and gamma for the above functions are also used:

$$\text{SSIM}(I, J) = l(I, J)^\alpha \cdot c(I, J)^\beta \cdot s(I, J)^\gamma.$$

For easier calculation of the SSIM formula, Wang et al. calculated the last constant as  $C_3 = C_2/2$  and they chose  $\alpha = \beta = \gamma = 1$ . Since the constants  $C_1, C_2$  depend on the given range  $L$  of pixel intensities of the compared images, they are calculated from the formula  $C_1 = (K_1 \cdot L)^2, C_2 = (K_2 \cdot L)^2$ , where  $K_1, K_2 \leq 1$ . After setting these constants, we calculate the similarity measure of the two images using the formula:

$$\text{SSIM}(I, J) = \frac{(2\mu_I \mu_J + C_1)(2\sigma_{IJ} + C_2)}{(\mu_I^2 + \mu_J^2 + C_1)(\sigma_I^2 + \sigma_J^2 + C_2)}. \quad (4)$$

In practice, it is more convenient to compare two images not globally (all at once), but locally [48], in separate pieces using a suitably chosen sliding window, which gradually passes through all the pixels  $(x, y)$  of the compared images  $I$  and  $J$ .

To calculate the average intensity  $\mu_I$  of pixel  $I(x, y)$  and  $\mu_J$  of pixel  $J(x, y)$  in a window of size  $(2k + 1) \times (2k + 1)$  with weights  $\omega(x, y)$ , we used the formulas

$$\mu_I = \sum_{r=-k}^k \sum_{s=-k}^k \omega(r, s) I(x - r, y - s), \text{ for } k = 1, 2, \dots$$

and

$$\mu_J = \sum_{r=-k}^k \sum_{s=-k}^k \omega(r, s) J(x - r, y - s), \text{ for } k = 1, 2, \dots$$

For the standard deviation, we used the formulas

$$\sigma_I = \left( \sum_{r=-k}^k \sum_{s=-k}^k \omega(r, s) (I(x - r, y - s) - \mu_I)^2 \right)^{1/2}$$

and

$$\sigma_J = \left( \sum_{r=-k}^k \sum_{s=-k}^k \omega(r, s) (J(x - r, y - s) - \mu_J)^2 \right)^{1/2}.$$

Finally, for the covariance, we used the formula

$$\sigma_{IJ} = \sum_{r=-k}^k \sum_{s=-k}^k \omega(r, s) (I(x - r, y - s) - \mu_I) (J(x - r, y - s) - \mu_J).$$

The weights  $\omega(x, y)$  were calculated using the Gaussian function

$$g_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-\mu_I)^2 + (y-\mu_J)^2}{2\sigma^2}}.$$

The weights chosen in this way ensure that the local values  $\text{SSIM}_{xy}(I(x, y), J(x, y))$  of neighboring pixels do not change abruptly.

The resulting  $\text{SSIM}(I, J)$  measure is then calculated as the standard mean of local  $\text{SSIM}_{xy}$  values:

$$\text{SSIM}(I, J) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \text{SSIM}_{xy}(I(x, y), J(x, y)). \quad (5)$$

In our experiments, we used  $K_1 = 1.0$ ,  $K_2 = 0.04$ ,  $L = 255$ , window size  $5 \times 5$  and  $\sigma = 1.0$  to calculate the similarity of the two images. Compared to the recommended value of  $K_1 = 0.01$ , in the work of Wang et al. [47], we chose a much higher value of  $K_1$  to disregard very small structural changes, such as dust grains or very small scratches, in the comparisons. The smaller window size and smaller  $\sigma$  value (compared to the standard values in Wang et al.) provided finer details of the drawing on the displayed SSIM difference map.

#### 4.3. Edge Strength Similarity Image Quality Metrics

The third measure we used to compare images scanned in different light modes was the ESSIM (edge strength similarity image quality metrics) by Zhang et al. [49]. This measure exploits the fact that the most significant features of an image are the boundaries (edges) of objects. The authors used Scharr's gradient operator (filter), which they modified to detect not only horizontal and vertical edges but also diagonal edges.

Zhang et al. [49] used the directional derivative of  $\partial I_j(x, y)$  in pixel  $(x, y)$  for direction  $j = 1, 2, 3, 4$  and define the edge "strength" in the diagonal up ( $j = 2$ ) and diagonal down directions ( $j = 4$ ):

$$E_{2,4}(I(x, y)) = \left| \partial I_2(x, y) - \partial I_4(x, y) \right|^p,$$

then the edge strength for the horizontal direction ( $j = 1$ ) and the vertical direction ( $j = 3$ )

$$E_{1,3}(I(x, y)) = \left| \partial I_1(x, y) - \partial I_3(x, y) \right|^p,$$

where  $p$  is a suitable scaling parameter that affects the edge strength.

Since the human visual system is sensitive to the direction of the edge that is more intense, Zhang et al. [49] define the total edge strength for pixel  $(x, y)$  in the image  $I$

$$E(I(x, y)) = \max\{E_{1,3}(I(x, y)), E_{2,4}(I(x, y))\}.$$

To keep the selected directions in a given pixel the same (consistent) for both compared images, the total edge strength for the second compared image  $J$  is defined by the formula:

$$E(J(x, y)) = \begin{cases} E_{1,3}(J(x, y)) & \text{if } E(I(x, y)) = E_{1,3}(I(x, y)) \\ E_{2,4}(J(x, y)) & \text{if } E(I(x, y)) = E_{2,4}(I(x, y)) \end{cases}$$

To compute the directional derivative of  $\partial I_j(x, y)$ , we use the convolution of the image function  $I(x, y)$  with the function  $\omega_j(x, y)$  (see Equation (2)) that highlights the edges. The matrices of the convolution functions  $\omega_j(x, y)$  for each direction  $j = 1, 2, 3, 4$  are as follows:

$$\omega_1 = \frac{1}{16} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 3 & 0 \\ 0 & -10 & 0 & 10 & 0 \\ 0 & -3 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \omega_2 = \frac{1}{16} \begin{pmatrix} 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 10 & 0 \\ -3 & 0 & 0 & 0 & 3 \\ 0 & -10 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 \end{pmatrix},$$

$$\omega_3 = \frac{1}{16} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 10 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & -10 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \omega_4 = \frac{1}{16} \begin{pmatrix} 0 & 0 & 3 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & -10 & 0 \\ 0 & 0 & -3 & 0 & 0 \end{pmatrix}.$$

The resulting value (index) of the similarity of the two images is expressed by the formula

$$\text{ESSIM}(I, J) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \frac{2E(I(x, y))E(J(x, y)) + C}{E^2(I(x, y)) + E^2(J(x, y)) + C}. \quad (6)$$

Parameter  $C$  has a similar role to the SSIM measure and Zhang et al. [49] proposed to calculate it from the formula  $C = (B \cdot L)^{2p}$ , where  $B$  is a predefined constant and  $L$  is range of pixel intensities of the compared images. For the standard 8-bit grayscale images is  $L = 255$ .

In our experiments, we used the same values as suggested by the authors, i.e.,  $B = 0.1$  and  $p = 1.0$  because these constants gave sufficiently good results.

#### 4.4. Index Maps

The UIQI, SSIM, and ESSIM measures make it possible to visualize the differences between the compared images using so-called index (difference) maps. These difference maps are calculated from the local index values for each pixel  $(x, y)$ , similar to the SSIM case where the local indices were  $\text{SSIM}_{xy}(I(x, y), J(x, y))$ . For better visualization, we added the power  $\alpha$  in the difference map values for SSIM and ESSIM indices

$$\text{SSIM}_{map}(I(x, y), J(x, y)) = (\text{SSIM}_{xy}(I(x, y), J(x, y)))^\alpha. \quad (7)$$

In our case, we chose  $\alpha = 1.2$  for SSIM and  $\alpha = 1.5$  for ESSIM. The UIQI measure did not provide good results, so we did not calculate difference maps for it.

#### 4.5. Image Entropy

The last measure that we used in our calculations to compare the two images was the entropy of the image. The image entropy uses the histogram  $h(i)$ , which is calculated from the pixel intensity values of the grayscale image and is expressed by the formula

$$\text{entropy}(I) = \sum_{i=0}^{255} p(i) \log_2 p(i), \text{ where } p(i) \neq 0, \text{ for } p(i) = \frac{h(i)}{\sum_{i=0}^{255} h(i)}. \quad (8)$$

By calculating the image entropy for all scanning modes (see Tables 1–6) of the three scanned objects, we found that the *LLa* scanning mode provides the “most information” in the image. The scan mode that would provide the lowest image entropy for all three scanned objects was not clearly determined. Therefore, we decided to use *LRFB* as a reference (ground truth) scanning mode, which excels in sufficiently homogeneous illumination of the entire scanned area of the object from all sides. In our experiments, we further compared all scan modes with the selected *LRFB* mode. Because the results of the 15-degree tilt scan were significantly different from the standard no-tilt scan, we divided the comparisons into two groups.

**Table 1.** Ambrotype. Comparison of images scanned in different scan modes with a reference image scanned in *LRFB* mode.

Ambrotype	UIQI	SSIM	ESSIM	Entropy
LRFB	1.0000	1.0000	1.0000	6.4305
LR	0.9903	0.9879	0.9962	6.4254
LFB	0.9881	0.9683	0.9887	6.6589
LF	0.9240	0.9573	0.9851	6.6550
LTx	0.9996	<u>0.9441</u>	<u>0.9818</u>	<u>7.0454</u>
LTx5cm	0.9903	0.9634	0.9875	6.6757
LTx10cm	0.9892	0.9667	0.9872	6.4510
RTx	0.9567	0.9453	0.9819	6.7601
RTx5cm	0.9881	0.9608	0.9852	6.5538
RTx10cm	0.9875	0.9643	0.9852	6.4417
LLa	0.9776	<b>0.9162</b>	<b>0.9743</b>	<b>7.0590</b>

**Table 2.** Ambrotype 15°. Comparison of images scanned in different scan modes with a reference image scanned in *LRFB* 15° mode.

Ambrotype 15°	UIQI	SSIM	eSSIM	Entropy
LRFB15deg	1.0000	1.0000	1.0000	6.9887
LR15deg	0.9721	0.9879	0.9963	6.9889
LFB15deg	0.9738	0.9680	0.9887	7.1650
LF15deg	0.9749	<b>0.9649</b>	<u>0.9875</u>	<u>7.1690</u>
L15deg	0.9993	<u>0.9660</u>	<b>0.9874</b>	<b>7.1763</b>

**Table 3.** Daguerreotype. Comparison of images scanned in different scan modes with a reference image scanned in *LRFB* mode.

Daguerreotype	UIQI	SSIM	ESSIM	Entropy
LRFB	1.0000	1.0000	1.0000	5.9694
LR	0.9969	0.9846	0.9951	5.9905
LFB	0.9993	0.9886	0.9965	5.9219
LF	0.9797	0.9871	0.9956	5.8905
LTx	0.9980	0.9858	0.9947	6.0465
LTx5cm	0.9979	0.9841	0.9946	5.9017
LTx10cm	0.9988	<b>0.9704</b>	<b>0.9899</b>	5.8276
RTx	0.9966	0.9844	0.9944	6.0112
RTx5cm	0.9986	0.9850	0.9952	5.9424
RTx10cm	0.9989	0.9806	0.9938	5.8953
LLa	0.9932	<u>0.9793</u>	<u>0.9918</u>	<b>6.0670</b>

**Table 4.** Daguerreotype 15°. Comparison of images scanned in different scan modes with a reference image scanned in *LRFB* 15° mode.

Daguerreotype 15°	UIQI	SSIM	eSSIM	Entropy
LRFB15deg	1.0000	1.0000	1.0000	6.0504
LR15deg	0.9930	<b>0.9843</b>	<u>0.9948</u>	6.0094
LFB15deg	0.9887	0.9873	0.9956	6.0540
LF15deg	0.9865	0.9868	0.9953	<u>6.0584</u>
L15deg	0.9952	<u>0.9847</u>	<b>0.9944</b>	<b>6.1109</b>

**Table 5.** Tintype. Comparison of images scanned in different scan modes with a reference image scanned in *LRFB* mode.

Tintype	UIQI	SSIM	eSSIM	Entropy
LRFB	1.0000	1.0000	1.0000	5.5269
LR	0.9969	<u>0.9783</u>	0.9928	5.5345
LFB	0.9995	0.9801	0.9930	5.5269
LF	0.9738	0.9817	0.9937	5.5560
LTx	0.9996	0.9792	0.9929	<u>5.6443</u>
LTx5cm	0.9973	0.9804	0.9935	5.5671
LTx10cm	0.9970	0.9784	0.9931	5.4833
RTx	0.9932	0.9797	<u>0.9928</u>	5.5510
RTx5cm	0.9983	0.9816	0.9938	5.4741
RTx10cm	0.9992	0.9785	0.9929	5.4091
LLa	0.9984	<b>0.9767</b>	<b>0.9920</b>	<b>5.7240</b>

**Table 6.** Tintype 15°. Comparison of images scanned in different scan modes with a reference image scanned in *LRFB* 15° mode.

Tintype 15°	UIQI	SSIM	eSSIM	Entropy
LRFB15deg	1.0000	1.0000	1.0000	5.5559
LR15deg	0.9888	0.9844	0.9949	5.5757
LFB15deg	0.9889	0.9840	0.9946	5.6118
LF15deg	0.9891	<b>0.9824</b>	<b>0.9938</b>	<u>5.6180</u>
L15deg	0.9997	<u>0.9837</u>	<u>0.9944</u>	<b>5.6524</b>

#### 4.6. Difference Image and Edge Accentuation

At the end of the experiment, we attempted to visualize the differences between the *LRFB* reference scan mode and the other modes to see how the individual images scanned in the different modes differed. The simplest way was to create a difference image using ordinary difference

$$D_{IJ}(x, y) = |I(x, y) - J(x, y)|, \text{ for } x = 0, 1, \dots, M - 1 \text{ and } y = 0, 1, \dots, N - 1.$$

The difference images did not have enough contrast, so we adjusted them by rescaling the intensity

$$D_{IJ}(x, y) = \frac{D_{IJ}(x, y) - D_{min}}{D_{max} - D_{min}},$$

where  $D_{max} = \max\{D_{IJ}(x, y)\}$  and  $D_{min} = \min\{D_{IJ}(x, y)\}$ . We then reduced the gamma value from 1.0 to 0.5 and inverted the intensities of the adjusted difference image.

Since the *LLa* scanning mode provided images with the largest difference and entropy compared to the *LRFB* reference mode (see tables below), we decided to highlight the edges directly in the scanned images to see if, in addition to the edges of the objects, damaged parts (scratches, pinholes, cavities, etc.) would also be visible in the image. In this case, any available edge filter (Sobel, Canny, Prewitt, Roberts, Laplace, etc.) could be used, but we chose the simplest Roberts edge detector. We modified this procedure slightly by using a coefficient of  $k = 4$  to make the edges more visible. We applied the convolution function  $\omega(x, y)$  to the image function  $I(x, y)$  and have

$$I'(x, y) = \left( k \cdot (G_x^2(x, y) + G_y^2(x, y)) \right)^{1/2},$$

where

$$G_x(x, y) = (\omega_x * I)(x, y), \quad \omega_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and

$$G_y(x, y) = (\omega_y * I)(x, y), \quad \omega_y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

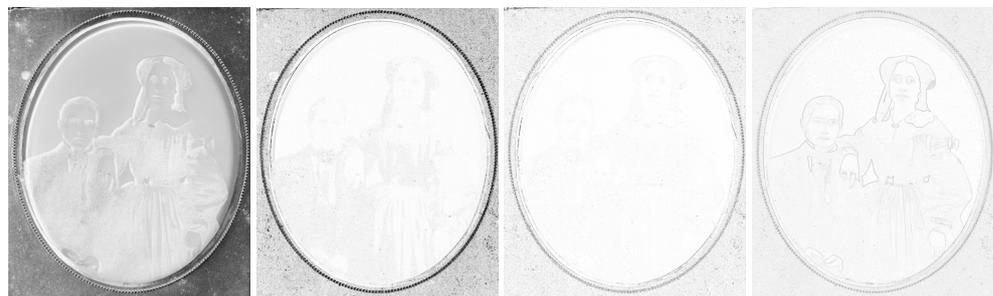
We can see the grayscale Figures 12 and 13, the adjusted difference images, difference maps and highlighted edges on images scanned in *LLa* mode in Figures 14–16.



**Figure 12.** Grayscale images of ambrotype, daguerreotype, and tintype scanned in *LRFB* mode.



**Figure 13.** Grayscale images of ambrotype, daguerreotype, and tintype scanned in *LLa* mode.



**Figure 14.** Difference images of ambrotype between *LLa* mode and *LRFB* reference mode. Left to right: difference image, SSIM map, ESSIM map, filtered edges of the image scanned in *LLa* mode.



**Figure 15.** Difference images of daguerreotype between *LLa* mode and *LRFB* reference mode. Left to right: difference image, SSIM map, ESSIM map, filtered edges of the image scanned in *LLa* mode.



**Figure 16.** Difference images of tintype between *LLa* mode and *LRFB* reference mode. Left to right: difference image, SSIM map, ESSIM map, filtered edges of the image scanned in *LLa* mode.

## 5. Results

In this section, we present the results we obtained by comparing images scanned in different scanning modes with images scanned in the *LRFB* reference mode. We used the

UIQI, SSIM, and ESSIM measures described earlier. For each image, we also calculated the image entropy.

The results presented in the Tables 1–6 show that *LLa* lighting mode generally provides the lowest similarity to *LRFB* lighting mode and has the highest image entropy. Values in bold indicate the lowest similarity to the image scanned in *LRFB* reference mode or the highest image entropy. Underlined values indicate the second-lowest similarity to the image scanned in *LRFB* reference mode or the second-highest image entropy. The largest entropy value in an image means that the image contains the most image information. This means that the surface damage of scanned objects is more visible in the *LLa* lighting mode than in the *LRFB* or other lighting modes. We can say that the shadows “added” by *LLa* mode are quite significant even when scanning flat objects. This information, together with difference images and difference maps, helps restorers locate damaged parts and determine the extent of surface damage to scanned objects. In conclusion, we can say that the *LLa* lighting mode is the best of all lighting modes when it is necessary to capture even minor damage to scanned objects.

## 6. Qualitative Evaluation of Selected Scans

Within this section, we cite the human expertise of experienced restoration professionals. They are responsible for the initial selection of three hard-to-photograph originals, the creation of the above-cited Europeana-style metadata, and the selection of the below commented DAG, AMB, and TIN scans or their processed versions in PNG image format. There are two levels of observation. First, the quality of the scan itself is for documentation and presentation. Second, the visual damage serves educational and possibly future restoration purposes.

### 6.1. Ambrotype

The grayscale image of ambrotype scanned with *LRFB* mode (see Figure 12 left): black and white flattened image of the hand-colored ambrotype allows us to look at the image as a whole without being disturbed by the coloring and the three-dimensionality of the image and glass surface. By viewing this image, we can focus more on the composition and tonality of the ambrotype.

The difference image of the *LRFB* vs. *LLa* scan mode (see Figure 14 left): does not give us much information needed for the correct identification or conservation treatment.

The SSIM map (see Figure 14 middle–left): does not give us much information needed for the correct identification or conservation treatment.

The ESSIM map (see Figure 14 middle–right): does not give us much information needed for the correct identification or conservation treatment.

The image of filtered edges of the picture scanned in *LLa* mode (see Figure 14 right): in this case, this type of rendering is useful mainly for showing the deterioration of the background of the image in form of little dots.

### 6.2. Daguerreotype

The grayscale image of daguerreotype scanned with *LRFB* mode (see Figure 12 middle): The difference image of the *LRFB* vs. *LLa* scan mode (see Figure 15 left): emphasize the imperfection of the daguerreotype (the plate) such as abrasions, scratches gouges, and corrosion.

The SSIM map (see Figure 15 middle–left): does not give us much information needed for the correct identification or conservation treatment.

The ESSIM map (see Figure 15 middle–right): does not give us much information needed for the correct identification or conservation treatment.

The image of filtered edges of the picture scanned in *LLa* mode (see Figure 15 right): all imperfections are here even more detailed visualized than on the Figure 15 left. It is hard to tell though if there are only imperfections of the daguerreotype plate showing here or also the small surface particles such as dust. This could be a little misleading at the end.

### 6.3. Tintype

The grayscale image of the tintype scanned with *LRFB* mode (see Figure 12 right): in the case of tintype, this rendering is not so much different from the original scan.

The difference image of the *LRFB* vs. *LLa* scan mode (see Figure 16 left): does not give us much information needed for the correct identification or conservation treatment.

The SSIM map (see Figure 16 middle-left): does not give us much information needed for the correct identification or conservation treatment.

The ESSIM map (see Figure 16 middle-right): does not give us much information needed for the correct identification or conservation treatment.

The image of filtered edges of the picture scanned in *LLa* mode (see Figure 16 right): this image is very helpful for determining the exact number of scratches in form of lines. It also shows the minor scratches on the photographs that otherwise might not be detected with the naked eye. Very helpful rendering in this case.

One can conclude that qualitative judgment offers support for rendering differences in addition to the primary look of a given digital twin of a cultural object. Moreover, there are indications for future research and the utility of tested methods for restoration practice.

## 7. Discussion

In general, we are aware of three types of uncertainty: vagueness (fuzziness), non-specificity, and controversy (strife). They cause certain experiences with scanning natural or cultural heritage items. In [22], we report on the missing data, limitations of sharpening or scanner precision, and lack of expertise when scanning such unusual originals, such as rare handmade papers for creative use in achieving real-media effects, resp., “ultra-realistic papers, canvases, and lithography stones developed in cooperation with academic professionals” [50]. The non-specificity of describing the original can be improved by clarifying metadata with respect to Europeana requirements or even to the international metadata standard CIDOC-CRM. The most challenging uncertainty seems to be the controversy. It is caused by multiple ground truths. Our InfoVis differentiation may help to support the decision-making in the time of scanning, and increase the accessibility of certain intuition for beginners or students, but this sort of machine perception does not suffice for restoration experts.

Our results can be seen in methodology, original problem statement, and workflow design. In particular, we contributed to the study of image formation with manual focusing enabled by CRUSE technology and postprocessing inspired by image quality methods and machine perception (feature points, novel coSIFT, evaluating SSIM, saliency, and other image quality properties). Last but not least, we confronted the automated processing with the perceptual experiment in the eye-tracking case study. Combining CRUSE scans with eye-tracking evaluation seems to be novel and not available in the given literature. The meaning of our approach can either influence the deeper study of already scanned and published images or even help to propose better or faster procedures in future digitization.

## 8. Conclusions

Our novel workflow for relatively systematic documenting of rare cultural objects can be characterized as a virtual museum creation. First, we chose the globally unique primary dataset of cultural objects. Second, we produced the secondary digital dataset. Third, we designed the scenarios for the presentation of their properties and comparisons. Fourth, we planned and implemented the integrated hardware and software solution. Fifth, we designed the scenarios for presentation, e.g., the web page [9] or this paper. Sixth, we integrated, verified, and tested our “virtual museum or exhibition”. Seventh, we publish and evaluate our findings, contribution, and data for target groups, resp. interests, including students at our faculties. As research on the CRUSE scanner properties is in its beginnings, we omitted previous work surveys, and except for the sharpening and coSIFT parts, we just mentioned relevant ideas at the appropriate places in the text. The sharpening

or manual focusing, coSIFT, and scanning mode comparisons are systematic, but certain parts, such as heatmap comparisons are, obviously, in the proof-of-concept state.

Our experiments with image similarity and entropy measures, as well as evaluations by photo restoration experts, confirmed that among all lighting modes, *LLa* is the most advantageous. In some cases, differential images and differential maps are also useful, helping to visualize even very subtle surface deterioration that is not visible at first glance.

Our methodological and practical contributions imply several future research avenues (e.g., precision, perception, presentation, statistic evaluation, and appropriate visualization), which we support via the published dataset [9] under CC BY-NC 4.0 license.

**Author Contributions:** Conceptualization, Z.B.H. and J.B.K.; methodology, R.B. and A.F.; software, Z.Č., P.Š. and R.B.; validation, J.B.K., J.H. and K.B.; formal analysis, R.B.; investigation, Z.Č. and P.Š. and R.B.; resources, J.B.K., J.H. and K.B.; data curation, R.B.; writing—original draft preparation, Z.B.H.; writing—review and editing, Z.B.H., Z.Č., P.Š., R.B., A.F., J.B.K., J.H. and K.B.; visualization, Z.Č. and P.Š. and R.B.; supervision, Z.B.H.; project administration, Z.Č.; funding acquisition, Z.B.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** The CRUSE scanner was funded by Comeniana grant no. ITMS 26240220077 (OPVaV-2011/4.2/07-SORO). The Tobii Pro Glasses was funded under the Operational Program Integrated Infrastructure for the project: Advancing University Capacity and Competence in Research, Development, and Innovation (ACCORD), co-financed by the European Regional Development Fund.

**Institutional Review Board Statement:** Not Applicable.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the eye-tracking study.

**Data Availability Statement:** The data supporting the reported results can be found at [9], available under the CC BY-NC 4.0 license.

**Acknowledgments:** This publication is the result of support under the Operational Program Integrated Infrastructure for the project: Advancing University Capacity and Competence in Research, Development and Innovation (ACCORD), co-financed by the European Regional Development Fund. The Cruse scanner was supported by Comeniana, grant no. ITMS 26240220077 (OPVaV-2011/4.2/07-SORO) in cooperation with EDICO SK and Comenius University. We thank the owners of the originals used. Finally, we express our gratitude to the anonymous reviewers for the ideas on how to improve the written presentation.

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

## Abbreviations

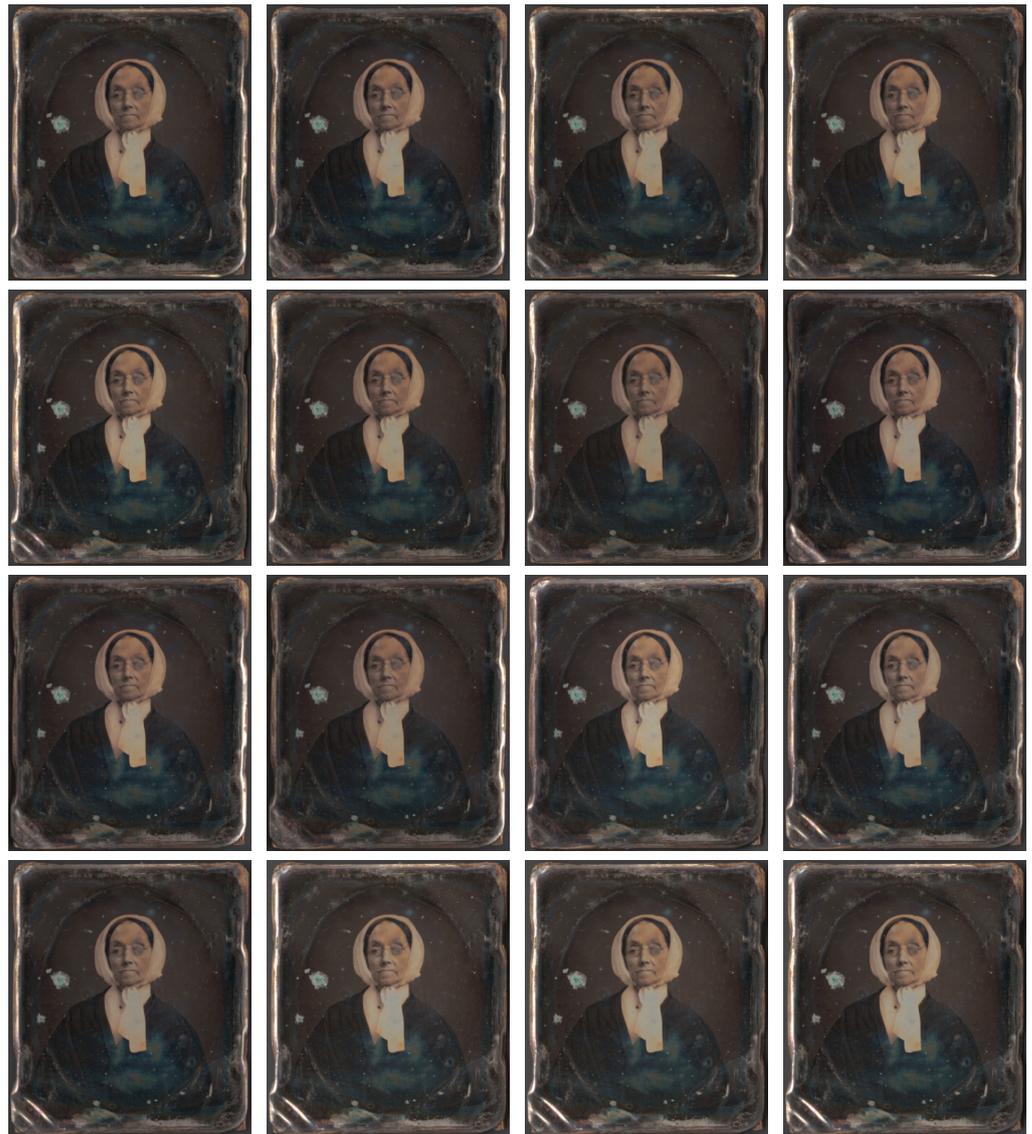
The following abbreviations are used in this manuscript:

AFAD	Academy of Fine Arts and Design
GAN	generative adversarial network
RNG	relative neighborhood graph
MST	Minimum Spanning Tree
AMB	ambrotypes
DAG	daguerreotypes
TIN	tintypes
DT	Delaunay triangulation
SIFT	stop, investigate, find, and trace
DFT	discrete Fourier transform
DCT	discrete cosine transform
PPI	pixels per inch
HVS	human visual system
MSE	mean square error
RMSE	root mean square error
UIQI	universal image quality index
SSIM	structural similarity index measure
ESSIM	edge strength similarity image quality metrics

## Appendix A



**Figure A1.** Images of ambrotypes scanned with different light modes. From left to right: *LRFB, LR, LFB, LF, LTx, LTx 5cm, LTx 10cm, RTx, RTx 5cm, RTx 10cm, LLa, LRFB 15deg, LR 15deg, LFB 15deg, LF 15deg, L 15deg.*



**Figure A2.** Images of daguerreotypes scanned with different light modes. From left to right: *LRFB*, *LR*, *LFB*, *LF*, *LTx*, *LTx 5cm*, *LTx 10cm*, *RTx*, *RTx 5cm*, *RTx 10cm*, *LLa*, *LRFB 15deg*, *LR 15deg*, *LFB 15deg*, *LF 15deg*, *L 15deg*.



**Figure A3.** Images of tintypes scanned with different light modes. From left to right: *LRFB, LR, LFB, LF, LTx, LTx 5cm, LTx 10cm, RTx, RTx 5cm, RTx 10cm, LLa, LRFB 15deg, LR 15deg, LFB 15deg, LF 15deg, L 15deg.*

## References

1. Bee, R. The importance of preserving paper-based artifacts in a digital age. *Libr. Q.* **2008**, *78*, 179–194. [CrossRef]
2. Google Arts and Culture. Available online: <https://artsandculture.google.com/> (accessed on 19 December 2022).
3. Europeana. Available online: <https://www.europeana.eu/sk> (accessed on 19 December 2022).
4. Digital Museums Canada. Available online: <https://www.digitalmuseums.ca/> (accessed on 19 December 2022).
5. Wikipedia: GLAM. Available online: <https://en.wikipedia.org/wiki/Wikipedia:GLAM> (accessed on 19 December 2022).
6. Slovakiana. Available online: <https://www.slovakiana.sk/> (accessed on 19 December 2022).
7. Hojstričová, J. *Renesancia Fotografie 19. Storočia*; VSVU: Bratislava, Slovakia, 2014.
8. Wan, Z.; Zhang, B.; Chen, D.; Zhang, P.; Chen, D.; Wen, F.; Liao, J. Old photo restoration via deep latent space translation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 2071–2087 [CrossRef] [PubMed]
9. CRUSE Scans of Ambrotype, Daguerrotype and Tintype. Available online: <https://flurry.dg.fmph.uniba.sk/scans> (accessed on 30 December 2022).
10. Albright, G.; Barger, M.S.; Baas, V.; Daffner, L.A.; Derby, D.; Edmondson, T.; Fischer, M.; Foster, C.; Gilliland, L.; Hunter, M.L., Jr.; et al. Cased photographs including daguerreotypes, ambrotypes and tintypes (Chapter 2). In *Photographic Materials Conservation Catalog American Institute for Conservation of Historic and Artistic Works (AIC)*; Photographic Materials Specialty Group (PMG): Washington, DC, USA, 1997.
11. Chen, J. A Photodocumentation Method: Taking the Negative Image of a Daguerreotype. 2003, *Unpublished Report*.
12. Daniel, M. Daguerre (1787–1851) and the Invention of Photography. Available online: [https://www.metmuseum.org/toah/hd/dagu/hd\\_dagu.htm](https://www.metmuseum.org/toah/hd/dagu/hd_dagu.htm) (accessed on 19 December 2022).
13. Institute, I.P. Graphic Atlas, Ambrotype. Available online: [http://www.graphicstlas.org/identification/?process\\_id=283](http://www.graphicstlas.org/identification/?process_id=283) (accessed on 19 December 2022).
14. Lavédrine, B.; Frizot, M.; Gandolfo, J.P.; Monod, S. *Photographs of the Past: Process and Preservation*; Getty Publications: Los Angeles, CA, USA, 2009.
15. Blaško-Križanová, J. Laboratory of Restoration of Photography. Available online: <https://www.vsvu.sk/en/departments/departments-of-conservation-and-restoration/laboratory-of-photography-restoration/> (accessed on 19 December 2022).
16. Hartley, R.I.; Gupta, R. Linear Pushbroom Cameras. In *Lecture Notes in Computer Science, Proceedings of the Computer Vision-ECCV'94, Third European Conference on Computer Vision, Stockholm, Sweden, 2–6 May 1994; Volume I*; Eklundh, J., Ed.; Springer: Berlin/Heidelberg, Germany, 1994; Volume 800, pp. 555–566. [CrossRef]
17. Chi, E. A taxonomy of visualization techniques using the data state reference model. In Proceedings of the IEEE Symposium on Information Visualization 2000, INFOVIS 2000, Salt Lake City, UT, USA, 9–10 October 2000; pp. 69–75. [CrossRef]
18. Harris, C.G.; Stephens, M. A combined corner and edge detector. In Proceedings of the 4th Alvey Vision Conference, Manchester, UK, 31 August–2 September 1988; pp. 23.1–23.6.
19. Stancelova, P.; Cernekova, Z.; Ferko, A. Novel Co-SIFT Detector for Scanned Images Differentiation. In Proceedings of the International Conference on Computer Vision and Graphics 2022; Springer Nature: Cham, Switzerland, 2023; pp. 85–98, ISBN 978-3-031-22025-8.
20. CRUSE Spezialmaschinen GmbH. *Cruse Software CSx 3.9 Manual 2018*; CRUSE Spezialmaschinen GmbH: Wachtberg, Germany, 2018.
21. Cerneková, Z.; Berger Haladová, Z.; Blaško Križanová, J.; Hojstričová, J.; Bohdal, R.; Ferko, A.; Zabadal, L. “Hungarian” Image (Differencing) Descriptor. In Proceedings of the DSAI 2022: International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion, NOVA-IMS, Lisboa, Portugal, 31 August–2 September 2022; ACM: New York, NY, USA, 2022.
22. Bohdal, R.; Ferko, A.; Fabián, M.; Bátorová, M.; Hrabovský, M.; Lúčan, L.; Zabadal, L. Adaptive Scanning of Diverse Heritage Originals like Synagogue Interior, Empty Rare Papers or Herbarium Items from the 19 th Century. In Proceedings of the 18th Conference on Applied Mathematics (APLIMAT 2019), Bratislava, Czech Republic, 5–7 February 2019; Curran Associates, Inc.: Red Hook, NY, USA, 2019; pp. 72–82.
23. Polak, M.; Pecko, M.; Bohdal, R.; Ferko, A. Experimental 3D Shape Reconstruction from High-precision 2D CRUSE Scans Using Shadows and Focusing. In Proceedings of the 19th Conference on Applied Mathematics (APLIMAT 2020), Bratislava, Slovakia, 4–6 February 2020; Curran Associates, Inc.: Red Hook, NY, USA, 2020; pp. 878–887.
24. Toussaint, G.T. The relative neighbourhood graph of a finite planar set. *Pattern Recognit.* **1980**, *12*, 261–268. [CrossRef]
25. Cao, Y. Hungarian Algorithm for Linear Assignment Problems (V2.3). MATLAB Central File Exchange. Available online: <https://www.mathworks.com/matlabcentral/fileexchange/20652-hungarian-algorithm-for-linear-assignment-problems-v2-3> (accessed on 19 December 2022).
26. Kuhn, H.W. The Hungarian Method for the Assignment Problem. In *50 Years of Integer Programming 1958-2008-From the Early Years to the State-of-the-Art*; Jünger, M., Liebling, T.M., Naddef, D., Nemhauser, G.L., Pulleyblank, W.R., Reinelt, G., Rinaldi, G., Wolsey, L.A., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 29–47. [CrossRef]
27. Budzakova, P.; Sikudova, E.; Berger Haladova, Z. Color object retrieval using local features based on opponent-process theory. In Proceedings of the International Conference on Computer Vision and Graphics, Warsaw, Poland, 17–19 September 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 275–286.
28. Perko, R.; Ferko, A.; Bornik, A.; Cech, P. Beyond Image quality comparison. In Proceedings of the European Association for Computer Graphics-Eurographics, Granada, Spain, 1–5 September 2003; pp. 271–276.

29. Tobii Connect. Available online: <https://connect.tobii.com/> (accessed on 30 December 2022).
30. Engel, S.; Zhang, X.; Wandell, B. Colour tuning in human visual cortex measured with functional magnetic resonance imaging. *Nature* **1997**, *388*, 68–71. [[CrossRef](#)] [[PubMed](#)]
31. Thornton, J.E.; Pugh, E.N. Red/green color opponency at detection threshold. *Science* **1983**, *219*, 191–193. [[CrossRef](#)] [[PubMed](#)]
32. Jost, T.; Ouerhani, N.; Von Wartburg, R.; Müri, R.; Hügli, H. Assessing the contribution of color in visual attention. *Comput. Vis. Image Underst.* **2005**, *100*, 107–123. [[CrossRef](#)]
33. Anderson, M.; Motta, R.; Chandrasekar, S.; Stokes, M. Proposal for a standard default color space for the internet—sRGB. In Proceedings of the Color and Imaging Conference. Society for Imaging Science and Technology, Scottsdale, AZ, USA, 19–22 November 1996; Volume 1996, pp. 238–245.
34. Swain, M.; Ballard, D. Color Indexing. *Int. J. Comput. Vis.* **1991**, *7*, 11–32. [[CrossRef](#)]
35. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
36. Batten, C.F. Autofocusing and Astigmatism Correction in the Scanning Electron Microscope. Mphil Thesis, University of Cambridge, Cambridge, UK, 2000.
37. Mir, H.; Xu, P.; Van Beek, P. An extensive empirical evaluation of focus measures for digital photography. In Proceedings of the Digital Photography X. SPIE, San Francisco, CA, USA, 2 February 2014; Volume 9023, pp. 167–177.
38. Yao, Y.; Abidi, B.; Doggaz, N.; Abidi, M. Evaluation of sharpness measures and search algorithms for the auto-focusing of high-magnification images. In Proceedings of the Visual Information Processing XV. SPIE, Orlando (Kissimmee), FL, USA, 17 April 2006; Volume 6246, pp. 132–143.
39. Pech-Pacheco, J.L.; Cristóbal, G.; Chamorro-Martinez, J.; Fernández-Valdivia, J. Diatom autofocusing in brightfield microscopy: A comparative study. In Proceedings of the 15th International Conference on Pattern Recognition. ICPR-2000, Barcelona, Spain, 3–8 September 2000; IEEE: Piscataway, NJ, USA, 2000; Volume 3, pp. 314–317.
40. Chern, N.N.K.; Neow, P.A.; Ang, M.H. Practical issues in pixel-based autofocusing for machine vision. In Proceedings of the 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164), Seoul, Republic of Korea, 21–26 May 2001; IEEE: Piscataway, NJ, USA, 2001; Volume 3, pp. 2791–2796.
41. Kristan, M.; Pernus, F. Entropy based measure of camera focus. In Proceedings of the Thirteenth Electrotechnical and Computer Science Conference ERK, Portorož, Slovenia, 27–29 September 2004; Citeseer: Princeton, NJ, USA, 2004; pp. 179–182.
42. Jeon, J.; Lee, J.; Paik, J. Robust focus measure for unsupervised auto-focusing based on optimum discrete cosine transform coefficients. *IEEE Trans. Consum. Electron.* **2011**, *57*, 1–5. [[CrossRef](#)]
43. Aljanabi, M.A.; Hussain, Z.M.; Shnain, N.A.A.; Lu, S.F. Design of a hybrid measure for image similarity: A statistical, algebraic, and information-theoretic approach. *Eur. J. Remote Sens.* **2019**, *52*, 2–15. [[CrossRef](#)]
44. Mahamud, S.T.; Rahmatullah, B. Image quality assessment based on properties of HVS and principle of image structure. In Proceedings of the 2015 International Conference on Image and Vision Computing New Zealand (IVCNZ), Auckland, New Zealand, 23–24 November 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1–6.
45. Ware, C. *Information Visualization: Perception for Design*; Morgan Kaufmann: Burlington, MA, USA, 2019.
46. Wang, Z.; Bovik, A.C. A universal image quality index. *IEEE Signal Process. Lett.* **2002**, *9*, 81–84. [[CrossRef](#)]
47. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]
48. Wang, Z.; Bovik, A.C.; Sheikh, H.R. Structural similarity based image quality assessment. In *Digital Video Image Quality and Perceptual Coding*; CRC Press: Boca Raton, FL, USA, 2017; pp. 225–242.
49. Zhang, X.; Feng, X.; Wang, W.; Xue, W. Edge strength similarity for image quality assessment. *IEEE Signal Process. Lett.* **2013**, *20*, 319–322. [[CrossRef](#)]
50. Escape Motions. Available online: <https://www.escapemotions.com/products/rebelle/papers.php> (accessed on 19 December 2022).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.