ANALYSIS OF NOISE IN MULTI-CAMERA SYSTEMS

Olgierd Stankiewicz, Marek Domański, Krzysztof Wegner

Poznań University of Technology, Chair of Multimedia Telecommunications and Microelectronics

ABSTRACT

This paper presents a methodology for analysis of noise in context of 3D video systems. The goal of such analysis is devoted to development of novel noise-aware algorithms and techniques. A simple yet valuable set of analytic tools is presented first. Then a novel proposal is shown, which employs disparity-based pixel correspondence and allows for comprehensive analysis of the noise and other phenomena in multi-camera systems. In both parts, experimental results attained for some existing modern 3D video systems are reported.

Index Terms **—** Noise analysis, noise-aware algorithms, pixel-correspondence, multiview.

1. INTRODUCTION

Nowadays, a variety of research is underway in context of multi-camera video systems for free-viewpoint television (FTV), three-dimensional television (3DTV) [1] and light-field [2]. In such systems, advanced processing algorithms are used for stereoscopic depth estimation [3], virtual view synthesis [4], super-resolution [5], light-field rendering [2], advance refocusing [6]. The operation of these algorithms is conditioned by presence of noise in the input video content. In formulation of such algorithms, typically, either the presence of noise is omitted [7] or the assumptions about the characteristics of noise are made only implicitly [8]. Even when presence of the noise in the video material is assumed explicitly, the most commonly solely Gaussian noise [9-11] is considered, often without any experimental verification [12].

This is very unfortunate, as in practical applications the devised algorithms work with very various video sequences, acquired with different cameras. Such exemplary test set of sequences is presented in Table 1. Those sequences are used by MPEG group for research [13] on standardization of new generation of 3D video coding technology.

If exact characteristics of the noise (e.g. spatial spectrum, temporal correlations, value distributions) are ignored, the performance of such algorithms may be degraded in unpredictable manner, which also lowers reproducibility of the research.

Table 1. Multiview video sequences acquired with various camera systems, which have been used for experimentation.

Sequence Name	Width x Height	Frame rate [frames/s]	Camera	Total frames	Views numbers	Depth data available for views
Poznan Carpark	1920 x 1088	25	Canon XH-G1, 3-CCD camera	250	0…8	3,4,5
Poznan Street						
Poznan Hall				200	0…8	5,6,7
Lovebird1	1024 x 768	30	Point Grey Flea camera (CCD), Moritex ML-0813 lenses	240	0..8	3,5,7
Newspaper			Point Grey Research Flea camera (CCD) with 1/3-inch Sony lenses	300	0…8	2,4,6
Balloons			XGA CMOS, 8-bit RGB-Bayer camera	300	0…6	1,3,5

Noise analysis and modelling itself also has gained a relatively small interest among researchers. In work [14], noise caused by usage of CMOS sensor is considered. Author analyze probability distribution of noise in form of Gaussian function but no experimental verification results are provided. In paper [15] it is proposed to analyze noise by acquiring a testing pattern, which comprises several patches of constant intensity values. Such methodology can be applied only if the camera system is available for testing and calibration and cannot be used in case when the working content is given as-is without any other specific camera system knowledge.

This work presents a simple methodology for noise extraction and verification of the assumptions about characteristics of noise in the multi-camera video material, without knowing any a priori information about the camera system. First, a simple set of useful analytic tools is shown which is concluded by original results for exemplary modern multi-camera video systems. Then, a novel analysis approach is shown, which uses disparity-based pixel-correspondence in multiview video and allows more robust observation of noise and other phenomena.

2. NOISE EXTRACTION

Analysis of noise characteristics can be done by noise extraction. Many algorithms known from literature could be used, but the method chosen for analysis should make as few assumptions about noise characteristics as possible. This is why a straightforward approach is used in this paper.

Therefore, extraction of the noise is done independently in each of the acquired videos. Only still, not moving regions of the scene are considered. As pixel values in those regions vary only due to noise, the average pixel values (denoted $Denoised(x,y)$) are used as denoised version of the image:

$$Denoised(x,y) = \frac{1}{N}\sum_{i=0}^{N-1} Frame_i(x,y) \quad , \qquad (1)$$

where $Frame_i(x,y)$ are pixel values (luminance and chrominance are processed independently) in frame i and given spatial coordinates.

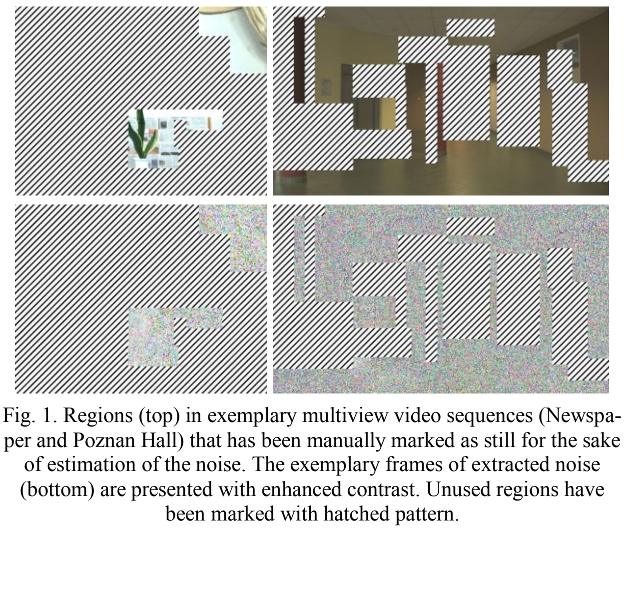


Fig. 1. Regions (top) in exemplary multiview video sequences (Newspaper and Poznan Hall) that has been manually marked as still for the sake of estimation of the noise. The exemplary frames of extracted noise (bottom) are presented with enhanced contrast. Unused regions have been marked with hatched pattern.

Sought noise is simply difference between the denoised and the original image:

$$Noise_i(x,y) = Frame_i(x,y) - Denoised(x,y) \quad . \qquad (2)$$

The selection of still, not moving regions has been done manually, with special care taken in order to represent the whole color and lightness dynamics of the scene (e.g. the selected content contains lit and dark regions) – Fig. 1 (top).

Exemplary noise frames are presented in Fig. 1 (bottom). The observed practical noise value range is $-8 \le Noise_i(x,y) \le 8$ and thus such a range has been chosen for further considerations.

3. NOISE INDEPENDENCY IN TIME DOMAIN

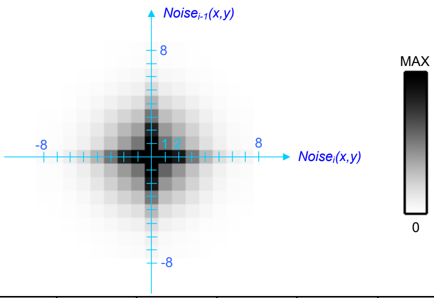
The usage of many techniques, including noise extraction techniques (also the mentioned above), require that the noise is statistically independent and not correlated in subsequent frames.

There are many known tests for statistical dependency and correlation, among which the strongest is statistical independency test, like chi-square. Thus chi-square has been chosen for experimentation in this paper. From a statistical point of view two variables (let's say α and β) are statistically independent if:

$$p(\alpha,\beta) = p(\alpha)\cdot p(\beta) \quad , \qquad (3)$$

Therefore, in order to evaluate whether $Noise_i(x,y)$ (pixel value of the noise in given pixel) is independent in consecutive frames $i-1$ and i, two-dimensional histograms of $Noise_{i-1}(x,y)$ vs. $Noise_i(x,y)$ have been measured.

An exemplary estimated histogram is presented in Fig. 2. If there would be any dependence between those two random variables (modeling extracted noise in $i-1$ and i, there would be an asymmetry in the graph, related to the fact that (3) is not meet.



Sequence	camera 0	camera 1	camera 2	camera 3	$\frac{\chi_{ind.}{}^2}{\chi_{ind.}{}^2{}_{critical}}$
Poznan Street					0.0145
Poznan Carpark					0.0249
Poznan Hall					0.0194
Lovebird1					0.0387
Newspaper					0.0269
Balloons					0.0307

Fig. 2. Calculated chi-square test values for tested sequences (right column in the table) and plots of two-dimensional histograms of $Noise_i(x,y)$ vs. $Noise_{i-1}(x,y)$ (zoomed example above).

Basing on the histogram analysis results (visualized in Fig. 2) chi-square independence test has been performed. The working null hypothesis is that the observed distributions are dependent (not independent). The working alternative hypothesis is that the observed distributions are independent.

The considered were bins representing integer noise values in range [-8;8] (and thus there were total $\varphi = 17$ bins). Basing on that, and the fact that the expected distribution (left side of Eq. 3) has been estimated from empirical data (it is not known from a theoretical model) the number of degrees of freedom is 256.

The confidence level has been assumed to be 0.05 and thus the corresponding left-tailed χ^2 distribution critical value is $\chi_{ind.}{}^2{}_{critical} = 294.32067$.

The results of the chi-square test are summarized in table in Fig 3. If the measured $\chi_{ind.}{}^2$ statistic is greater/equal than $\chi_{ind.}{}^2{}_{critical}$ there is no statistically significant reason to rejected null-hypothesis and thus the observed distributions of the noise values may be dependent. Otherwise, there is statistically significant reason (at the given confidence level) to reject the null hypothesis and thus the observed distributions of the noise values are independent.

From the results presented in table in Fig. 2, it can be seen that the ratio is definitely below 1 (ranges from 0.0145 to 0.0387, which is negligible). This leads to a conclusion that the null hypothesis has to be rejected. This provides evidence that the noise in subsequent frames is independent.

4. GAUSSIAN DISTRIBUTION TEST

In many video processing algorithms Laplace or Gaussian distribution of the noise is assumed. In order to justify whether such assumptions are valid, distributions of the $Noise_i(x,y)$ values have to be estimated. In this paper, we have estimated noise distributions basing on noise histograms.

For sufficient accuracy, bin size of $^1/_{16}$ of the smallest quantization step of the luminance values (the smallest representable luminance value difference) has been used. Such subpixel bin size was possible due to the fact that $Denoised(x,y)$ values (Eq. 1) are real (not rounded to integers) numbers.

The exemplary results obtained on some of sequences from Table 1 are presented in Fig. 3. As it can be seen, clearly, the distributions are not Laplace. The estimated noise distributions resemble Gaussians and thus assumption about such distribution will be further evaluated (among others by chi-square test).

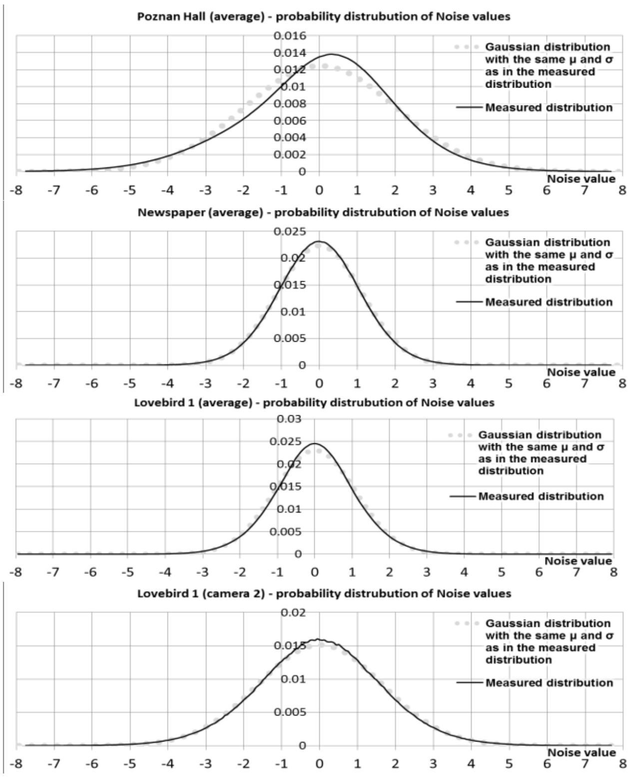


Fig. 3. Measured probability distribution of noise values (averaged over all cameras).

First, simple parameters of the estimated distributions have been estimated (Table 2). In the case of Poznan Street, Poznan Carpark and Poznan Hall sequences, the measured distribution is slightly skewed in such a way, that the maximum point of the distribution is at position of about 0.4, related to the expected value. This may be a results of internal noise reduction algorithm implemented in the Canon XH-G1 camera (used for recording of this sequence – see Table 1) or a results of internal non-linear processing of data from the camera sensor.

Table 2. Summary of parameters of the noise distributions in the test sequences.

Sequence Name	Standard deviation	Position of maximum point related to expected value of distribution	Notes
Poznan Street	2.45	0.41	Measured distribution is skewed
Poznan Carpark	2.28	0.42	
Poznan Hall	2.01	0.51	
Lovebird1, without camera 2	0.66	0.02	Camera 2 of Lovebird1 sequence has vastly different noise profile
Lovebird1, camera 2	1.65	0.01	
Newspaper	1.11	-0.02	
Kendo	1.01	0.01	Kendo is a moving sequence – values taken basing on Balloons sequence only
Balloons			

In the case of Lovebird1 sequence, standard deviations are the lowest in the whole test set and are very similar across all of the cameras – at level of about 0.66. The only exception is camera 2, where the standard deviation is about 2.5 times higher (about 1.65). This might be evidence that this particular view has been acquired with different parameters – e.g. the exposure time has been shorter, which has been corrected with higher amplification gain, which also amplified the noise.

In other cases, there are no anomalies, and the distributions resemble Gaussians - are well-symmetric and centered at value of 0 (corresponding to the expected value of distribution).

Although it can be noticed that the measured distributions are visually very similar to shape of Gaussian function, the statistically significant similarity has to be confirmed. This will be done with chi-square goodness-of-fit test. The working null hypothesis is that the observed distribution is normal (Gaussian). The working alternative hypothesis is that the observed distribution is not normal (Gaussian).

The number of degrees of freedom has been calculated as $df_{gof} = 253$, as bin size of ${}^{1}/_{16}$ has been used for measurement of noise values in range -8 to 8 (see: previous page), and as the standard deviation and the mean of expected distribution has been estimated (are not known from a theoretical model).

The confidence level has been assumed to be 0.05 and thus the corresponding critical value, calculated from right-tailed χ^2 distribution, is ${\chi_{gof}}^2{}_{critical} = 291.10174$.

Table 3. Chi-square results for all views of the tested sequences. Values that are less than 1.0 (gray) indicate that given cases pass the χ^2 test.

Sequence \ Camera index		0	1	2	3	4	5	6	7	8
Name	Multiplier	$\frac{{\chi_{gof}}^2}{{\chi_{gof}}^2{}_{critical}}$, scaled by the multiplier								
Poznan Street (cam. 0..8)	$10^1\times$	7.93	7.65	6.71	6.82	7.00	4.90	5.54	5.51	5.11
Poznan Carpark (c. 0..8)	$10^2\times$	3.89	3.56	3.03	3.18	3.03	3.33	3.31	2.02	1.89
Poznan Hall (cam. 0..8)	$10^3\times$	2.12	1.66	1.84	1.75	1.64	2.08	1.76	1.55	1.28
Lovebird1 (cameras 0..8)	$10^2\times$	0.50	1.49	0.46	1.84	1.95	1.56	1.08	0.86	1.33
Newspaper (cam. 0..8)	$10^1\times$	1.30	1.38	1.03	2.07	1.92	1.24	2.03	1.84	2.65
Balloons (cameras 0..6)	$10^0\times$	1.03	1.42	1.16	0.88	0.94	1.90	0.69	-	

The results of the goodness-of-fit test are gathered in Table 3. It can be noticed that for the most of the cases, the ratio between test statistics ${\chi_{gof}}^2$ and critical value ${\chi_{gof}}^2{}_{critical}$ is of magnitude of about $10^1 - 10^2$ proving that there are no statistically significant reasons for assumption that estimated distributions are Gaussians. The only exception is the Balloons sequence, where for some cameras (marked in gray in Table 3) assumption about Gaussian is statistically significant (the presented multiplied showing the level of magnitude of 10^0).

Therefore, in spite of the visual impression that the observed probability distributions are Gaussian-like, generally it can be concluded that for most of the sequences, the null-hypothesis must be rejected and almost none of them is Gaussian (at given confidence level).

5. DISPARITY-BASED PIXEL CORRESPONDENCE NOISE ANALYSIS

In previous sections, a simple noise analysis techniques have been presented along with original results for commonly known multiview video sequences. The presented techniques have an important drawback - they require noise analysis to be performed, which confines either in need to select an arbitrary noise extraction algorithm or to arbitrarily select still regions. Also, in some material, still regions may be unavailable.

In this section a novel analysis method is shown, aimed at overcoming this drawback. The method is based on disparity-based pixel correspondence between different views of the same scene. For the sake of brevity, only two views (X and Y) are used (Fig. 4).

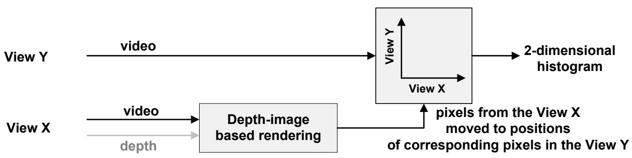


Fig. 4. Scheme of the experiment for disparity-based pixel correspondence noise analysis..

The main idea of the method is to compare pixel values from view Y and corresponding pixels values from view X with use of 2-dimensional histogram. For that, pixel correspondence between the views has to be known, which in the paper is attained by use of depth maps (whose are assumed to be available). Depth-Image Based Rendering is used to warp pixels from the view X to position of the corresponding pixels in the view Y. After the warping, two-dimensional histogram is created from all corresponding pixels.

Such 2-D histograms have been estimated for each combination of the views available for the sequences mentioned in Table 1. The results have been presented in Fig. 5. The graphs have been row-wisely normalized with respect to the occurrence of given pixel values in the view X (estimated basing on histogram of the view X solely) for the sake of clarity of presentation.

In order to objectively assess the proposed pixel-correspondence-based approach and compare its results with the simple method which uses noise extraction, noise characteristics have been estimated (similarly to the ones presented in Fig. 3). This has been done by aligned averaging the rows of estimated 2-dimensional histograms (Fig. 5) which led to creation of 1-dimensional noise distribution. The results, presented in Table 3, show standard deviations of attained distributions, which are 1.5×-2× higher that those presented previously in Table 2. Apart from this, it can be said that, roughly, the measured noise amplitudes are correct. Importantly, the results have been attained also for sequences in which previous methodology could not be used (computer-generated sequences without noise and moving sequences). This results from the fact, that such analysis method also yields with other phenomena. In Fig. 5 it can be noted that in the case of natural sequences, appearance of the pixels in both views is not exactly the same (luminance values do not lay strictly on the diagonal of the graph). The peak curve is slightly distorted which suggests that the color profiles of the cameras used were not perfectly calibrated and moreover this mismatch was not corrected perfectly afterwards.

What can also be noted in Fig. 5 is that the width of the peak is changing, which suggests that the amplitude of the noise varies with the luminance level or can be an indication of a fact, that the observed appearance of the given point in the scene changes with position of the camera and thus cannot be strictly resulting from Lambertian model of reflectance. This is important notice for inter-view matching problems, like depth estimation or super-resolution.

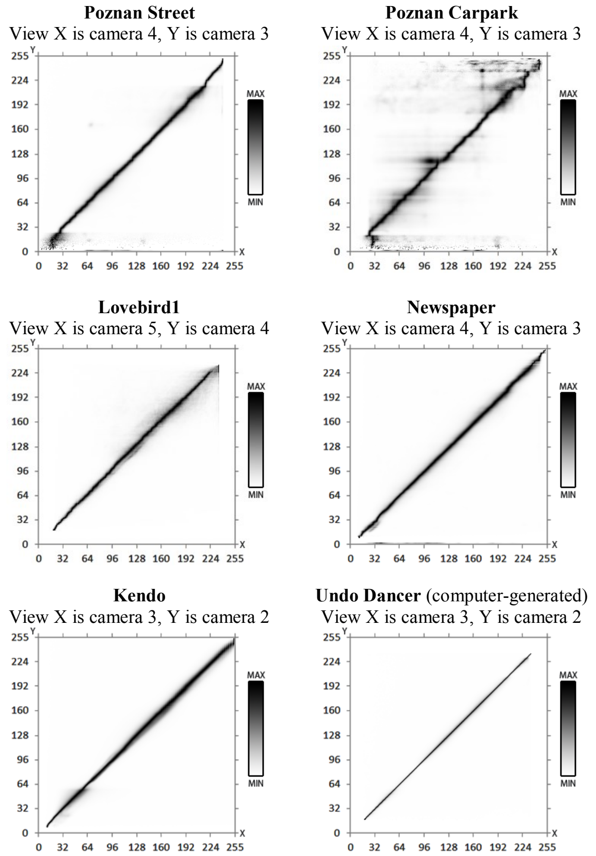


Fig. 5. Exemplary graphs of row-wise normalized 2-dimensional histograms of luminance values (in logarithmic gray-level scale) of corresponding pixels in views X and Y.

Table 4. Summary of the parameters of the noise distributions in the test sequences, measured with use of the proposed pixel-correspondence-based analysis method.

Sequence Name	Standard deviation	Sequence Name	Standard deviation
Poznan Street	4.21	Newspaper	2.84
Poznan Carpark	4.73	Kendo	2.01
Poznan Hall	4.30	Balloons	2.24
Lovebird1, without camera 2	1.84	GT Fly	0.53
Lovebird1, camera 2	3.42	Undo Dancer	0.72

6. CONCLUSSIONS

In this paper, first it has been shown that usage of simple methods can lead to interesting and novel conclusions about presence of noise in a multi-camera systems. It has been confirmed that noise is independent in time domain, which justifies usage of most noise reduction techniques. Then, it has been verified, that in the case of the tested sequence, the most commonly made assumptions about presence of Laplace distribution of noise is not valid. Then, although the noise distributions are similar to Gaussians it has been shown that similarity is not statistically significant. It is noted that an important drawback of the presented methodology is requirement to perform noise extraction.

In order to overcome this drawback, a novel analysis method is presented later. Instead of noise extraction, the method exploits depth maps in order to perform disparity-based pixel correspondence noise analysis. It has been shown that the proposed method yields with slightly different results that the simple methods presented first, but at the same time, allows for comprehensive analysis of the noise, like order of magnitude of noise amplitude and, which is important, also other phenomena in multi-camera systems: in the tested sequences effects of color miscalibration and some indication of non-Lambertian reflectance in the scenes has been noticed.

This work was partially supported by National Science Centre, Poland, according to the decision DEC-2013/07/N/ST6/02267 and partially co-financed by European Union funds as a part of European Social Funds.

7. REFERENCES

[1] C. Fehn, R. de la Barr´e, S. Pastoor, “Interactive 3-DTV – concepts and key technologies”, *Proc. IEEE*, Vol. 94, No. 3, pages 524–538, Mar. 2006.

[2] T. E. Bishop, P. Favaro. “The light field camera: Extended depth of field, aliasing, and superresolution” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, v. 34, 2012.

[3] D. Scharstein, R. Szeliski, „A taxonomy and evaluation of dense two-frame stereo correspondence algorithms”, *Int. Journal of Computer Vision* 2002.

[4] D. Tian, P.-L. Lai, P. Lopez, C. Gomila, “View Synthesis Techniques for 3D Video”, *Digital Image P*roc. XXXII. Proc, of the SPIE, Vol. 7443, 2009.

[5] S. Farsiu, D. Robinson, M. Elad, P. Milanfar, "Fast and Robust Multi-frame Super-resolution", *IEEE Trans. on Image Processing*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.

[6] T. Georgiev, G. Chunev, A. Lumsdaine, “Superresolution with the focused plenoptic camera” *SPIE Electronic Imaging,* Jan 2011.

[7] M. Joachimiak, D. Rusanovskyy, M. M. Hannuksela, M. Gabbouj “Multiview 3D Video Denoising in sliding 3D-DCT Domain”, *European Signal Processing Conference, EUSIPCO 2012*, Budapest, Romania, 2012.

[8] Jian Sun, Nan-Ning Zheng, Heung-Yeung Shum, „Stereo Matching Using Belief Propagation*”, IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 25, Issue: 7, pages 787 - 800, 2003.

[9] S. V. Vaseghi, “Advanced Digital Signal Processing and Noise Reduction (Third Edition)”, *John Wiley & Sons* 2006, ISBN: 978-0-470-09495-2, 2006.

[10] P. Leclercq, J. Morris, “Robustness to noise of stereomatching,” *in Proc. of the 12th Int. Conf. on Image Analysis and Processing*, pp. 606–611, Sept. 2003.

[11] G. Varghese, Z. Wang, “Video denoising based on a spatiotemporal gaussian scale mixture model,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 7, pp. 1032–1040,July 2010.

[12] C. Boncelet "Image Noise Models" *In Alan C. Bovik. Handbook of Image and Video Processing. Academic Press.* ISBN 0-12-119792-1, 2005.

[13] “Common Test Conditions for AVC and HEVC-based 3DV”, *ISO/IEC MPEG Doc.*N12560, San Jose, USA, 2012.

[14] Hui Tian. “Noise Analysis in CMOS Image Sensors”. *A Dissertation of Stanford University*, 2000.

[15] R. C. Gonzalez, R.E. Woods “Digital Image Processing” *2nd Ed. Bovik: Elsevier Academic Press*, 2002.