A Reduced-Reference Method for Characterizing Color Noise in Natural Images Captured by Digital Cameras

Mikko Nuutinen[1], Olli Orenius[2], Timo Säämänen[2], Pirkko Oittinen[1]; [1]Department of Media Technology, Aalto University; Espoo, Finland; [2]Department of Psychology, Helsinki University; Helsinki, Finland

Abstract

Noise is one fundamental quality attribute in digital cameras. Traditionally, noise has been measured from solid patches of artificial test targets. In image quality research, it has been difficult to find connections between a test target and subjective test data. In addition, image quality algorithms computed from natural images are not well correlated. In this paper, we propose a novel approach for measuring color noise from natural images. With the proposed method, the suitable surfaces for noise calculations are located from the scene using a reference camera image. It is now possible to use the same image files for subjective and objective measurements and correlations are easier to find. The results show that the method is promising. Its performance was better for predicting subjective noise compared to the visual noise metric that is the state-of-the-art test target method for digital cameras.

Introduction

Digital cameras produce different noise types in images. For example, noise can be high-frequency achromatic noise, low-frequency red-green or yellow-blue color noise or a combination of both. In this study, a new method to measure and characterize color noise directly from a natural image is described. The proposed method is based on a reference camera. The reference camera shoots a natural scene, and the appropriate areas are identified from the image for measurement purposes. The method has been developed for camera benchmarking studies. The method requires that the images of the reference camera and the cameras to be benchmarked are produced from the same scene.

The study of color noise in the literature can be divided in two distinct areas. In the first area, the goal is to describe the noise model and weighting factors of its chrominance components. In these studies, noise level has often been measured from solid patches of specific test targets. For example, Kuang et al. [2] fit the parameters of the noise model based on empirical data. They also implemented a function incorporated in the noise model that described the effect of luminance level. In another study, Kelly and Keelan [3] described new weighting factors of the chrominance component for the signal-to-noise ratio calculation.

In the second area of color noise study, the goal is to find the noise level or noisy areas from natural images for noise reduction purposes. Gheorghe et al. [5] proposed a method to reduce color noise from a natural image. Their method was based on a hybrid multi-scale spatial dual tree adaptive wavelet filter in hue-saturation-value color space. Lee [4] proposed a method to detect color noise areas from natural images. His method was based on correlation between the R/G/B color channels. In addition, a noise metric for luminance channel has been proposed [10]. These methods are based on the no-reference (NR) approach. The measurements are performed without the original noiseless images. The problem with using NR methods with digital cameras is that these methods are often sensitive to other image distortions. For example, NR noise metrics can interpret image details as noise energy. In addition, NR metrics are often highly image-content specific.

The proposed method differs from the earlier methods discussed in the literature. The method is based on the reduced-reference (RR) approach. It utilizes information from a reference image, but it does not need a pixel-wise equivalence as the full-reference (FR) approach does. Pixel-wise comparison is not even possible when digital cameras are benchmarked. When images are produced from a given scene using different digital cameras, there is always rotation, scaling and 3D-projection between the images.

We can find an analogy between the test target method and the proposed method. With the test target method the properties of the solid patches are known. With the proposed method, the suitable areas are located for measurements from the scene using a reference camera image. The selection is based on distortion. In this study, we describe how surfaces for noise calculations can be selected. In addition, we show how noise type can be characterized and noise level can be measured from these surfaces.

The benefit of the proposed method compared to test target methods is that the same images can be used for subjective and objective measurements. It has been difficult to find correlations between test target computations and subjective test data such as MOS using conventional image quality research methods. We believe that these relationships are easier to find if both measurements are made using the same natural images. Compared to NR methods, the benefit of the proposed method is that at least some features from the reference (noiseless) image and scene are known. With these features, the problems related to the other image distortions and image content can be avoided.

Method

The proposed method is based on blocks that are located from the scene using a reference camera image. The block selection is based on three features: chromatic energy, achromatic energy and brightness of the block. The chromatic energy of the blocks should be low. The blocks can have achromatic structural energy, but this structure should be composed more of random texture than edges. There are two reasons why random texture in a scene can be beneficial for noise measurements. The first reason is that achromatic texture-like surfaces in scenes are sensitive to color noise in digital camera images. The second and more important reason is that texture-like surfaces present challenges for noise reduction algorithms in cameras. If the structure is edge-like, then a noise reduction method can easily filter the noise away from the neighboring smooth area of the edges. If the structure is a random texture, then it is difficult to separate the noise energy from the image structure energy using computational methods.

In addition, the intensity of the selected blocks should not be too low or too high. If a block is too bright, then it becomes saturated for images produced by low-end cameras. If the block is too dark, then it is possible that a low-end camera does not detect its structure energy and that the camera image processing software applies strong noise reduction to it.

The method was applied in the YCbCr space. With opponent color space, it is possible to separate achromatic information from chromatic information. The method operates on the principle that the control blocks are initially symmetrically located on reference image (Figure 1a). The method searches for new locations for the blocks on the limited neighborhood in Cb and Cr channels by maximizing the homogeneous metric value. Figure 1b shows the blocks that are located on the new places for the Cb channel. The homogeneous metric used was the co-occurrence matrix energy feature CO_E of the blocks calculated by Equation (1):

$$CO_E(\phi, d) = \sum_{a,b} P^2_{\phi,d}(a,b) \tag{1}$$

where $P_{\phi,d}(a,b)$ describes the probability that two pixels with intensity levels a,b appear in the window separated by a distance d in direction ϕ [1]. Feature CO_E gets a higher value in scale 0 - 1 if the block is homogenous. The homogenous metric was calculated as an average of directions in the 0, 45 and 90 degrees. In this study, the distance d was 30 pixels. The image size was 1600 x 1200 pixels and the block size was 100 x 100 pixels in all cases. The limited neighborhood area used for minimizing block homogenous values was 120 pixels for all directions measured from original block center point to the new block center point.

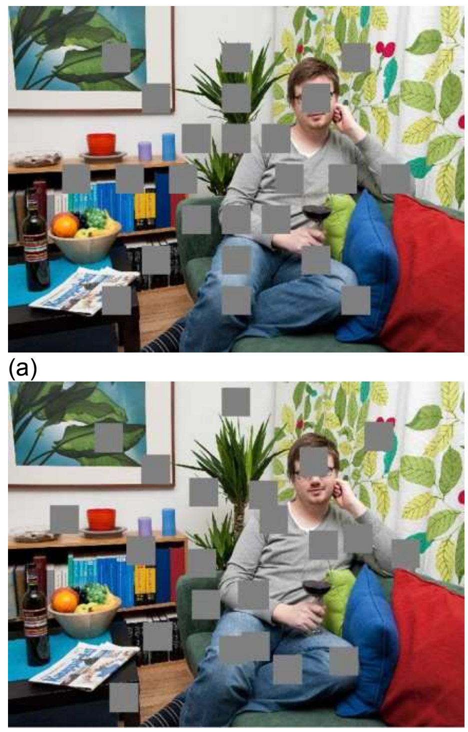
Next the energy values of the Y channel were calculated for homogenous blocks. The used energy metric was the co-occurrence matrix inverse difference moment feature CO_{IDM} of the blocks calculated by Equation (2):

$$CO_{idm}(\phi, d) = \sum_{a,b;a\neq b} \frac{P^{\lambda}_{\phi,d}(a,b)}{|a-b|^{\kappa}} \tag{2}$$

where constants λ and κ were set to 1 and 2 [1]. The metric was calculated as an average of directions in the 0, 45 and 90 degrees. In this study, the distance d was 10 pixels. Feature CO_{IDM} gets a higher value in scale 0 - 1 if the block pixel intensity values are close to each other. This means that if the CO_{IDM} value is small the intensity structure in the block is more texture like than smooth. The brightness value was calculated in addition to the homogenous and energy values. The brightness value was the mean intensity of the Y channel. CO_E and CO_{IDM} co-occurrence features were chosen because they are well-known and widely used texture descriptions.

In this study, six blocks were selected as candidate blocks for color noise calculation based on the CO_E values. The blocks with the lowest and highest CO_{IDM} values from the six candidate blocks with brightness values between 100 and 200 levels were then selected for the analyses.

Once the blocks have been selected from the scene for analysis, corresponding blocks should be searched from images produced by the cameras to be benchmarked. With camera images, there are differences in features such as rotation, scaling, 3D-projection and brightness. Noise levels and types also naturally differ. We expect that correspondence block searching using only pixel coordinates in images is not precise enough. We also expect that searching using the block correlation method lacks the necessary precision. The proposed method applies the popular SIFT method (Scale Invariant Feature Transform) developed for object recognition [7]. SIFT-based methods are invariant to scaling, translation and rotation. In addition, it is partially invariant to brightness changes and 3D-projection.



(a)

(b)

Figure 1. Reference image with blocks in predetermined symmetric order (a), Reference image with blocks when homogeneous metric was maximized for the Cb channel (b).

The method finally calculates noise values from the blocks located and cropped from the distorted images. Noise type was characterized using power spectra and noise level was measured using wavelet transformation. Power spectra were calculated as rotational average of intensities after discrete Fourier transform.

For measuring the visually perceived noise level, the metric should use different weighting factors for achromatic and chromatic noises. The spatial frequency of noise energy should also be taken into account. For example, Johnson and Fairchild [6] studied the effect of frequency and noise color channel on image quality. This empirical study showed that the distracting level depended on the noise frequency, and it resembles the human visual contrast sensitivity. For example, luminance noise was more distracting at higher frequencies than chrominance noise.

The noise model used in this study combines both achromatic and chromatic components of noise linearly as follows:

$$N = k_1 c_y + k_2 c_{cb} + k_3 c_{cr} \tag{3}$$

where k_i are weighting factors and c_y, c_{cb} and c_{cr} are mean standard deviations of the vertical, horizontal and diagonal wavelet coefficients for the blocks. The achromatic noise energy was calculated using higher frequencies than those for chrominance noise. Wavelet coefficients for chrominance channels were calculated from the second scale and for achromatic channels from the first scale of wavelet decomposition. Wavelet decomposition was done by the Matlab wavelet toolbox using Haar wavelets. Figure 2 shows image decomposition and scales and orientations from where noise energy is extracted for different components of the noise model.

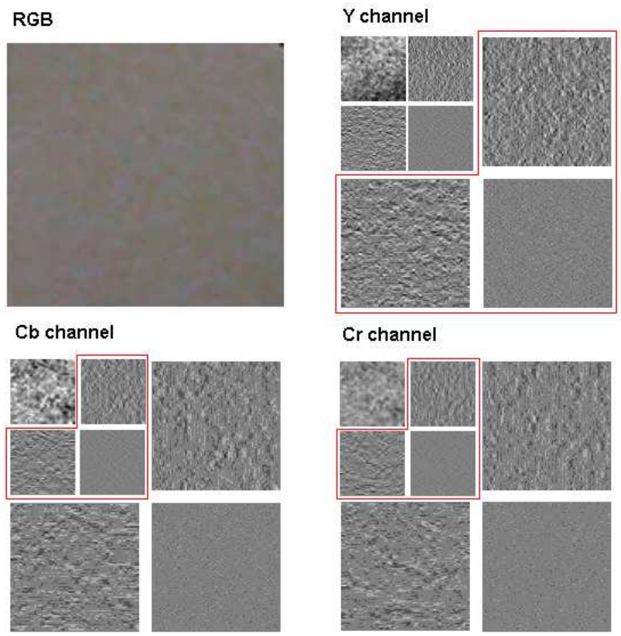


Figure 2. Noise model uses first scale coefficients for achromatic Y component and second scale coefficients for chromatic Cb and Cr components (wavelet coefficients are scaled for visualization purposes)

The next section describes how noise values are calculated from the images produced by 13 cameras. Twelve cameras were mobile phone cameras, and one was a digital still camera. The scene structure used in the study was a normal portrait under a low illumination level. The illumination level was 100 lux, and the color temperature was 2800 K. These values equal the values of a living room environment. Table 1 lists the indices and pixel count of the cameras benchmarked. The reference camera was Canon 5D with Canon EF 24-70/2.8 L USM lens. Image processing of the reference image was done by Canon Digital Photo Professional software. The cameras to be benchmarked were set to the automatic mode.

Table 1. Indices and pixel counts of cameras used in the study

Camera index	Camera type	Mpix
ref	SLR	12.00
1-3	mobile	3.10
4	mobile	4.90
5-9	mobile	5.00
10	mobile	8.00
11-12	mobile	12.00
13	DSC	10.00

Block selection for noise metric

Figure 1a shows the reference image with blocks in predetermined symmetric order. Figure 1b shows the new locations of the blocks after the homogeneous metric was maximized for the Cb channel. Figure 3a shows the six most homogenous blocks in the Cb channel, and Figure 3b shows the six most homogenous blocks in the Cr channel. The six most homogenous blocks are indexed in Figures for analysis purposes.

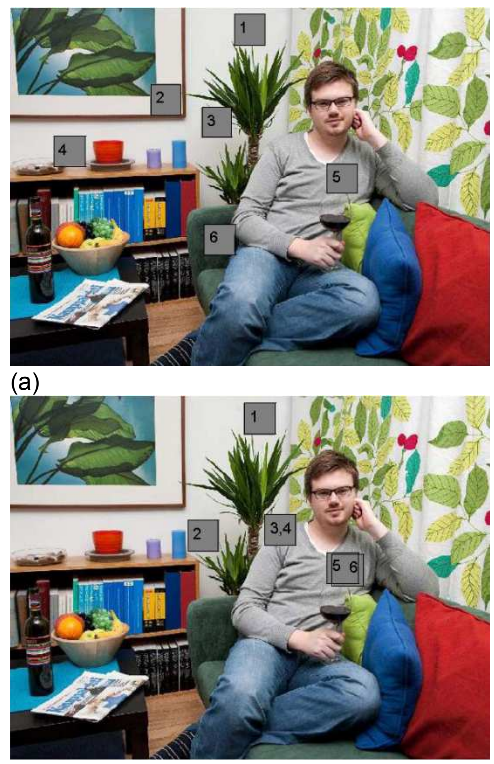


Figure 3. Reference image with the six most homogeneous blocks in the Cb (a) and in the Cr (b) channels.

Table 2 lists the values of the CO_E and CO_{IDM} features and brightness for the six most homogenous blocks in the Cb and Cr channels. The blocks that we selected for noise measurements had low or high achromatic energy. With high achromatic energy blocks (low CO_{IDM} value), we wanted to study color noise in the texture region. With low achromatic energy blocks (high CO_{IDM} value), we wanted to study color noise in the smooth region. Another constraint is that the brightness of the block should be in the 100 – 200 scale. Based on these criteria, blocks 2 and 5 were selected from both the Cb and Cr channels for subsequent analyses. In both channels, the values of the CO_{IDM} feature were the highest in block 2 and the lowest in block 5 for brightness values between 100 and 200.

Table 2. Characterization values for the six most homogeneous blocks of the Cb and Cr channels (selected blocks marked in bold)

	Cb channel			Cr channel		
Block number	CO_{IDM} (Y)	CO_E (Cb)	Brightness (Y)	CO_{IDM} (Y)	CO_E (Cr)	Brightness (Y)
1	0.637	0.968	210	0.665	0.994	210
2	**0.564**	**0.217**	**199**	**0.371**	**0.538**	**194**
3	0.366	0.696	194	0.173	0.383	183
4	0.592	0.262	201	0.172	0.382	182
5	**0.117**	**0.301**	**174**	**0.113**	**0.462**	**174**
6	0.129	0.294	55	0.117	0.469	172

The corresponding blocks of the reference and distorted images were searched for using the SIFT algorithm [8]. The SIFT algorithm was applied so that the 10 nearest correspondence features of blocks in the reference image were used for correspondence block searching in the distorted images. The block center in the distorted image was located by calculating the angle and length information of the vectors between the block center and features in reference image. Figure 4 shows an example where 10 correspondence features near Block 5 are used for the correspondence block search in a distorted image. The block center in the distorted image is the average of the vector end-points from the correspondence features.

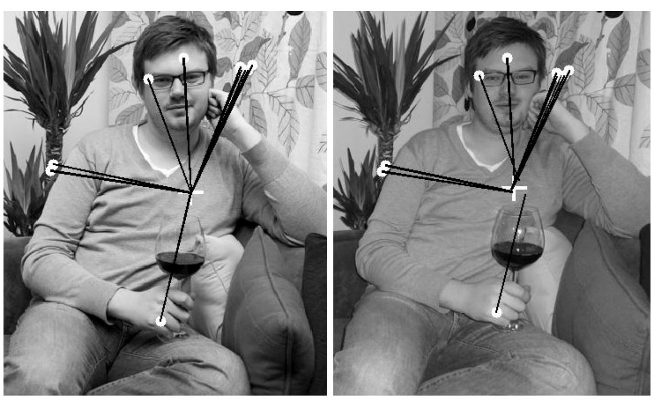


Figure 4. Correspondence features (circles) between the reference image (left) and distorted image (right) are located using the SIFT algorithm. Correspondence block centers (crosses) are located for distorted images using average end-point values of vectors derived from reference image.

Method validation

We used subjective data to confirm that the proposed method is suitable for measuring color noise of natural images captured by digital cameras. The data set utilized in this study is a portion of a large-scale subjective test set. University students were used as observers (n=25). They were all naïve regarding image quality. The test images were scaled to a 1600 x 1200 pixel size. In addition, black borders were added to images to reconcile image file resolution to the display resolution (1920 x 1200). The test setup included two Eizo ColorEdge CG241W displays. The test image was shown on one display, and a high quality reference image was shown on the other. The observers evaluated picture graininess level on a 0 to 100 scale, where 0 was the worst and 100 was the best score.

The prediction accuracy of the proposed metric was compared to the visual noise metric [9]. The visual noise is a state of the art test target metric. In this study, the visual noise was measured using the neutral patches of the Gretag Macbeth test target. The illuminance of the test target was 100 lux, and the color temperature was 3200 K. Parameters for the visual noise was set for the subjective test environment: 100 % view on display, 0.8 m viewing distance and 94 ppi resolution. We used IE Analyzer v4.0.5 software for calculations, and reported values were an average of 10 images. The proposed metric was calculated from a single image that was also used in the subjective tests.

Characterization Results

One novel feature of the proposed method is the possibility of characterizing color noise of achromatic texture regions. Table 3 shows intensity distributions of the RGB, Cb and Cr channels of texture regions (Block 5) from Cameras 1 – 13 and the reference camera. The intensity of the Cb and Cr block images is scaled for visualization purposes. The Cb and Cr blocks of the reference camera are smooth, but there is a texture-like intensity structure in the achromatic channel. However, the Cb and Cr blocks of distorted cameras have intensity distributions at different frequencies. Power spectra were used next to characterize the color noise types.

Figure 5 shows the power spectra of the Cb channel for Cameras 1 – 3. Figure 6 shows the power spectra of the Cr channel for Cameras 4 – 9. The power spectrum of the reference camera is shown in both Figures 5 and 6. Cameras 1 – 3 are low-end products, and Cameras 4 – 9 are moderate-level products. Figure 5 shows that Camera 1 produces less noise in the Cb channel than Cameras 2 and 3. The power spectrum of Camera 1 is closer to the reference camera than the power spectra of Cameras 2 and 3. However, a visual inspection shows that the noise energy of Camera 1 is not random (see Table 3). The power spectrum values come from the color blotches. Color blotches can be visually more distracting than random Gaussian type noise. Figure 6 indicates that the noise level of Camera 9 was low compared to the other cameras. The noise level for Camera 4 was also low.

Figure 7 compares noise types between Cameras 8 and 12. Camera 8 produces strong noise in the Cb channel compared to the Cr channel. The Cb channel power spectrum has a peak in the medium frequency area. Camera 12 produces high frequency noise energy for both channels.

Table 3. Intensity distribution for the texture blocks

Camera	RGB	Cb	Cr
Ref			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			

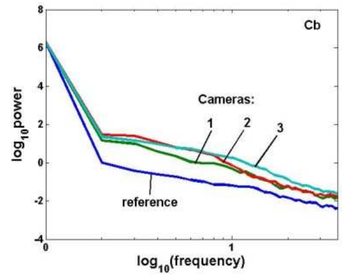


Figure 5. Power spectrum of block 5 for Cameras 1 – 3 in the Cb channel.

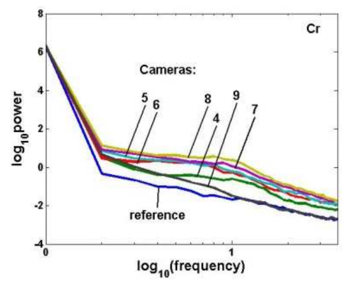


Figure 6. Power spectrum of block 5 for Cameras 4 – 9 in the Cr channel.

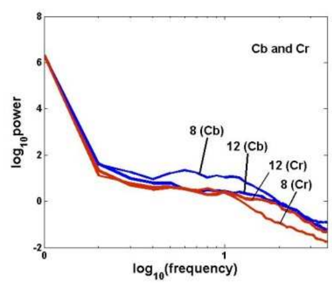


Figure 7. Power spectra for Cameras 8 and 12 in the Cb and Cr channels.

Prediction Results

Power spectra can provide valuable information about noise properties. However, power spectra do not give a single number for noise level related to the perceived image quality. Instead, we used the noise model expressed in Equation (3) to measure the noise level of benchmarked cameras. The metric values were calculated for blocks 2 and 5. Block 2 of the Cb and Cr channels were used for smooth area noise component, and block 5 was used for texture area noise component. Total noise was calculated as the sum of both blocks. Single component values were always normalized before summation because of the scale differences between components.

The noise model in Equation (3) includes weighting factors for chromatic and achromatic components that should be specified. In this study factors were specified based on the empirical data. Weighting factors k_2 and k_3 for chromatic components were set to 5. Weighting factor for achromatic component was set to 1 for block 2 and 0 for block 5. The weighting factor of 0 was selected for block 5 because the block in the reference image had achromatic structure energy. This factor made it possible to avoid interpreting the structure energy as noise energy.

Figures 8 and 9 show scatter plots for the proposed total noise metric and visual noise test target metric. Linear and rank ordered correlation coefficients between subjective and objective data are shown in Table 4. The results demonstrate that the proposed method is a promising approach for the measurement of visually perceived noise levels from natural images. The performance of the texture noise component was only moderate, but performances of the smooth area component and total noise were rather high.

The scatter plot of the visual noise includes two or three clear outliers that decrease the linear correlation coefficient. However, the rank ordered correlation coefficient of visual noise was higher. The scatter plot of the total noise includes one clear outlier. Prediction of the total noise level was too low for Camera 1. The noise energy of Camera 1 was not random, but more like color blotches that are visually distracting. The proposed noise metric needs to be fine-tuned to find and handle these types of distractions.

Conclusions

The proposed method is a promising approach to color noise characterization. This study shows that the proposed method is better than test target methods when the reasons behind the subjective studies are analyzed because now both subjective and objective measurements can be done from same natural images. The proposed method is better than NR methods because it takes into account image content and other types of image distortion for calculations.

Future efforts should be directed at further developing the method. For example, this study utilized constant block sizes for noise analysis. The block size could vary based on the energy structure and properties of the neighborhood.

Only one scene was used for the analyses in this study. The method should be applied for different scenes. The scene-specific threshold values of structural energy or brightness for blocks should be studied in future work.

Table 4. Linear correlation and rank ordered correlation coefficients between subjective noise and objective metrics

Metric	Linear correlation, LCC	Rank ordered correlation, ROCC
Proposed texture	-0.635	-0.676
Proposed smooth	**-0.837**	-0.775
Proposed total	-0.800	**-0.786**
Visual noise	-0.471	-0.709

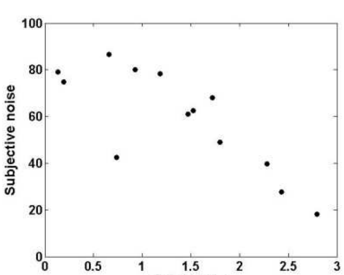


Figure 8. Total noise as a function of subjective noise.

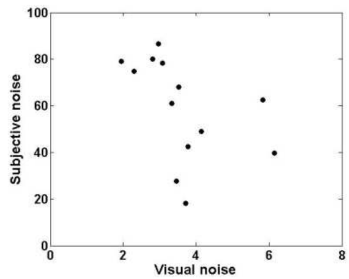


Figure 9. Visual noise as a function of subjective noise.

Acknowledgements

This work was partially financed by Nokia Mobile Solutions / Symbian Smartphones. The authors thank Fredrik Hollsten and Jussi Turunen for the test images.

References

[1] M. Sonka, V. Hlavac, R. Boyle, "Image Processing, Analysis, and Machine vision" Brooks/Cole Publisihing Company, 1999, 2nd. ed. 770 p.

[2] J. Kuang, X. Jiang, S. Quan, A. Chiu, "Perceptual Color Noise Formulation", Proc. SPIE vol. 5668, p. 90-97 (2005).

[3] S.C. Kelly, B.W. Keelan, "ISO 12232 revision: determination of chrominace noise weights for noise-based ISO calculation", Proc. SPIE vol. 5668, p. 139-147 (2005).

[4] S. Lee, "Blind Cross Color Noise Detection and Reduction", IEEE Transaction on Consumer Electronics, vol. 54, nro. 4, p. 1825-1829 (2008).

[5] R. Gheorghe, S. Goma, M. Aleksic, "An image noise filter with emphasis on low frequency chrominance noise", Proc. SPIE vol. 7250, [72500B] (2009).

[6] G.M. Johnson, M.D. Fairchild, "The Effect of Opponent Noise on Image Quality", Proc. SPIE vol. 5668, p. 82-89 (2005).

[7] D.G. Lowe, "Object Recognition from Local Scale-Invariant Features", Proc. IEEE Int. Conf. of Computer Vision, p. 1150-1157 (1999).

[8] A. Vedaldi, "SIFT for Matlab", http://www.vlfeat.org/~vedaldi/code/sift.html, [april 6, 2010].

[9] ISO 15739:2003(E), Photography – Electronic still-picture cameras – Noise measurements. ISO 2003. 26 s.

[10] J. Immerkaer, "Fast Noise Variance estimation", Computer Vision and Image Understanding, vol. 64, nro. 2, p 300-302 (1996).

Author Biography

Mikko Nuutinen received a M. Sc. degree in automation and system technology in 2004 and a Lic. Sc. (tech) in 2007 from the Helsinki University of Technology. He is currently a Ph.D. student at the department of Media Technology at Aalto University School of Science and Technology. His research interests include color image processing, image device measurements and image quality measurements.