

RESEARCH

Open Access

A new quality assessment and improvement system for print media

Mohan Liu^{1*}, Luliu Konya², Jan Nandzik³, Nicolas Flores-Herr³, Stefan Eickeler² and Patrick Ndjiki-Nya¹

Abstract

Print media collections of considerable size are held by cultural heritage organizations and will soon be subject to digitization activities. However, technical content quality management in digitization workflows strongly relies on human monitoring. This heavy human intervention is cost intensive and time consuming, which makes automatization mandatory. In this article, a new automatic quality assessment and improvement system is proposed. The digitized source image and color reference target are extracted from the raw digitized images by an automatic segmentation process. The target is evaluated by a reference-based algorithm. No-reference quality metrics are applied to the source image. Experimental results are provided to illustrate the performance of the proposed system. We show that it features a good performance in the extraction as well as in the quality assessment step compared to the state-of-the-art. The impact of efficient and dedicated quality assessors on the optimization step is extensively documented.

1 Introduction

1.1 Significance of quality assessment in print media domain

Print media collections of considerable size are held by cultural heritage organizations (e.g., libraries and archives) and other content owners (e.g., publishers, financial institutions, hospitals or insurances) have been or will soon be subject to digitization activities. These organizations typically aim at

- digitally archiving their print media collection and/or;
- making the content available for end-users at a grand scale.

As the former is cost-intensive [1] and the latter may involve machine-based media analysis (e.g., text indexing for search or semantic clustering of texts) next to usability considerations [2], content owners face the challenge to safeguard the information contained within the print assets during transference from the analog to the digital domain. This involves measuring, interpreting, and, if required, optimizing the quality of each digital object

alongside

- (1) analog-to-digital conversion,
- (2) media processing workflow and
- (3) final image formats,

especially if the latter were produced using lossy compression.

In this article, technical content quality is understood as the amount of information contained within analog and digital media respectively. To optimally preserve information of the analog media

- Color fidelity
- Spatial resolution
- Contrast/brightness
- Image geometry
- Sharpness
- Noise

and many other parameters [3] have to be faithfully conveyed by the analog-to-digital converter (e.g., camera or scanner) [4-6] and further processing steps in the digital domain. In addition, the detection of unwanted objects like dust, fingers (belonging to the digitization operator) etc. is of similar importance. Recent research approaches in this field are almost exclusively carried

* Correspondence: mohan.liu@hhi.fraunhofer.de

¹Image Processing Department, Fraunhofer Institute for Telecommunications - Heinrich Hertz Institute, Berlin, Germany

Full list of author information is available at the end of the article

out by the industrial sector [2-9] or the cultural sector [1,10-13] while there are only few publications from the field of basic research dating back to the 1990s [14-16] including a review of results related to image quality research [17].

Current research carried out in the industry sector focuses on methodologies to measure the analog-to-digital conversion quality of a digitization device using various test targets [2-9]. Among other goals industry research aims at identifying evaluation methodologies and quantifiable parameters (like modulation transfer function, noise, sharpness etc.) that determine the quality of the digitization device. Hence, by using a defined set of parameters that are critical for a faithful analog-to-digital-conversion a quantitative method to standardize the measurement of the quality of digitization devices and to certify digitization service providers may be established.

Present research from the cultural heritage sector mainly addresses technologies and most of all best practices to improve the quality of the digitization process [10,13,18] and subsequent optical character recognition (OCR) of archived print media. In contrast to the industrial approach—where the quality of digitization devices is primarily addressed—research of the cultural heritage sector focuses on the detection of device-independent errors that typically appear in digitization projects (e.g., detection of missing pages, unwanted objects, irregular illumination of the book or document page).

Benefits of automatic quality assessment and quality optimization

With respect to the management of technical content quality, today's digitization workflows heavily rely on human monitoring and decision making [7]. The major disadvantages of manual quality assessment are:

- Not scalable and costly.
- Subjectively estimated technical content quality parameters are prone to errors [8].
- Quality levels cannot be objectively benchmarked or standardized [9].
- Quantitative content quality policies cannot be applied.

For example, using a state-of-the-art book scanner^{a,b,c} more than 10,000 pages per machine and day can be digitized, which is considerably low when compared to the throughput of document scanners. Given an operator may be able to check the quality of 1,000 digitized pages per day, the workflow would require 10 operators for 100% coverage. Thus, considering the throughput of a set of simultaneously operating machines rather than just one scanner, manual quality assessment makes

efficient and cost-effective mass digitization almost impossible.

Vice versa, when using algorithms for quality assessment highly automated digitization workflows can be established and quantitative content quality thresholds or, more general, content quality policies can be introduced. This is of great importance as the quality requirements of an institution may vary with respect to the purpose a collection and even to individual print media items, for example: reproduction (like art prints), automated media analysis like optical character recognition (e.g., books) or viewing on a computer monitor (like invoices).

As of today, quality requirements for various usage scenarios are implemented by rule of thumb rather than quantitative criteria. Not matching the quality requirements may lead either to high costs (for processing and storage if the digital object is larger than necessary) or poor results (for automated media analysis like OCR, long-term preservation or reproduction). Thus, automated tools will lower the costs by greatly improving the degree of automation as well as lowering the efforts to implement the content owner's quality requirements.

In addition, as mass digitization is carried out by a growing number of organizations and by various service providers, general content quality policies cannot be applied. This in turn hampers for example a quality-based cross-institutional content aggregation, a problem that is very common to the cultural sector [10], and prevents the establishment of quantitative quality guidelines for specific sectors (like medical files in the health care sector).

To sum up, tools for automated quality analysis are a prerequisite for the management of technical content quality in mass digitization environments—not only in the cultural heritage domain, but across various types of organizations that still depend on printed analog media like for example the financial sector, jurisprudence, and health care sector.

Existing quality management systems for print media

Existing systems for automatic quality assessment are mostly based on *reference based* measurement methods. Thus, they can only be used in combination with their respective targets and are not able to measure quality independently in a *no-reference* scenario. Most of the existing software for automatic quality assessment of digitized print media rely heavily on expertise of the operator and are not suitable for mass processing.

The Software iQ-ANALYZER 5^d by German company *Image Engineering* allows for the automatic assessment of standardized test targets (*Image Engineering* have contributed to the development of the *Universal Test Target* for this specific purpose). The output of the iQ-

ANALYZER provides a detailed analysis of many aspects of image quality, yet offers only minimal user assistance for batch processing. The software is specifically tailored for characterization of a machine's capabilities, not for continuous performance monitoring.

The US-based company *certifi media* offers an integrated software *Certifi Pedigree QP[®]* that also uses specialized test targets for image analysis. Pedigree QP offers the creation of quality profiles for batch processing, but is not able to operate no-reference analysis of files that do not include scanned reference targets.

2 Overall system

The algorithms described in this article could be used as building blocks for an integrated quality assessment and processing system for still images that originate from a digitization of printed media.

Some of the automated quality assessment (QA) algorithms for still images as proposed by the authors rely on the presence of reference targets within every digitized image file. This is achieved by digitizing the source media and reference target by the same digitization machine in one single pass so that both are contained in one single image file. The color reference target and source media could be located anywhere within the raw digitized image—with the prerequisite that the reference target is contained fully within the raw image files and is not overlapping the source media.

A system encompassing a compilation of the algorithms presented in this article would then be able to handle the following processing and analysis steps:

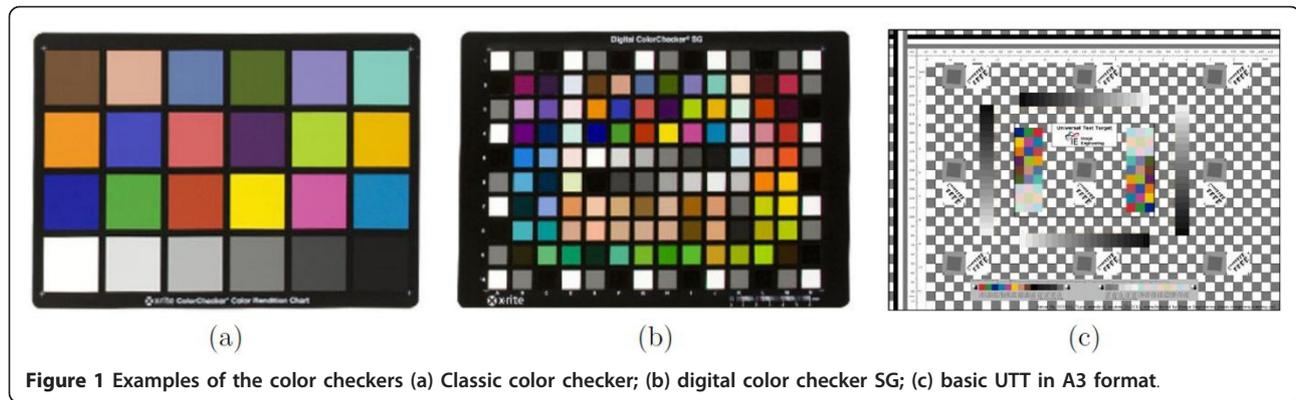
- *1. Image segmentation* The spatial analysis of the raw digitized images, leading to detection of the digitized media (essence), the color reference target and the scanner background. The bounding boxes obtained by segmentation can then be used for an independent analysis of these elements (cf. Section 3).
- *2. Reference-based assessment of quality* Measurement of *color rendering accuracy* by analysis of the extracted reference targets (cf. Section 4.1).
- *3. No-reference assessment of quality* Measurement of *in-picture contrast, brightness, sharpness, compression block artifacts* and *overall quality* within the extracted source image (cf. Section 4.2).
- *4. Quality optimization* Application of quality analysis to efficient image optimization (cf. Section 4.3).
- *5. Presentation of layer derivatives* The separation of targets and source material can be used for automatic image cropping to create presentation ready derivatives from the raw scans containing both media and targets.

3 Color reference target detection

The employment of reference targets in the process of professional digitization or photography is ubiquitous today. Practically all current digitization guidelines aimed at the preservation of cultural heritage material highly recommend the inclusion of reference targets in each of the originals being scanned. Many agencies go even further, advocating the use of several targets in order to allow for a more accurate quantization of the variables involved in the digitization process. A prominent example are the numerous federal agencies from the US adhering to the Federal Agencies Digitization Guidelines Initiative (FADGI) [11]. The FADGI suggests as a minimal requirement the use of photographic gray scale as a tone and color reference, as well as the utilization of an accurate dimensional scale. Color reference targets, also known as color checkers are therefore of central importance in any mass digitization process.

Depending on their type, reference targets allow a precise measurement of many different parameters influencing the digitization. Examples of such parameters are: the scale, rotation and any distortions present in the digitized asset, as well as the color and illumination deviation/uniformity. In the following we briefly present a few of the most popular color reference targets in use today:

- The classic color checker [19], initially commercialized starting from 1976 as the “Macbeth” color checker [20]. It contains 24 uniformly-sized and -colored patches printed on a 8.5” × 11.5” cardboard. The colors are chosen so that they represent many natural, frequently occurring colors such as human skin, foliage, and blue sky. Nowadays it is still the most common tool employed for color comparison due to its small size and ease of use (cf. Figure 1a).
- The digital color checker SG [21] contains an extended color palette in the form of 140 quadratic patches. It is tailored to offer a greater accuracy and consistency over a wide variety of skin tones, as well as the provision of more gray scale steps ensuring a finer control of the camera balance and the ability to maintain a neutral aspect regardless of light source (cf. Figure 1b).
- The Universal Test Target (UTT) [12] is one of the most recent open-source efforts for the development of a single reference target covering a large array of scanning parameters. The development of the UTT is an ongoing process directed by the National Library of the Netherlands as part of *Metamorfoze* [22], the Dutch national program for the preservation of article heritage. It is available with various options in the DIN sizes A3 to A0 and has as main



purpose a general applicability in all kinds of digitization projects, preservation and access, carried out by libraries, archives and museums (cf. Figure 1c).

By using the information extracted with the help of the reference targets, a human operator is able to correct any inaccuracies of the scanning procedure on-the-fly. In case of a fully automated digitization process, a computer algorithm has the possibility of performing accurate corrections on the digitized assets for a better reproduction of the original item. Very little research has been done in the area of automatic color reference target detection. To the best of the authors' knowledge there currently exist no fully automatic solutions to this problem. A step in this direction was recently done by Tajbakhsh and Grigat [23], who introduced a semi-automatic method for color target detection and color extraction. Their method focuses on images exhibiting a significant degree of distortion (e.g., as caused by perspective, mechanical processes, camera lens). In the process of mass document digitization however, such pronounced distortions are extremely seldom mainly due to the cooperation with professional scan service providers. A commercial software which is also capable of a semi-automatic color target detection is the X-Rite color checker Passport Camera Calibration Software [24]. The X-Rite software ultimately relies on the human operator to manually mark/correct the detected reference target in order to be able to perform any subsequent color correction. Human intervention is of course not practical in any mass digitization process, as it would cause far too large disruptions in the process.

In the following section, we present a fully automatic and robust algorithm for the detection of classic color checker targets in digital document images. Our main focus is its applicability in mass document processing, where robustness and flexibility are of paramount importance. The proposed algorithm can readily be extended to other types of color reference targets, including the digital color checker SG as well as the

UTT (cf. Figure 1). An evaluation on a set of 239 real-life document and photograph scans is done to investigate the robustness of our algorithm.

Proposed algorithm

As mentioned in the previous section, our algorithm targets professional scans, as normally found in any mass digitization project. As such, in order to ensure a fully automatic and robust operation, we make a few assumptions. The first assumption is that the scans exhibit no or low perspective distortions. In this respect, the method of Tajbakhsh and Grigat [23] complements well the proposed algorithm in case of larger distortions. The second assumption is that the scanning resolution is known (exactly or approximately) for each scanned image, which is virtually always true in case of professional scans. A last but very important requirement is that the lighting is approximately constant on the whole image. Note that the last restriction is not specific to our system, but it applies to all methods employing color targets. In case of uneven lighting conditions (e.g., shadows, multiple light sources possibly having different spectral distributions), it is generally not possible to obtain a meaningful automatic color difference measurement/adjustment without a priori knowledge of the lighting conditions. One may easily see this by considering the following exemplary basic situations: uneven lighting solely on the color target or uneven lighting restricted to the scanned object. In the former case, an automated color evaluation would match the (possibly correct) colors from the scanned object to the (partially) wrong ones from the color target, thus reporting large color differences and performing wrong color corrections. The latter case would result in no correction being applied to the object (possibly exhibiting large color shifts), due to the perfect lighting in the region of the color target.

Our algorithm consists of the four main steps (cf. Figure 2) presented in detail below, followed by the automatic color quality assessment described in Section 4.

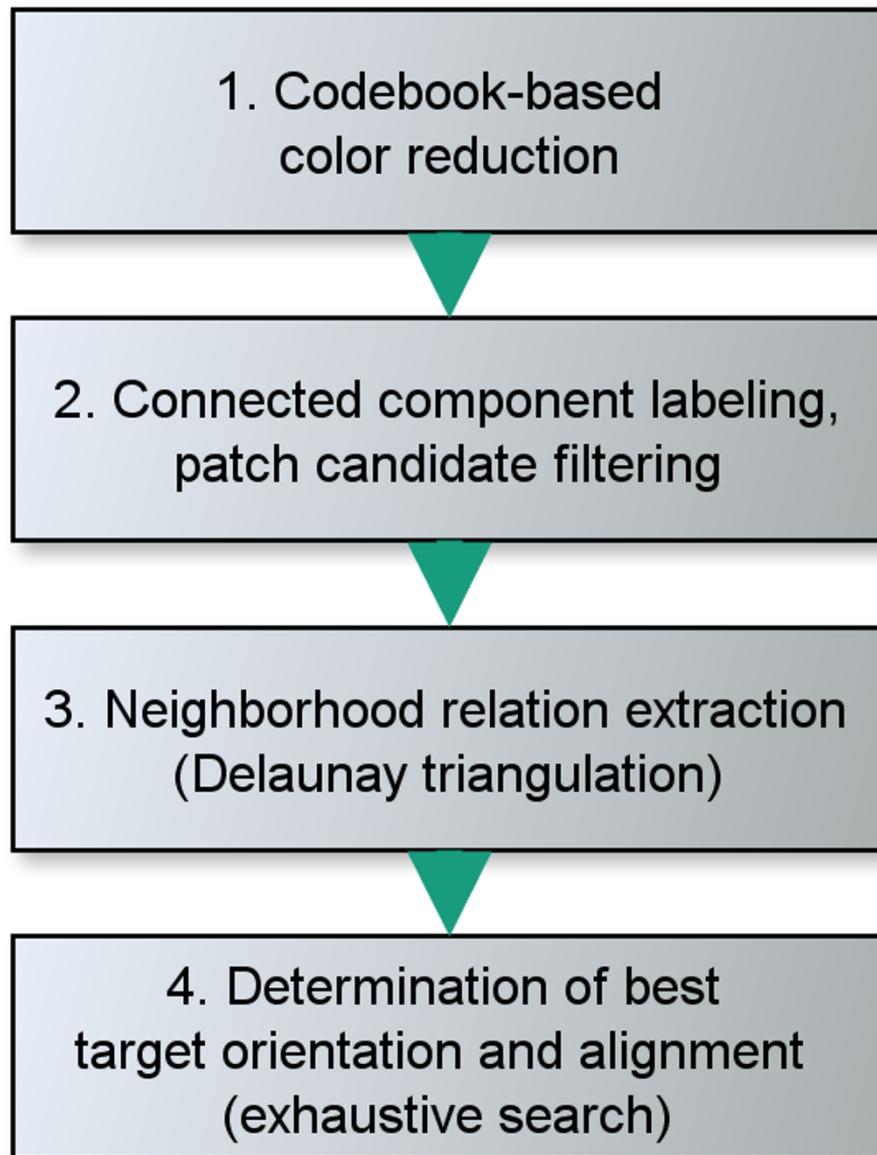


Figure 2 Block diagram of proposed color target detection algorithm.

The first step is the application of a codebook-based color reduction (cf. Figure 3b). More specifically, the color C_p of each pixel p is replaced with the value of the nearest codebook color:

$$C_p \leftarrow \text{Codebook}_{\arg \max_i D(C_p, \text{Codebook}_i)}, \quad (1)$$

where

$$D(C_1, C_2) = \sqrt{(r_1 - r_2)^2 + (g_1 - g_2)^2 + (b_1 - b_2)^2} \quad (\text{cf.}$$

Figure 3e,f). The simple Euclidean distance in the *sRGB* color space has performed well enough in our tests, however, in order to obtain more perceptually accurate

color reduction results, one may use any more advanced color measure, such as CIEDE2000 [25]. The codebook consists of the set of colors existing on the color target. Note that all color components for each patch are precisely known, being specified by the reference target manufacturer as both *sRGB* and CIE $L^* a^* b^*$ triplets [26]. In case of the classic color checker this step results in a color-reduced image having exactly 24 colors.

In the next step, a connected component analysis [27] is performed on the color-reduced image (cf. Figure 3c). In practice, connected component analysis is extremely fast even for large, high-quality scans because the complexity of the algorithm is constant in the number of

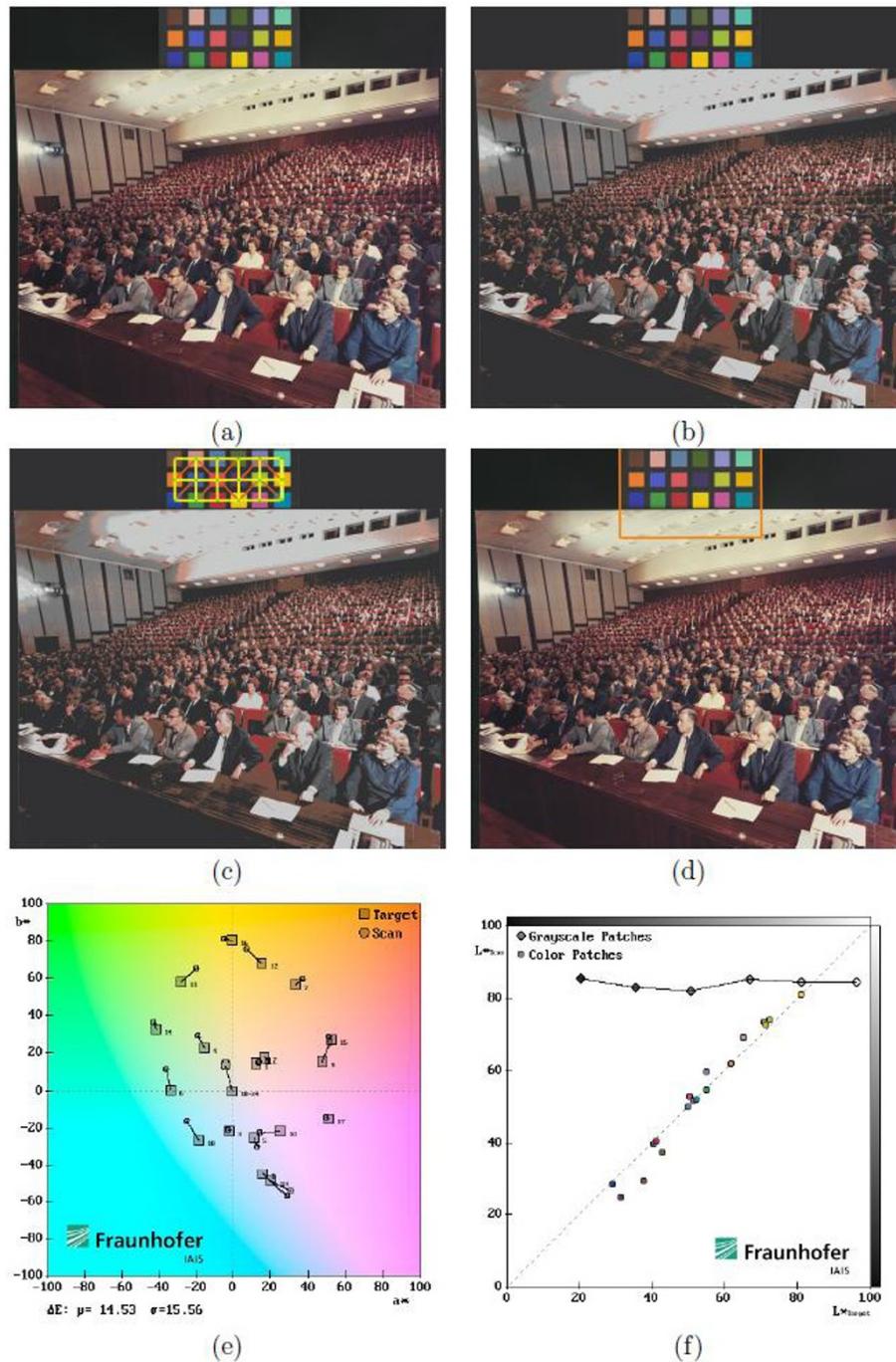


Figure 3 Illustration and results of the color checker detection (a) Original scan with a completely occluded last color checker line (i. e., no grayscale patches visible); (b) after color quantization; (c) with superimposed Delaunay triangulation (orange-painted edges were discarded from the adjacency list used for matching); (d) correctly detected color target; (e), (f) visualizations of the automatically extracted color quality results (chrominance, luminance).

pixels in the image. Subsequently we make use of the known scanning resolution to perform a filtering of the potential patch candidates based on their size, namely we discard all connected components having a width or height deviating more than 20% off the expected patch

size. Since the shapes as well as the average distances between the color patches on the reference target are also known in advance for each color target model, they are used next as a refinement to the initial filtering. For the classic color checker our algorithm uses the

following restrictions:

- the (roughly) square aspect of each color patch, i.e., width and height, are within 20% of each other;
- the size uniformity between the patches, i.e., area of each bounding box, deviates less than 30% from the median area;
- the average distance between direct horizontally or vertically neighboring patches, i.e., distance to closest patch candidate must lie within 15% of the median minimum distance.

All previously mentioned thresholds have been experimentally determined via an independent training image set consisting of a random sample of images with a similar provenience as the images from the test set. The thresholds allow our algorithm to successfully cope with minor perspective distortions, image blur, as well as lens distortions (e.g., chromatic- and spherical aberrations). The remaining connected components constitute the final list of patch candidates. For each candidate we now determine its dominant color as the mean color of the pixels from the original scan located within its connected component. Since the original colors within each patch are relatively close to each other (they were assigned to the same cluster center by the color reduction), such a mean can be computed safely even in the *sRGB* color space. It is important to observe that generally computing a simple mean is not possible because of resulting color interpolation artifacts, which are highly-dependent on the employed color space, such as color bleeding. Another possibility for assigning a single representative color to each patch would be the use of the median, computed either channel-by-channel or by considering each color as a vector.

A third step consists of the determination of all direct neighborhood relations between the final list of patch candidates (cf. Figure 3d). This is accomplished via a Delaunay triangulation [28] using as seed points the centers of the patch candidates. Next, the obtained triangulation is pruned of the edges diverging significantly from the horizontal or vertical, as regarded from a coordinate system given by the main axes of the color target. Discarding edges which deviate more than 20% from the median edge length represents an efficient pruning method. Finding one of the axes of the color target can at this point be readily accomplished by determining the median skew angle from the remaining edges. The other axis of the reference target is always considered to be perpendicular to the determined axis.

As a final step, we may now determine the exact orientation of the color target by employing the direct neighborhood relations extracted, as well as the dominant color for each patch candidate (cf. Figure 4). For

this purpose, we employ an exhaustive search over all four possible orientations (0° , 90° , 180° , 270°) of the target in order to compute the best matching one. The optimization criterion used is the minimization of the sum of the per-patch color distances under the neighborhood restrictions extracted in the previous step. Note that the search algorithm used in this step has a relatively small importance with respect to the running time in case of the classic color checker, as the size of the candidate list and the number of neighborhood relations is generally low. However, for large/complex color targets one may wish to use a more sophisticated search algorithm such as A^* [29] or one of its variants.

Experimental results

Our test set consists of 239 test images including photographs and various printed documents (newspaper excerpts, book pages) digitized by the German National Library as part of the use case Contentus in the Theseus program [30]. The test images were scanned using different resolutions ranging from 300 dpi to 600 dpi. The position of the color target varied considerably as the human scanning operator was allowed to put the color target anywhere in the vicinity of the item to digitize. Table 1 depicts the color target orientations in the dataset. As can be seen, our detection algorithm yields an average precision of 97.1%. We can thus conclude that it is robust to color checker orientations.

From the analysis of the test data, we have identified two main causes for the detection failures. The majority of the cases were caused by errors in the metadata of the input scan, namely the units for the scanning resolution were incorrectly specified as being dots per centimeter instead of dots per inch. The resulting grossly different scan resolution value caused the candidate patch filtering process to fail and further prevented the correct recognition of the color reference targets. Since all scans have the same provenience and were taken in the same time interval, it seems most likely that these inaccuracies are simply glitches in the image metadata. Such errors are practically unavoidable when large amounts of data are involved. The other cause of failure was a very high or complete occlusion of the color checker. In case less than a single row of color patches is visible on the scan (cf. Figure 5), our algorithm fails because of its inability to find enough initial patch candidates required for establishing a reliable orientation match. It is interesting to observe that in such extreme situations with very few reference color patches visible, the identification of a color target may not even be desirable because of the inherent inability to perform a subsequent meaningful color quality evaluation and/or correction.

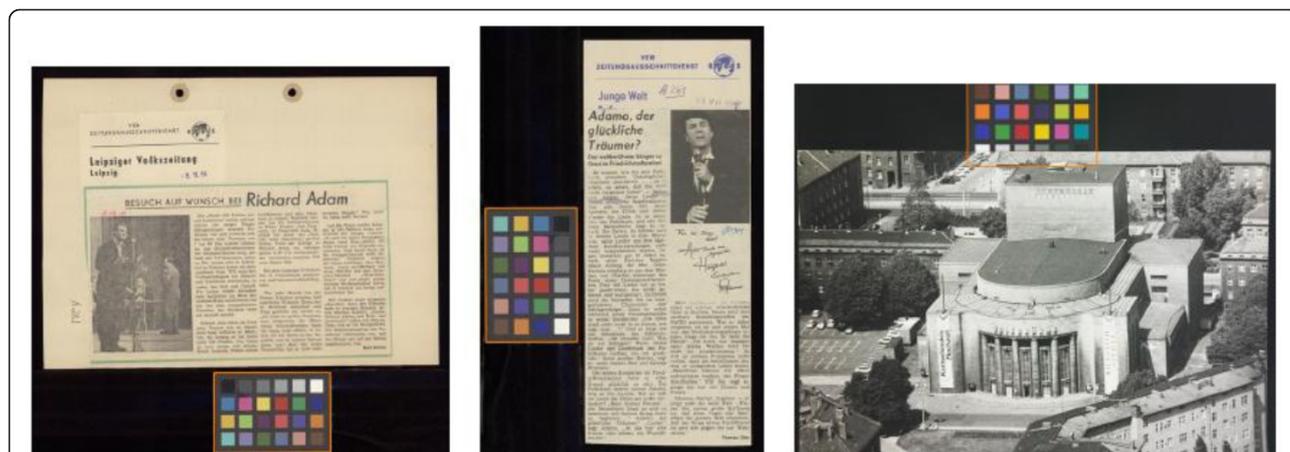


Figure 4 Examples of identified color reference targets, illustrating correct target detection for multiple orientations and robustness to partial occlusion.

4 Objective quality assessment and optimization

Although human observers can effectively and easily judge the quality of an image, this is highly time and resource consuming. Therefore, subjective measurements are not suitable in most cases. Three measure classes, i.e., full-reference (FR), reduced-reference (RR) and no-reference (NR) exist in the area of objective image quality assessment. The FR metrics predict the quality of an image based on differences to a reference image. Mean square error (MSE), peak signal-to-noise ratio (PSNR), and structure similarity (SSIM) [31] are popular representatives of FR metrics. They are also often used as benchmarks for evaluating RR and NR metrics. As print media must be digitalized before objectively assessing their quality, little reference information can be exploited. Hence, only RR and NR measurements can be used in this context. On the contrary to RR metrics, which use indirect information as a reference, NR metrics predict the quality of images by extracting and modeling prior knowledge on specific distortions. Thus, an objective NR metric will typically be designed for measuring a particular distortion type [32].

In the following sections, the proposed quality assessment system will be presented in two steps. First, color deviations via an RR metric will be analyzed. Second,

the distortions of the scanned image will be assessed via NR metrics. We will thereby distinguish between structural distortions, which distort the structures of textures like blocking artifacts and blurriness, and non-structural distortions such as brightness and contrast [31]. These NR metrics will additionally be combined into an overall quality metric.

4.1 Color-checker-based quality assessment

Commonly used professional scanners, either flatbed scanners or high-resolution DSLR cameras, capture the RGB components of the scanned image. These RGB components are then calibrated by some specific color calibration programs [33]. However, even after the calibration step, the color distribution of the scanned document may still significantly deviate from that of the original because of different photosensitive materials, refraction and reflection of different materials or the illumination conditions. The degree of color shift is an important a priori information, which should be assessed before the quality evaluation of the scanned document.

The digitization of print media is based on embedded ICC (International Color Consortium) profiles. The CIE $L^*a^*b^*$ is used, instead of the sRGB color space, to evaluate the color deviation between the print media and their scanned version, because the CIE $L^*a^*b^*$ color space is applied to most ICC profiles for color management of scanners and printers [34]. The ΔE_{ab}^* , defined by the *International Commission on Illumination* (CIE), describes the difference between the original and the scanned color checkers. The ΔE_{ab}^* between two color samples $(L^*, a^*, b^*)_1$ and $(L^*, a^*, b^*)_2$ is defined as,

$$\Delta E_{ab}^* = \sqrt{(L_1^* - L_2^*)^2 + (a_1^* - a_2^*)^2 + (b_1^* - b_2^*)^2} \quad (2)$$

Table 1 Color reference target detection results on a heterogeneous dataset consisting of books, newspaper excerpts, and photographs

	Color checker Orientation				Total
	Horizontal		Vertical		
	0°	180°	90°	270°	
Total Scans	115	13	109	2	239
Failed Detections	2	0	5	0	7
Precision (%)	98.4		95.5		97.1

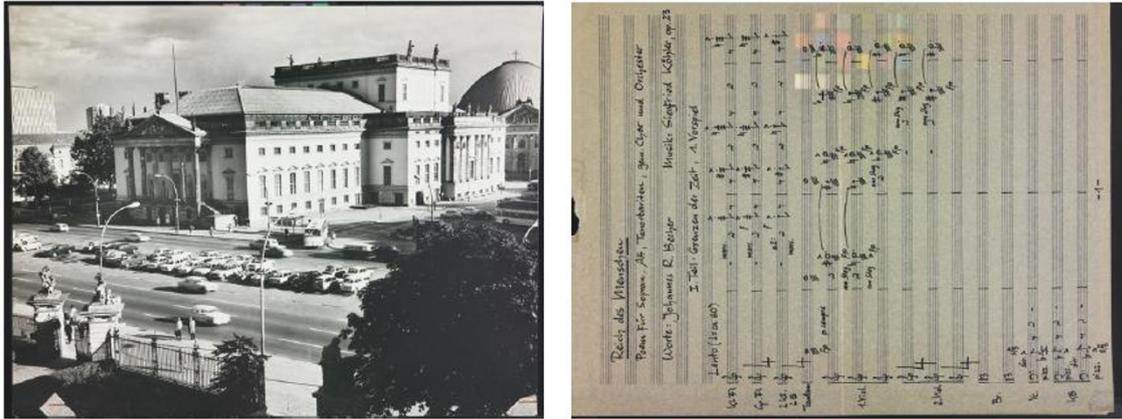


Figure 5 Examples of failed color target identifications caused by (partial) occlusion.

Based on the CIE76 [35] criteria, if $\Delta E_{ab}^* > 2.3$, the difference is already noticeable. However, if $\Delta E_{ab}^* > 5.0$, the difference can be evaluated as a different color [36] and the ICC profile of the analog-to-digital device needs to be re-customized.

4.2 No-reference image quality assessment

In this section, major structural and non-structural distortions are measured by different NR metrics. The single NR metrics are finally integrated into an overall quality metric.

4.2.1 Sharpness

Sharpness is one of the most important properties of image quality. Many reasons may lead to the distortion of image sharpness such as compression, smoothing/denoising, defocusing, motion, discoloration of photos, etc., [37]. Most of the objective NR sharpness metrics are proposed in the frequency or the spatial domains [38]. Although the metric just noticeable blur (JNB) [37] and its improved version cumulative probability of blur detection (CPBD) [39] feature a better performance than other NR sharpness metrics, their runtime costs are very high. We therefore propose a two-stage sharpness metric that shows better runtime performance than JNB and CPBD. First, the probability of local blurriness is fragmentarily measured. Then, the visually salient high frequency components are measured to extend the probability of local blur to the whole image. The calculation of the blur probability of an edge pixel follows the concept of the JNB proposed in [37] and is defined as,

$$p(e_i) = 1 - \exp\left(-\left|\frac{w(e_i)}{w_{JNB}(e_i)}\right|^\beta\right), \quad (3)$$

where w_{e_i} is the width of edge e_i [40] and $w_{JNB}(e_i)$ is the width of the JNB. If the standard contrast (cf.

Equation 10) of an edge region is larger than 50, the corresponding $w_{JNB}(e_i)$ is measured to be 3, otherwise to be 5 [37]. The parameter β controls the curvature of the psychometric probability function. β values are chosen between 3.4 and 3.8 with a median at 3.6 by means of least-square fitting [37].

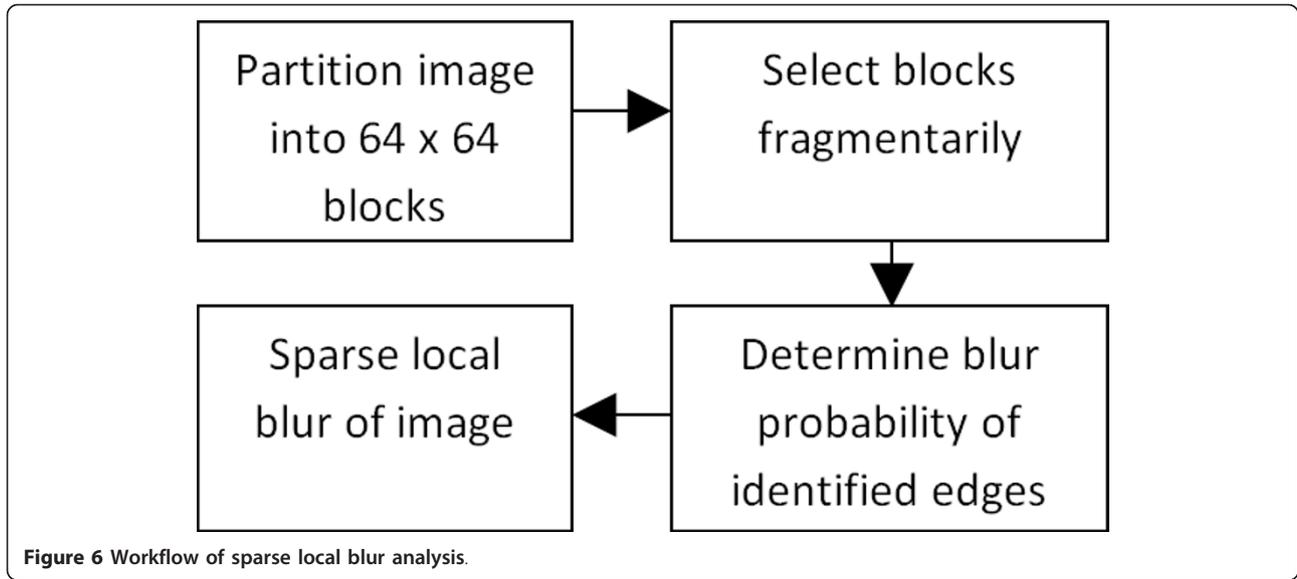
Sparse local analysis Figure 6 illustrates the computation of the local blurriness probability. The image is first partitioned into 64×64 blocks. However, we assume that evaluation of dominant structures and textures may not require to consider all blocks. Thus, we suggest selecting the blocks fragmentarily to increase processing speed. Our experiments have shown that best performance is achieved if the total area of selected blocks is not less than 25% of the image area independently of the content. The second step consists in the block selection. The blocks may be selected at $m\%2 = 0$ and $n\%2 = 0$, where m and n correspond to coordinates at block resolution. Therefore, we name this approach *Sparse Local Analysis*. Thirdly, a block is classified as an edge block when at least 0.2% of its pixels are marked as edge pixels [37]. The blur probability of an edge block is defined as the sum of blur probabilities of each edge pixel within the block [37]. The overall local blur probability is defined as,

$$S_L = \frac{N_b \Gamma_n}{\Gamma}, \quad (4)$$

where N_b is the number of processed edge blocks, Γ is the sum of the blur probabilities of all processed edge blocks and Γ_n is a normalization factor defined as,

$$\Gamma_n = \left(\sum_b \left|\frac{64 \times 64 \times 0.2\%}{w_{JNB}(b)}\right|^\beta\right)^{\frac{1}{\beta}}, \quad (5)$$

which denotes the sum of the blur probabilities of all edge blocks with the best sharpness values. $w_{JNB}(b)$ is



the edge width of the JNB based on the contrast of the block b .

Global analysis According to the curves and surfaces theory, edges of objects, which have distinctive shapes, are normally continuous in a small area [41]. Therefore, it can be supposed that the edges in an unmeasured edge block are similar to the edges in its analyzed neighbor edge blocks. The proportion between the edge and non-edge blocks is approximately equal to the proportion between the high and low frequency components in an image. Thus, the global sharpness of the image can be estimated from the sparse local sharpness by an extension with the weight of the high frequency components. We suggest using higher order statistics (HOS) to extract the high frequency components from images, as they can suppress Gaussian noise and preserve non-Gaussian information [42]. The fourth-order moments of the HOS are calculated for all pixels within the luminance channel I of an $M \times N$ image,

$$\text{HOS}(y, x) = \frac{1}{4(\epsilon + 1)^2} \sum_{i=y-\epsilon}^{y+\epsilon} \sum_{j=x-\epsilon}^{x+\epsilon} (I(i, j) - \mu(y, x))^4, \quad (6)$$

where $\mu(y, x)$ is the mean luminance of the pixel set $\{0 \leq y \pm \epsilon < M, 0 \leq x \pm \epsilon < N\}$ and ϵ is set to the mean w_{JNB} of all measured edge blocks. The range of HOS values is limited to $[0, 2^{2n-1}]$, where n denotes the bit-depth of color space. Thus, the relationship between the low and high frequency components can be formulated as the following equation,

$$m_L = \sum_{i=0}^T i \cdot h(i) = \eta \cdot \sum_{j=T}^{2^n-1} j \cdot h(j) = \eta \cdot m_H, \quad \eta > 0, \quad (7)$$

where m_L/m_H denotes the mass of the low/high frequency components in the histogram h of the corresponding HOS map. The parameter η controls the ratio of the low/high frequency components. For example, $\eta = 1$ means $m_L = m_H$. The unknown threshold T , which distinguishes between the low and high frequency components of the histogram, can be calculated from the Equation 7. Then, the weight of the high frequency components is defined as,

$$W = \frac{m_{HC}}{2^{n-1}} = \frac{1}{2^{n-1}} \frac{m_H}{\sum_{i=T}^{2^n-1} i}, \quad (8)$$

where m_{HC} denotes the center of mass of the high frequency components in the histogram h .

However, the sparse local sharpness analysis does not measure global image properties. The weight of the high frequency components W , which is estimated as the distribution of edges-of-interest within the image, can be used to extend the local metric to the whole image area. Thus, the new proposed metric LGS based on both *Local and Global Sharpness* analysis is defined as,

$$S_q = W \cdot S_L. \quad (9)$$

In practice, the areas near the boundaries of an image are less important for human observers than the rest areas within the image [43]. They are thus ignored to save runtime costs in this work.

4.2.2 Contrast and brightness

Contrast, in the terms of vision, is the difference in luminance of the background and the objects of interest. Visually apparent contrast is an important perceptual attribute of an image for human observers to distinguish objects from their background. The study of contrast

sensitivity [44] shows that contrast is mainly dependent on the brightness and the dynamic range of the image.

The contrast of an image is referred to as the Euclidian distance between the mass centers of the upper and lower parts of the luminance histogram projected onto the bin axis [45].

Each part has the same mass as the other. The standard contrast metric is defined as

$$C = \frac{|L_u - L_l|}{L_a}, \quad (10)$$

where L_u/L_l is the average mass of the upper/lower part and L_a is the average mass of the luminance channel. L_u , L_b , and L_a are projected onto the bin axis.

However, if the average luminance is too large or too small, the standard metric should be changed to the *Weber* contrast [45]. Depending on different image properties, different algorithms for measuring the contrast are applied accordingly. For practical applications, we design a new contrast quality metric by normalizing the standard contrast metric to avoid the inconsistency of the standard metric. The proposed normalization factor is defined as,

$$\text{norm} = \frac{|L_u - L_l|}{R_l}, \quad (11)$$

where R_l denotes the range of luminance intensity. This factor represents the ratio of the luminance difference and the luminance intensity. Thus, the final metric of contrast quality assessment (CQA) is

$$C_q = C \cdot \text{norm}. \quad (12)$$

The brightness measure corresponds to L_a in this work. Brightness can be seen as a secondary attribute of contrast. If the image has good contrast but poor brightness, it means that, although the image has some bright areas, most areas are dark.

4.2.3 Overall quality

Human observers judge an image subjectively based on its overall quality [31]. The quantification of single image properties by objective metrics is not enough to simulate human perception. In contrast to the common FR metrics, which can be used to measure generic distortions of an image, state-of-the-art NR metrics will typically be applied to specific distortions [46]. We thus propose a NR metric for overall quality assessment that integrates the structural and non-structural distortion [31] probabilities.

The Blocking Artifacts (BA) and Blurriness (Bl) are the major Structural Distortions (SD). Blocking and blur results from compression are not totally uncorrelated if a hybrid block based video codec is used [47]. Non-

structural distortions (ND) is solely defined as the contrast (C_o) measure, as the brightness is a component of at measure. Our overall quality metric can be thus formulated as,

$$Q_o = C_q \cdot (a \cdot B_q + b \cdot S_q), \quad (13)$$

where $a + b = 1$; $a, b > 0$. The higher Q_o is, the better quality the image has. All NR metrics are limited and normalized to 0[1].

B_q corresponds to a metric for blocking artifacts as defined in [48]. This NR blocking metric is modeled in the spatial domain. It is based on three characteristics of perceivable blocking artifacts, the strength of the block boundaries, the discontinuities across the block boundaries and the flatness of the image. Compared to other NR blocking metrics, this metric has a good balance between complexity and performance. But the range of this metric depends on its parameter sets. Thus, we normalize this metric to achieve a range of 0[1].

4.2.4 Experimental results

The print materials used in our experiment were digitized by *Hasselblad H4D-50MS^f* to 8-bit TIFF format or compressed to 8-bit JPEG formats. The images were converted to the *sRGB* color space based on embedded ICC profiles. The results of all NR metrics were normalized to 0[1] to become resolution invariant. Outliers, which were smaller than 0 or larger than 1 after normalization, were also truncated to 0[1]. Two common image databases, the CSIQ [49] and LIVE [50] databases, were selected to evaluate the metrics. The CSIQ database contains 30 reference images. There are a total 886 distorted images generated from six types of distortions with 4-5 corresponding levels per image. All images were rated by 25 observers. The subjective ratings were recorded by differential mean opinion scores (DMOS). The resolution of each image is 512×512 . The LIVE database contains 29 reference images and 26-29 distorted versions of each reference image. For all images of the LIVE database, subjective quality rates generated by 20-25 observers are provided in form of DMOS. The resolution of each image is 768×512 . Figures 7 and 8 show some exemplar images of the databases.

The Pearson (CC) and the Spearman (SROCC) coefficients [51] were used to measure the correlation of the proposed metrics with mean opinion scores (MOS) from subjective ratings. High CC score and SROCC score relate to high accuracy, monotonicity and consistency of the metric under test [52]. 95% confidence intervals of CC and SROCC were calculated based on Fisher's transformation [53]. The computational complexity of the metric is also a very important property for practical applications. Therefore, the complexity of each metric was evaluated by measuring the mean



Figure 7 Examples of the CSIQ database (a) Gaussian blur; (b) JPEG2000; (c) contrast; (d) JPEG.

runtime in *seconds per image (s/img)* over all datasets. The metrics evaluated in this section are all based on luminance information. Thus, the sRGB images were transformed into 8-bit grayscale images [54]. The best performances of the evaluated NR metrics were highlighted in the corresponding tables. The operating system used for the experimental environment was Windows 7 64-bit Professional with i7-M620 2.67 GHz CPU and 8.00 GB memory. Two full-reference objective metrics, PSNR and SSIM [31], were also evaluated as benchmarks.

Sharpness 150 Gaussian blurred and 150 JPEG2000 compressed images of the CSIQ database as well as 175 Gaussian blurred and 288 JPEG2000 compressed images of the LIVE database were used to evaluate the sharpness metrics. In this experiment, 1/16 pixels from each image boundary were cut off. 3/4 of all partitioned blocks were ignored by the proposed sharpness metric LGS in its fragmentary local analysis. The metric CPBD, is extended from CPBD, which also ignores 3/4 of all blocks of the image as a benchmark. The parameter η was set to 1, since we assume that the proportions of the low/high frequency components are the same.

The performances of the evaluated sharpness metrics on the CSIQ database are summarized in Table 2. Considering the evaluation, it is shown that the proposed

metric (LGS) and CPBD offer the best results in comparison with other NR metrics. LGS features higher CC respectively lower SROCC than CPBD with Gaussian blur. CPBD could be successfully used to measure JPEG2000 distortions. According to the mean CC and SROCC, LGS performs similarly to CPBD. In spite of this fact, LGS has a speed-up of over three times, compared to the metric JNB or CPBD. Further in-depth information is shown in Figure 9a,b. It can be seen that the differences of CC and SROCC between LGS and CPBD are not significant. The confidence intervals of JNB have noticeable deviations from other evaluated sharpness metrics. It is observed that JNB is not suitable for measuring low-depth-of-field [55] images. We supposed similar results were obtained for the LIVE database (cf. Figure 9c,d). Generally, the new metric performs closely to the FR metrics. LGS features some slight advantages compared to the FR metrics for Gaussian blur. The performance of LGS on JPEG2000 images is comparable to PSNR. However, the FR metrics are not suitable for the proposed framework, since no reference is available.

Contrast/brightness Table 3 shows the performance of the evaluated contrast metric with the contrast dataset (150 images) of the CSIQ database. It is shown that the proposed metric (CQA) is a good alternative to the



Figure 8 Examples of the LIVE database (a) Gaussian blur; (b) JPEG2000; (c) JPEG.

Table 2 Performance of sharpness measures using the CSIQ database

Metrics	Dataset	CC	SROCC	Mean CC	Mean SROCC	Complexity (s/img)
PSNR	Gaussian blur JPEG2000	0.824	0.862	0.836	0.878	0.004
		0.849	0.894			
SSIM [31]	Gaussian blur JPEG2000	0.850	0.925	0.864	0.923	0.083
		0.878	0.921			
JNB [37]	Gaussian blur JPEG2000	0.588	0.717	0.679	0.746	1.447
		0.769	0.776			
CPBD [39]	Gaussian blur JPEG2000	0.830	0.885	0.846	0.869	1.479
		0.862	0.853			
CPBD _r	Gaussian blur JPEG2000	0.830	0.881	0.848	0.869	1.434
		0.866	0.856			
LGS	Gaussian blur JPEG2000	0.887	0.872	0.861	0.849	0.459
		0.836	0.825			

standard metric. Slight gains are visible, but they are not statistically significant. However, the standard metric (Equation 10) cannot be directly integrated into our system, because its interval does not lie within $[0,1]$. The proposed normalization does not cost much. Therefore, the computation speed of the CQA is very close to that of the standard metric, and faster than that of the FR metrics. CQA has good correlation and complexity for practical applications. The performance of the FR metrics on contrast is lower than on other structural distortions. Figure 9e shows that there is no significant difference among the evaluated metrics.

Overall quality The TID2008 database [56], which contains 25 reference images and 1700 distorted images, was used for training the parameters of the proposed overall metric. The training shows that the combination of the parameters leads to a good performance if $a \geq b$. $a = 0.6$ and $b = 0.4$ were thus used in the experiment.

Tables 4 shows the performance of the evaluated overall metric using the CSIQ and the LIVE databases. In our evaluation, we use not only the single datasets but also integrate both to a single database.

Figure 10 shows the 95% confidence interval of these metrics with the CSIQ and LIVE database. There is no significant deviations of CC between the proposed NR overall metric (OQA) and SSIM. However, a small deviation of SROCC can be observed in measuring the JPEG2000 distortion. Compared to the FR metrics, OQA still has some advantages in measuring contrast and Gaussian blurred distortion. The performance of these three metrics is close to, in the measurement of the blocking artifacts caused by the JPEG compression. There is also no significant difference between OQA and the proposed single NR metrics. The proposed NR overall metric is robust of measuring different unspecified distortions.

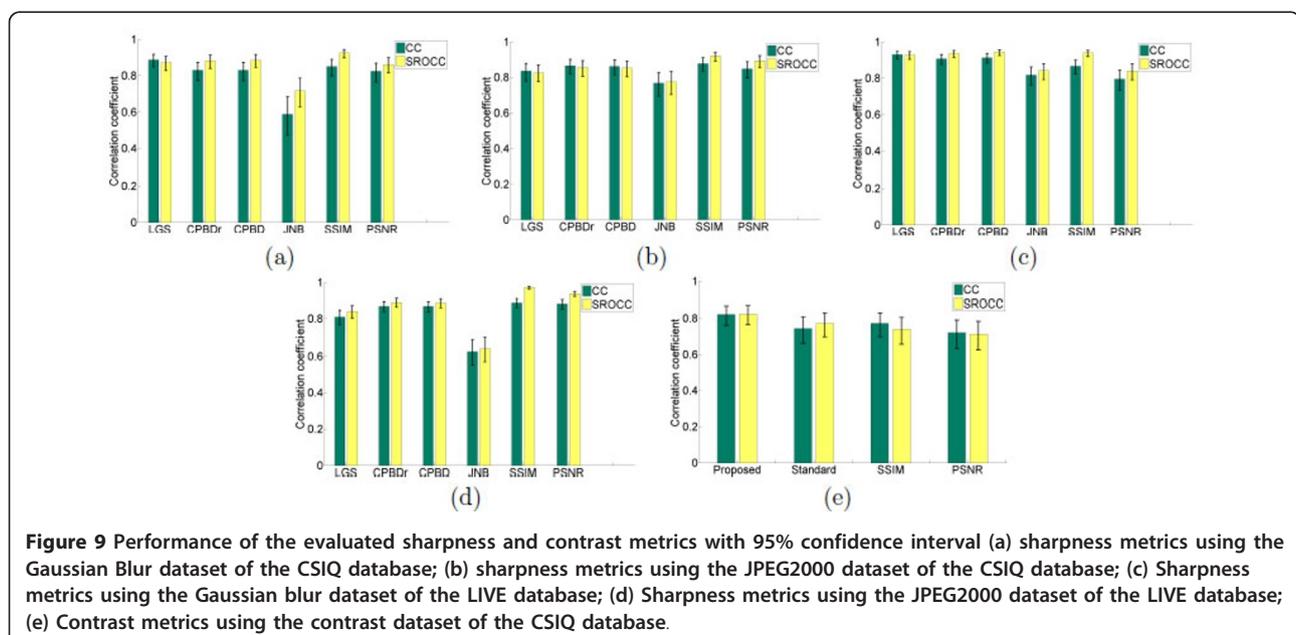


Figure 9 Performance of the evaluated sharpness and contrast metrics with 95% confidence interval (a) sharpness metrics using the Gaussian Blur dataset of the CSIQ database; (b) sharpness metrics using the JPEG2000 dataset of the CSIQ database; (c) Sharpness metrics using the Gaussian blur dataset of the LIVE database; (d) Sharpness metrics using the JPEG2000 dataset of the LIVE database; (e) Contrast metrics using the contrast dataset of the CSIQ database.

Table 3 Performance of contrast measures using the CSIQ database

Metrics	CC	SROCC	Complexity (s/img)
PSNR	0.721	0.713	0.005
SSIM [31]	0.772	0.740	0.086
Standard	0.744	0.772	0.001
CQA	0.822	0.824	0.001

4.3 Quality optimization

Automated quality analysis methods can be applied for efficient image optimization, both manual and automated. A prerequisite is the return of relevant image parameters (as shown in this article) as metadata by the quality analysis algorithm.

If a digital image does not match given quality benchmarks (like insufficient contrast for printed text), an operator may improve the image by using image processing software. Alternatively, an automated image quality optimization algorithm may adjust the image such as to match insufficient image parameters to given quality benchmarks—i.e., the amount of image optimization is not fixed but rather depends on the actual image quality. In this case, images with sufficient quality benchmarks will not be further processed. With digitized text material, where OCR quality or human readability depends on certain image parameters like contrast and sharpness, the use of quality metadata will provide good automated optimization results without overprocessing the digitized text (like halos around letters that occur due to oversharpening) [57]. This approach can be compared to program dependent compression of an audio signal in the domain of broadcasting where the amount of signal processing is adjusted depending on the audio source and the required sound quality.

Quality metadata may also be used in reporting systems that support operators to quickly detect images below the quality threshold out of a large population. In practice, the operator may be able to halt the digitization process in case of scanner-related errors.

Table 4 Performance of the proposed overall metric using the CSIQ and the LIVE database

Dataset	Database	CC	SROCC	Complexity (s/img)
JPEG	CSIQ	0.889	0.870	0.524
	LIVE	0.839	0.881	0.698
Gaussian blur	CSIQ	0.884	0.850	0.488
	LIVE	0.875	0.911	0.668
JPEG2000	CSIQ	0.811	0.799	0.513
	LIVE	0.796	0.821	0.703
Contrast	CSIQ	0.769	0.758	0.585
All	CSIQ	0.825	0.808	0.527
	LIVE	0.815	0.873	0.689

While it is obvious that, instead of checking the quality of thousands of scanned images per day manually, operators making use of automated quality analysis algorithms will save a tremendous amount time, it is also important to consider the response time of error detection. In case of information loss due to scanning or image processing errors—including all types of defects which cannot be corrected using automated quality optimization—books or documents have to be digitized once again. If a significant error is detected with a low response time, the book may have already been brought back to the archive and thus, digitizing it once more will consume even more human resources in addition to the re-digitizing step itself.

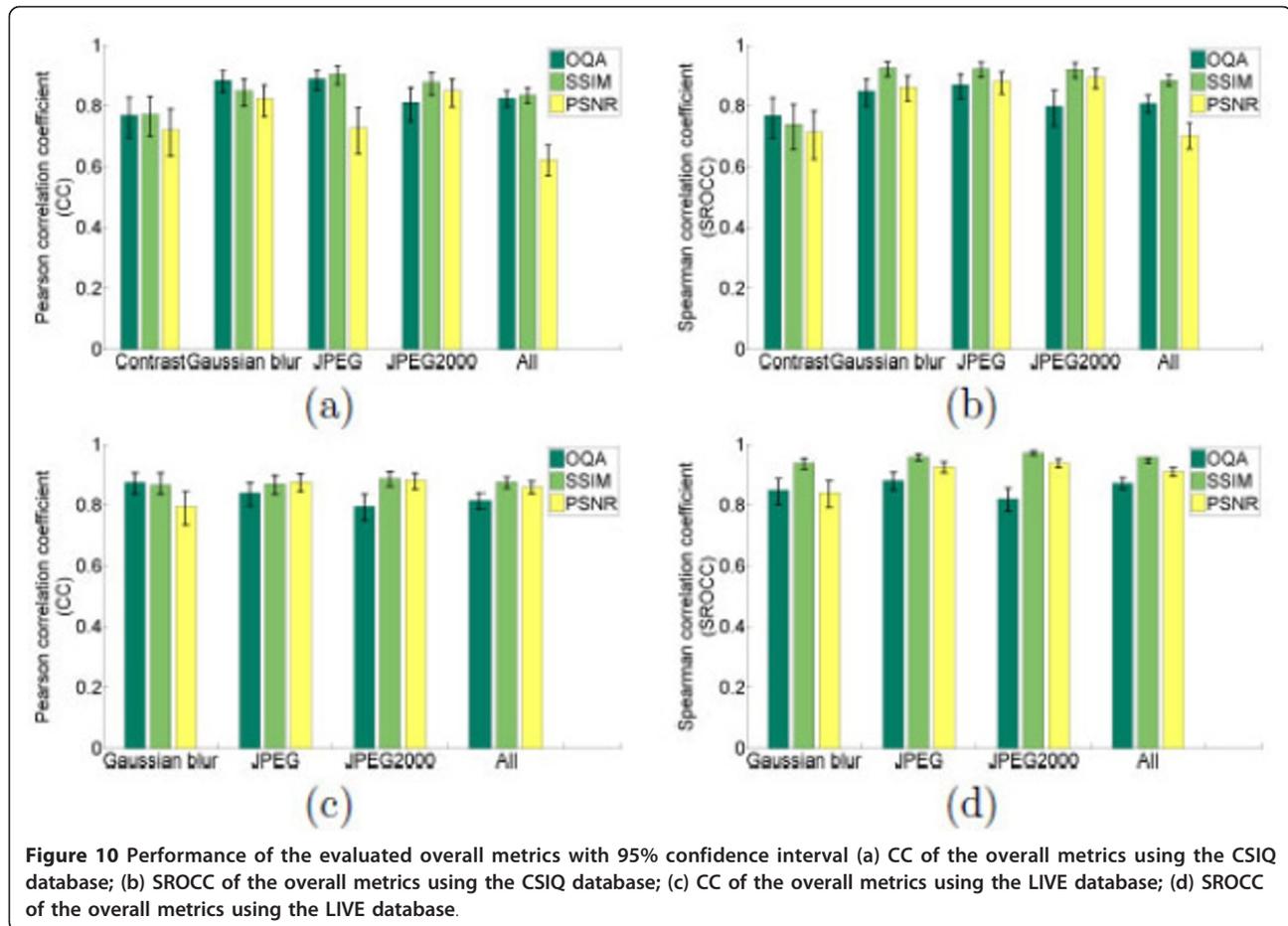
In another case, images may be below quality thresholds due to systematic scanner-related errors (e.g., offset in color temperature due to deteriorated scanner light bulbs). The larger the time lag between the occurrence of a systematic error and its detection, the more books or documents have to be re-digitized.

As a consequence, a high-throughput workflow that quickly detects errors will elicit a rapid halt of production and/or trigger an immediate redigitization. To sum up, mass digitization workflows heavily rely on automated tools as described in this article for the image quality to be instantaneously optimized.

4.4 Further work: integration of technologies in a productive overall system

Future study will deal with the adaption of the technologies described in this article to production environments of mass digitization workflows. In order to achieve this, two challenges—one related to the function and another relating to technical integration—have to be taken into account. The former can be further subdivided into

- 1). Reliability:
 - a. Implementation of further metrics, as noise and picture dynamic.
 - b. Correct measurement of quality parameters.
 - c. Quality analysis should be sufficiently robust and configurable to account for variations appearing in analog print media collections;
- 2). Performance:
 - a. Optimize the metric performances and the respective complexity.
 - b. Optimize speed to analyze a given set of quality parameters.
 - c. Reduce required hardware resources to analyze a given set of quality parameters.
 - d. Automated target detection when used;
- 3). Significance:
 - a. Providing the user meaningful interpretation of measurement results.



- b. Hiding irrelevant data, i.e., information not immediately related to image quality;
- 4). Usability:
 - a. Displaying quality parameters in a graphical way enabling the user a quick overview for a large set of digitized images.
 - b. Usage has to take into account the qualification level of digitization operators.
 - c. Quality reports should be both human- and machine-readable and concise.
 - d. Quality metadata should be provided in an achievable format,

while the latter challenge addresses requirements related to implementation costs, hardware technology, workflow framework, overall architecture etc., that are important but beyond the scope of this article.

On one hand, best practice guidelines from the cultural sector [58-60] have provided detailed image quality metrics as a basis for a quantifiable quality management in digitization workflows (for overview see [7]). On the other hand, a considerable subset of image quality parameters which cultural heritage organizations require for

quality assessment in mass digitization workflows are measured by technologies described in this article. Hence, the results for mass digitization workflows comply with “real world” requirements from cultural heritage organizations and content holders such as publishers likewise.

5 Conclusions

This article presents a new quality assessment and improvement system for print media. After the digitization of print media, an automatic segmentation is applied to separate the scanned print media and the color checker. The color disparities between the original and the scanned color checkers are first measured to give a priori quality assessment of the digitization. No-reference quality metrics measure the eventual distortions of the scanned print media. The no-reference metrics are also integrated to an overall quality metric. An automatic image quality optimization algorithm is then applied to adjust the image to match given quality benchmarks. The LIVE and the CSIQ databases were used to evaluate the performance of these no-reference metrics. The evaluation shows that the proposed no-

reference metrics are robust in measuring the corresponding distortions.

Acknowledgements

This study was funded by the German ministry of industry and is part of the Contentus project [13], which is itself part of the Theseus program [30].

Endnotes

^a04/15/2011: http://www.roboticbookscan.com/index/products/rbspro_tt_a3.
^b04/15/2011: <http://www.kirtas.com/kabislll.php>.
^c04/15/2011: http://www.treventus.com/bookscanner_pageturner.html.
^d04/15/2011: http://www.image-engineering.de/index.php?option=com_content&view=article&id=530&Itemid=50.
^e04/15/2011: <http://www.certifi-media.com/QandPProducts.aspx>.
^f09/20/2011: <http://www.hasselbladusa.com/products/h-system/h4d50ms.aspx>.

Author details

¹Image Processing Department, Fraunhofer Institute for Telecommunications - Heinrich Hertz Institute, Berlin, Germany ²NetMedia Department, Fraunhofer Institute for Intelligent Analysis and Information System -IAIS, Sankt Augustin, Germany ³Acosta Consult GmbH, Frankfurt, Germany

Competing interests

The authors declare that they have no competing interests.

Received: 27 May 2011 Accepted: 15 May 2012 Published: 15 May 2012

References

1. M Baker, M Shah, DSH Rosenthal, M Roussopoulos, P Maniatis, T Giuli, P Bungale, A fresh look at the reliability of long-term digital storage, in *Proc ACM SIGOPS/EuroSys European Conference on Computer Systems*, vol. 40. Leuven, pp. 221–234 (2006)
2. D Williams, PD Burns, Preparing for the image literate decade, in *Proc IS&T Archiving Conference*, vol. 6. Arlington, pp. 124–127 (2009)
3. D Williams, PD Burns, L Scarff, Imaging performance taxonomy, in *Proc SPIE-IS&T Electronic Imaging Symposium*, vol. 7242. San Jose, pp. 724208–724208-7 (2009)
4. M Stelmach, D Williams, When good scanning goes bad: a case for enabling statistical process control in image digitizing workflows, in *Proc IS&T Archiving Conference*, vol. 3. Ottawa, pp. 237–243 (2006)
5. PD Burns, D Williams, Ten tips for maintaining digital image quality. *IS&T Soc Imag Sci Technol.* **4**, 16–22 (2007)
6. PD Burns, D Williams, Identification of image noise sources in digital scanner evaluation, in *Proc SPIE-IS&T Electronic Imaging Symposium*, vol. 5294. San Jose, pp. 114–123 (2004)
7. PW Jones, CW Honsinger, Image quality assurance for the real world, in *Proc IS&T Archiving Conference*, vol. 7. Den Haag, pp. 90–95 (2010)
8. D Williams, Debunking of specsmanship: progress on ISO/TC42 standards for digital capture imaging performance, in *Proc IS&T PICS Conference*, vol. 7. Rochester, pp. 77–81 (2003)
9. D Williams, PD Burns, Diagnostics for digital capture using MTF, in *Proc IS&T PICS Conference*, vol. 4. Montreal, pp. 227–232 (2001)
10. KB national library of the Netherlands, IMPACT project 2008 <http://www.impact-project.eu/home/> (2011)
11. Group SIW, Technical guidelines for digitizing cultural heritage materials: creation of raster image master files <http://www.digitizationguidelines.gov/guidelines/digitize-technical.html> (2011)
12. D Wueller, H van Dormolen, V Jansen, Universal Test Target Technical Specification, (National Library of the Netherlands, 2011)
13. German Federal Ministry of Economics and Technology, Contentus project 2007 <http://www.contentus-projekt.de> (2011)
14. H Baird, Document image defect models and their uses, in *Proc IEEE Document Analysis and Recognition*, Tsukuba Science City, pp. 62–67 (1993)
15. D Ittner, D Lewis, D Ahn, Text categorization of low quality images, in *Proc SDAIR 4th Annual Symposium on Document Analysis and Information Retrieval*, Las Vegas, pp. 301–315 (1995)
16. LR Blando, Evaluation of page quality using simple features, Master thesis, (Department of Computer Science, University of Nevada, Las Vegas, 1994)
17. BW Keelan, *Handbook of Image Quality: Characterization and Prediction*, (Marcel Dekker, Inc., New York, 2002)
18. The European Union, SCAPE project 2009 <http://www.scape-project.eu/> (2011)
19. X-Rite, Color Checker Classic 2011 http://xritephoto.com/ph_product_overview.aspx?id=1192&catid=28&action=overview (2011)
20. C McCamy, H Marcus, J Davidson, A color-rendition chart. *J Appl Photogr Eng.* **2**(3), 95–99 (1976)
21. X-Rite, Color checker digital SG, 2011 http://xritephoto.com/ph_product_overview.aspx?id=938&catid=28&action=overview (2011)
22. National Library of the Netherlands, Metamorfoze Programme, 2011 <http://www.metamorfoze.nl/en/programma/index.html> (2011)
23. T Tajbakhsh, R Grigat, Semiautomatic color checker detection in distorted images, in *Proc Signal Processing, Patt Recognition and Applications (SPPRA)*, Innsbruck, pp. 347–352 (2008)
24. X-Rite, Color checker passport camera calibration software, 2011 http://xritephoto.com/ph_product_overview.aspx?ID=1257 (2011)
25. G Sharma, W Wu, EN Dalal, The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Res Appl.* **30**, 21–30 (2005). doi:10.1002/col.20070
26. RS Hunter, Photoelectric color-difference meter. *JOSA.* **38**, 651 (1948)
27. MB Dillencourt, H Samet, M Tamminen, A general approach to connected-component labeling for arbitrary image representations. *J ACM.* **39**(2), 253–280 (1992). doi:10.1145/128749.128750
28. L Guibas, J Stolfi, Primitives for the manipulation of general subdivisions and the computation of Voronoi diagrams. *ACM Trans Graphics.* **4**(2), 74–123 (1985). doi:10.1145/282918.282923
29. P Hart, N Nilsson, B Raphael, A formal basis for the Heuristic determination of minimum cost paths. *IEEE Trans Syst Sci Cybern.* **4**(2), 100–107 (1968)
30. German Federal Ministry of Economics and Technology, Theseus program, 2007 <http://theseus-programm.de/en-US/home/default.aspx> (2011)
31. Z Wang, A Bovik, H Sheikh, E Simoncelli, Image quality assessment: From error visibility to structural similarity. *IEEE Trans Image Process.* **13**, 600–612 (2004). doi:10.1109/TIP.2003.819861
32. Z Wang, AC Bovik, *Modern Image Quality Assessment*, (Morgan & Claypool, California, US, 2005)
33. PR Chang, CC Chang, Color correction for scanner and printer using B-spline CMAC neural networks, in *Proc IEEE Asia-Pacific conference on Circuits and Systems, Taipei*, pp. 24–28 (1994)
34. International Color Consortium, Image technology colour management: architecture, profile format, and data structure, 2004 <http://www.color.org> (2011)
35. CIE, CIE 1976 L*a*b Colour space draft standard, 2007 http://cie.co.at/index.php?i_ca_id=485 (2011)
36. Sharma G (ed.), *Digital Color Imaging Handbook* (CRC Press, Florida, US, 2003)
37. R Ferzli, LJ Karam, A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb). *IEEE Trans Image Process.* **18**, 717–728 (2009)
38. R Hassen, Z Wang, M Salama, No-reference image sharpness assessment based on local phase coherence measurement. in *Proc IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Dallas 2434–2437 (2010)
39. ND Narvekar, LJ Karam, A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection. in *Proc IEEE International Workshop on Quality of Multimedia Experience (QoMEX)*, San Diego 87–91 (2009)
40. P Morziliano, F Dufaux, S Winkler, T Ebrahimi, Perceptual blur and ringing metrics: Application to JPEG2000. *Signal Process Image Commun.* **19**, 163–172 (2004). doi:10.1016/j.image.2003.08.003
41. W Kühnel, *Differential Geometry Curves-Surface-Manifolds*, 2nd edn. (American Mathematical Society, Rhode Island, US, 2006)
42. G Gelle, M Colas, G Delaunay, Higher order statistics for detection and classification of faulty fanbelts using acoustical analysis, in *Proc IEEE Signal Processing Workshop on Higher-Order Statistics*, vol. 2. Banff, pp. 43–46 (1997)
43. B Clements, D Rosenfeld, *Photographic Composition*, (Van Nostrand Reinhold, New York, US, 1979)
44. MD Fairchild, A victory for equivalent background-on average, in *Proc Seventh Color Imaging Conference: Color Science, Systems and Applications*, vol. 7. Scottsdale, pp. 87–92 (1999)
45. E Peli, Contrast in complex images. *J Opt Soc Am.* **A7**(10), 2032–2040 (1990)

46. M Nuutinen, O Orenius, T Säämänen, P Oittinen, Reference image method for measuring quality of photographs produced by digital cameras, in *Proc Image Quality and System Performance VIII*, vol. 7867. San Francisco, pp. 78670M–14 (2011)
47. T Wiegand, GJ Sullivan, G Bjøntegaard, A Luthra, Overview of the H.264/AVC Video Coding Standard. *IEEE Trans Circ Syst Video Technol.* **13**, 560–576 (2003)
48. Z Wang, HR Sheikh, AC Bovik, No-reference perceptual quality assessment of JPEG compressed images, in *Proc IEEE International Conference on Image Processing (ICIP)*, vol. 1. New York, US, pp. 477–480 (2002)
49. EC Larson, DM Chandler, Most apparent distortion: full-reference image quality assessment and the role of strategy. *J Electron Imag.* **19**(1), 1–21 (2010)
50. HR Sheikh, MF Sabir, AC Bovik, A statistical evaluation of recent full reference image quality assessment algorithm. *IEEE Trans Image Process.* **15**, 2440–2451 (2006)
51. M Kendall, J Gibbons, *Rank Correlation Methods*, 5th edn. (Edward Arnold, London, 2002)
52. VQEG: Final report from the video quality experts group on the validation of objective models of video quality assessment. Tech rep, VQEG (2003)
53. MF Triola, *Essentials of Statistics*, 4th edn. (Pearson, New Jersey, 2011)
54. WK Pratt, *Digital Image Processing*, (John Wiley & Sons, Inc., New York City, US, 2007)
55. HM Merklinger, *The Ins and Outs of Focus: An Alternative Way to Estimate Depth-Of-Field and Sharpness in the Photographic Image*, (Seaboard Printing Limited, Nova Scotia, 1992)
56. N Ponomarenko, V Lukin, K Egiazarian, J Astola, M Carli, F Battisti, Color image database for evaluation of image quality metrics. in *Proc IEEE Workshop on Multimedia Signal Processing*, Cairns 403–408 (2008)
57. D Williams, Measuring and managing digital image sharpening. in *Proc IS&T 2008 Archiving Conference*, Bern 89–93 (2008)
58. Federal Agencies Digitization Guidelines Initiative (Still Image Working Group): Digital Conversion - Documents and Guidelines: A Bibliographic Reference, 2009 http://www.digitizationguidelines.gov/guidelines/Guidelines_Bibliography-2009rev.pdf (2011)
59. S Puglia, J Reed, E Rhoads, Technical guidelines for digitizing archival materials for electronic access: creation of production master files - raster images, (U.S. National Archives and Records Administration (NARA), 2004, 2011) <http://www.archives.gov/preservation/technical/guidelines.pdf>
60. H Dormolen, R Gillesse, Metamorfoze preservation imaging guidelines, in *Proc IS&T Archiving Conference*, vol. 5. (Bern: Koninklijke Bibliotheek (National Library of the Netherlands), 2008), pp. 162–165

doi:10.1186/1687-6180-2012-109

Cite this article as: Liu et al.: A new quality assessment and improvement system for print media. *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:109.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
