# Automatic Image Enhancement from a Mobile Synthetic Vision System

Dmitry A. Kolchaev, Aleksey I. Efimov, Dmitry I. Ustukov
Department of Electronic Computers
Ryazan State Radio Engineering University
Ryazan, Russia
kolchaevda@mail.ru

*Abstract*— **The work discusses basic image contrasting algorithms and noise compensation methods, an algorithm for estimating image quality based on an integral quality indicator, as well as approaches for estimating noise values in images. The results of contrasting algorithms work (with a numerical estimation) and the most prominent image filtering methods are presented. A description is given for an automatic image enhancement algorithm from a mobile synthetic vision system based on a choice of contrasting algorithms using an integral quality indicator, and a space-time filter using a pyramidal version of Lucas-Kanade optical flow algorithm is proposed.**

***Keywords - image enhancement, mobile synthetic vision system, contrasting algorithms, noise compensation algorithms, image quality estimation, noise estimation, optical flow.***

## I. INTRODUCTION

The current state of mobile synthetic technical vision systems (MSTV) development allows obtaining sufficiently high-quality information about the surrounding space in a wide range of external conditions (adverse weather conditions: rain, snow, fog; twilight) [1]. However, the problem of obtaining high-quality images in low light (at night) remains relevant. To solve this problem, contrasting algorithms and noise reduction are used, which allow obtaining an image suitable for further processing and visualization.

Existing contrasting algorithms give different results which depend on the original data set [2]. This leads to the need to choose a contrasting algorithm from a certain set, taking into account MSTV work at night [3].

The larger part of the known approaches to image filtering provide poor noise reduction or results in blurring of significant areas in images [4], or are suitable for use only in stationary video surveillance systems.

The above problems make it necessary to search for new methods of contrasting and noise filtering in video sequences from moving MSTVs, and to form a complex processing algorithm suitable for use in various external conditions.

The work was performed as part of the grant of the President of the Russian Federation for young Russian scientists state support MK-661.2018.9, grant agreement No. 075-02-2018-683 of 11/16/2018.

978-1-7281-1740-9/19/$31.00 ©2019 IEEE

## II. BASIC CONTRASTING ALGORITHMS

As the set used for the selection problem, the most prominent contrasting algorithms were selected, showing different results on the same set of input data.

- *Linear contrasting* - linear stretching of the processed image histogram is performed. Histogram sections with frequency values of more than 5% of the maximum value on the entire histogram are stretched proportionally. The calculation of the new brightness for the pixel is made using the following formula [5]:

$$I_{new}(x,y) = \frac{(I(x,y) - \Delta\min_j)\cdot L_j}{\Delta\max_j - \Delta\min_j},\ j \in \overline{1,k}, \quad (1)$$

where $I(x,y)$ - pixel brightness of the input image, $L_j$ size of the new brightness range for the *j*-th interval (brightness range of pixels on the histogram to be stretched), $\Delta\min_j$, $\Delta\max_j$ - left and right border of the *j*-th interval on the histogram of the source image, *k* - the number of intervals.

- *Histogram equalization* - non-linear histogram stretching is performed, at which the source image real brightness range (from the minimum to the maximum intensity value) is displayed on the [0, b] range. This ensures the “alignment” of the number of image pixels with different brightness values. The conversion formula is as follows [5, 6]:

$$I_{new}(x,y) = \frac{\sum_{i=0}^{I(x,y)} H(i)}{\sum_i^b H(i)}. \quad (2)$$

In cases where all gradations are more or less evenly present in the [$I_{min}$, $I_{max}$] range, the visual effect of equalization is comparable to the effect of histogram normalization. However, in the case when a significant part of brightness gradations is absent, equalization allows using the [0, *b*] range more evenly for a more contrast display of the gradations being presented in the image.

- *Multi Scale Retinex with Color Restoration (MSRCR)* is a variation of Single Scale Retinex (SSR) algorithm. The idea of SSR algorithm is to obtain an approximate luminance map by low-pass filtering using the formula $\bar{I}(x,y) = G \cdot I(x,y)$ where $G$ - Gaussian filter [7]. Then the image brightness is restored by the formula:

$$\bar{r}(x,y) = \log(I(x,y) - \log(\bar{I}(x,y)). \quad (3)$$

The difference between the Multi Scale Retinex (MSR) algorithm and SSR is that several variants $\bar{I}(x,y)$ are calculated using the $G$ operator with a different standard deviation parameter. The brightness restoration is reduced to the formula

$$MSR = w_1 SSR_1 + w_2 SSR_2 + ... + w_n SSR_n,$$

and also $w_1+w_2+...+w_n=1$. In practice, $n = 3$ is usually taken.

- *Gamma Correction* - Correction of brightness function $I(x,y)$. The conversion formula is:

$$I_{new}(x,y) = c \cdot I(x,y)^{\gamma},$$

where $c$ is the scaling coefficient, $\gamma$ - gamma value.

- *Logarithmic correction* is made according to the following formula

$$I_{new}(x,y) = c \cdot \log(1 + I(x,y)).$$

## III. Image Quality Estimation Algorithm

The task for image quality estimation is multicriterial; therefore, an additive integral quality criterion $K$ of the following shape is introduced:

$$K = \sum_{i=1}^{p} \beta_i f_i,$$

where $\beta_i$ - weighting coefficients, $\sum_{i=1}^{p} \beta_i = 1$ - normalization condition $K$, $f_i$ - partial normalized criteria, $p$ - the number of partial criteria.

The main difficulty in applying partial indicators is the choice of weights, which take into account the effect of the corresponding partial indicators on the generalized criterion as a whole. To select the values of these coefficients, the method of expert estimates is usually used. Initial weights can be determined by the Fishburn criterion [1].

As a result, for a generalized numerical image quality estimation, an integral quality indicator (*IQI*) is used [1, 5]:

$$IQI = 0{,}33\bar{L}_n + 0{,}27\sigma_n + 0{,}20K_n + 0{,}13N_n + 0{,}07\varepsilon_n.$$

Where $\bar{L}_n$ - average image brightness, $\sigma_n$ - RMS image brightness deviation, $K_n$ - contrast indicator, $N_n$ - number of information levels, $\varepsilon_n$ - image entropy.

A visual comparison of the basic contrasting algorithms result is presented in Figure 1.

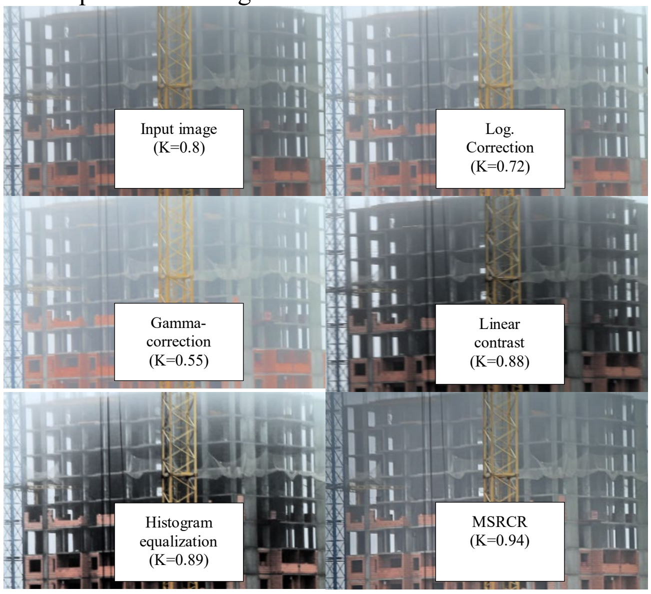


Figure 1. A visual comparison of the contrasting algorithms work results on the image from the test sequence

The highest value of the integral quality indicator ($K$ = 0.94) in Figure 1 corresponds to the MSRCR algorithm, the lowest — to the gamma correction algorithm.

## IV. Noise Compensation Algorithms

### A. Noise estimation

For noise estimation, a classic peak signal-to-noise ratio (*PSNR*) is usually used. To determine it, the mean square error (*MSE*) is used. *MSE* and *PSNR* are calculated using the following formulas:

$$MSE = \frac{1}{WH} \sum_{y=0}^{H-1} \sum_{x=0}^{W-1} \left| I(x,y) - GT(x,y) \right|^2,$$

$$PSNR = 10 \log_{10} \left( \frac{I_{\max}}{MSE} \right),$$

where $W$, $H$ is the width and height of the image in pixels, $x$, $y$ are the pixel coordinates, $I$ is a noisy image, $GT$ is the reference image, $I_{\max}$ - the maximum pixel brightness (with a bit depth of 8 bits $I_{\max} = 255$).

The use of *MSE* and *PSNR* estimates is only possible when the reference image (*GT*) is known. For individual frames, *GT* can be obtained by changing the shooting conditions on a fixed scene (exposure, interframe processing) [8]. In this case, estimation can be considered fairly objective, since there is information about the reference image. Regarding the noise estimation on the frame sequence (video image) for a mobile sensor that is in motion, it is difficult to create a representative data set for *PSNR* based estimation. Although the work [9] is known, where special masks are used to estimate the noise, which allow finding the noise estimate in local areas of the image, and on the basis of these data, construct a generalized noise estimate without using *GT*. This approach can be used to estimate the noise in the video image if the nature of the noise is known in advance (in order to use an appropriate set of masks).

In practice, for mobile sensors, the nature of noise varies and depends on the shooting conditions (internal camera parameters: matrix quality, exposure, matrix resolution) and external conditions (light, object composition, dynamics of scene and sensor movement the scene and the sensor). Noise compensation on images from mobile sensors is necessary, firstly, to visually improve the image quality, and secondly, as a preliminary stage in object recognition tasks. In the second case, it is possible to estimate the noise, based on the estimation of the recognition result, on test data sets, however, it is necessary to take into account the error introduced by the recognition algorithm itself, its robustness. As for the noise estimation in the task of visual image improvement, then the best way is peer review.

Figure 2 presents a comparison of work results of the known spatial filters, as well as the characteristics of *MSE*, *PSNR* for each result.


a) Noisy image b) Reference image

c) Median filtering d) Gaussian filter

e) Box filter f) Bilateral filter

Figure 2. Comparison of the results of spatial filtering by known algorithms on the image from the test set [8]

As can be seen from Figure 2, most filters lead to a “blurring” of objects of interest (for example, text), which is undesirable in real synthetic vision systems

## B. *Spatio-temporal filtering for the mobile platform*

Spatial filters use only the information that is available in the current frame, which limits filtering possibilities. For systems that form a sequence of frames (video), temporal or spatio-temporal filters have a great advantage. They allow compensating the error (noise) due to the many available measurements (frames). One of the most prominent technologies of spatio-temporal filtering is the 3DNR technology from Vivotek [10], which is used in stationary video surveillance systems.

Since most of the time the camera is stationary, the frame area is divided into fixed sections (background) and moving parts. For fixed sections, temporal filtering is performed on a sufficiently large number of frames, which makes it possible to almost completely suppress the random noise component. The use of temporal filtering for moving objects is difficult since the movement of objects between frames will lead to the appearance of so-called “loops” of objects during normal temporal filtering. Therefore, 3DNR technology uses spatial filters for moving parts in the image, which significantly reduces the quality of filtering in these areas. Also, the use of 3DNR technology is difficult for moving platforms equipped with a surveillance system, since both the background and objects in the frame will almost always be mobile, which will only lead to the work of spatial filters.

The solution to the problem of dynamically changing background, as well as moving objects, is the use of optical flow [11].

Optical flow allows determining the shift (correspondence) of the same points on an image pair (frames taken at neighbouring moments of time) arising from the movement of objects in the field of view of the camera or the movement of the camera itself. There's a number of known methods for finding the optical flow. The least sensitive to noise is Lucas-Kanade algorithm.

The original algorithm does not allow to find large point shifts since the search is limited by the window size. To solve this problem, the pyramidal version of Lucas-Kanade algorithm is used [11].

The main ideas of the algorithm:

1. Using pyramids of images;

2. Transformation of coordinates of images in accordance with the levels of the pyramid;

2. Calculation of gradient matrices;

3. Finding the differences between the images;

4. Calculation of optical flow;

5. Multiple iterations by the pyramid levels (points 2 to 4), and optical flow calculation (points 3 to 4).

Figure 3. Visual representation of a loose optical flow

The calculation of the optical flow over a set of key points is performed iteratively for each point. The result of the work is a set of shift vectors *V*, a visual representation of which is shown in Figure 3.

The proposed filtering algorithm is based on the formation of a set of shift vectors at each frame (using the current frame and the previous one), which allows forming a tuple $P$ consisting of pixel $p^i$ brightness values on the *i*-th frame, (*i-1*)-th frame, ..., (*i*-$n_p$)-th frame, where $n_p$ sets the length of temporal filter:

$$P = (p^{(i-n_p)}, p^{(i-n_p+1)}, ..., p^{(i-1)}, p^{(i)}).$$

Moreover, $p^{(i-1)}_{(x-v_x,\, y-v_y)} = p^{(i)}_{(x,y)}$, where $v_x$ and $v_y$ - coordinates of the shift vector (optical flow) for a pixel $p$ found on the *i*-th frame. Similarly, for $p^{(i-n)}_{(x-v_x,\, y-v_y)} = p^{(i-n+1)}_{(x,y)}$.

Based on the data from the tuple *P*, a new (after filtering) value $\hat{p}^{(i)}$ is formed. Similarly, for the remaining pixels belonging to the image. The median can be used as a temporal filter after ordering *P* tuple.

The use of a median filter in the time area is due to a sufficient sample length of $n_p > 15$ and resistance to random outbreaks, which make significant changes in the result, for example, using the arithmetic average.

The pixel value after filtering $\hat{p}^{(i)}$ is formed as the median from the $P$ tuple. For this, the ordering of elements $p \in P$ occurs in increasing intensity and $p$ is selected with an index equal to $\left[\frac{n_p}{2}\right]$ (average element).

## V. The General Scheme of the Automatic Image Enhancement Algorithm

Based on the approaches proposed in sections 2 and 3, an algorithm was developed for automatically enhancement of images from a mobile synthetic vision system, which is presented in Figure 4.

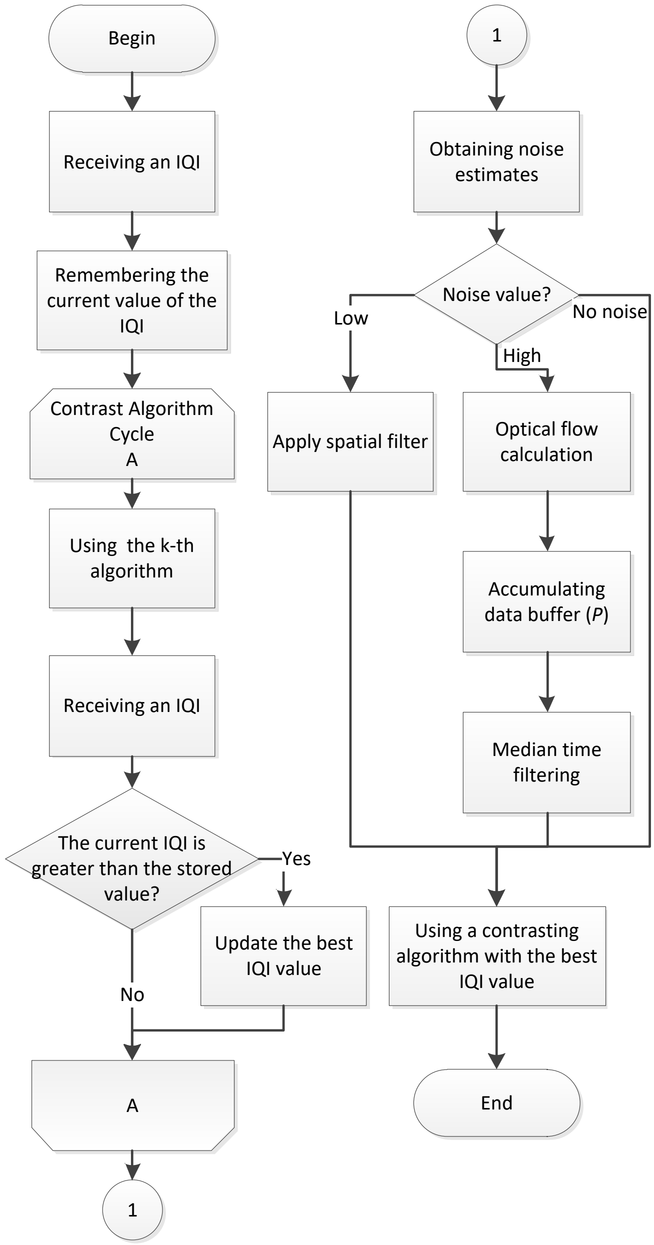


Figure 4. Automatic image enhancement algorithm scheme

The algorithm consists of two stages, the first one - finding the best contrasting algorithm using *IQI*, the second one - applying a spatial or temporal filter depending on the noise level in the image.

To speed up the work of the software implementation, it is proposed to use the phase of finding the best contrasting algorithm after a certain number of frames $t$. $t$ - characterizes the amount of delay, the period through which the system will be adapted to the incoming video sequence. It is also proposed to save the result of processing by the $k$-th algorithm, which currently has the best $IQI$ value, in order to avoid time-consuming, to re-apply the contrasting algorithm.

The proposed approach suggests the possibility of changing and complementing a set of new contrasting algorithms that show the best results.

For noise estimation, the classic version of the intensity comparison of diagonal pixels with the central one (in a certain window) can be used to integrate over the entire image, or the methods proposed in [9] can be used. Depending on the estimation received, the appropriate mode of operation is selected. Noise thresholds are determined experimentally and are set in advance. As a spatial filter, it is recommended to use a bilateral filter, which gives a fairly good result in the quality of the final image or to use the spatial filters proposed by the authors [4, 12]. The pyramidal version of Lucas-Kanade optical flow [11] is used in the “optical flow calculation” block, which forms the data buffer $P$ (a tuple of intensity values of the same point). Based on the tuple $P$, temporal median filtering is applied.

Approbation of the algorithm took place on various video sequences. The main criterion for comparison was the visual estimation of algorithm results by experts. In most cases, with the proposed algorithm, the result of the work was estimated higher in comparison with the methods illustrated in Figures 1 and 2. According to the results of visual comparison on the test video sequence, the spatial median filter demonstrates poor filtering quality, and an increase of the window size leads to a deterioration in image clarity. The proposed approach of spatio-temporal filtering gives a qualitative filtering result, both in static and dynamic areas.

## VI. CONCLUSION

The application of the proposed automatic image enhancement algorithm from a mobile synthetic vision system allows to automatically find and apply the best contrasting algorithm from a predetermined set, based on the above numerical criterion, and also to perform spatial-temporal image filtering. The algorithm provides the ability to change the used contrasting methods, and the image estimation method, due to its modular structure. This allows introducing both new contrasting algorithms and new methods for image quality estimation.

The proposed approach that performs automatic image filtering from a video sequence coming from a mobile synthetic vision system, by linking a series of frames in the time area with optical flow, can significantly improve the quality of the generated video image.

## REFERENCES


[1] L. N. Kostyashkin, M. B. Nikiforov, “Processing of images in aviation systems of machine vision,” - M.: PHYZMATLIT, 2016. – 240p.

[2] Z. Chen, T. Jiang and Y. Tian, “Quality Assessment for Comparing Image Enhancement Algorithms,” IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 3003-3010.

[3] A. I. Efimov, D. A. Kolchaev, M. B. Nikiforov, A. I. Novikov, “Algorithm of Geometrical Transformation and Merging of Radar and Video Images for Technical Vision Systems,” in Proceedings of the 7th Mediterranean Conference on Embedded Computing (MECO), Budva, Montenegro, 2018, pp. 292-295.

[4] D. I. Ustukov, I. A. Orlova, A. I. Efimov, “Features of the implementation of spatial image filters on FPGA,” // News of TSU, technical sciences, 2018, № 2, pp. 195-206.

[5] V. S. Gurov, D. A. Kolchaev, Y. R. Muratov, M. B. Nikiforov, “Adaptive System of Image Processing,” in Proceedings of the 6th Mediterranean Conference on Embedded Computing (MECO), Bar, Montenegro, 2017, pp. 141-144.

[6] R. P. Singh, M. Dixit, “Histogram Equalization: A Strong Technique for Image Enhancement,” International Journal of Signal Processing, Image Processing and Pattern Recognition – 2015 – Vol.8, No.8, 2015, pp. 345-352.

[7] A. B. Petro, C. S. Sbert, J. M. Morel, “Multi scale Retinex,” Image Processing On Line, 4, 2014, pp. 7188.

[8] A. Abdelrahman, S. Lin, M. S. Brown, “A High-Quality Denoising Dataset for Smartphone Cameras,” // IEEE Computer Vision and Pattern Recognition (CVPR'18), June 2018.

[9] Amer, Aishy & Dubois, Eric, “Fast and reliable structure-oriented video noise estimation,” // Circuits and Systems for Video Technology, IEEE Transactions on. 15, 2005, pp. 113 - 118.

[10] Vivotek, “2 Dimensional, 3 Dimensional Noise Redcution in Video Surveillance Cameras,” New Taipei City, Taiwan, Vivotek, White Paper, 2014. [Online]. Available: http://download.vivotek.com/downloadfile/downloads/whitepaper/2dimensional_3dimensional_noise_reduction_in_video_surveillance_cameras_en.pdf

[11] J. Y. Bouget, “Pyramidal Implementation of the Lucas Kanade Feature Tracker: Description of the algorithm,” Microprocessor Research Labs, Intel Corporation (2002).

[12] A. I. Novikov, D. I. Ustukov, “Interference suppressor operators researching in a discrete white noise,” Dynamics of Complex Systems – XXI century, Radiotekhnika, 2017, №1, pp. 26-32.