Resolution analysis of archive films for the purpose of their optimal digitization and distribution

Karel Fliegel[a,b], Stanislav Vítek[a,b], Petr Páta[a,b], Jiří Myslík[b], Josef Pecák[b], and Marek Jícha[b]

[a]Czech Technical University in Prague, Technická 2, 166 27 Prague 6, Czech Republic
[b]Film and TV School of Academy of Performing Arts in Prague, Smetanovo nábřeží 2, 116 65 Prague 1, Czech Republic

ABSTRACT

With recent high demand for ultra-high-definition (UHD) content to be screened in high-end digital movie theaters but also in the home environment, film archives full of movies in high-definition and above are in the scope of UHD content providers. Movies captured with the traditional film technology represent a virtually unlimited source of UHD content. The goal to maintain complete image information is also related to the choice of scanning resolution and spatial resolution for further distribution. It might seem that scanning the film material in the highest possible resolution using state-of-the-art film scanners and also its distribution in this resolution is the right choice. The information content of the digitized images is however limited, and various degradations moreover lead to its further reduction. Digital distribution of the content in the highest image resolution might be therefore unnecessary or uneconomical. In other cases, the highest possible resolution is inevitable if we want to preserve fine scene details or film grain structure for archiving purposes. This paper deals with the image detail content analysis of archive film records. The resolution limit in captured scene image and factors which lower the final resolution are discussed. Methods are proposed to determine the spatial details of the film picture based on the analysis of its digitized image data. These procedures allow determining recommendations for optimal distribution of digitized video content intended for various display devices with lower resolutions. Obtained results are illustrated on spatial downsampling use case scenario, and performance evaluation of the proposed techniques is presented.

Keywords: Digital cinema, archive records, film restoration, image resolution, film scanning, digitization, image quality.

1. INTRODUCTION

A large number of motion pictures is stored on film material, which often has a high historical and artistic value, and must, therefore, be preserved for future generations. Film material, however, might have a relatively short life, since it is subject to chemical decomposition, wear and other kinds of deteriorations. Besides, it is strongly prone to mechanical damage, which also degrades the recorded image information. Degradation can be mitigated by storage in suitable conditions of film archives with controlled temperature and humidity. However, as a consequence, valuable archive content might become confined to the general public.

Given that the quality of moving image information recorded on film material has been met only recently by digital technology, archival film materials are also a valuable source of entertainment programs for ultra-high-definition (UHD) projection in digital cinemas and using modern home entertainment systems. To preserve valuable image information and also to make it accessible to the public, it is necessary to digitize the film

Further author information: (Send correspondence to Karel Fliegel)
Karel Fliegel: E-mail: fliegek@fel.cvut.cz, Telephone: +420 224 352 026
Stanislav Vítek: E-mail: viteks@fel.cvut.cz, Telephone: +420 224 352 232
Petr Páta: E-mail: pata@fel.cvut.cz, Telephone: +420 224 352 248
Jiří Myslík: E-mail: jiri.myslik@famu.cz, Telephone: +420 234 244 326
Josef Pecák: E-mail: josef.pecak@famu.cz, Telephone: +420 234 244 326
Marek Jícha: E-mail: marek.jicha@famu.cz, Telephone: +420 234 244 323

records stored in film archives. The digitized material can also be processed with a variety of advanced image reconstructing methods that will contribute to a better viewing experience.

The choice of scanning resolution and also the resolution of the distribution copies is closely related to the quality of preserved image information. It may seem that scanning film material at the highest possible resolution is the most sensible choice. However, the useful image information content of the film material is limited, and it is further reduced also due to various degradations. In particular, distribution of such content might, therefore, be unnecessary or uneconomical.

This paper deals with the useful image information content analysis of the film material. The theoretical limit of useful image information is discussed as well as the factors that make the actual measured values lower. Finally, there are methods proposed that could determine the information content of each film material individually based on the analysis of its digitized picture information and following its outcome provide requirements for resolution of potentially downsampled distribution copies so that the image quality is not degraded.

More details on the related methodology of digitally restored authorizate (DRA), which defines a procedure and set of tools to achieve the audio and visual appearance of the digitized film as close as possible to the original author's concept, can be found in our previous papers.[1–4]

The content of this paper is organized as follows. Available studies on achievable image resolution of film are briefly summarized in Section 2. The issue of digitized image downsampling and proposed techniques for objective analysis of its impact on obtained image resolution and perceived image fidelity or quality are presented in Section 3. Preliminary performance verification of the proposed techniques and its results are summarized in Section 4. Section 5 concludes the paper and discusses possible future work.

2. IMAGE RESOLUTION IN CINEMATOGRAPHIC FILM CHAIN

This section deals with the question of the appropriate choice of resolution, which is affecting the quality of digitized archive film materials and summarizes related work in the field. At first, we focus on the analysis of the image resolution of a typical 35 mm film chain. Then we shortly deal with the resolving ability of the film stock itself. This analysis then implies the basic requirements for the technical specifications of the film scanner.

2.1 Image resolution analysis in theatrical presentation

There were several studies published since the digital technology onset in the film industry trying to determine requirements for the digital projection. The goal of Baroncini et al.,[5] in one of the first published studies on this topic, was to find out what was the standard resolution common for conventional analog movie theaters. The focus of this study was not on the measurement of the maximum achievable resolution of the film material. The presented results are conditional on the characteristics of the particular material used and other standard technologies, including the projection, and thus describe the limits of resolution that are achieved in the case of the best effort practice. The process was carried out in the state-of-the-art facilities. The modulation transfer function (MTF)[6] was measured using a conventional approach with the test pattern and microdensitometer. Further details on the methodology can be found in the study together with the detailed results obtained for subjective resolution limit evaluation of the images on the cinematographic screen. As pointed out by Morton et al.,[7] the study of Baroncini et al.[5] presents the results that may be misleading when misinterpreted, e.g. the study is focused on the analysis of a system working with a specific image height, the test pattern used has a sinusoidal profile of brightness, and the authors contemplate a limiting visual resolution for a drop to 10% in MTF, etc.

Vitale[8*] published another approach in his study to estimate the resolution of images captured on historical film materials. This approach is based on the equation determining the overall resolving power of the imaging system, taking into account the characteristics of the individual components, in particular, the camera lens and the film material. The study presents tabular and graphical estimates of imaging capabilities for film materials and lenses typical of a given historical period and the resulting estimate of the resolving power of the system.

[*] Timothy Vitale Art Conservation (Accessed: July 23, 2017): http://vitaleartconservation.com/

It should be noted here that, even beyond the resolution limits determined according to the methodologies described above (MTF drop to 10% or 5%), there are image details captured on film material. These are not observable under normal conditions by human sight. However, such details, as far as uncompromising archiving is concerned, are often desirable to capture during digitization.

2.2 Basic resolution requirements for the film scanner

The main goal of this paper is not to present the actual requirements for the resolution of the film scanners or the digital distribution copies but to propose and evaluate techniques to obtain this characteristic from the image data. The following paragraph is included only for the completeness.

Analyzing the information image content captured by the 35 mm film in respect to the film scanner resolution is not a trivial question.[5,9] The limiting resolution calculations are usually based on a simplified Shannon sampling condition,[10] where spatial sampling is sufficient for two image pixel detectors (sampling sites) per line pair, corresponding to the so-called Nyquist frequency of 0.5 cycles per pixel (cy/px). This simplification is based on the assumption that the scanning system has at its input an ideal optical anti-aliasing filter and the sampled image signal is therefore limited in the spatial frequency domain to 0.5 cy/px.

From the simplified analysis,[5,8,11] it follows that the open gate resolution of 4K ($4\,096 \times 3\,112$) can be considered as the lowest resolution limit for scanning the modern 35 mm original camera negative, especially for 35 mm formats with a smaller area, e.g. cropped format of 1.85 : 1. At this point, it is worth recalling that the DCI (Digital Cinema Initiatives)[12] consortium defines the size of specific resolution containers. For larger and higher-resolution formats, there is a certain risk of loss in recorded image information while scanning at 4K. Another significant problem is the possible distortion of the sampled image by aliasing, which appears as a moiré in the acquired image data when the optical anti-aliasing filter[6,13] is imperfect or missing. To prevent this distortion and reliably capture image information carried by the 35 mm film, scanning in a higher format, such as 6K or better 8K, is required. Higher resolution when scanning the original camera negative (OCN) with subsequent downsampling, such as to 4K using a suitable interpolation algorithm and a digital anti-aliasing filter, can better suppress the artifacts associated with sparse sampling. The filter prevents moiré distortion that might occur when sampling high-quality film material with a fine-textured image. This problem can, on the other hand, happen when digitizing using a lower resolution scanner, e.g. 4K and especially 2K, which has an improper optical anti-aliasing filter.

2.3 Film grain and its relation to image resolution requirements

The optimal resolution for scanning must be such as to capture the finest detail of recorded image information. The ultimate goal would be to capture also the film grain to maintain the distinct film look. Although the film grain is essentially noise, it belongs to the film characteristic, and it adds to its perceived naturalness. On the contrary, the unwanted artifact is the noise created by the scanner image sensor during the digitization of the film. Useful image information about the captured scene, film grain and noise are the three distinct properties that should be analyzed separately. Detailed analysis of the three properties is beyond the scope of this paper.

In the digital world, pixel is the finest detail that can be shown in the image. The film material is different in that the fundamental particles are silver crystals in black and white, and color dye clouds in color film. For more details refer to Vitale.[8,11] Thus, theoretically, the resolution of the film should be determined by the size of its fundamental particles. Fundamental particles are dispersed randomly in the emulsion of film material, but with a high density. The emulsion is relatively thick in comparison to the particle size, and therefore the particles are scattered in all three of its dimensions. There is overlapping and clustering of the particles that leads to the known phenomenon of film material, which is its grain. The grain is not an elementary unit constituting the picture information. It only arises as a result of the visual perception when one observes a large number of fundamental particles distributed in a relatively thick film of grain versus resolution. It is the product of the limited resolution of the vision, which is unable to distinguish the actual fundamental particles. The grain is not a particle, and because it is so indistinct, it is difficult to measure.[11]

The film grain can be preserved into the digitized image if the grain characteristics are known, and the digitization takes them into account. Based on the study of Hepper,[14,15] each film record (depending on e.g. material type, exposure, development) has its own grain structure. Film grain, as other noises, negatively influences the performance of image compression algorithms, refer to Oh et al.[16]

3. ANALYSIS OF RESOLUTION LOSS DUE TO SPATIAL DOWNSAMPLING

The image resolution of cinematographic chain plays an important role, especially if the archive film records are digitized and subsequently distributed in a variety of formats targeting the full range of devices, from hand-held devices, e.g. smartphones and tablets, through ultra-high definition (UHD) home entertainment systems, to high-end digital cinematographic theaters. In each of these environments, spatial resolution plays the crucial role and influences the quality of experience (QoE)[17] as perceived by the human observer.

Given the scanner's resolution required for the digital footprint of the archive analog film, it is always necessary to conduct a detailed analysis of the analog film media before scanning, and then perform scanning using the best available technology. Scanning the original camera negative (OCN) for 35 mm film at 8K resolution with the subsequent conversion to lower formats, especially for distribution needs, seems to be optimal. There are suitable formats for storing image data for such workflows. Regarding archiving, it is desirable to preserve the scanned master data without any modifications in the original resolution and bit depth. It should be noted that the film strip degrades with time, and when the image is re-digitized, the image information may be significantly more disturbed or may not be available for digitization anymore. A suitable long-term storage format for this data is MAP (Master Archive Package), which allows up to 16K resolution ($16\,384 \times 8\,640$ pixels), the bit depth of up to 32 bits per color channel, and lossless compression. The next step in this workflow is the IAP (Intermediate Access Package) profile, which represents a high-quality digital copy of the archived film with 4K ($4\,096 \times 2\,160$ pixels), 2K ($2\,048 \times 1\,080$ pixels) resolutions, and subsequent production of distribution formats.[18]

It is beyond the scope of this paper to analyze QoE loss prediction due to spatial subsampling in details for the whole range of possible techniques when image data scanned by high-quality film scanner are used or even objectively assess physical resolution requirements for the film scanner. Thus a particular use case scenario was selected for further analysis.

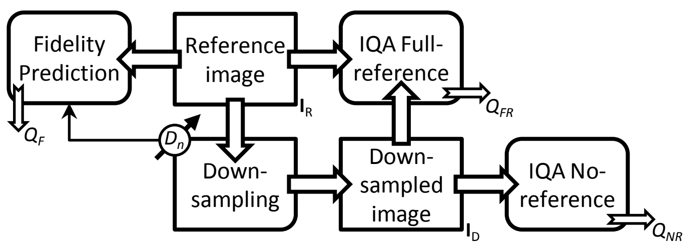


Figure 1: Schematic diagram of analyzed use case scenario. Reference images are downsampled, and their quality is analyzed using image quality assessment (IQA) techniques based on full-reference (FR), no-reference (NR) IQA or image fidelity prediction based on analysis of the reference image.

3.1 Use case scenario

The ultimate goal would be to develop a technique which would provide requirements for the resolution of the film scanner to be used for reliable digitization of the image details captured in the film image, see Section 2 for the overview. Instead, here we consider a simplified use case scenario.

The best practice for critical archive records is that the film is scanned with the highest resolution available to the state-of-the-art technology, e.g. 8K, and then for distribution purposes, lower resolution versions are derived. It is also the use case scenario for further analysis in this paper. The main idea, also captured in Fig. 3.1, can be summarized as follows.

- Most of the useful image details and preferably also the film grain is captured in the high-resolution scan of the film.
- Based on the analysis of the derived downsampled version of the original image or based on the analysis of the original image itself, the deterioration of the downsampled version is predicted.
- The prediction of the image fidelity or quality in the derived downsampled version allows the user to determine the lowest limit of the resolution of the downsampled version so that the image fidelity or quality is not deteriorated.

As mentioned above, fidelity or quality of the downsampled version in respect to the original can be predicted based on the analysis of available image data before and after downsampling. Various options for such technique are discussed in the following paragraph.

3.2 Description of the analyzed techniques

The main goal of the discussed techniques, see Fig. 1, is to predict the fidelity or quality of the downsampled version $\mathbf{I}_D$ of the reference image $\mathbf{I}_R$. The downsampling is performed by a defined algorithm with the downsampling factor D_n. The quality or fidelity of the downsampled image $\mathbf{I}_D$ with respect to the reference $\mathbf{I}_R$ can be predicted using one of the three following approaches.

(a) Full-reference (FR) image quality/fidelity assessment, where both images, the downsampled version $\mathbf{I}_D$ and the reference image $\mathbf{I}_R$, are available and compared, while measure Q_{FR} is obtained.
(b) No-reference (NR) image quality assessment, where only the downsampled version $\mathbf{I}_D$ is available, while measure Q_{NR} is obtained.
(c) Fidelity prediction based only on the reference image $\mathbf{I}_R$, properties of the processing, i.e. downsampling with the downsampling factor D_n, while measure Q_F is obtained.

Each of the three general approaches listed above can be further analyzed based on the particular measure used for the calculation of the resulting quality/fidelity criterion. Selected relevant techniques for each of the three classes are briefly discussed in the following paragraphs. More details can be found in the literature.[19–21]

Full-reference (FR) image quality/fidelity assessment

Full reference measures require the presence of the original/reference image. This type of criteria measure a fidelity of the processed version in respect to the reference, i.e. how similar the two versions are. The FR algorithms provide the most reliable estimates of quality, and thus it is the most common class of measures used in the field of near-lossless image processing techniques, where the fidelity plays very important role.

Since the FR measures are the oldest and the most developed class, there are numerous approaches, ranging from the signal or pixel-based metrics to human visual system based (HVS) metrics. Since the pixel-based metrics, e.g. mean square error (MSE), peak signal to noise ratio (PSNR),[19] despite being computationally efficient, do not correspond well with the subjective perception, these are not further considered. The HVS based techniques provide a better estimation of the perceived image quality. In this class the most popular techniques are e.g. structural similarity index (SSIM)[21] with its derivatives (e.g. MS-SSIM, IW-SSIM), visual information fidelity (VIF),[22] feature similarity index (FSIM),[23] and visible difference predictor (VDP) with its derivatives (e.g. HDR-VDP, DHR-VDP-2, HDR-VDP-2.2).[24]

For our study, the well established and computationally efficient SSIM,[21] was used in the framework, see Fig. 1, to predict the fidelity of the downsampled image $\mathbf{I}_D$ with respect to the reference $\mathbf{I}_R$, while obtaining the measure Q_{FR}.

No-reference (NR) image quality assessment

The full-reference measure requires the original reference image to assess the quality of the processed version. This restriction is eliminated with no-reference measures, where there is no need for the information about the original image.[19] There is a large group of specialized, distortion-aware, measures, designed specifically for certain image processing distortions, e.g. lossy compression, blur/sharpness, contrast, and colorfulness.

For the selected use case scenario, see Fig. 1, the blur/sharpness distortion-aware metrics are the most relevant since the downsampling with low-pass antialiasing filter deteriorates high spatial infrequences, thus introduces blur to the processed image and depreciates its sharpness. Based on our previous work[25] two measures were selected for further assessment within the analyzed framework, i.e. S3 metric proposed by Vu et al.[26] and LPC-SI metric proposed by Hassen et al.[27] S3 algorithm was designed to measure the local perceived sharpness in images. An application of derived S3 measures was demonstrated for no-reference image quality assessment of blurred images[27] or derived techniques augmented for image sharpening.[25] No-reference image sharpness assessment based on local phase coherence measurement (LPC-SI) proved to correlate well with subjective evaluations on blur image database and also for the image sharpening assessment.[25,27]

In our study, the well established no-reference measures, S3[26] and LPC-SI,[27] were used in the framework, see Fig. 1, to predict the quality of the processed image $\mathbf{I}_D$ considering its sharpness deterioration[25] due to downsampling, while obtaining the measure Q_{NR}.

Fidelity prediction

Full-reference and no-reference techniques, as described in the previous two paragraphs, use the processed image $\mathbf{I}_D$, see Fig. 1, to predict its fidelity or quality. The processing responsible for the image quality deterioration is downsampling in our use case scenario. Thus it might be reasonable to predict the fidelity of the processed image. A technique for such image fidelity prediction is described in this paragraph.

Image fidelity of $\mathbf{I}_D$ can be predicted based on the analysis of the original reference image $\mathbf{I}_R$ and known characteristics of the processing, i.e. downsampling factor D_n. Here we introduce a simple fidelity measure based on power spectral density (PSD) characteristics of the reference image in respect to the prediction of the PSD of the distorted image. The PSD describes the distribution of the image signal power with respect to its spacial frequency.[10] The PSD of an image can be calculated as the squared magnitude of its Fourier transform

$$PSD_{\mathbf{I}}(k,l) = |\mathcal{F}\{\mathbf{I}(m,n)\}|^2, \tag{1}$$

where $\mathcal{F}$ represents the Fourier transform (here two-dimensional discrete Fourier transform), $\mathbf{I}(m,n)$ is the image intensity, and (m,n) are rows and columns of the 2D image.

The process of downsampling can be described to some extent by its optical transfer function (OTF) and modulation transfer function (MTF).[6] The shape of resulting MTF will depend mostly on the interpolation kernel used and the shape of antialiasing low-pass filter.[10] If the proper antialiasing filter is used then the power spectral density PSD_D of the downsampled image $\mathbf{I}_D$ is related to the PSD_R of the reference image $\mathbf{I}_R$ as follows

$$PSD_D(k,l) = |H(k,l)|^2 PSD_R(k,l), \tag{2}$$

where $|H(k,l)|$ is the MTF of the downsampling process.

The simplest measure of fidelity of the processed image $\mathbf{I}_D$ in respect to the reference one $\mathbf{I}_R$ can be defined as a ratio between the total power carried by the processed image P_D and the reference one P_R. This measure denoted in Fig. 1 as Q_F can be calculated as

$$Q_F = \frac{P_D}{P_R} = \frac{\sum_k \sum_l PSD_D(k,l)}{\sum_k \sum_l PSD_R(k,l)} = \frac{\sum_k \sum_l |H(k,l)|^2 PSD_R(k,l)}{\sum_k \sum_l PSD_R(k,l)}, \tag{3}$$

where PSD_D and PSD_R are power spectral densities of the downsampled and reference image, and $|H(k,l)|$ is modulation transfer function associated with the downsampling process.

The shape of the MTF of the particular downsampling process, with defined interpolation and antialiasing kernels, depends mostly on the downsampling factor D_n. Then the fidelity measure Q_F can be evaluated from the Eq. (3) as

$$Q_F(D_n) = \frac{\sum_k \sum_l |H_{(D_n)}(k,l)|^2 PSD_R(k,l)}{\sum_k \sum_l PSD_R(k,l)}, \tag{4}$$

where $|H_{(D_n)}(k,l)|$ is the downsampling MTF associated with the downsampling factor D_n. Since the downsampling MTF has a shape of the low-pass filter, and $|H_{(D_n)}(k,l)| \leq 1$, the fidelity measure $Q_F(D_n)$ value is between 0 and 1. The value of 1 is only reachable in the case $|H_{(D_n)}(k,l)| = 1,\ \forall(k,l)$, where $PSD_R(k,l) \neq 0$.

The performance of the selected techniques from each discussed class, i.e. full-reference image quality/fidelity assessment Q_{FR}, no-reference image quality assessment Q_{NR}, and fidelity prediction Q_F, is evaluated for the particular use case scenario in the following section.

4. EXPERIMENTAL RESULTS

The three approaches for the use case scenario of scanned film image resolution loss analysis due to spatial downsampling are described in Section 3, see Fig. 1. In this section, the performance of the selected techniques from the three described classes is analyzed. The performance evaluation of the discusses techniques is based on simulation framework for the selected use case scenario. Description of the simulation framework, selected test image contents, and performance evaluation results are presented in the following paragraphs.

4.1 Simulation framework

The performance evaluation of the selected techniques was done using real test images and simulated use case scenario with spatial downsampling, see Fig. 1. As this paper is focused on preliminary evaluation, only the results of objective image fidelity/quality assessment are presented here. Performance evaluation based on the results from the subjective experiment with human observers is beyond the scope of the paper, and it is one our future work items.

The particular steps associated with the selected use case scenario, see Fig. 1, i.e. the simulation of the applied processing and performance evaluation of the tested techniques, can be summarized in the following basic steps.

- (i) Suitable test image contents selection, with characteristics relevant to the use case scenario, e.g. image contents originating from film media, required resolution, and scene variability. Here a set of twenty-four test images covering typical natural scenes was selected. More details on image contents selection can be found in Section 4.2.
- (ii) Simulation of the use case scenario, with selected image postprocessing technique, i.e. spatial downsampling. Spatial downsampling using popular Lanczos-3 interpolation kernel[28] was applied to spatially upsampled reference images (with the factor of two) for subsequent downsampling factors $D_n = \{2, 3, 4, 6, 8, 12, 16\}$. Using this approach a set of distorted images was created, resulting in 168 images. More details on distorted image contents dataset creation can be found in Section 4.2.
- (iii) Image fidelity/quality measures are calculated for the dataset created in the previous steps. Measure values Q_{FR} for full-reference (i.e. SSIM fidelity measure), Q_{NR} for no-reference (i.e. S3 and LPC-SI sharpness/blur measures), and Q_F (i.e. power spectral density analysis) for fidelity prediction were calculated. More details on selected objective measures can be found in Section 3.2.
- (iv) In the last step, the performance of the tested techniques was evaluated. The evaluation was based on the calculation of correlation coefficients between the full-reference image fidelity measure Q_{FR} and the results from the two other approaches, i.e. no-reference Q_{NR}, and especially the power spectral density based image fidelity prediction analysis Q_F. More details on performance evaluation results analysis can be found in Section 4.3.

The idea behind this procedure, especially the step (iv), lies in the fact that full-reference image fidelity criterion Q_{FR} can be considered here as the ultimate measure of distorted image $\mathbf{I}_D$ fidelity in respect to the reference image $\mathbf{I}_R$. It is not very practical, due to computational demands, to calculate Q_{FR} for the whole video sequence, since also the processed video sequence frames $\mathbf{I}_D$ has to be calculated for each configuration of the tested processing, i.e. spatial downsampling with downsampling factor D_n.

It is more convenient, especially for the sake of computational simplicity, to predict image fidelity Q_F of the distorted image $\mathbf{I}_D$ based on the features of the reference one $\mathbf{I}_R$. The technique considers also known properties and parameters of the processing, i.e. downsampling factor D_n, see Section 3.2 for more details. The performance of the fidelity prediction Q_F has to be evaluated concerning the full-reference measure output Q_{FR}. On the other hand, there are cases, where the reference image $\mathbf{I}_R$ is not available, and still, we would like to assess the quality loss of the processed image $\mathbf{I}_D$, sometimes in respect to higher quality processed version. Here comes the importance of the no-reference measure Q_{NR} and evaluation of its performance in respect to the full-reference fidelity measure Q_{FR}.

4.2 Test image contents

For the preliminary experimental verification of the techniques proposed in Section 3, twenty-four test images were chosen from the publicly available Kodak PhotoCD dataset†. The images were downloaded in BMP file format‡ with the original resolution of $3\,072 \times 2\,048$ pixels. This resolution, considering 24×36 mm full-frame, is equivalent to the scanner resolution of 2 167.5 ppi.

These test scenes, see Fig. 2, were selected because of the image contents variability, covering typical natural scenes. Images available in Kodak PhotoCD dataset are suitable for the test also because they were originally captured on 35 mm film using an analog still camera and then scanned in high resolution.

(a) [01] Stone building; (b) [02] Red door; (c) [03] Hats; (d) [05] Motocross bikes; (e) [06] Sailboat at anchor; (f) [07] Shuttered windows; (g) [08] Market place; (h) [11] Sailboat at pier; (i) [12] Couple on beach; (j) [13] Mountain stream; (k) [14] Water rafters; (l) [15] Painted face; (m) [16] Tropical key; (n) [20] P51 Mustang; (o) [21] Portland light; (p) [22] Barn and pond; (q) [23] Two macaws; (r) [24] Mountain chalet; (s) [04] Girl in red; (t) [09] Sailboats; (u) [10] Sailboat race; (v) [17] Monument; (w) [18] Model in black; (x) [19] Lighthouse

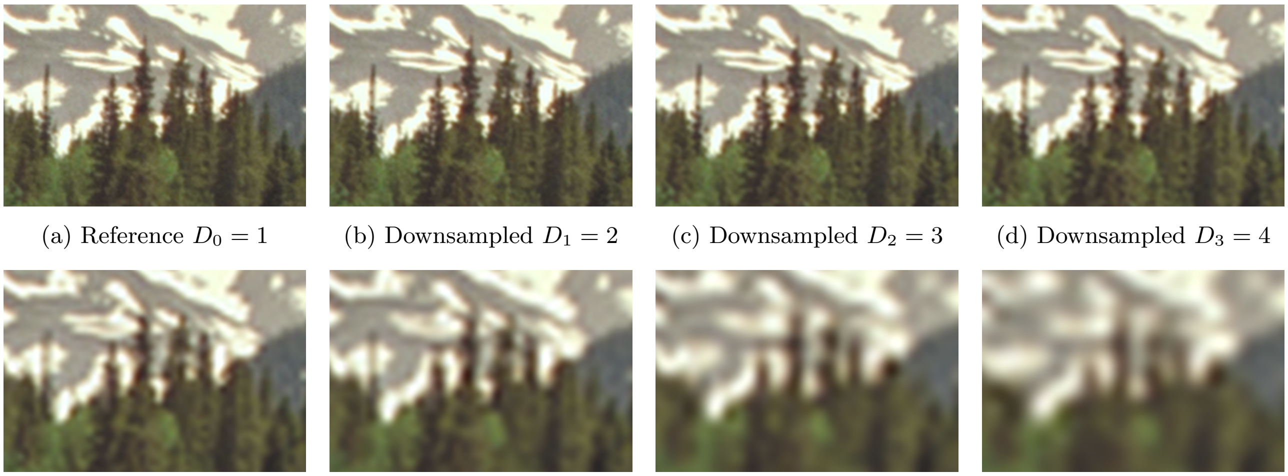
Figure 2: Thumbnails of test images from Kodak PhotoCD. Note the modified order of the test images in the figure.

The original test images with the resolution of $3\,072 \times 2\,048$ pixels (6.29 Mpix) where upsampled with the factor of two obtaining a set of twenty-four reference images with the resolution of $6\,144 \times 4\,096$ pixels (25.2 Mpix). The upsampling was performed using popular Lanczos-3 interpolation kernel.[28] From each reference image, a set of seven downsampled images was created again with Lanczos-3 interpolation kernel and anti-aliasing filter for downsampling factors $D_n = \{2, 3, 4, 6, 8, 12, 16\}$. These factors were selected to cover a broad range of downsampled derivatives of the reference image with common 2^n based and also other downsampling factors. The resolutions of the downsampled images are $3\,072 \times 2\,048$ pixels (6.29 Mpix), i.e. equal to the resolution of the

†Kodak PhotoCD dataset (Accessed: July 23, 2017): `http://r0k.us/graphics/kodak/`

‡B. J. Lucier's Kodak PhotoCD dataset (Accessed: July 23, 2017): `http://www.math.purdue.edu/~lucier/PHOTO_CD/`

original image, $2\,048 \times 1\,366$ (2.80 Mpix), $1\,536 \times 1\,024$ (1.57 Mpix), $1\,024 \times 683$ (0.70 Mpix), 768×512 (0.39 Mpix), 512×342 (0.18 Mpix), and 384×256 (0.10 Mpix). These resolutions, considering 24×36 mm full-frame, would be equivalent to the scanner resolution between 4 334.9 ppi and 270.9 ppi. To allow calculation of basic full-reference pixel metrics the set of downsampled images was in the end upsampled back to the resolution of the reference image, i.e. $6\,144 \times 4\,096$ pixels. The upsampling was performed again using Lanczos-3 interpolation kernel.



(a) Reference $D_0 = 1$ (b) Downsampled $D_1 = 2$ (c) Downsampled $D_2 = 3$ (d) Downsampled $D_3 = 4$

(e) Downsampled $D_4 = 6$ (f) Downsampled $D_5 = 8$ (g) Downsampled $D_6 = 12$ (h) Downsampled $D_7 = 16$

Figure 3: Closeups (315×210 pixels) of "[13] Mountain stream" (see Fig. 2(j)) test image for the eight available versions obtained after downsampling $D_n = \{2, 3, 4, 6, 8, 12, 16\}$ including the reference version.

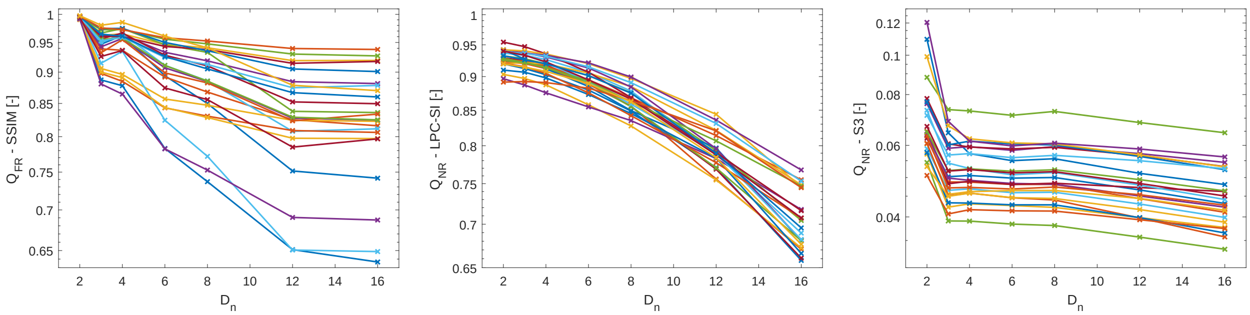
4.3 Results

Simulation and performance evaluation results for the techniques described in Section 3 are presented and discussed in the following paragraphs. At first, the measures Q_{FR}, Q_{NR}, and $Q_F(D_n)$ were calculated for the whole set of twenty-four reference images $\mathbf{I}_R$, downsampled with the factor $D_n = \{2, 3, 4, 6, 8, 12, 16\}$ while obtaining a set of downsampled images $\mathbf{I}_D$.

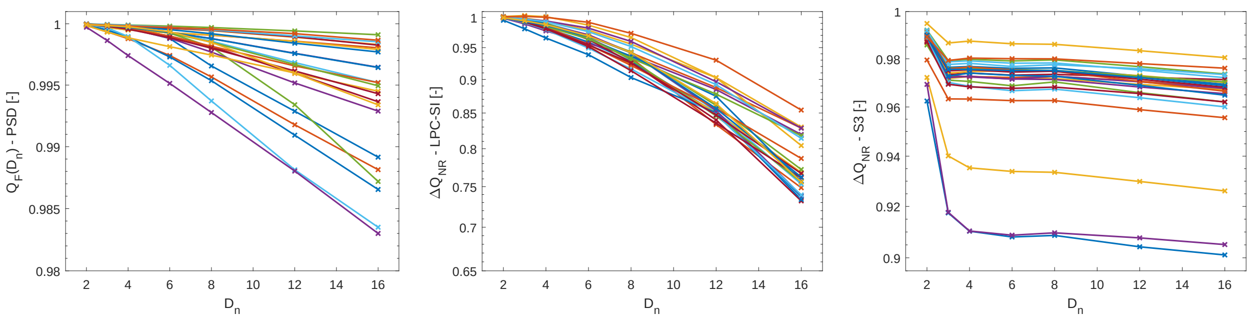
For the case of full-reference measure Q_{FR}, here SSIM, the downsampled versions were upsampled using the same algorithm, see Section 4.2, to the original resolution of the reference image for pixel-wise calculation of the respective measure. The MTF of the downsampling process $|H_{(D_n)}(k, l)|$, see Eq.(4), was for the sake of simplicity approximated with ideal low-pass filter mask in the spatial frequency domain with the cut-off frequency equal the Nyquist spatial frequency for the downsampled image.[10] Resulting dependencies of the calculated objective measures for the whole dataset and range of the downsampling factor can be seen in Fig. 4. In the case of no-reference measures Q_{NR} also the difference of the measure for the reference image and the distorted image was calculated as $\Delta Q_{NR} = 1 - Q_{NR}(\mathbf{I}_R) + Q_{NR}(\mathbf{I}_D)$, actually turning the no-reference measure to a reduced-reference case.

The ability of the tested no-reference measures Q_{NR}, and especially fidelity prediction measure Q_F predicting the image fidelity provided by the full-reference measure Q_{FR}, can be evaluated based on the graphs depicted in Fig. 5. It is clear that in the tested use case scenario the selected no-reference techniques, i.e. Q_{NR} based on LPC-SI (Fig. 5(b)) and S3 (Fig. 5(c)), are not suitable for fidelity prediction. These measures are highly content dependent. This dependency can be alleviated if the reduced reference variant of the measures ΔQ_{NR} is used, see Fig. 5(e, f). It is evident that the proposed simple fidelity prediction Q_F based on PSD analysis, see Section 3.2, is the most efficient in this use case scenario and can predict well the processed image fidelity as assessed by the full-reference technique Q_{FR}. The fidelity prediction ability of the proposed technique, see Eq. (4), can be seen in Fig. 5(a), and Fig. 5(b), where the five-parameter logistic regression is used.

There are various metrics used to evaluate the performance of an objective measure. Prediction accuracy can be evaluated using the Pearson linear correlation coefficient (PLCC) denoted as r, whereas prediction monotonicity through the Spearman rank order correlation coefficient (SROCC), referred to as Spearman's ρ.[19–21]



(a) Q_{FR} vs. D_n for SSIM. (b) Q_{NR} vs. D_n for LPC-SI. (c) Q_{NR} vs. D_n for S3.



(d) $Q_F(D_n)$ vs. D_n for PSD measure. (e) ΔQ_{NR} vs. D_n for LPC-SI. (f) ΔQ_{NR} vs. D_n for S3.

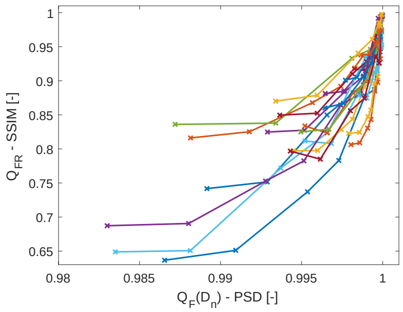
Figure 4: Selected objective measures, see Section 3.2, evaluated for twenty-four test images, see Fig. 2, from the selected dataset downsampled with the downsampling factor $D_n = \{2, 3, 4, 6, 8, 12, 16\}$.

The prediction accuracy of both no-reference techniques Q_{NR} based on LPC-SI and S3 measures is very limited, reaching $r = 0.587$ and $r = 0.260$. The prediction monotonicity of S3 based no-reference measure is very low $\rho = 0.264$, while for LPC-SI based measure is slightly better $\rho = 0.587$. The results of both measures improve for their reduced-reference modification ΔQ_{NR}, reaching $r = 0.651$, $\rho = 0.770$ for LPC-SI, and $r = 0.653$, $\rho = 0.673$ for S3. While the simple fidelity prediction $Q_F(D_n)$ achieves better prediction accuracy $r = 0.835$ and much better prediction monotonicity $\rho = 0.920$.

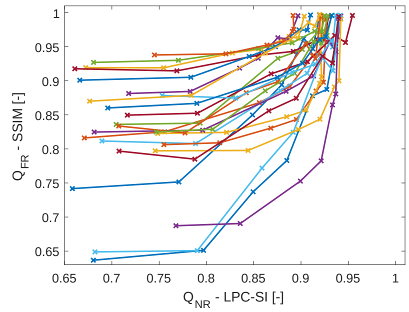
5. CONCLUSIONS AND FUTURE WORK

In this paper an overview of achievable image resolution in archival records is presented, associated resolution requirements for the film scanner are discussed, followed by brief analysis on the relation between the resolution and film microstructure properties. The primary goal of the paper is not to give conclusive answers on what is the optimal resolution of the film scanner for high-quality film digitization but propose and evaluate the performance of efficient techniques for fidelity prediction in use case scenario of deriving spatially downsampled versions of original high-resolution recording.

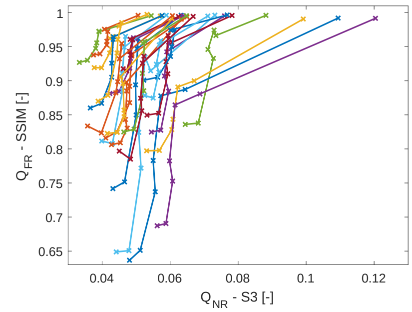
It was shown that simple fidelity prediction technique could achieve high prediction accuracy and monotonicity while compared to state-of-the-art full-reference image quality/fidelity assessment methods, reaching Pearson linear correlation coefficient $r = 0.835$ and Spearman rank order correlation coefficient $\rho = 0.920$, while compared with SSIM. The proposed approach is based on the power spectral density analysis of the reference image while considering characteristics and basic parameters (downsampling factor) of the used downsampling process. On the contrary, the tested no-reference techniques did not prove to be suitable for the tested use case scenario. Their computational complexity is much higher than for the proposed technique and achieved prediction accuracy, and monotonicity is inferior. It has to be noted that the performance evaluation was based on the comparison of the outcome of the tested measures with the full-reference image fidelity/quality measure. Following the preliminary results presented in this paper, a proper subjective image quality assessment experiment has to be performed to



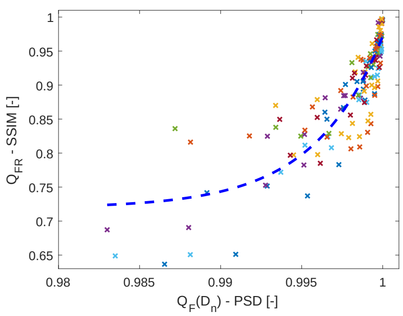
(a) $Q_F(D_n)$ - PSD vs. SSIM. $r = 0.835$, $\rho = 0.920$



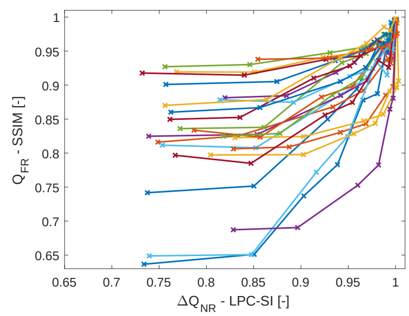
(b) Q_{NR} - LPC-SI vs. SSIM. $r = 0.587$, $\rho = 0.642$



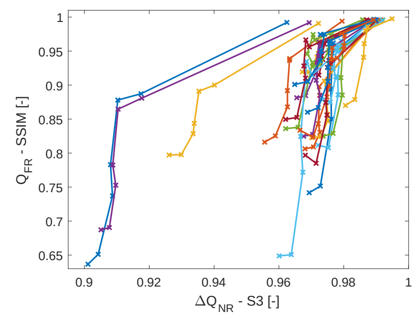
(c) Q_{NR} - S3 vs. SSIM. $r = 0.260$, $\rho = 0.264$



(d) $Q_F(D_n)$ - PSD vs. SSIM (fit). $r = 0.835$, $\rho = 0.920$



(e) ΔQ_{NR} - LPC-SI vs. SSIM. $r = 0.651$, $\rho = 0.770$



(f) ΔQ_{NR} - S3 vs. SSIM. $r = 0.653$, $\rho = 0.673$

Figure 5: Scatter plots between the fidelity measure Q_{FR} (SSIM), and other tested no-reference objective measures, see Section 3.2, evaluated for twenty-four test images from the selected dataset downsampled with the downsampling factor $D_n = \{2, 3, 4, 6, 8, 12, 16\}$, see Fig. 2. Relationship between the fidelity full-reference measure Q_{FR} (SSIM) and fidelity prediction $Q_F(D_n)$, see Fig. 5(d), is fitted with five parameter logistic regression.

draw conclusive answers. Also, the performance evaluation should be done with more realistic test image data obtained from a high-end film scanner. The items mentioned above are planned to be treated in our future work.

ACKNOWLEDGMENTS

This work was supported by the project NAKI DF13P01OVV006 "Methodics of digitizing of the national film fund" of the Ministry of Culture of the Czech Republic, which was conducted at the Film and TV School (FAMU) of Academy of Performing Arts in Prague[§]. The authors would like to thank Iveta Kostelníčková for her help with preliminary analysis of tested techniques. Our thanks belong also to Miloslav Novák for his valuable consultations.

REFERENCES

[1] Fliegel, K., Krasula, L., Páta, P., Myslík, J., Pecák, J., and Jícha, M., "System for objective assessment of image differences in digital cinema," *Proc. SPIE* **9217**, 92170I–92170I–14 (2014).

[2] Fliegel, K., Vítek, S., Páta, P., Janout, P., Myslík, J., Pecák, J., and Jícha, M., "Evaluation of color grading impact in restoration process of archive films," *Proc. SPIE* **9971**, 997121–997121–18 (2016).

[3] Fliegel, K., Vítek, S., Páta, P., Novák, M., Myslík, J., Pecák, J., and Jícha, M., "Set of methodologies for archive film digitization and restoration with examples of their application in ORWO region," *Archiving Conference* **2017**, 62–67 (2017).

[§]University website (Accessed: July 23, 2017): `https://www.famu.cz/eng`

[4] Jícha, M. and Šofr, J., "Digitální restaurování památek filmového umění. Metoda DRA," *Zprávy památkové péče* **76**(1), 76–90 (2016). (in Czech).

[5] Baroncini, V., Mahler, H., and Sintas, M., "The image resoulution of 35mm cinema film in theatrical presentation," *SMPTE Motion Imaging Journal* **113**(2-3), 60–66 (2004).

[6] Boreman, G., [*Modulation Transfer Function in Optical and Electro-optical Systems*], SPIE tutorial texts, SPIE Press (2001).

[7] Morton, R., Cosgrove, A., and Masson, A., "Letter to the editor: The image resolution of 35mm cinema film in theatrical presentation," *SMPTE Motion Imaging Journal* **113**(4), 102–102 (2004).

[8] Vitale, T., "Estimating the resolution of historic film images: Using the resolving power equation (RPE) and estimates of lens quality." `http://vitaleartconservation.com/PDFgallery.htm` (2009). (Accessed: July 23, 2017).

[9] FIAF Technical Commission, "Choosing a film scanner." `http://www.fiafnet.org/pages/E-Resources/Technical-Commission-Resources.html` (2016). (Accessed: July 23, 2017).

[10] Gonzalez, R. C. and Woods, R. E., [*Digital Image Processing (3rd Edition)*], Prentice-Hall, Inc., Upper Saddle River, NJ, USA (2006).

[11] Vitale, T., "Film grain, resolution and fundamental film particles." `http://vitaleartconservation.com/PDFgallery.htm` (2009). (Accessed: July 23, 2017).

[12] "DCI Specification, Version 1.2 with Errata." Digital Cinema Initiatives (2012).

[13] Holst, G. C., [*CCD Arrays, Cameras and Displays*], SPIE, Bellingham, 2nd ed. ed. (1998).

[14] Hepper, D., "Investigating properties of film grain noise for film grain management," *Proceedings 2013 IEEE 3rd International Conference on Consumer Electronics - Berlin, ICCE-Berlin 2013* , 185–188 (2013).

[15] Hepper, D., "Film grain noise superimposition for film grain management," *Digest of Technical Papers - IEEE International Conference on Consumer Electronics* , 252–255 (2014).

[16] Oh, B., Lei, S.-M., and Kuo, C.-C., "Advanced film grain noise extraction and synthesis for high-definition video coding," *IEEE Transactions on Circuits and Systems for Video Technology* **19**(12), 1717–1729 (2009).

[17] Le Callet, P., Möller, S., Perkis, A., et al., "Qualinet white paper on definitions of quality of experience." `http://www.qualinet.eu/images/stories/QoE_whitepaper_v1.2.pdf` (2012). (Accessed: July 26, 2017).

[18] "ISO/IEC 15444-1:2004/Amd 2:2009 Extended profiles for cinema and video production and archival applications." International Organization for Standardization (2009).

[19] Wu, H. and Rao, K., [*Digital Video Image Quality and Perceptual Coding*], Signal Processing and Communications, Taylor & Francis (2005).

[20] Winkler, S., [*Digital Video Quality: Vision Models and Metrics*], Wiley (2013).

[21] Wang, Z. and Bovik, A., [*Modern Image Quality Assessment*], Synthesis lectures on image, Morgan & Claypool Publishers (2006).

[22] Sheikh, H. R. and Bovik, A. C., "Image information and visual quality," *IEEE Transactions on Image Processing* **15**(2), 430–444 (2006).

[23] Zhang, L., Zhang, L., Mou, X., and Zhang, D., "FSIM: A Feature Similarity Index for Image Quality Assessment," *IEEE Transactions on Image Processing* **20**(8), 2378–2386 (2011).

[24] Narwaria, M., Mantiuk, R., Da Silva, M., and Le Callet, P., "Hdr-vdp-2.2: A calibrated method for objective quality prediction of high-dynamic range and standard images," *Journal of Electronic Imaging* **24**(1) (2015).

[25] Krasula, L., Le Callet, P., Fliegel, K., and Klíma, M., "Quality assessment of sharpened images: Challenges, methodology, and objective metrics," *IEEE Transactions on Image Processing* **26**(3), 1496–1508 (2017).

[26] Vu, C. T., Phan, T. D., and Chandler, D. M., "S3: A spectral and spatial measure of local perceived sharpness in natural images," *IEEE Transactions on Image Processing* **21**(3), 934–945 (2012).

[27] Hassen, R., Wang, Z., and Salama, M. M. A., "Image sharpness assessment based on local phase coherence," *IEEE Transactions on Image Processing* **22**(7), 2798–2810 (2013).

[28] Burger, W. and Burge, M., [*Principles of Digital Image Processing: Core Algorithms*], Undergraduate Topics in Computer Science, Springer London (2010).