

**INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG2014/m34305
July 2014, Sapporo, Japan**

Source Poznan University of Technology
Status Input
Title Noise in 3D video sequences
Author Olgierd Stankiewicz (ostank@multimedia.edu.pl),
Marek Domański,
Krzysztof Wegner

1 Introduction

This document presents results characterizing noise existing in 3D video sequences used by MPEG (Table 1) for research and standardization purposes in areas of 3DTV and FTV. In such systems, advanced processing algorithms are used, like depth estimation, view synthesis or light-field rendering. Operation of these algorithms is conditioned by presence of noise in the input video. Unfortunately, in formulation of such algorithms, typically, the presence of noise is either omitted or the assumptions about the characteristics of noise are made only implicitly. Even when presence of the noise in the video material is assumed explicitly, the most commonly solely Gaussian noise is considered, often without any experimental verification.

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.

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. Results show proof of independency of noise in time and characteristics of noise value distributions, which are similar to Gaussians.

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

Sequence Name	Camera	Width x Height	Frame rate [frames/s]	Total number of frames	Views numbers	Depth data available for views
Poznan Carpark	Canon XH-G1, 3-CCD camera	1920 x 1088	25	250	0...8	3,4,5
Poznan Street				200	0...8	5,6,7
Poznan Hall						
Lovebird1	Point Grey Flea camera (CCD), Moritex ML-0813 lenses	1024 x 768	30	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

2 Noise extraction

The noise existing in a video sequence can be simply attained as a difference between the original and a denoised version of the given sequence. Of course, for denoising any advanced technique could be used, but for the sake of simplicity and robustness of the results, the simplest and the most straight-forward one has been used.

In the used simple analysis method, only fragments of sequences representing still fragments of the scene (without any movement) are considered. Such regions have been marked manually (Fig.1). It has been taken care, to include as representable set of regions as possible, for example to cover fragments of the scene with different color and brightness.

It is assumed that all frames of given sequence represents the same image altered only by presence of noise. Therefore, the image without noise can be easily retrieved, as an average of the frames:

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

where $Frame_i(x, y)$ express luminance value at coordinates x, y in i -th frame of the sequence. The sought noise value $Noise_i(x, y)$, can be thus simply calculated as:

$$Noise_i(x, y) = Frame_i(x, y) - Denoised(x, y)$$

3 Probability distribution estimation

For estimation of probability distribution of the extracted noise values $Noise_i(x, y)$ we have used histograms. In our experiments, histograms are being constructed with bin size was equal to $1/16$ of quantization step of the luminance values (e.g. in $Frame_i(x, y)$). This was possible due to fact, that the noise values $Noise_i(x, y)$ can be fractional numbers due to averaging process while denoising. Therefore, given value of the noise $Noise_i(x, y)$ is counted in histogram in bin identified by index $NoiseBinIdx_i(x, y)$ identified as follows:

$$NoiseBinIdx_i(x, y) = [Noise_i(x, y) \cdot 16]$$

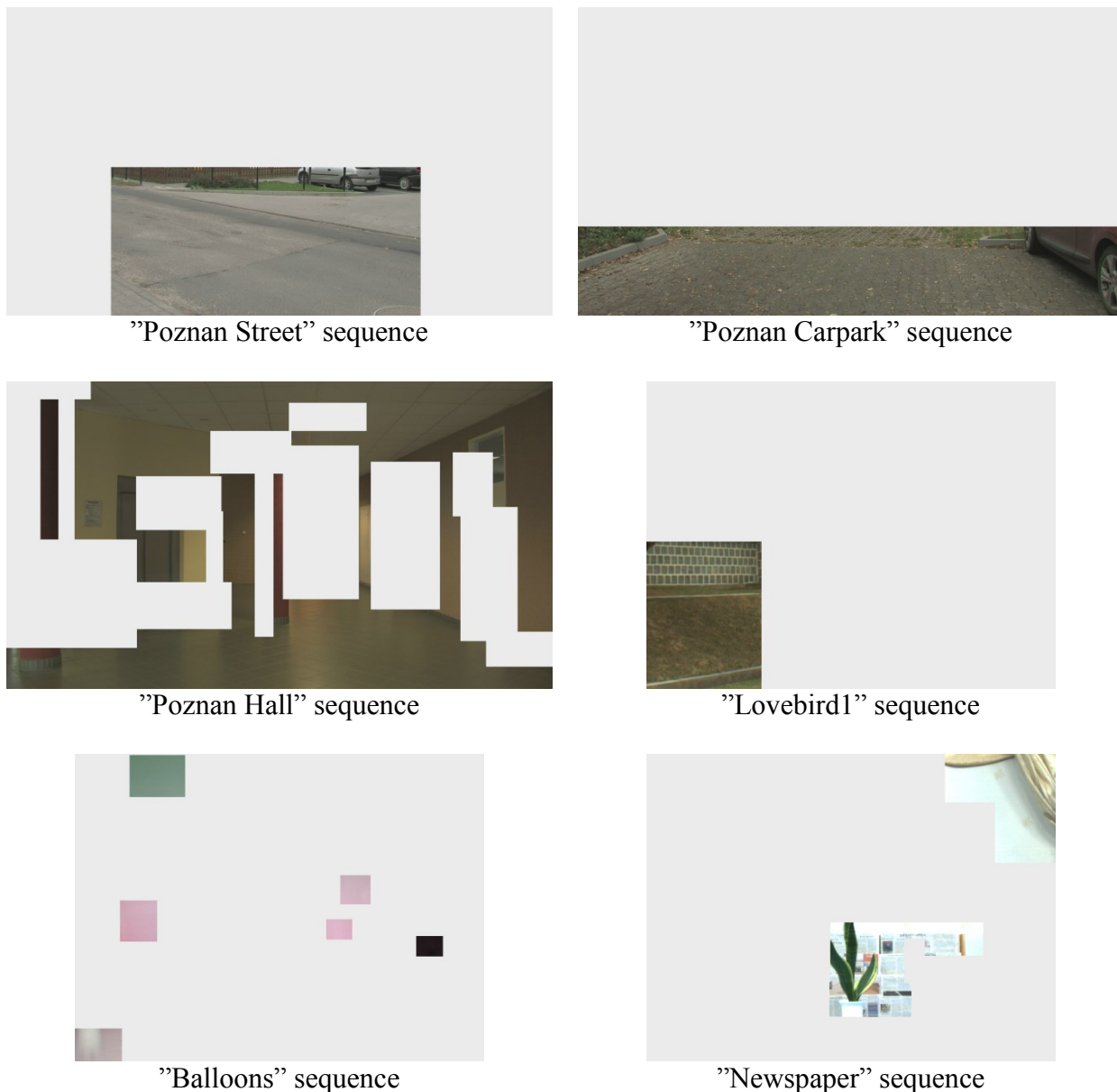


Fig.1. Regions in the multiview video sequence set that has been manually marked as still for the sake of estimation of the noise. Regions have been marked in gray.

4 Time independency of the observed noise

First, time dependency of noise values has been assessed. A simple and robust method of verifying whether there is any dependence has been used. If between two random variables, e.g. α and β there is any dependence, the following equality is not true:

$$p(\alpha, \beta) = p(\alpha) \cdot p(\beta)$$

In order to perform such verification, the two-dimensional distribution of the noise in two consecutive frames has to be estimated. Two-dimensional distributions have been estimated with use of histograms of $Noise_i(x, y)$ vs. $Noise_{i-1}(x, y)$. Histograms have been averaged over all stable frames in the sequence. An exemplary histogram has been presented in Fig. 2. One axis of

the histograms relates to the values of noise in the i -th frame, and the other axis relates to the values of noise in $(i - 1)$ -th frame. The analogous histograms for other sequences have been gathered in Fig. 3 (in the same presentation form as in Fig. 2). If there would be any dependence between distribution of the noise in i -th frame and $(i - 1)$ -th frame there would be an asymmetry in the graph.

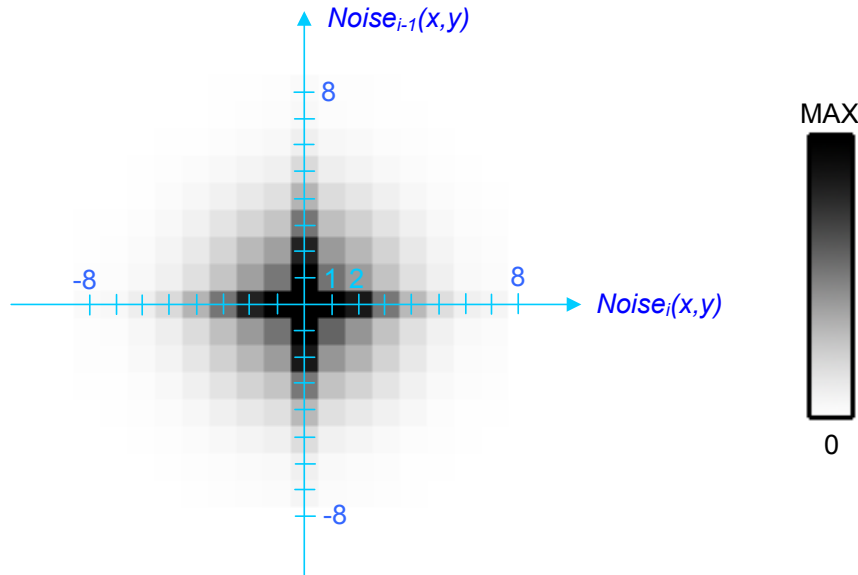


Fig. 2. Exemplary plot of two-dimensional histogram of $Noise_i(x,y)$ vs. $Noise_{i-1}(x,y)$ for Poznan Street test sequence, camera 0, frame 0. One axis of each histogram relates to the values of noise in the frame i , and the other axis relates to the values of noise in frame $i - 1$. The same visualization method has been used in Fig. 3.

Sequence	Camera Index									
	0	1	2	3	4	5	6	7	8	
Poznan Street (cameras 0..8)										
Poznan Carpark (cameras 0..8)										
Poznan Hall (cameras 0..8)										
Lovebird1 (cameras renumbered as 0..8)										
Newspaper (cameras 0..8)										
Balloons (cameras 0..6)								no data		

Fig. 3. Plots of two-dimensional histogram of $Noise_i(x,y)$ vs. $Noise_{i-1}(x,y)$ for various test sequences and cameras, averaged over frames. The plots are presented in the same way as in Fig. 2 but for the sake of brevity, visualization of axes has been omitted.

As can be noticed in Fig. 3, the graphs are separable, which indicates that the random variables are not dependent (are independent) and thus that the noise in subsequent frames is independent. And thus proves time independence of the observed noise.

Visual verification is not sufficient to draw the above mentioned conclusion so apart from such visual verification, the independence has been tested mathematically. Given normalized 2-dimensional histogram (depicted by $H[\cdot]$ operator) of noise values in subsequent pair of frames i and $i - 1$:

$$H[Noise_i(x, y), Noise_{i-1}(x, y)](\alpha, \beta) = h_{i,i-1}(\alpha, \beta),$$

histogram of noise values in frame i :

$$H[Noise_i(x, y)](\alpha, \beta) = h_i(\alpha)$$

and histogram of noise values in frame $i - 1$:

$$H[Noise_{i-1}(x, y)](\beta) = h_{i-1}(\beta),$$

we assume that those normalized histograms are estimates of probability distributions of noise in the corresponding cases. If the noise distribution are independent between the frames, then the expected distribution of $h_{i,i-1}'(\alpha, \beta)$ will be:

$$h_{i,i-1}'(\alpha, \beta) = h_i(\alpha) \cdot h_{i-1}(\beta).$$

The energy of difference, between the expected distribution $h_{i,i-1}'(\alpha, \beta)$ and the observed one $h_{i,i-1}(\alpha, \beta)$ has been used in order to perform **chi² independence test**.

$$\chi_{ind.}^2 = \sum_{\alpha \in \Phi} \sum_{\beta \in \Phi} \frac{|h_{i,i-1}'(\alpha, \beta) - h_{i,i-1}(\alpha, \beta)|^2}{h_{i,i-1}'(\alpha, \beta)} .$$

Φ is a set of all possible noise values. Range $[-8..8]$ has been selected in order to cover the whole range on significant noise values and at the same time, to avoid small number of samples in histogram bins, which is desired in case of chi^2 test. Bin size have been selected to be equal to 1 quantization step of the luminance, therefore, range of $-8 \leq \alpha, \beta \leq 8$, results in total $\varphi = 17$ bins.

The working null hypothesis is that the observed distributions are dependent.

The working alternative hypothesis is that the observed distributions are independent.

Now, number of degrees of freedom will be calculated, which is equal to the number of bins in two-dimensional histograms $\varphi \cdot \varphi$, minus the reduction in degrees of freedom df_{red} . As the expected distribution has been estimated (it is not known from a theoretical model), the number of degrees of freedom have to be reduced by $df_{red} = \varphi + \varphi - 1$ (the number of rows and cols is equal to φ). Finally, the number of degrees of freedom is:

$$df_{ind.} = \varphi \cdot \varphi - (\varphi + \varphi - 1) = 256.$$

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

$$\chi_{ind.}^2_{critical} = 294.32067.$$

For each of the sequences and each of views $\chi_{ind.}^2$ statistic has been calculated and compared to $\chi_{ind.}^2_{critical}$ value:

$$\chi_{ind.}^2_{ratio} = \frac{\chi_{ind.}^2}{\chi_{ind.}^2_{critical}}.$$

As left-tailed χ^2 distribution is used, value of $\chi_{ind.}^2_{ratio}$ which is greater or equal than 1 (measured $\chi_{ind.}^2$ statistic is greater/equal than $\chi_{ind.}^2_{critical}$) means that the null hypothesis cannot be rejected and thus the observed distributions of the noise values may be dependent. Value of $\chi_{ind.}^2_{ratio}$ which is smaller than 1 (measured $\chi_{ind.}^2$ statistic is lesser than $\chi_{ind.}^2_{critical}$) means that the null hypothesis must be rejected and thus the observed distributions of the noise values are independent (at the given confidence level).

From the results presented in Table 2, it can be seen that $\chi_{ind.}^2_{ratio}$ 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.**

Table 2. Results of χ^2 independence test, for pairs of successive frames of the test sequences. The results have been averages over time and over cameras.

Sequence name	$\chi_{ind.}^2_{ratio}$
Poznan Street	0.0145
Poznan Carpark	0.0249
Poznan Hall	0.0194
Lovebird1	0.0387
Newspaper	0.0269
Balloons	0.0307

Interesting observation can be noticed in Fig. 3. In the case of Lovebird 1 sequence, camera 2 of has slightly different graph than other cameras. It can be supposed that this particular view has been acquired with different camera settings. Similar phenomenon can be observed in results presented in the next Subsection.

5 Probability distributions of the noise

In the previous point it has been proven that the realizations of noise in subsequent frames of the tested sequences are independent. Therefore the sought probability distributions of the noise can be estimated with use of histograms calculated over all frames of each sequence. If the noise was not independent between the frames, averaging over the frames would be statistically incorrect.

For the reasons stated before, the histogram analysis of the noise is performed with use of bins $NoiseBinIdx_i(x, y)$ defined with bin size of $1/16$ of the normal quantization step of the luminance values (the smallest representable luminance value difference).

The results have been presented in Figs. 4-10 in form of average (over all cameras) for visualization.

As can be noticed, none of measured distribution of the noise, extracted from the test sequence set, represents a Laplace distribution. In many algorithms Laplace distribution of noise values is equivalent to matching with use of SAD metric (Sum of Absolute Differences). **As assumption about Laplace noise is not justified, also usage of SAD matching metric cannot be justified in such sequences.**

In general, it can be said, that the measured noise distributions are visually very similar to Gaussian (normal) distribution. For the visual comparison, in Figs. 4-10, apart from the measured data (marked in continuous blue line), on the same figures, also Gaussian (normal) distribution has been depicted (marked in dotted-red line). The visualized Normal distribution has the same parameters: μ and σ (Table 3). There are some exceptions for this mentioned “similarity” to Gaussian distribution though, described below.

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 of the distribution is at position of about 0.4. This may be a results of internal noise reduction algorithm implemented in the Canon XH-G1 camera or a results of internal non-linear processing of data from the camera sensor. Standard deviations are very similar among the views (Table 3), but there are little differences among the sequences. Those are 2.45 (Poznan Street,), 2.28 (Poznan Carpark) and 2.01 (Poznan Hall).

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 (renumbered index; in original numbering this camera has index 3), where the standard deviation is about 2.5 times higher – it has been measured to be about 1.65. **This might be evidence that this particular view has been acquainted with different parameters – e.g. the exposure time has been shorter, which has been corrected with higher amplification gain, which also amplified the noise.** Apart from that anomaly, the Gaussians are well-symmetric and centered at value of 0. This means that the distribution of the noise in such example is well-centered.

The probability distribution of the noise in **Newspaper** sequence is very similar to Gaussian distribution in all of the cameras. The standard deviations are very similar among the views at a level of about 1.23.

The distributions of **Balloons** sequence strictly follow Gaussian “bell” shape. Also here, standard deviations are very similar among the views, at level of about 1.01.

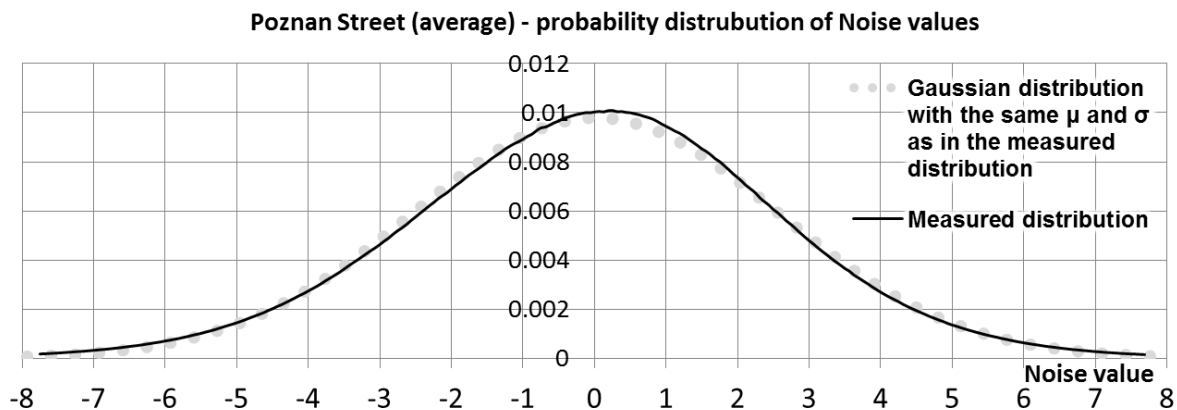


Fig. 4. Measured probability distribution of noise values in Poznan Street sequence (averaged over all views), estimated with histogram bin size of $1/16$.

Poznan Carpark (average) - probability distribution of Noise values

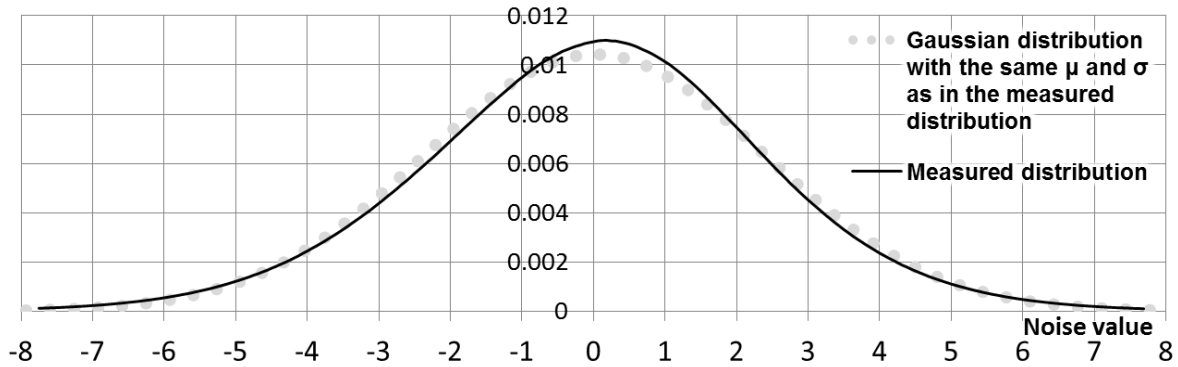


Fig. 5. Measured probability distribution of noise values in Poznan Carpark sequence (averaged over all views), estimated with histogram bin size of $1/16$.

Poznan Hall (average) - probability distribution of Noise values

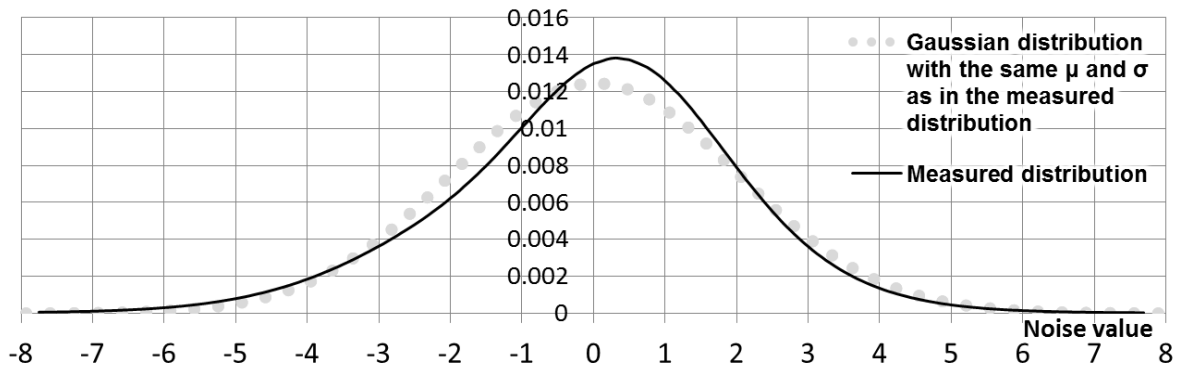


Fig. 6. Measured probability distribution of noise values in Poznan Hall sequence (averaged over all views), estimated with histogram bin size of $1/16$.

Newspaper (average) - probability distribution of Noise values

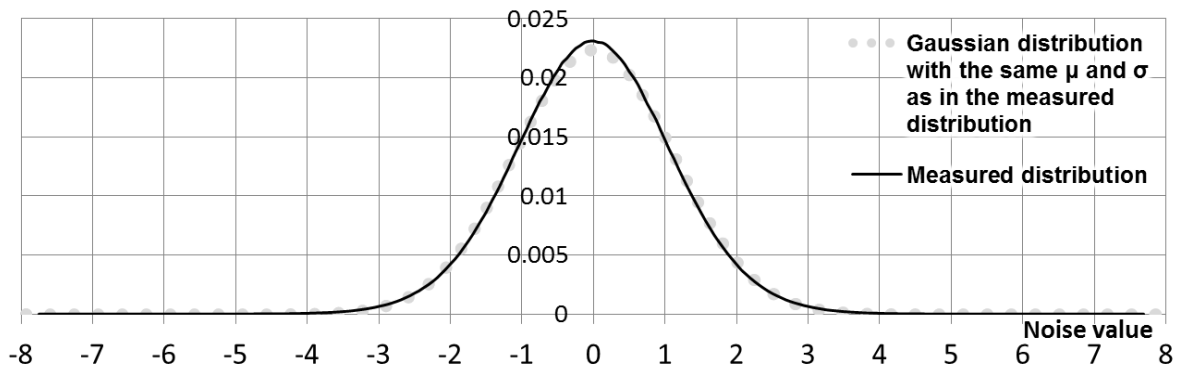


Fig. 7. Measured probability distribution of noise values in Newspaper sequence (averaged over all views), estimated with histogram bin size of $1/16$.

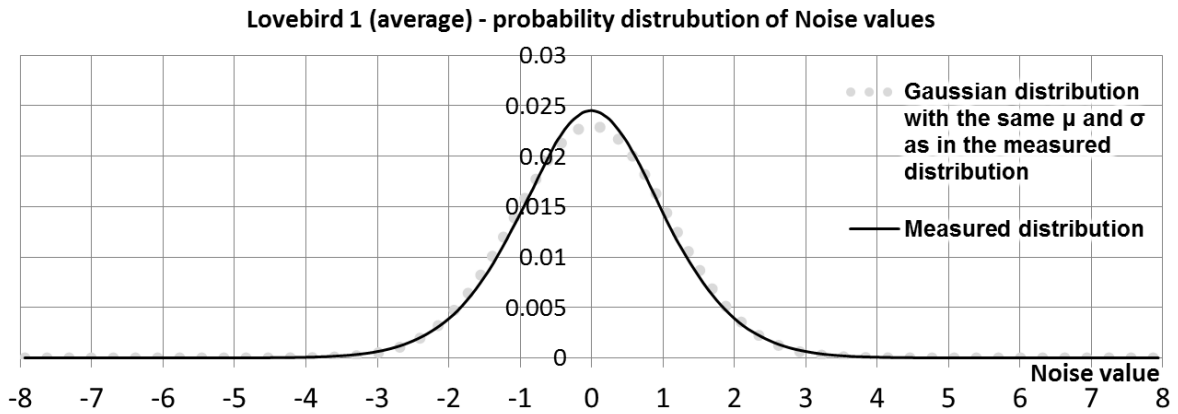


Fig. 8. Measured probability distribution of noise values in Lovebird 1 sequence (averaged over all views), estimated with histogram bin size of $1/16$.

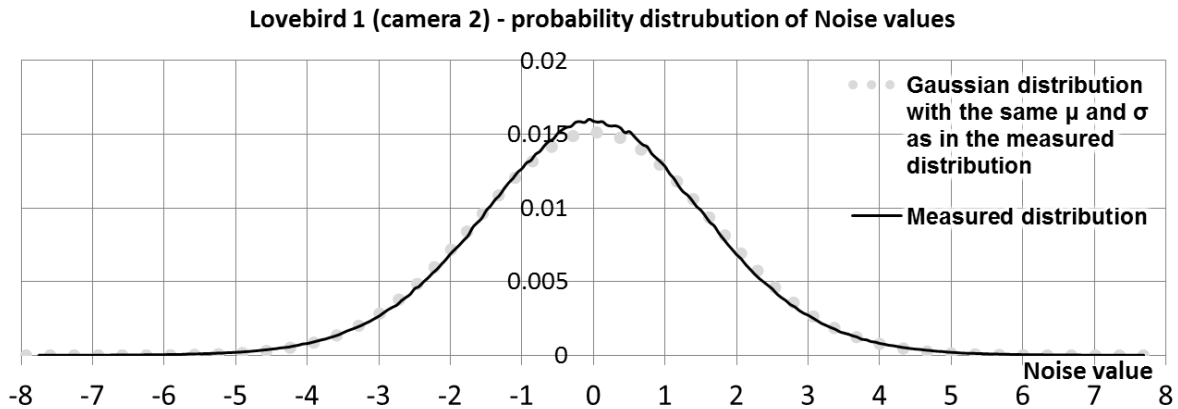


Fig. 9. Measured probability distribution of noise values in Lovebird 1 sequence (renumbered camera index 2), estimated with histogram bin size of $1/16$. In the case of this camera, the standard deviation is about 2.5 times higher than in other cameras of Lovebird 1 sequence..

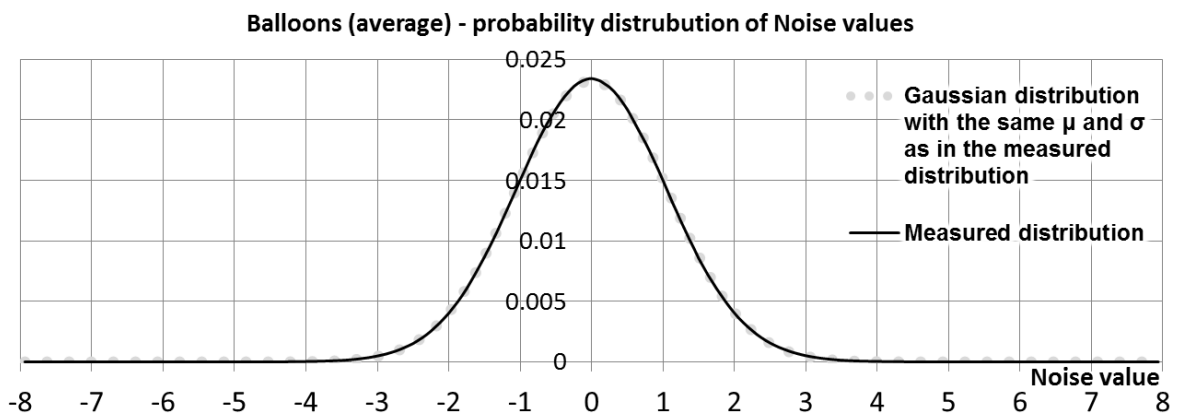


Fig. 10. Measured probability distribution of noise values in Balloons sequence (averaged over all views), estimated with histogram bin size of $1/16$.

Table 3. Summary of the Gaussian probability model matching
- parameters of the noise distributions in the test sequences.

Sequence Name	Standard deviation	Maximum point of distribution, related to EX	Notes
Poznan Street	2.45	0.41	Measured distribution is skewed
Poznan Carpark	2.28	0.42	
Poznan Hall	2.01	0.51	
Lovebird1, w.o. cam.2	0.66	0.02	Camera 2 (renumbered index) 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			

6 Chi-square test for Gaussian probability distribution of the noise

In previous point it has been shown that the distributions presented in Figs. 4-10 undoubtedly are not Laplace distributions, but in general follow the shape of Gaussian function. Yet it has not been proven whether those distributions are indeed Gaussians or not. First of all, the shape of noise distribution slightly varies among the views. Also, in some of the sequences (Poznan Street, Poznan Carpark, Poznan Hall) the distribution is skewed, such that its most probability (maximum) point is displaced in relation to the expected value (mean) (Table 2).

Therefore, to provide a proof, a statistical test has to be performed. As measured histograms of the considered distributions are available, **statistical χ^2 goodness-of-fit statistical test χ^2_{gof} has been used.**

In fact, the following reasoning will show that, in spite of the visual similarity, the measured distributions are not Gaussians:

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).

As stated in earlier the histogram analysis of the noise is performed with use bin size of $1/16$ of the normal quantization step of the luminance value (which correspond to $1/16$ of the smallest representable luminance value difference). For the practical reasons, each of the distributions (for all sequences and all views) has been observed with use of 256 bins which covers noise value range of [-8; 8]. Experiments have revealed that most of the noise samples are in the selected value range.

In the analyzed case of χ^2 goodness-of-fit, the observed distribution thus will be histogram of noise in given view/sequence and the expected distribution is Gaussian. The standard deviation and the mean of expected distribution has been estimated (are not known from a theoretical model) the number of degrees of freedom is:

$$df_{gof} = 256 - 1 - 2 = 253.$$

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

$$\chi_{gof}^2_{critical} = 291.10174.$$

For each of the sequences and each of views χ_{gof}^2 statistic has been calculated and compared to $\chi_{gof}^2_{critical}$ value.

$$\chi_{gof}^2_{ratio} = \frac{\chi_{gof}^2}{\chi_{gof}^2_{critical}}$$

As right-tailed χ^2 distribution is used, value of $\chi_{gof}^2_{ratio}$ which is smaller than 1 (measured χ_{gof}^2 statistic is lesser than $\chi_{gof}^2_{critical}$) means that the null hypothesis cannot be rejected and thus the observed distribution may be Gaussian. Value of $\chi_{gof}^2_{ratio}$ which is greater or equal than 1 (measured χ_{gof}^2 statistic is greater/equal than $\chi_{gof}^2_{critical}$) means that the null hypothesis must be rejected and thus the observed distribution is not Gaussian.

The results of $\chi_{gof}^2_{ratio}$, calculated for the test sequences are gathered in Table 4. It can be noticed that for the most of the cases, the ratio between χ_{gof}^2 and $\chi_{gof}^2_{critical}$ is of magnitude of about $10^1 - 10^2$ proving that the distributions are not Gaussians. The only exception is the Balloons sequence, where $\chi_{gof}^2_{ratio}$ fluctuates around 1 (the presented multiplied showing the level of magnitude of 10^0). Thus, depending on a particular camera of the Balloons, the hypothesis that the distributions are Gaussians must be rejected (marked in white in Table 4) in or may not be rejected (marked in gray in Table 4).

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).

Table 4. χ^2_{ratio} results for all views of the tested sequences. Values that are less than 1.0 (marked in gray) indicate that given cases pass the χ^2 test.

Sequence		Camera index									
		0	1	2	3	4	5	6	7	8	
Name	Multiplier	χ^2_{ratio} , scaled by the multiplier									
Poznan Street (cameras 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 (cameras 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 (cameras 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 (cameras 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	-	-	

Both the graphs presented in Figs. 4-10 and Table 4, refer to the question, whether the probability distributions of the noise in tested multiview video sequences are Gaussians. A comparison of the visual impressions that can be done, basing on the mentioned figures (that the distributions are similar to Gaussians), and the results of statistical analysis (that almost none of the distributions are Gaussian) show that there is discrepancy between those two methods. This

discrepancy (between the visual impressions and results of χ^2_{gof} test) can be explained on the basis of number of observed samples. As number of samples increase, the χ^2 test becomes more discriminating. With a large number of observed samples (millions in the performed experiment), the measured distribution should be almost exactly Gaussian in order to pass through the χ^2 test, while the measured distributions still have slight variations.

As almost none of the test sequences have passed the performed χ^2 test and that their noise distributions are not Gaussians (the null-hypothesis has been rejected). In many 3D algorithms assumption about Gaussian distribution of noise values allows to use SSD metric for image matching (Sum of Squared Differences). An example of such algorithm is Depth Estimation Reference Software (DERS). **As assumption about Gaussian noise is not justified, also usage of SSD matching metric cannot be strictly justified in such sequences. On the other hand, the presented distributions are very similar to Gaussians, and thus usage of SSD formula might be heuristically reasonable.**

7 Conclusions

In this document a simple methodology for measurement of noise characteristics has been presented. On the considered set of test sequences (Poznan Street, Poznan Carpark, Poznan Hall, Lovebird1, Newspaper, Balloons) three important conclusions has been drawn.

Firstly, the noise is independent in time domain, which justifies usage of many algorithms, like the denoising technique used in this work.

Secondly, the measured noise value distributions undoubtedly are not Laplace distributions. In many algorithms Laplace distribution of noise values is equivalent to usage of SAD metric (Sum of Absolute Differences) and in such cases, usage of SAD is not justified.,

Lastly, the measured noise value distributions visually resemble Gaussians but statistical chi-square goodness-of-fit test has shown that in fact they are not Gaussians. In algorithms, where Gaussian distribution of noise values leads to usage of SSD matching metric (Sum of Squared Differences), usage of SSD is not strictly justified and at least heuristically reasonable. An example of algorithm which uses SSD for image matching is Depth Estimation Reference Software (DERS). Therefore, matching metric in DERS should be reconsidered.

More detailed results can be found in [1].

Acknowledgement

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.

References

- [1] O. Stankiewicz, Marek Domański, K. Wegner, „Analysis of noise in Multi-camera systems”, 3DTV-Conference 2014 The True Vision Capture, Transmission and Display of 3D Video, Budapest, Hungary, 2-4 July 2014.
- [2] M. Domański, T. Grajek, K. Klimaszewski, M. Kurc, O. Stankiewicz, J. Stankowski, K. Wegner, „Poznań Multiview Video Test Sequences and Camera Parameters”, ISO/IEC JTC1/SC29/WG11 Doc. M17050, Xian, China, October 2009.
- [3] Yo-Sung Ho, E.-K. Lee, C. Lee “Video Test Sequence and Camera Parameters”, ISO/IEC MPEG M15419, Archamps, France, April 2008.
- [4] I. Feldmann, A. Smolic, T. Wiegand et.al., „HHI Test Material for 3D Video”, ISO/IEC JTC1/SC29/WG11, Doc M15413, Archamps, France, April 2008.
- [5] Gi-M. Um, G. Bang, N. Hur, J. Kim and Yo-S. Ho, „Video Test Material of Outdoor Scene”, ISO/IEC JTC1/SC29/WG11, MPEG/ M15371, Archamps, France, April 2008.
- [6] M. Tanimoto, T. Fujii, N. Fukushima, “1D Parallel Test Sequences for MPEG-FTV”, MPEG M15378, Archamps, France, April 2008.
- [7] J. Zhang, Ri Li, H. Li, D. Rusanovskyy, M.M. Hannuksela, “Ghost Town Fly 3DV sequence for purposes of 3DV standardization”, ISO/IEC JTC1/SC29/WG11, Doc. M20027, Geneva, Switzerland, March 2011.
- [8] D. Rusanovskyy, P. Aflaki, M. M. Hannuksela, “Undo Dancer 3DV sequence for purposes of 3DV standardization”, ISO/IEC JTC1/SC29/WG11, Doc. M20028, Geneva, Switzerland, March 2011.