Comparison of Image Quality Assessment Methods for Multi-Focused Image Fusion

Andrey Noskov, Andrey Priorov

Yaroslavl State University

Yaroslavl, Russia

noskoff.andrey@gmail.com, andcat@yandex.com

Abstract—Image fusion is a process of obtaining one image from multiple. The resulting image carries more information about the photographed scene, than each of the originals. Such an image can be more useful when we deal with human or image processing system. Algorithms that performed this task are used in a wide applying in practical: computer vision, robotics, medicine, forensics, etc. Most popular quality assessment measure for multi-focused image fusion are discribed. Expert image quality assessment experiment was performed. Different kinds of image quality assessment were proposed for scenes with various characteristics.

I. Introduction

Image fusion is used in computational photography. Image fusion is a sub-field of image processing in which two or more images of a scene are combined into a single composite image that is more informative and is more suitable for visual perception and for digital processing. We categorize the fusion methods based on the input data of the fusion process and also based on directly pixels processing. In general, image obtaining methods can be classified into four main groups:

- multi-camera image obtaining;

- image getting via panorama scene;

- at different times survey in order to detect changes between them or to synthesize realistic images of objects;

- with various focal lengths image capturing (multi-focus method).

Multi-focused image is a combination of several images of the same scene taken with different focal lengths. The first and the most important stage of all image fusion techniques is to compute focus value of original images or the parts of them. At once, pixels with greater values of this measurement, when source images are compared, are considered to be in focus and selected as the pixels of the fused image. Once the focus measure is done, there are different fusion rules to fuse the images. One of them is selecting the sharp pixels with maximum of focus value in the spatial domain to Multi-Scale decomposition (MSD) transform image information in the high-frequency via multiscale approximation. In general, there are two main stages of image fusion. The first stage contains a search and assessment of the most interesting areas in image. It can be, for example, the most sharpness or the most brightness areas of image. At the second stage this areas are merging according to selected fusion rule. The ordinary scheme of image fusion is shown in Fig. 1. After image fusion,

it is necessary to perform some automated evaluation of the quality of the resulting image. There are many algorithms of fused image quality assessment, most popular of them will be considered below.
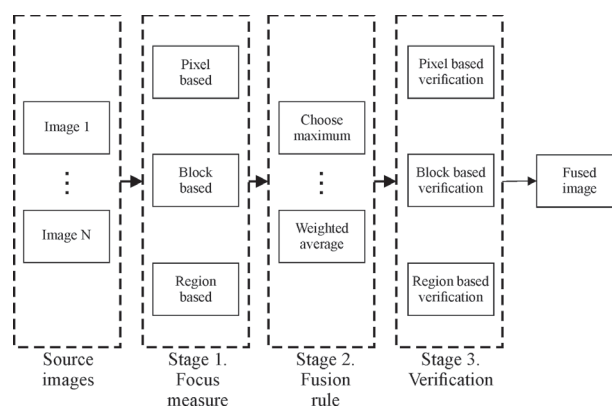

Fig. 1. Scheme of multi-focused image fusion

Estimation the depth of sharpness on the field of image is a key problem in the computational photography in general and the main task of multi-focused images construction in particular. This problem arises at the time of transition from the three-dimensional perception to a two-dimensional projection of the image.

Important thing that needs to be said is that the depth of image can be restored via binocular (trinocular) systems in case of absence of physical interaction with the captured scene as well as with a few shots taken at different settings of monocular system [1], [2], [3], [4].

II. Image quality assessment methods

Image quality describes how the image has deteriorated compared to some reference image. Image processing systems can bring some distortion and artifacts in the resulting images, hence — an assessment of their quality is an important task. All methods of image quality assessment can be divided into two groups: methods based on comparison with reference image and methods without use of the reference image. In the first case, the image is compared with the reference image, which is considered to be the ideal quality [5]. When the reference image is not available — it is necessary to use methods that do not rely on the knowing of the ideal values of the respective pixels. For this referenceless metric can be used [6], [7]. There are many ways to evaluate image

quality without using any standard. The most commonly used evaluation are shown below.

1) Entropy: Entropy — a measure of information content of the message. Increasing the value of image entropy obtained after image fusion, as compared with the source image entropy value indicates that the obtained image carries more information. Entropy is defined as follows:

$$E = -\sum_{i=0}^{G} p(i) log_2 p(i),$$

where G — number of grey levels in image histogram (255 for 8-bit image), $p(i)$ — normalized frequency of $i-th$ gray level. Informativeness of the image is expressed in bits per pixel. It should be noted that the metric is sensitive to noise and other sharp fluctuations in the intensity of pixels [8].

2) Dispersion: This metric is most effective in the absence of noise. It allows you to evaluate the contrast of a fused image. Images with high contrast, have a higher value metric.

$$D = \frac{1}{MN} \sum_{i}^{M} \sum_{j}^{N} [I(i,j) - \mu]^2,$$

where μ — mean intensity value of the image.

3) Spatial frequencies: Spatial frequencies characterize the intensity of the changes taking place in the fused image [9]. The metric is defined as follows:

$$SF = \sqrt{RF^2 + CF^2},$$

where RF CF — vertical and horizontal frequencies, respectively, defined as follows:

$$RF = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=2}^{N} [I(i,j) - I(i,j-1)]^2},$$

$$CF = \sqrt{\frac{1}{MN} \sum_{i=2}^{M} \sum_{j=1}^{N} [I(i,j) - I(i-1,j)]^2}.$$

Entropy, Dispersion and Spatial Frequencies metrics do not use information contained in source images. In general, this metrics are significantly dependent on the content of the scene and can not be used as a universal metric. Thus, these metrics will not participate in further research.

4) Mutual fused information (MI): Let A and B — source images, F — fused image. Mutual fused image is defined as follows:

$$\mathrm{MI}(A,B) = \mathrm{M}(F,A) + \mathrm{M}(F,B),$$

where $\mathrm{M}(X,Y)$ is joint information of X and Y defined as follows:

$$\mathrm{M}(X,Y) = \sum_{x,y} \left(P_{XY}(x,y) \frac{log_2(P_{XY}(x,y))}{P_X(x)P_Y(y)} \right),$$

where $P_{XY}(x,y)$ — the joint probability distribution, $P_X(x)$ and $P_Y(y)$ — the probability distribution on the images X and Y, respectively.

The metric represents similarity of the fused image and the source images. There is also a high metric value corresponds to a higher quality of the algorithm. There are more accurate methods for calculating the similarity between two images based on the idea of calculating the mutual entropy between two images, however, they are generally have higher computational cost.

To calculate the metrics mentioned above, it is necessary to calculate the $p_{i,j}$ — elements of the joint probability density (JPD) of brightness of the image, which can be estimated using an image histogram.

For the two 8-bit single-channel image X and Y, each of them has a size of $M \times N$, values of $p_{i,j}$ can be represented as values of bins of two-dimensional histogram with size 256×256.

$$p_{i,j} = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \begin{cases} 1, & \text{if } I(X_{m,n}) = i \text{ and } I(Y_{m,n}) = j \\ 0, & \text{else} \end{cases},$$

where $I(X_{m,n})$ — pixel intensity at (m,n) of the image X, $I(Y_{m,n})$ — of the image Y respectively. Examples of pairwise joint probability density for the three images are shown at Fig. 2. The histogram is shown in a logarithmic scale by value.

5) $Q^{AB/F}$-measure: The metric proposed by Xydeas and Petrovic at the works [10], [11]. The basic idea behind the metric is the assumption that most of the borders should be transferred from the source images to the fused image. This method uses a Sobel operator for computing strength $g(n,m)$ and orientation $\alpha(n,m)$ of the gradient in each pixel at the image. For the two source images A and B and the fused image F, Sobel operator is used as follows:

$$g_A(n,m) = \sqrt{s_A^x(n,m)^2 + s_A^y(n,m)^2},$$

$$\alpha_A(n,m) = tan^{-1} \left(\frac{s_A^y(n,m)^2}{s_A^x(n,m)^2} \right),$$

where $s_A^y(n,m)$ $s_A^x(n,m)$ — values of vertical and horizontal Sobel patterns, convolute with the area centered at the point with coordinates (n,m) at the image A. The relative values of the strength and orientation of the image gradient between A and F are formed as follows::

$$\left(G_{n,m}^{AF}, A_{n,m}^{AF} \right) = \left[\left(\frac{g_{n,m}^F}{g_{n,m}^A} \right)^M, 1 - \frac{|\alpha_A(n,m) - \alpha_F(n,m)|}{\pi/2} \right],$$

where

$$M = \begin{cases} 1, & \text{if } g_A(n,m) > g_F(n,m) \\ -1, & \text{else} \end{cases}.$$

Information about the borders can be calculated by follows:

$$Q_{n,m}^{AF} = \frac{\Gamma_\alpha}{\left(1 + e^{k_\alpha \left(A_{m,n}^{AF} - \sigma_\alpha \right)}\right)} \frac{\Gamma_g}{\left(1 + e^{k_g \left(G_{m,n}^{AF} - \sigma_g \right)}\right)},$$

where Γ_α, Γ_g, k_α, k_g and σ_α, σ_g — constants determined by the curvature and amplitude of the sigmoid function. Authors at [12] propose to use the following values for these variables:
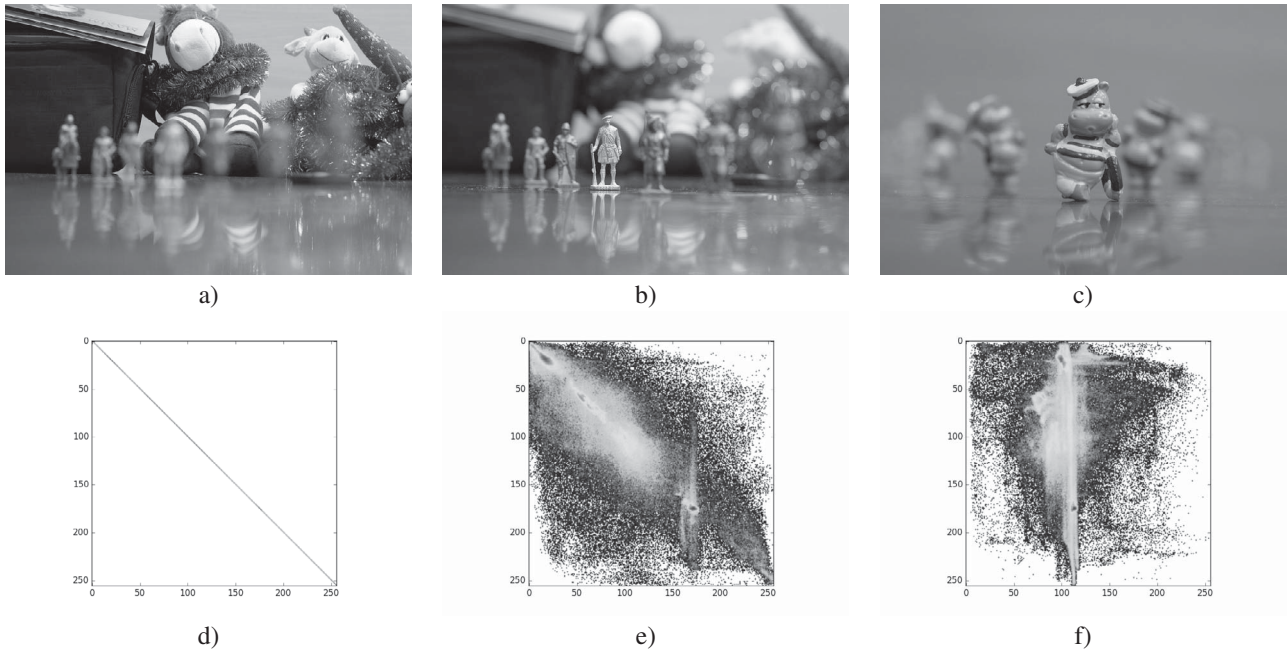
Fig. 2. Joint probability density for different pairs of images: a-c) images; d) for (a) and (a); e) for (a) and (b); f) for (a) and (c)

$\Gamma_\alpha * \Gamma_g = 1$, $k_\alpha = 0.8$, $k_g = 0.7$ and $\sigma_\alpha = 24$, $\sigma_g = 11$. Thus the value of the metric $Q^{AB/F}$ is calculated as:

$$Q^{AB/F} = \frac{\sum_{n,m} \left(Q_{n,m}^{AF} w_{n,m}^A + Q_{n,m}^{BF} w_{n,m}^B\right)}{\sum_{n,m} \left(w_{n,m}^A + w_{n,m}^B\right)},$$

it is a weighted sum of the values $Q_{n,m}^{AF}$ $Q_{n,m}^{BF}$ w^A weights w^B, which determine the importance of the pixel (m, n). Weights are determines as follows:

$$w_{n,m}^A = d_A(m,n) C_A(m,n) P_A(m,n) g_A(n,m)^L,$$

where $d(m,n)$ — detection flag is 1 if pixel with coordinates (m, n) belongs to border, and 0 otherwise. $C(m,n)$ — the coefficient that determines the correlation by direction with neighboring boundaries, tends to 1 if the neighboring border co-directed with a given border, and tends to 0 otherwise. $P_A(m,n)$ — coefficient depending on the position of the considered pixel, is generally determined by the fusion method, for example, linearly decreases from the center to the edges of the image. The last factor depends on the value of the gradient at this point, since the larger gradient value, the greater focus, where L — some constant [12], which the authors propose to take equal $L = 1.5$.

For ideal fusion $Q^{AB/F} = 1$. For this metric the next rule is true: the higher value for better result. The general scheme of calculating the metric is shown in Fig. 3.

6) Modified universal quality index (WaB-Measure): This metric is proposed by Wang and Bovik in [13]. Let there are two images A and B size $M \times N$ each. Let $\overline{A}$ and $\overline{B}$ — the average image intensity values of A and B, respectively, and σ_A^2, σ_B^2 — dispersion. Also denote the covariance of the image as a σ_{AB}. Thus we can write:

$$\sigma_A^2 = \frac{1}{1 - MN} \sum_{m=1}^M \sum_{n=1}^N \left(A(m,n) - \overline{A}\right)^2,$$

$$\sigma_B^2 = \frac{1}{1 - MN} \sum_{m=1}^M \sum_{n=1}^N \left(B(m,n) - \overline{B}\right)^2,$$

$$\sigma_{AB} = \frac{1}{1 - MN} \sum_{m=1}^M \sum_{n=1}^N \left(A(m,n) - \overline{A}\right) \left(B(m,n) - \overline{B}\right).$$

Define

$$Q_0(A, B) = \frac{4 \sigma_{AB} \overline{AB}}{\left(\overline{A}^2 + \overline{B}^2\right) \left(\sigma_A^2 + \sigma_B^2\right)}, \tag{1}$$

which can be represented as follows:

$$Q_0(A, B) = \frac{\sigma_{AB}}{\sigma_A \sigma_B} \cdot \frac{2\overline{AB}}{\overline{A}^2 \overline{B}^2} \cdot \frac{2\sigma_A \sigma_B}{\sigma_A^2 + \sigma_B^2}. \tag{2}$$

In the original article, the authors describe the Q_0 as universal index of image quality, and use it to quantify the structural distortions between images A and B. We can consider Q_0 as a measure of similarity of images A and B, which takes values in the range of $[-1; 1]$. It should be noted that the components of the expression 2 have some meaning. The first component is the coefficient of correlation between the images. Second one — characterizes the average distortion of the brightness and is in the range of $[0, 1]$. The third component determines the contrast distortion between images and also has a value of 0 and 1. Thus, if $Q_0 = 1$ image A and B are identical.

Since the image is non-stationary signal it is necessary to limit the area in which Q_0 evaluates, and then merge the data to evaluate the image as a whole. In the original article, authors suggest the use of this method of sliding window. Sliding window w with size $m \times n$ passes the image from the upper left corner to the right and down in increments of one pixel in each direction. Thus the estimated local value Q_0 is within the current window position. The total value Q_0 is defined as

Fig. 3.   Scheme of $Q^{AB/F}$-measure calculating

the average value of the obtained local values, according to the expression 3, where W — the set of all the windows, and $|W|$ — number of windows.

$$Q_0(A,B) = \frac{1}{|W|} \sum_{w \in W} Q_0(A,B|w).\qquad(3)$$

7) Fusion quality index: Based on the results of [13], particularly in expression 3, in [14], suggested the expansion of the quality index for the case of fusion images. Let A and B — the source images, and F — image obtained as a result of the fusion algorithm. Thus, the value of the metric $Q(A,B,F)$ should describe the quality of the algorithm by fusing images A and B.

Denote $s(A|w)$ as certain quantitative characteristics defining image A in some window w. Value $s(A|w)$, may, for example, depend on the contrast, the dispersion, entropy or other characteristics of the image, or combinations thereof. Having determined the $s(A|w)$ and $s(B|w)$ for a particular window w, we introduce the value of $\lambda_A(w)$, varying in the range of $[0,1]$ and characterize the relative importance of the image A as compared with the image B to describing the content of w. In the original article, it is proposed to use the expression. 4.

$$\lambda_A w = \frac{s(A|w)}{s(A|w) + s(B|w)}.\qquad(4)$$

A similar expression can be written for $\lambda_B(w)$. Thus we can write the expression determines the fusion quality index $Q(A,B,F)$ as follows:

$$Q(A,B,F) = \frac{1}{|W|} \sum_{w \in W} (t_A + t_B).\qquad(5)$$

where

$$t_A = \lambda_A(w)Q_0(A,F|w),$$
$$t_B = \lambda_B(w)Q_0(B,F|w).$$

Thus, in those areas of the image, where the most important to preserve information from the image A, the value for the most part of the metric is determined by the value of $Q_0(A,F|w)$ and accordingly in areas where B more fully describes the content of the scene value $Q_0(B,F|w)$ is included with more weight.

The author also proposes a number of modifications proposed by the image fusion quality metrics. Firstly, it is

suggested to consider the importance of a strong image in some areas compared to others. For this proposed use weights $C(w)$, describing the significance of the window w, thus the expression 5 is transformed to the following form:

$$Q_W(A,B,F) = \frac{1}{|W|} \sum_{w \in W} c(w)(t_A + t_B),\qquad(6)$$

where $c(w) = \frac{C(w)}{\sum_{w' \in W} C(w')}$ — relative weight of window w. Another refinement of the metric associated with a features of human vision: when human analyzing an image, the human brain extracts the most information from the boundaries of objects in the scene. If we replace the original images in expression 3 for their boundaries A', B', F', which can be calculated, for example, using the Sobel operator, combining the obtained expressions, we can write:

$$Q_E(A,B,F) = Q_W(A,B,F)^{1-\alpha} \cdot Q_W(A',B',F')^{\alpha},$$

where α — parameter varying in the range $[0;1]$, describing the importance of the images with borders in a final metric value.

All of the metrics $Q(A,B,F)$, $Q_W(A,B,F)$ and $Q_E(A,B,F)$ are in range $[-1;1]$, the closer the value of any of the metrics is to the 1, the more quality is the result of fusion.

III.   COMPARISON OF IMAGE QUALITY MEASURES

For a numerical comparison of the quality assessment methods formed its own image database. The database contains images containing four different scenes: "Robot", "Toys", "Soldiers" and "Numbers". Series characteristics are presented in Table I. Each of these series contains a scene saturated by areas with lots of detail. In turn, these areas are located at different distances from the acquisition device. Acquisition device is a digital SLR camera Sony Alpha 37 with lens Sony 50mm $1.8/f$ (fixed SAL   50F18).

TABLE I.   CHARACTERISTICS OF SOURCE IMAGE SERIES

Characteristic	Value			
Name	"Toys"	"Robot"	"Soldiers"	"Numbers"
File extension	.jpg			
Color schema	RGB			
Color depth	8 bit/channel			
Resolution	786×523	786×523	786×523	892×592
Number of images	5	8	8	7

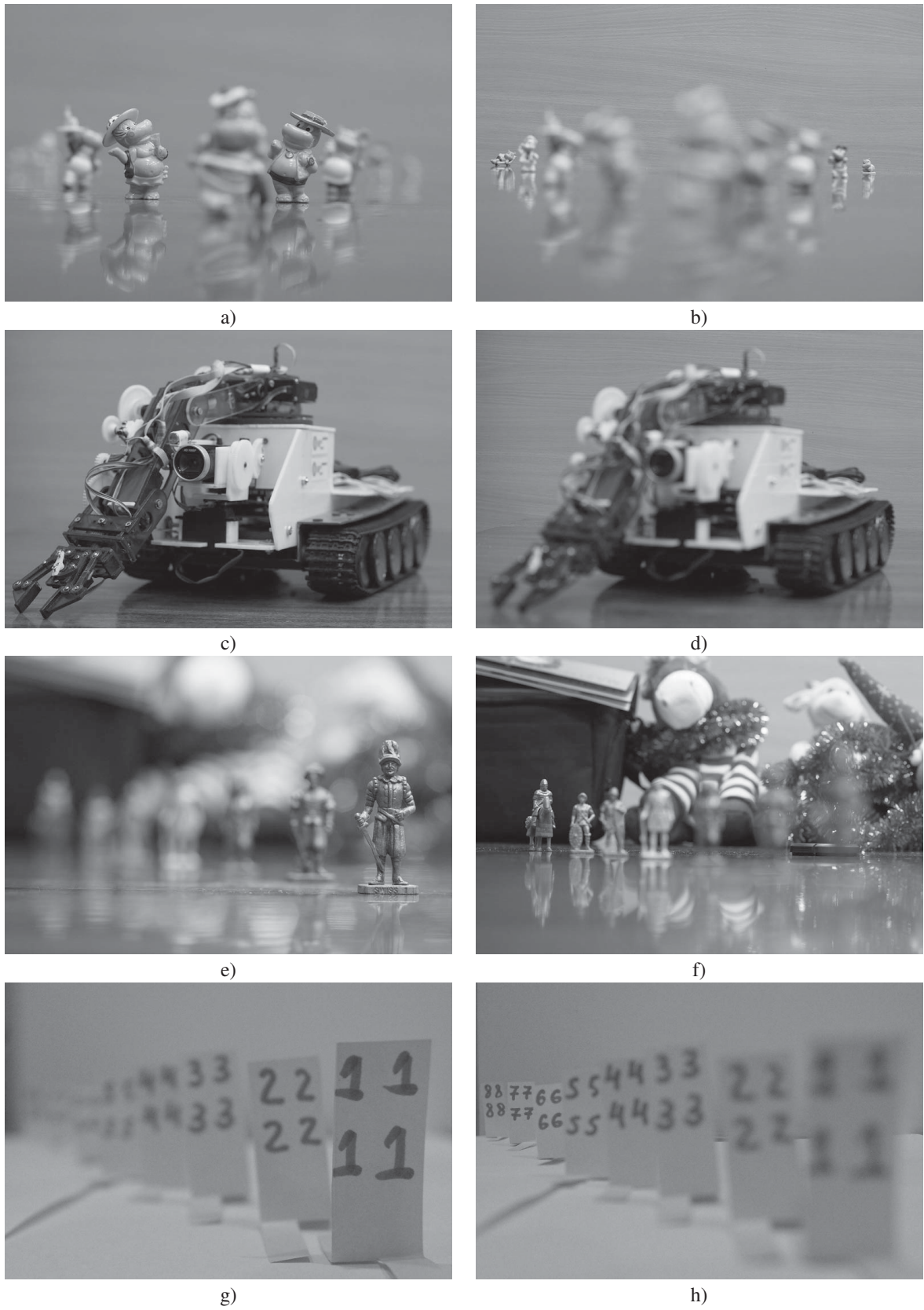Examples of source images series are shown at Fig. 4.

Fig. 4. Examples of source image series: a, b) "toys"; c, d) "robot"; e, f) "soldiers"; g, h) "numbers". Different source images from the same scene are focused on various areas of image
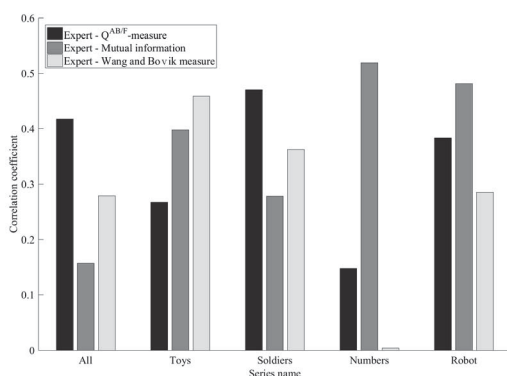
Fig. 5. Spearman's rank correlation coefficients

To compare methods of assessing the quality three fusion algorithm was selected: block algorithm [7], algorithm based on cellular automaton and algorithm based on cellular automaton with a pyramidal fusion [15], [16].

Each algorithm was run with eight different values of its parameters. Thus the total number of ranked images $4(scenes) \times 3(algorithms) \times 8(parameters) = 96$. 545 expert assessments were carried out on a scale from zero to nine points. Zero points corresponds to the worst quality, nine points to the best quality. One image has an average of 5.6 evaluations. Expert assessment is considered to be the average value of the experts of this image.

Spearman's rank correlation coefficient is selected as a measure of the correspondence between the expert estimations and values which are calculated according to the quality assessment measure. The correlation coefficients were calculated for the entire set of images, as well as separately for each scene. This approach will be concluded according to the effectiveness of the quality assessment measures from the content of the scene. Graphical representation of the calculated correlation coefficients is shown in Fig. 5.

Obtained values characterize the relationship between the expert and the calculated values as a medium or low powered link. However, in considering separately each scene, we can select the metrics that have a strong correlation with experts estimation. Thus, we can conclude that joint mutual information measure is good to be used for assessing the scene with a low number of sharp boundaries (scene "Numbers"). While in scenes with lots of details (scene "Soldiers") $Q^{AB/F}$-measure works well.

IV. CONCLUSION

Image quality assessment methods are very important part of the image fusion algorithm development process. Most popular referenceless methods are discribed. $Q^{AB/F} - measure$, Joint Mutual information and Wang and Bovik measure are selected for quality assessment experiment. The experiment established that different quality assessment methods work better in different types of scenes. The results of the work can be used to automate the process of evaluating the quality of the full-focused image fusion algorithms and adjusting their

parameters. This algorithms can be used in various practical applications: medicine, quality control of products, computer vision, robotics and others.

REFERENCES

[1] W. Huang, Z. Jing, "Evaluation of focus measures in multi-focus image fusion"*Pattern Recognition Letters*, vol. 28, 2007, pp. 493-500.

[2] C.Pohl and J.L. van Genderen, "Multisensor image fusion in remote sensing: Concepts, methods and applications", *International Journal of Remote Sensing*, Vol. 19, 1998, pp. 823–854.

[3] Huafeng Li, Yi Chai, Hongpeng Yin, Guoquan Liu., "Multifocus image fusion and denoising scheme based on homogeneity similarity", *Optics Communications*, vol. 285, 2012, pp. 91-100.

[4] Chun-Hung Shen and Homer H.Chen, "Robust Focus Measure for Low-Contrast Images", *Consumer Electronics*, Jan 2006, ICCE '06, Digest of technical Papers, pp. 69–70.

[5] Kekre H. B., Sarode T., Dhannawat R., "Implementation and Comparison of different Transform Techniques using Kekres Wavelet Transform for Image Fusion", *International Journal of Computer Applications*, 2012. Vol. 44, no. 10. pp. 41–48.

[6] Kekre H. B., Athawale A., Sadavarti D., "Algorithm to Generate Kekres Wavelet Transform from Kekres Transform", *International Journal of Engineering Science and Technology*, 2010. Vol. 2. pp. 756–767.

[7] Stathaki T., Image Fusion: Algorithms and Applications, *Academic Press*, 2008.

[8] Du-Yih Tsai, Lee Y., Matsuyama E., "Information Entropy Measure for Evaluation of Image Quality". *J Digit Imaging*, 2008. no. 21(3), pp. 338–347.

[9] Guangyao Cao, Luhong Liang, Siwei Ma, Debin Zhao, "Image Quality Assessment Using Spatial Frequency Component". *Lecture Notes in Computer Science*, 2009, vol. 5879, pp. 201–211.

[10] Xydeas C., Petrovic V., "Objective image fusion performance measure", *Electronics Letters*, 2000. no.36. pp. 308–309.

[11] Petrovic V., Xydeas C., "Sensor noise effects on signal-level image fusion performance", *Information Fusion*, 2003. Vol. 4. pp. 167–183.

[12] Petrovic V., Xydeas C., "Objective evaluation of signal-level image fusion performance", *Optical Engineering*, Aug. 2005. no. 44.

[13] Wang Z., Bovik A. C., "A universal image quality index", *IEEE Signal Processing Letters*, Mar. 2002. Vol. 9, no. 3. pp. 81–84.

[14] Piella G., "New quality measures for image fusion", *Proceedings of the Seventh International Conference on Information Fusion*, Jun. 2004. pp. 542–546.

[15] Noskov A., Aminova E., Volokhov V., "Comparison of image focus assessment methods for multi-focused image construction", *Proceeding of the 18th conference of FRUCT association* 2016. pp. 226–232.

[16] Noskov A., Volokhov V., Aminova E., "Multi-focus image fusion based on cellular automata method", *Proceeding of the 17th conference of FRUCT association*, 2015. pp. 136–141.