Review of Noise Reduction Methods and Estimation of their Effectiveness for Medical Endoscopic Images Processing

Natalia Obukhova, Alexandr Motyko, Alexandr Pozdeev
Saint Petersburg Electrotechnical University "LETI"
Saint Petersburg, Russia
natalia172419, motyko.alexandr@yandex.ru, puches4@gmail.com

Boris Timofeev
Saint Petersburg State University of Aerospace Instrumentation
Saint Petersburg, Russia
timofeev-boris36@mail.ru

Abstract—**The paper is devoted to the analysis of the efficiency of different algorithms for noise reduction on medical images in endoscopy. The algorithms of adaptive median filtering, KNN filtering, block matching filtering (Non local means and 3DBM) are analyzed. The conclusions and the recommendations for the implementation of these algorithms in clinical decision support systems in endoscopy are given.**

I. Introduction

Clinical Decision Support Systems (CDSS) are the modern trend of medical video systems development. Systems of this type realize the integration of automatic image analysis results with the results obtained by the physician, and also use the information of system database. This integration of visual and automatic analysis allows the sensitivity and specificity of the diagnosis to be higher than in the case of diagnosis by the physician alone, or independently by the system.

Thus, the visual analysis, realized by physician is important part in CDSS. This statement demands to ensure high quality of the images for physician and, as a consequence, to realize high efficiency of visual analysis.

The objects of interest in this paper are CDSS using endoscopic images. Endoscopic images is a very wide class of medical images obtained by examining various organs and body cavities by via the special optoelectronic device - endoscope. The types of endoscopy are: gastroscopy - examination of stomach, colposcopy - examination of cervix, laparoscopy examination of abdominal cavity, etc. There are more than 30 types. During endoscopy, physician performs a visual examination of organ surface and makes decision about diagnosis, possible biopsy; also he can perform surgical interventions (endosurgery). The wide application of endoscopic examinations and the complexity of the tasks solved during endoscopic examination determine the high relevance of new processing and analysis methods developing for endoscopic images. New effective enhancement methods of image for visual analysis must be used in clinical decision support system based on endoscopy images, but also they are useful in any modern endoscopy device.

The need to form a high-quality image of the organ under investigation requires the implementation of pipelining of the video stream. The main procedures of pipeline are: noise reduction, brightness correction, contrast correction, blur compensation and sharpening.

The crucial point for quality representation and processing of endoscopic images is noise reduction. Because the noise is one of the most common problems in medical image processing. Even a high-resolution medical video is bound to have some noise in it. As for the endoscopic video it is a constant problem, because the lack of light forces to use high gain that leads to high level of noise.

The various techniques for noise reduction can be included in endoscopy images processing pipeline for CDSS.

It is necessary to evaluate the possibility and effectiveness of modern noise reduction methods for the medical endoscopic images quality improving. For this aim the basic requirements for methods and the features of the solved task must be formulated.

In medical endoscopic images, rather big light areas and rather dark areas can simultaneously be present with high probability (see Fig. 1).

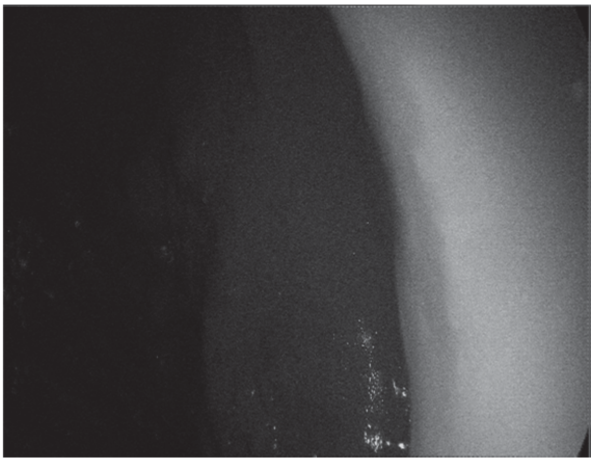

Fig. 1. The simultaneously presence of light and dark areas on medical image

It is connected with the complex shape of the object under endoscopic investigation. Thus, in the case of endoscopic images, local non-linear enhancement methods must be used to effectively correction the brightness and contrast characteristics, and for edges sharpening. Examples of such

ISSN 2305-7254

methods are the Multi Scale Image Contrast [1], Adaptive and Integrated Neighborhood Dependent Approach for Nonlinear Enhancement (AINDANE) [2], Locally Tuned Nonlinear Technique for Color Image Enhancement (LTSNE) [3]. All these methods are based on the operation of image differentiation and their application significantly increases the noise level in the image, especially in areas of the image with low detail.

Our investigation showed that the signal-to-noise ratio of the image processed by LTSNE and AINDANE methods deteriorates on average by 10-15 dB. Figures are obtained for low detail fragments of the image. This phenomena determines the high requirements for the effectiveness of noise reduction - the procedure of the pipeline, which imposes the procedure of brightness and contrast correction. So the first requirement for noise reduction method for endoscopic image processing pipeline is the high effectiveness allowing to use nonlinear methods of contrast enhancement without significant artifacts.

The implementation of processing pipeline in endoscopic systems demands to realize enhancement procedures with very high speed. For the modern endoscopic system, the following performance must be ensured: all pipeline procedures should be implemented in real time for FHD video with FPS more than 100 frames per second and for 4K video with FPS 30 frames per second. So the high speed of processing is the second requirement for noise reduction method.

There are two additional special requirements, besides the high effectiveness and speed, for the method of noise reduction in medical endoscopic image. They are next:

- it is necessary to preserve all relevant information present in the raw image: the image after noise removing should not have smoothed area borders and texture changes (blur);
- adaptability to the main feature of endoscopic video, such as sufficiently low level of detail and the presence of significant areas without detail in initial endoscopic images.

II. Methods and Algorithms for Noise Reduction on Medical Endoscopic Images

Modern methods of noise reduction can be divided into two groups. The methods of the first group work with the signal of image. These methods include smoothing filters, morphological filters, median and rank filters, the adaptive median filter, the K nearest neighbor filter (KNN), and the non-local mean filter (NLM). The methods of the second group are based on the decomposition of the image signal into a basis, followed by decomposition transformants processing.

As a transformation, a discrete cosine transform (DCT), a discrete Fourier transform (DFT), and wavelet-based processing are used. The actual trends in the development of this group methods are aimed at eliminating detail suppression and artifacts caused by the basic decomposition. New methods of this group such as 3D block matching method use for the processing of static images a three-dimensional DCT and a sliding window.

In the first group, the adaptive median filter, the KNN filter, and the NLM filter are the most effective in terms of noise reduction and the original image preservation. In the second group, the 3D block matching method based on effective filtering in 3D transform domain by combining sliding-window transform processing with block-matching is one of the most promising.

Thus, to study the methods of noise reduction for their using in endoscopic pipeline were chosen next methods: adaptive median filter, KNN filter, NLM filter, bilateral filter and 3D block matching method (BM3D).

An adaptive median filter is the improving of median filter aimed to reduce impulse high density noise while preserving details. Unlike the usual median filter the size of the analyzed area (the size of filter structural element) is variable. The basis of the filter is the analysis of median value *Lmed* found in the processed area. If the condition

$$L_{min} < L_{med} < L_{max} \tag{1}$$

where *Lmin* and *Lmax* - the maximum and minimum brightness value of the analyzed area is true, then the filter response is formed.

The filter response is *Lc* - the brightness value of pixel in the center of processed area if the condition

$$L_{min} < L_{c} < L_{max} \tag{2}$$

is true, otherwise the filter response is *Lmed*. If the found median value *Lmed* does not satisfy the condition (1), then the size of the processed area is increased.

If the size of processed area has reached the maximum possible, then the filter response is the value *Lc*.

The KNN filter was designed to reduce white noise and it is basically a more complex Gaussian blur filter. Let $\mathbf{u}(x)$ - the original noisy image in the point x, and $f(y)$ is the result produced by the KNN filter with parameters h and r in some point y of image. Ω is the spatial neighborhood of a certain size surrounding pixel under analysis, i.e. Ω is a block of pixels with size NxN, so that pixel under analysis (point y) is the central pixel of Ω. Then the filter response can be obtained next way:

$$f(y) = \frac{1}{C(x)} \sum_{\Omega(x)} \mathbf{u}(y)\, e^{-\left(\frac{(y-x)^2}{r^2}\right)}\, e^{\frac{(\mathbf{u}(y)-\mathbf{u}(x))^2}{h^2}} \tag{3}$$

where $C(x)$ – normalizing coefficient.

The NLM is the modern algorithm for image denoising. Unlike "local mean" filters, which take the mean value of a group of pixels surrounding a target pixel to smooth the image, non-local means filtering takes a mean of all pixels in the image, weighted by how similar these pixels are to the target pixel. This results in much greater post-filtering clarity, and less loss of detail in the image compared with local mean algorithms [4].

NLM [5], [6] has the following background. The main idea is to define the filtered pixel value by averaging the pixels with the similar neighborhoods (see Fig. 2).

Let $\mathbf{f}(p)$ be the restored image at the point p, $\mathbf{u}(q)$ – unrestored image at the point q, $\boldsymbol{\Omega}$ is the area of the image, $\mathbf{S}(p)$ and $\mathbf{S}(q)$ – the spatial neighborhoods of a certain size surrounding pixels p and q. It will be considered as a block of pixels of size NxN, so that, for

example, q is the central pixel of $\mathbf{S}(q)$. Then in discrete form:

$$\mathbf{f}(p) = \frac{1}{\mathrm{C}(p)} \sum_{\mathrm{q} \in \Omega} \mathbf{u}(q) \exp -\left(\frac{|\mathbf{S}(p) - \mathbf{S}(q)|^2}{h^2}\right) \tag{4}$$

where C(p) – is a normalized coefficient.

$$C(p) = \sum_{q \in \Omega} \exp -\left(\frac{|\mathbf{S}(p) - \mathbf{S}(q)|^2}{h^2}\right) \tag{5}$$

h is the filtering parameter (i.e., standard deviation) and $\mathbf{S}(p)$ is the local mean value of the image point values surrounding p. The second addendum of color distance $\mathbf{S}(q)$ defined as

$$\mathbf{S}(q) = \frac{1}{R(q)} \sum_{i \in R(q)} \mathbf{u}(i) \tag{6}$$

where $R(q)$ – is the area of $\mathbf{S}$. Thus, the square of $\mathbf{S}(p)$ - $\mathbf{S}(q)$ represents a normalized sum of the absolute differences between blocks around pixel $\mathbf{f}(p)$ and around pixel $\mathbf{u}(q)$.

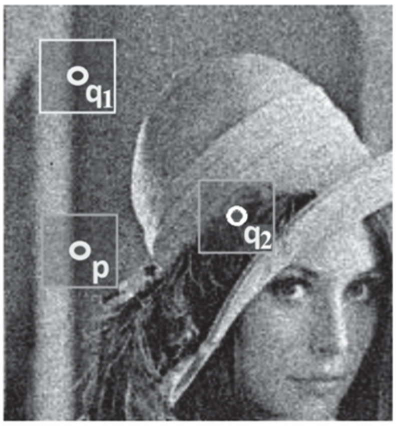

Fig. 2. The illustration to NLM algorithm. The pixel q-white influences on the averaging for pixel p more than q-black, because its similarity is bigger

The 3D block matching method [7] is a new approach to noise reduction developed for static images. It is based on the using of a three-dimensional DCT. The most often digital image processing uses a two-dimensional DCT applied to a square block (for example, an 8 by 8-pixel block is used in the JPEG standard). To ensure the possibility of applying a three-dimensional DCT in 3D block matching method, each processing block is associated with an array of blocks. The array includes image blocks that have a high level of correlation with the block being processed. For the formed array, a three-dimensional DCT is performed, followed by threshold processing of decomposition transformant. Threshold processing of decomposition transformants provides noise reduction. The described procedure is repeated for all blocks of image, and the blocks processing is performed in accordance with the principle of the sliding window. The resulting brightness estimation for each pixel of the processed image is formed as a weighted average, found over all blocks that overlap in a given pixel.

To find the image blocks $Z_{\mathbf{x}i}$, the most similar to the block being analyzed $Z_{\mathbf{x}R}$ the procedure of block matching is used. The procedure of block matching is well developed in in the video compression standards for the motion compensation. The original image has noise component, which makes it difficult to realize the block matching according to correlation function. To reduce the effect of noise, the correlation function estimating the measure of blocks similarity is formed on the basis of two-dimensional DCT transformants with threshold processing. In accordance with this, the correlation function has the next form:

$$d(Z_{\mathbf{x}R}, Z_{\mathbf{x}i}) = \left\| \Psi\left(\mathbf{T}_{2D}(Z_{\mathbf{x}R}, \mu)\right) - \Psi\left(\mathbf{T}_{2D}(Z_{\mathbf{x}i}, \mu)\right) \right\|_2 \tag{7}$$

$$\mu = \lambda_{thr2D} \sigma \sqrt{2\log(N^2)} \tag{8}$$

where and $Z_{\mathbf{x}R}$ and $Z_{\mathbf{x}i}$ are square blocks of the image with the coordinates of the upper left corner at the points $\mathbf{x}R$ and $\mathbf{x}i$, respectively; $\mathbf{T}_{2D}$ - operator of two-dimensional DCT; the norm is Euclidian. The operator determines the threshold processing of the two-dimensional DCT transformants:

$$\Psi_{2D}(\lambda, \lambda_{thr}) = \begin{cases} \lambda, if\ |\lambda| < \lambda_{thr} \\ 0, otherwise \end{cases} \tag{9}$$

As a result of correlation matching, a set of blocks S_x with different value of the similarity measure d ($Z_{\mathbf{x}R}$, $Z_{\mathbf{x}i}$) with the processed block is formed. For further analysis, select the set of blocks for which the similarity measure value is less than the specified threshold value.

$$S_x = \{x \in X, d\ (\mathrm{Z}\mathbf{x}R, \mathrm{Z}\mathbf{x}i) < \tau_{match}\} \tag{10}$$

The blocks included in the three-dimensional array for processing are ordered according to similarity measure value d ($Z_{\mathbf{x}R}$, $Z_{\mathbf{x}i}$). Three-dimensional DCT is applied to the obtained three-dimensional array. To reduce the noise level, the resulting transformants of three-dimensional are subjected to threshold processing. Then the inverse DCT transformation is performed.

$$\hat{\mathbf{Y}}_{\mathbf{x}R} = \mathbf{T}_{3D}^{-1}\left(\Psi\left(\mathbf{T}_{3D}(\mathbf{Z}_{\mathbf{x}R}), \lambda_{thr3D} \sigma \sqrt{2\log(N^2)}\right)\right) \tag{11}$$

where $\hat{\mathbf{Y}}_{\mathbf{x}R}$ the square block of the image after the inverse DCT transformation, $\mathbf{T}_{3D}$, $\mathbf{T}_{3D}^{-1}$- the operator of the three-dimensional DCT and inverse DCT; Ψ - the operator determines the threshold processing of DCT transformants, - λ_{thr3D} - fixed threshold parameter.

For each block obtained after applying the inverse three-dimensional DCT, the weighting coefficients $w_{\mathbf{x}R}$ are calculated. Its value is determined by the number of non-zero transformants N_{NzT} after the threshold processing.

$$w_{\mathbf{x}R} = \begin{cases} \dfrac{1}{N_{NzT}}, if\ N_{NzT} \geq 1 \\ 1, otherwise \end{cases} \tag{12}$$

After performing the described procedures for all image blocks, each pixel of the initial image will be associated with a set of brightness estimations. This is ensured by overlapping blocks (principle of sliding window). To calculate the result pixel brightness value on the restored image, the obtained estimations are averaged taking into account the found weight coefficients w_{xR}.

The basic approach can have extension with Wiener filtering. The linear Wiener filter replaces the nonlinear hard-thresholding operator on the step of denoising in three-dimensional transform domain.

The bilateral filter is well known technique for image denoising. The main idea is to replace the intensity of each pixel with a weighted average of intensity values from nearby pixels. The weights are based on a Gaussian distribution. The weights depend not only on Euclidean distance of pixels, but also on the radiometric differences (e.g., range differences, such as color intensity, depth distance, etc.) [1].

Bilateral filter has a strong advantage – it preserves sharp edges. Despite this fact, this filter is not suitable for medical image processing, because of its disadvantages, that are crucial for medical observing process:

- staircase effect – intensity plateaus that lead to images appearing like cartoons;
- gradient reversal – introduction of false edges in the image [9].

Nevertheless, we included this filter to our research – because of its popularity it was very interesting to compare its effectiveness with other algorithms.

III. THE RESULTS OF RESEARCH

The methods discussed above were implemented in special software and tested on real endoscopic images. The images were divided in three groups:

- initial images – whole endoscopic pictures, obtained by sensors in various medical devices. They have various resolution according to the specific sensor. On the Fig. 3 and 4 the examples of images with relative high resolution (1280x1024) are shown. They were obtained with colposcopic device;
- the set of images with rich number of details obtained by cropping initial images (see Fig. 6);
- the set of images with low number of details obtained by cropping initial images (see Fig. 7);

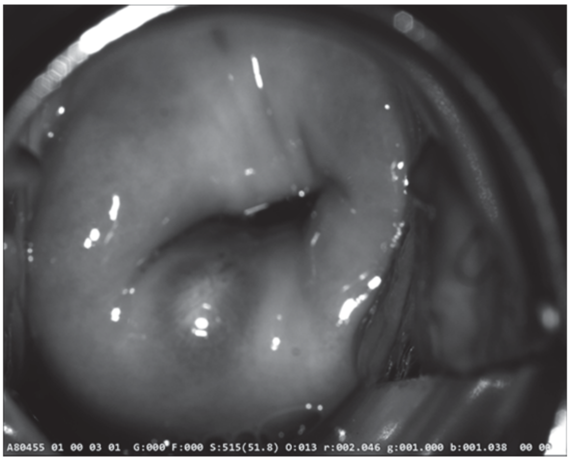

Fig. 4. Another example of the image with high resolution

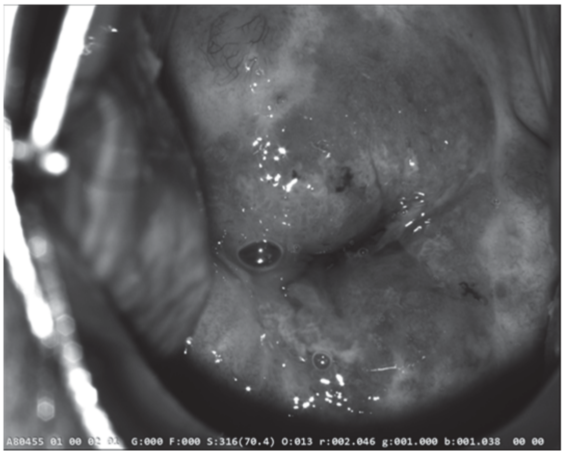

Fig. 4. Another example of the image with high resolution

The images from video gastroenteroscopy are presented on Fig. 5 They have relative small resolution (320x320)

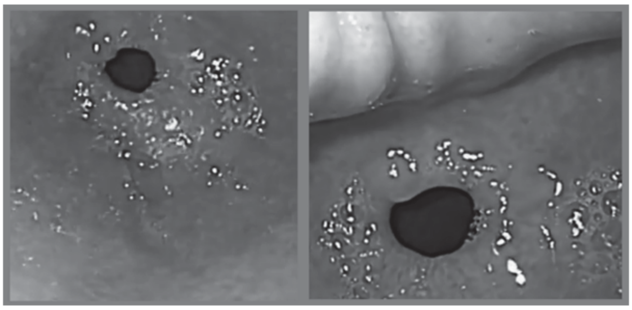

Fig. 5. The images from gastroenteroscopy with small resolution

On the Fig. 6 the samples from the set with high detailed fragments are represented.

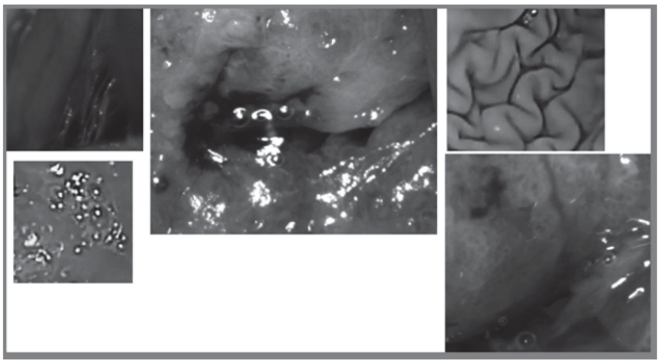

Fig. 6. The fragments with rich number of details

On the Fig. 7 the examples from the set with low detailed fragments are shown.

The reason of this dividing was to check the performance of the denoising algorithms in various circumstances.

The low detailed images (or large areas of images) are typical case in endoscopic observations of the organs with smooth tissues. That's why it is very important for denoising algorithms to cope with this kind of flat surfaces.

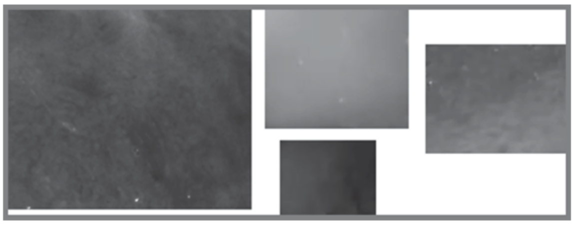

Fig. 7. The fragments with low number of details

However, the vascular patterns, different wrinkles and other high detailed parts of images represent often the most interesting part of the picture for physician. Of course the algorithm for noise reduction should do its best on these fragments.

In our research we used the described three groups of images (30 different images in each group) with added Gaussian noise with three different standard deviations (σ = 10, 20, 30 for 24 bits RGB images). On the Fig. 8 the image with different level of noise is shown.

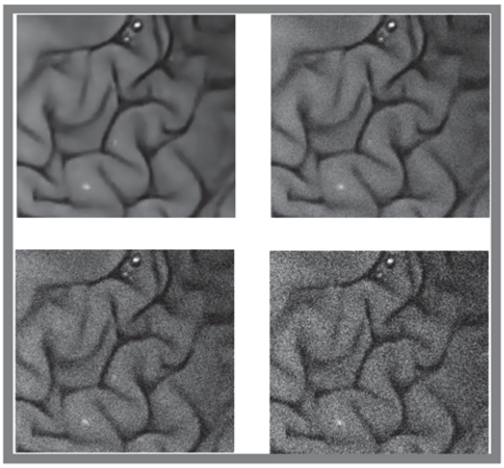

Fig. 8. The image with different presence of noise. Top row left to right: exact and σ = 10, bottom row left to right σ = 20, σ = 30

The Tables I, II and III represents the results. As a measure we used the standard PSNR metric.

TABLE I. THE RESULTS OF NOISE REDUCTION ALGORITHMS FOR INITIAL IMAGES (DB)

σ	noised	adaptive median	bilateral	KNN	NLM	BM3D
10	28,23	33,69	38,16	35,74	34,84	41,47
20	22,31	29,15	29,54	31,04	31,6	38,09
30	18,94	26,03	23,01	26,49	27,79	35,95

TABLE II. THE RESULTS OF NOISE REDUCTION ALGORITHMS FOR FRAGMENTS WITH HIGH DETAILS

σ	noised	adaptive median	bilateral	KNN	NLM	BM3D
10	28,17	32,32	35,87	32,9	32,36	39,11
20	22,23	28,33	28,77	29,66	30,13	35,58
30	18,86	25,44	22,65	25,76	26,93	33,52

TABLE III. THE RESULTS OF NOISE REDUCTION ALGORITHMS FOR FRAGMENTS WITH LOW DETAILS

σ	noised	adaptive median	bilateral	KNN	NLM	BM3D
10	28,16	34,86	41,02	35,49	35,51	43,56
20	22,2	29,4	29,97	31,07	32,06	40,94
30	18,8	25,97	22,94	26,49	28,12	39,21

The Fig. 9, 10 and 11 illustrates the data from tables.

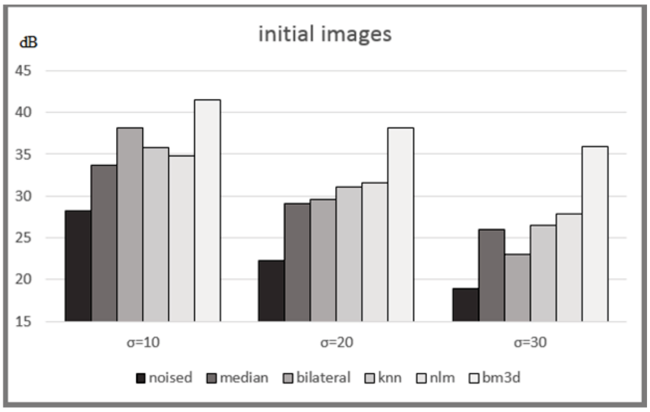

Fig. 9. The averaged performance of the algorithms for noise reduction for the set of initial images

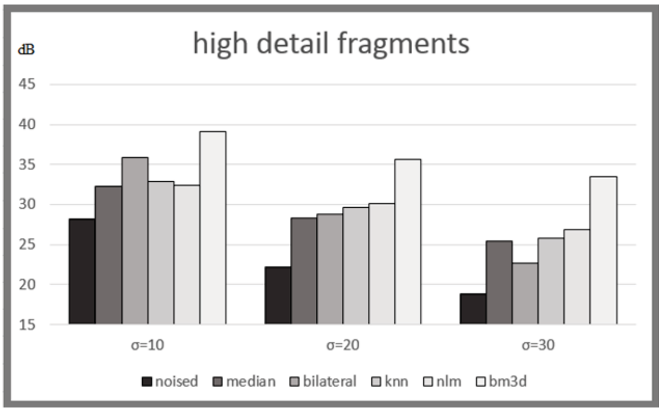

Fig. 10. The averaged performance of the algorithms for noise reduction for the set of high detailed images

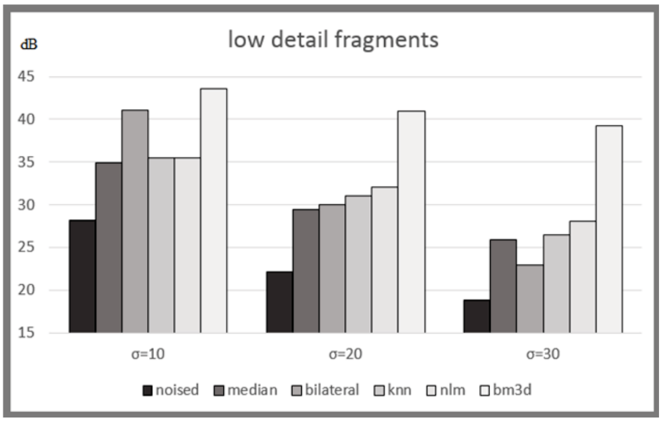

Fig. 11. The averaged performance of the algorithms for noise reduction for the set of low detailed images

On the Table IV and the corresponded Fig. 12 the overall data for all noise levels from previous tables is shown.

TABLE IV. THE OVERALL RESULTS OF NOISE REDUCTION ALGORITHMS (DB)

images	noised	adaptive median	bilateral	KNN	NLM	BM3D
initial	23,05	30,07	31,31	31,02	31,9	41,24
high detailed	23,16	29,62	30,24	31,09	31,41	38,5
low detailed	23,09	28,7	29,1	29,44	29,81	36,07

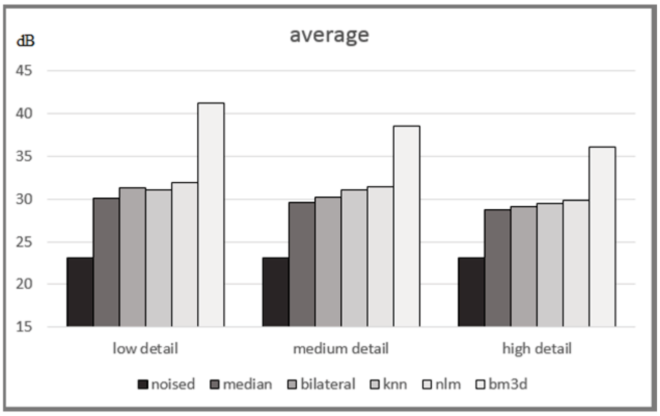

Fig. 12. The overall performance of the algorithms for noise reduction for all images

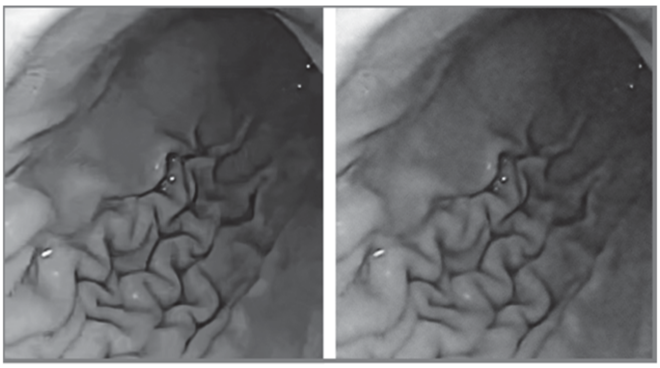

Fig. 13. The result of BM3D algorithm (left) and NLM (right)

IV. Conclusion

The main results of the research are the following.

1. For general video processing pipeline for endoscopic CDSS system it is not useful to implement some complex algorithms with large processing time. The well-known, relative simple and fast adaptive median filter provides very close results for the most cases of endoscopic images.

2. For the case, when the presence of noise is very high the implementation of the modern BM3D algorithm can make sense. But the developer should keep in mind the specific of BM3D algorithm. It is relative slow, has many parameters and can provide some artifacts on the image. The implementation of this algorithm in video endoscopic system should be guided by physicians, who will be able to estimate the allowable level of imposed distortions

3. The approach of complex technique: to use resource-consuming BM3D on some important (high-detailed) fragments and to use relative quick adaptive median on large flat areas seems very profitable.

Acknowledgement

The work was supported by the Russian Foundation for Basic Research, grant No. 17-07-00045.

References

[1]. V. Vonikakis and I. Andreadis, *Multi-Scale Image Contrast Enhancement,* Proceedings 10th Intl. Conf. on Control, Automation, Robotics and Vision Hanoi, Vietnam, 17–20 December 2008.

[2]. R. Maini and H. Aggarwal, *A Comprehensive Review of Image Enhancement Techniques*, Journal of computing, volume 2, issue 3, march 2010.

[3]. Cheng-Yi Yu, Chi-Yuan Lin, Sheng-Chih Yang and Hsueh-Yi Lin, *Eight-Scale Image Contrast Enhancement Based on Adaptive Inverse Hyperbolic Tangent Algorithm*, Algorithms 2014, v.7, pp.597-607, www.mdpi.com/journal/algorithms.

[4]. Buades, Antoni (20–25 June 2005), *A non-local algorithm for image denoising,* Computer Vision and Pattern Recognition, 2005. 2: 60–65.

[5]. A. Buades, B. Coll, and J. Morel, *Neighborhood Filters and PDE's*, Technical Report 2005-04, CMLA, 2005.

[6]. Qi J. *An Improved Approximate K-Nearest Neighbors Nonlocal-Means Denoising Method with GPU Acceleration*. In: Yang J., Fang F., Sun C. (eds) Intelligent Science and Intelligent Data Engineering. IScIDE 2012. Lecture Notes in Computer Science, vol 7751. Springer, Berlin.

[7]. M. Maggioni and A. Foi, *Nonlocal transform-domain denoising of volumetric data with groupwise adaptive variance estimation*, Proc. SPIE Electronic Imaging 2012, Computational Imaging X, 8296-22, Burlingame (CA), USA, January 2012. DOI:

[8]. Banterle, F.; Corsini, M.; Cignoni, P.; Scopigno, R., *A Low-Memory, Straightforward and Fast Bilateral Filter Through Subsampling in Spatial Domain,* Computer Graphics Forum. 31 (1), 2011.

[9]. He, Kaiming; Sun, Jian; Tang, Xiaoou, *Guided Image Filtering* (PDF), Retrieved 7 May 2017.