

VISUAL EXPERIENCE OF NOISE IN DIGITAL IMAGES

Ivana Jurič, Uroš Nedeljković, Dragoljub Novaković, Ivan Pinčjer

Original scientific paper

Noise in the image is always present, the only question is, to what extent it is visible to the human visual system. Some images could have a significantly high level of noise, but the human eye will not perceive that. The aim of this study was to investigate the influence of image content on the perception of image noise. We used three images with different contents (highlights, many details, and uniform surfaces with midtones). Samples were simulated by adding six levels of noise in MATLAB software, using the built-in function. Those samples were measured with image quality IQ metrics: SCIELab, SSIM and IQM2. For psychophysical scaling, we used the paired-comparison method (*PC-index*). The results obtained in this study showed that image content influenced the perception of image noise. The image with many details was the most difficult to estimate. Those paired samples looked the same, even though they had different levels of noise.

Keywords: *image quality metrics; noise; pictorial images*

Vizualni doživljaj buke u digitalnim slikama

Izvorni znanstveni članak

Buka na slici uvijek postoji, pitanje je samo koliko je vidljiva ljudskom oku. Na nekim slikama postoji značajno visoka razina buke, ali ljudsko oko to neće primijetiti. Cilj ovoga rada je ispitati utjecaj sadržaja slike na percepciju buke na slici. Upotrijebili smo tri slike različitog sadržaja (najsvidjetlija mjesta, mnogo detalja, jednolike površine blagih boja). Uzorci su simulirani dodavanjem šest razina buke u MATLAB softveru, primjenom ugrađene funkcije. Uzorci su mjereni IQ metrikom kvalitete slike: SCIELab, SSIM i IQM2. Za određivanje psihofizičke razine primijenili smo metodu komparacije para (paired-comparison method - *PC-index*). Rezultati su pokazali da je sadržaj slike utjecao na percepciju buke na slici. Najteže je bilo procijeniti sliku s mnogo detalja. Takvi uspoređeni uzorci izgledali su isto iako su imali različitu razinu buke.

Ključne riječi: *buka; metrika kvalitete slike; umjetničke slike*

1 Introduction

Many people will evaluate the quality in the way such as: "I know good quality when I see it". This subjective approach is neither helpful nor acceptable in the engineering and manufacturing environment. However, without human evaluation, we would not know whether the measured values correspond to the human visual system. Hence, both subjective (psychophysical) and objective (physical) methods should be included in defining quality and setting boundaries and categories of quality. Speaking about the quality, we should mention the quality attributes (QA). In addition to colour as the essential attribute of quality assessment, there are many others. Pedersen et al. [1] proposed six groups of QAs for the evaluation of print and image quality: colour, lightness, contrast, sharpness, artifacts (noise, contouring, banding) and physical parameters (paper properties, gloss).

In this paper, we focused only on one image quality attribute: image noise. Image noise is equivalent to film grain in analog cameras. Noise is a crucial parameter for image quality evaluation. It occurs as a random signal variation for each pixel [2].

According to the ISO 15739, noise is defined as "unwanted variations in the response of an imaging system" [3]. Noise in the image can be ranged from almost imperceptible specks to entirely noisy image. However, some degree of noise is always present in any electronic device that transmits or receives a "signal" [4].

Noise has many sources [5]. Despite the constant improvement of the imaging digital technology, noise cannot entirely be kept away because it has an inherent nature and statistical, random characteristic [6]. Sensitivity setting in the camera, length of the exposure, temperature, and different camera models could increase noise in the image. An example taken at various ISO sensitivity with the same camera is shown in Fig. 1.

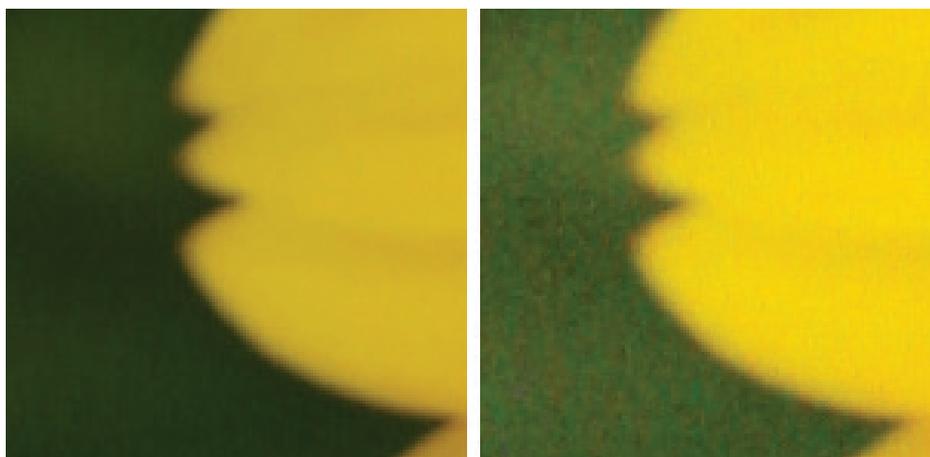


Figure 1 Example of random noise. The left image was shot at 100 ISO, the right part at 1600 ISO.

We can conclude that noise in the image is always present, but the main question is: "How much noise does the observer actually see in the picture?" What affects the perception of image noise? The image details, the illumination or the colour of the target? In order to answer some of these questions, this paper investigates the impact of image content on image noise. Since we used pictorial images in this study, the quality of samples was measured using image quality metrics (IQ metrics). Pedersen [7] defines IQ metric as an "objective mathematical way to calculate the quality without human observers". There are fundamentally three different types of IQ metrics: no-

reference (also referred to as blind-reference), reduced-reference, and full-reference [7].

The algorithm for almost all image metrics is the same, as shown in Fig. 2. In the first step, there are images (usually original and reproduction, or only reproduction), that need to be transformed to simulate the human visual system (HSV). The next step includes calculation of the quality, usually for every pixel in the image. After calculation, some metrics also involve pooling of obtained values. The final results are values that define quality of the image.



Figure 2 A general flow-chart of objective IQ assessment methods

In a report by Pedersen and Hardeberg more than 100 image quality metrics have been reviewed [8]. Some of the image quality metrics are based on the Human Visual System, and some are not, and also some of the image quality metrics are suitable just for digital images. In this paper, we used three IQ metrics: SCIELab [9], SSIM [10] and IQM2 [11].

In the study [6] it was found that the SCIELab metric quantifies noise based on visual perception and shows satisfying results. They also stated that the visual noise measurement model can only evaluate uniform colour patches in images while the S-CIELab model can be used for images with spatial content as well [6]. This metric is fully described in [12].

In the study [13] was investigated the influence of image quality attributes on the ability to use IQ metrics. It was found that SSIM (The Structural SIMilarity Index) metric is suitable for image quality artefacts. The SSIM index proposed and described in the study [14] attempts to quantify the visible difference between the distorted image and a reference image. The similarity between two

images can be measured with this metric. It is a quality measure that compares two images, regarding one as perfect quality. This metric is an improved version of the universal image quality index [15].

One of the latest IQ metric suggested by [11] is a full-reference objective image quality measure - IQM2, based on structural similarity index and steerable pyramid wavelet transform. This metric was compared with twelve commonly used full-reference objective measures. Results showed that the proposed IQM2 measure provides good correlation with the results of subjective evaluation [11].

2 Materials and methods

The experiment was conducted in two phases. In the first phase, we produced the samples with varying degrees of damage (different levels of noise) and in the second phase human observers classified those samples. The first phase also included quantification of reproduced samples using image quality metrics: SCIELab, SSIM and IQM2.

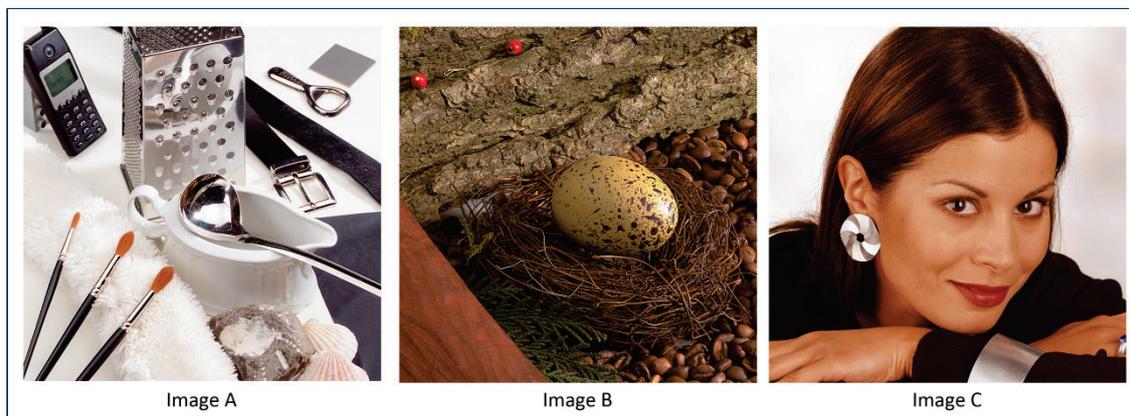


Figure 3 Pictorial images used in the experiment

Phase 1. In order to simulate different degrees of damage to the pictures, we applied different levels of noise to them. We used built-in function in MATLAB software, Gaussian white noise with constant mean and different variance ($v_1 = 0,001$; $v_2 = 0,003$; $v_3 = 0,005$; v_4

$= 0,007$; $v_5 = 0,009$; $v_6 = 0,01$) [16]. Standard pictorial images (Altona test suite) shown in Fig. 3 were used, in order to determine whether the content of the picture affects the perception of noise. These particular images were chosen because they contain the full range of

characteristics such as highlights (Image A), many details (Image B) and continuous mid-tones (Image C). Using three images with different content and six levels

of noise, we had a total of 18 samples. Further in the text, samples are marked as A1-A6, B1-B6 and C1-C6 depending on the image used.

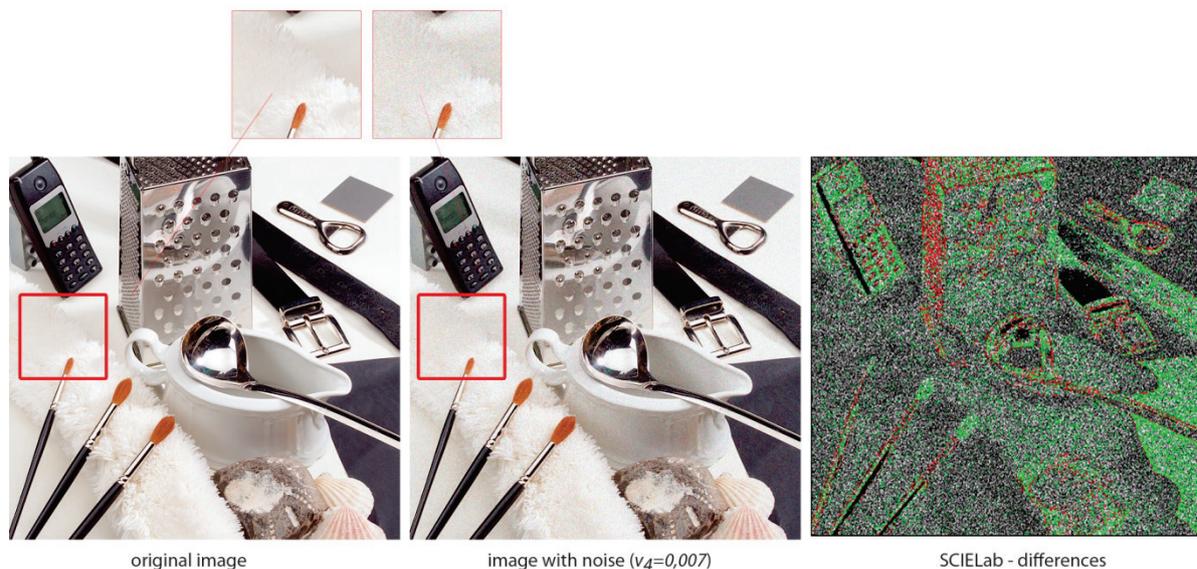


Figure 4 Comparison of a "noise-free" and a "noisy" image

In this first phase, we also quantified quality of samples using image quality metrics: SCIENLab, SSIM and IQM2.

Image quality metrics were performed in MATLAB software, in which samples were compared with the original image, without noise. In Fig. 4 is presented result obtained using SCIENLab image metric for Image A. Beside SCIENLab value we also got a picture of differences between two compared samples.

Phase 2. The 21 samples were used in a psychophysical experiment. Beside 18 samples used for measurement, in this stage we included samples without noise (three original images).

Participants:

29 participants (16 male and 13 female) estimated quality (noise) of samples. All of them had normal or corrected to normal visual acuity and normal colour vision. The participants were students of graphic engineering, but they did not have any prior knowledge and experience in this field.

Apparatus and stimuli for visual evaluation:

Samples (600×600 px), were presented on the calibrated monitor (1920×1080 px) opposite the observer. Brightness of monitor (107 cd/m²) and temperature (6978 K) were measured with i1 spectrophotometer in software BabelColorCT&A. The samples were randomly ordered for each participant. Observers were positioned at viewing distance (40 cm from the monitor). Experiment was carried out in a dark room eliminating the influence of ambient illumination.

Procedure:

The samples were set before the participant's arrival. The participants were asked to place their head on a chin rest, in order to maintain a constant distance from the samples during the experiment. Then the task was described to them also allowing for a couple of minutes to adapt to viewing conditions in the room.

We used the paired-comparison method because the relatively small number of samples made this method most suitable. According to this method, samples were presented to the observers in pairs. The observer task was then to select the preferred sample, relative to the noise level. This procedure was repeated for all possible combinations of sample pairs (63 in total).

The preferred sample (with lower level of perceived noise) got the score 2, and the other sample got 0. If they could not distinguish the samples, both got the score 1. When all observers have performed an evaluation, the *PC-index* (Paired-comparison index) was calculated. The *PC-index* could be a number from 0 to 200. High values correspond to samples perceived as having better quality while low values correspond to samples seen as having a poorer quality (larger noise in the image). The *PC-index* was calculated according to Eq. (1) [17].

$$PC - index = \frac{\sum_{i=1}^n v_i}{2n \cdot (m - 1)} \cdot 200, \quad (1)$$

where n is the number of observers, m is the number of samples and v_i is the value given by the observer i .

Paired-comparison is an expert based judgment estimation technique. The method reduces individual decision errors by requiring multiple pair-wise comparisons of relative sizes. Observations made in an academic environment showed that the paired-comparison produces more consistent results across estimators compared to ad-hoc and planning poker estimation techniques [18].

3 Results and discussions

In Tab. 1 are summarized results obtained using IQ metrics. All samples were compared with original noise-

free image. Lower values of SCIELab and higher values of SSIM and IQM2 corresponded to better image quality, i.e. to a lower degree of distortion. According to the results, we can notice that level of distortion increases linearly from samples 1 to 6, regardless of the content of the image (A, B or C). After physical measure of samples, we wanted to establish how humans perceive noise in pictorial images. Visual assessment was conducted as described in Materials and Methods section. As a result, we got *PC-index* values for each sample. Higher value of *PC-index* relates to better quality, lower noise in the

image. Results are presented in Fig. 5. Participants did not rate the samples as it was obtained by measuring them. Depending on the image content, the samples were evaluated differently. For participants, Image C was the easiest for recognizing any difference in noise level between samples. Image C does not have many details and has the most uniform surfaces (such as the uniform background and skin). Image with a lot of details (image B) was the most difficult for the evaluation of image noise. The content of the image was obstructing the perception.

Table 1 Results obtained using three IQ metrics: SCIELab, SSIM and IQM2

Image A	A1	A2	A3	A4	A5	A6
SCIELab	1,9922	3,4167	4,3927	5,1614	5,8457	6,1650
SSIM	0,9873	0,9660	0,9483	0,9324	0,9189	0,9119
IQM2	0,8682	0,7064	0,5924	0,5080	0,4410	0,4109
Image B	B1	B2	B3	B4	B5	B6
SCIELab	2,0887	3,6138	4,6367	5,4708	6,1736	6,4920
SSIM	0,9951	0,9858	0,9770	0,9685	0,9606	0,9565
IQM2	0,9374	0,8396	0,7589	0,6883	0,6299	0,6019
Image C	C1	C2	C3	C4	C5	C6
SCIELab	1,6487	2,9344	3,8916	4,7120	5,4176	5,7472
SSIM	0,9542	0,9016	0,8628	0,8306	0,8035	0,7905
IQM2	0,8330	0,6401	0,5144	0,4203	0,3528	0,3256

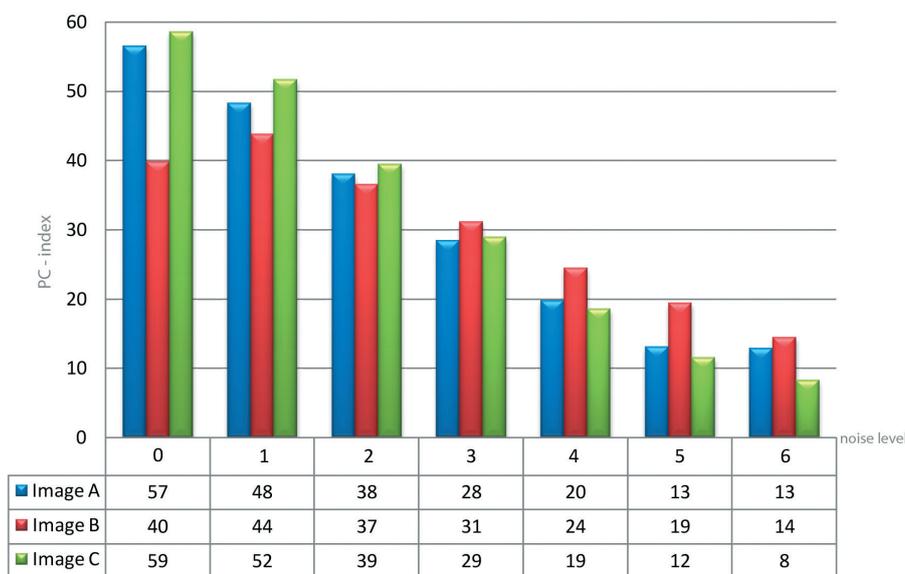


Figure 5 PC – index for all samples

Table 2 Pearson correlation coefficient (ρ) between measured values and visual estimation

Pearson's correlation coefficient (ρ)	SCIELab	SSIM	IQM2
<i>PC-index</i> for Image A	-0,98791	0,99642	0,99626
<i>PC-index</i> for Image B	-0,88963	0,97503	0,95622
<i>PC-index</i> for Image C	-0,98887	0,99597	0,98801

In order to compare physical and psychophysical measure, we calculated the Pearson's correlation coefficient presented in Tab. 2. Correlation is very high for all samples and all IQ metrics. The smallest coefficient was obtained between visual ranking for

Image B and SCIELab metric, what might be expected based on the *PC-index* for that image.

4 Conclusions

Through this research, we came to the conclusion that the noise is not affected only by settings and camera model, but it can also vary within an individual image and content of the image. In darker regions and parts with a lot of details, noise is harder to perceive (ρ is the smallest when comparing visual and measured values). Noise appears as random speckles, and it is more apparent on the smooth surfaces such as skin and uniform surfaces. Noise also becomes less pronounced as the tones become brighter. The regions with highlights have a stronger signal due to more light, so noise is less visible in those parts.

In this study, we used three images with different contents, at which was added the same levels of noise. Different noise level of the samples was confirmed by measurement using image quality metrics (SCIELab, SSIM and IQM2). However, participants did not perceive that difference in the same way as it was obtained by measurement. The content of the image influenced the visual assessment of the participants.

Acknowledgments

This work was supported by the Serbian Ministry of Science and Technological Development, Grant No.: 35027 "The development of software model for improvement of knowledge and production in graphic arts industry".

5 References

- [1] Pedersen, M.; Bonnier, N.; Hardeberg, J. Y.; Albrechtsen, F. Attributes of a New Image Quality Model for Color Prints. // 17th Color Imaging Conference Final Program and Proceedings / Albuquerque, New Mexico, 2009; p. 204-209
- [2] Artmann, U. Noise: SNR vs. visual noise.2013. <http://www.image-engineering.de/library-m/technotes/659-snr-vs-visual-noise> (02.01.2015.)
- [3] Kitoh, S; Hung, P. C. Development of an ISO 15739 Noise Measurement Tool for Digital Still Cameras. // Konita Minolta Technology Report, 2(2005)
- [4] Cambridge in Color. Digital Camera Image Noise- Part 1. <http://www.cambridgeincolour.com/tutorials/image-noise.htm> (2.1.2015)
- [5] Kidawa, N.; Simon, K. Perception of noise depending on spatial frequency and contrast with the aid of the CIE-Lab color space // University of Applied Sciences Cologne, July 2003.
- [6] Kleinmann, J.; Wuelle, D. Investigation of two Methods to Quantify Noise in Digital Images based on the Perception of the human Eye. // SPIE-IS&T, 6494(2007)
- [7] Pedersen, M. Image quality metrics for the evaluation of printing workflows. // PhD thesis. Faculty of Mathematics and Natural Sciences, University of Oslo, No. 1124, 2011
- [8] Pedersen, M.; Hardeberg, J. Y. Survey of full-reference image quality metrics. // Høgskolen I Gjøviksrapportserie, 2009 nr. 5, ISSN:1890-520X, 2009
- [9] Zhang, X. S-CIELAB: A Spatial Extension to the CIE L*a*b* DeltaE Color Difference Metric. // The Web version (1998). <http://white.stanford.edu/~brian/scielab/scielab.html> (20.12.2014)
- [10] Wang, Z.; Bovik, A. C.; Sheikh, H.R.; Simoncelli, E. P. The SSIM Index for Image Quality Assessment. The web version (2003). <https://ece.uwaterloo.ca/~z70wang/research/ssim/> (20.12.2014)
- [11] Dumic, E.; Grgic, S; Grgic, M. IQM2: new image quality measure based on steerable pyramid wavelet transform and structural similarity index. // SIViP, 8, 6(2014), pp. 1159-1168. DOI: 10.1007/s11760-014-0654-3
- [12] Johnson, G. M. Measuring images: differences, quality and appearance. // PhD thesis, Chester F. Carlson Center for Imaging Science of the College of Science, Rochester Institute of Technology, 2003
- [13] Pedersen, M.; Zheng, Y; Hardeberg, J. Y. Evaluation of Image Quality Metrics for Color Prints. // Lecture Notes in Computer Science, 6688(2011), pp. 317-326. DOI: 10.1007/978-3-642-21227-7_30
- [14] Wang, Z.; Bovik, A. C.; Sheikh, A. C.; Simoncelli, E. P. Image quality assessment: From error visibility to structural similarity. // IEEE Transactions on Image Processing, 13, 4(2004), pp. 600-612. DOI: 10.1109/TIP.2003.819861
- [15] Wang, Z. Universal Image Quality Index. // The web version (2001). https://ece.uwaterloo.ca/~z70wang/research/quality_index/demo.html (22.12.2014)
- [16] The MathWorks. Imnosie function in Matlab. <http://www.mathworks.com/help/images/ref/imnosie.html> (18.12.2014)
- [17] Norberg, O. Paper whiteness and its effect on the reproduction of colors. // Proceedings of the SPIE, Human Vision and Electronic Imaging XII, 6492, Sweden, 2007. DOI: 10.1117/12.703013
- [18] Miranda, E. An Evaluation of the Paired Comparisons Method for Software Sizing. // Proceedings of the 22nd International Conference on Software Engineering, 2000. DOI: 10.1145/337180.337477

Authors' addresses

Ivana Jurič, MSc

University of Novi Sad, Faculty of Technical Sciences,
Department of Graphic Engineering and Design,
Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia
Phone: +38121 485 2592
E-mail: rilovska@uns.ac.rs

Uroš Nedeljković, Mr

University of Novi Sad, Faculty of Technical Sciences,
Department of Graphic Engineering and Design,
Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia
Phone: +38121 485 2624
E-mail: urosned@uns.ac.rs

Dragoljub Novaković, PhD

University of Novi Sad, Faculty of Technical Sciences,
Department of Graphic Engineering and Design,
Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia
Phone: +38121 485 2620
novakd@uns.ac.rs

Ivan Pinčjer, Mr

University of Novi Sad, Faculty of Technical Sciences,
Department of Graphic Engineering and Design,
Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia
Phone: +38121 485 2624
E-mail: pintier@uns.ac.rs