Comparison of Image Focus Assessment Methods for Multi-focused Image Construction

Andrey Noskov, Elena Aminova, Vladimir Volokhov
Yaroslavl State University
Yaroslavl, Russia
noskoff.andrey@gmail.com, {lena, volokhov}@piclab.ru

***Abstract*—Image merging 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. The most popular image focus assessment for usage in the task of forming a multi-focused image were considered. The classification of metrics evaluation was done. Experiments on suggested image focus metrics construction were minutely described. The forms of referenced characteristics were proposed. Using correlation analysis to select the best from the point of view of the problem, metrics and estimated time required for their computation were discussed.**

I. Introduction

A. Multi-focus image fusion

Image merging 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 capture (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 are to compute focus value of original images or the parts of them. At once, pixels with greater values of these measurements, 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 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. The ordinary scheme of image fusion is shown in Fig. 1.

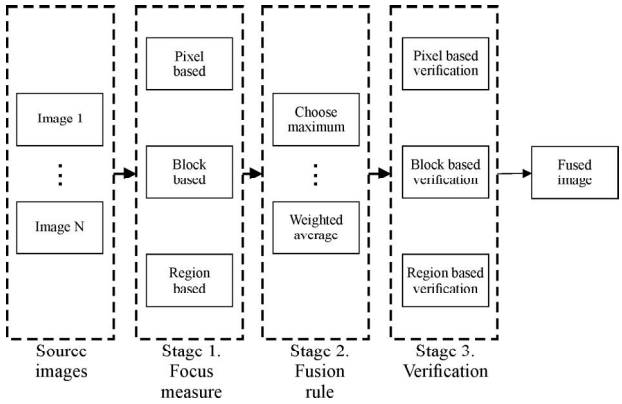


Fig. 1. Scheme of multi-focused image merging

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 of a two-dimensional projection of the image.

The first thing that needs to be said is the depth of image can be restored via binocular (trinocular) systems in case absence of physical interaction with the captured scene as well as with a few shots taken at different settings of monocular system.

B. Classification of image focus assessment methods

Evaluation of the image depth is the basis of many important applications of computational photography. For example, the task of object shape reconstruction on images is abundant evidence of that. Moreover, this problem is deterministically related to robot control system, nondestructive inspection in manufacturing production, renovation of models and in many other industries. It should be noted that few images focus measurement is made to determine the most informative image pixels, which will continue to play increasingly a vital role in the multi-focused image formation [1], [2], [3], [4], [5]. All focus assessment methods of whole image or different areas of image can be safely divided into several large groups [6]. The scheme of the mentioned above division is shown in Fig. 2.

1) The methods based on the calculation of the gradient. This group of methods is based on the calculation of

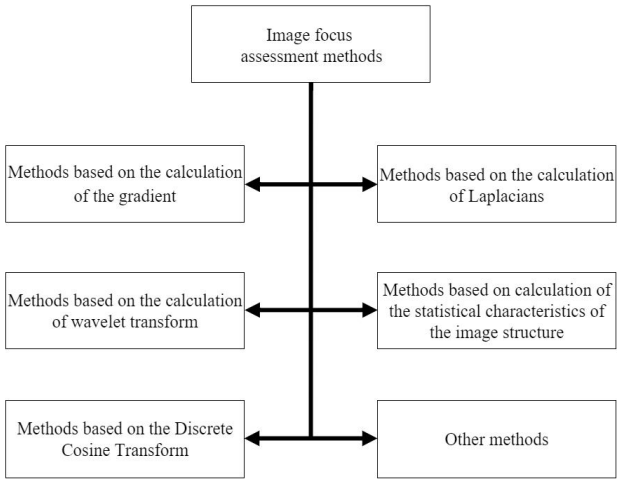


Fig. 2. Classification of images and its fields focus assessment

the image gradient or the first derivative of the image. The basic idea in the methods of this group lies in the assertion that the more pronounced the boundary object of interest on image, so it is more focused. Thus, the described metric possesses the highest value in the sharp areas of image than in the blurred.

2) The approaches based on the calculation Laplacians like the previous class propose boundaries separation on image, however, with the difference that they are calculated by the second derivative.
3) The algorithms used wavelet transform. This group of ways to assess the image focus value based on the coefficients of discrete wavelet transforms to describe the spatial and frequency characteristics of the image. Therefore, the coefficient of conversion and their various relationships may be used for obtaining estimates of non-focus.
4) The methods distinguished statistical characteristics. These approaches propose the calculation of the statistical characteristics of the image structure.
5) The ways are outlined Discrete Cosine Transform as well as methods based on the wavelet transform, the coefficients of the corresponding transformation coefficient to calculate the spatial and frequency characteristics of the image in order to assess image focus value.
6) Other methods. This group classified the methods that are not included in any of the previous ones. As a rule, they are based on any a priori information on the scene of interest, and apply in the case of images with specialized content.

C. Thin lens formula

To describe the distortions arising in the image the model of an ideal single lens system is used. Most real-world optical system may be reduced to the model. The described scheme is shown in Fig. 3.

In this case p — a point object. The image p' of p is constructed in optical lens with a focal length f. The distance from the object to the lens is U, whereas V is the distance from

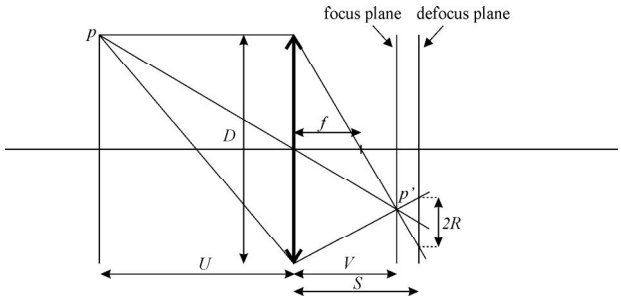


Fig. 3. Scheme of the ideal single lens system

the lens to the image. Therefore, the lens formula is defined as:

$$\frac{1}{f} = \frac{1}{U} + \frac{1}{V}$$

The camera has got an aperture with the size D. The charge-coupled camera located at a distance S from the lens. Hence, if $V \neq S$, image of the object on the matrix will be a blur with radius, which is calculated as:

$$R = \left| \frac{1}{2} DS \left(\frac{1}{f} - \frac{1}{U} - \frac{1}{V} \right) \right| \quad (1)$$

Sign of the module needs for generalization to the case of the formula $V > S$, i.e. matrix closer to the lens than the image. The value R may be selected, for instance, based on the pixel size of the sensor matrix. Thus, according to the value of R far and near border and depth of field (DOF) can be found as:

$$U_{far} = U \frac{f\left(1 - 2\frac{R}{D}\right)}{1 - 2\frac{RU}{D}}$$

$$U_{near} = U \frac{f\left(1 + 2\frac{R}{D}\right)}{1 + 2\frac{RU}{D}}$$

$$DOF = U_{far} - U_{near}$$

When $D \to \infty$, accordingly $DOF \to 0$, since $U_{far} \to U$ and $U_{near} \to U$. Otherwise, when the diaphragm drops significantly decreases the illumination image, that a variety of practical applications, such as microscopy, is unacceptable. Besides, in addition to directly blurring of p', it shifts radially from the main optical axis of the system. Even so, mentioned effect doesnt influence essentially to quality of purposed method and visual perception of the scene in vast majority of cases [7].

II. Methods of image focus assessment

For choice the best image focus estimation method investigation was done, the aim of which was to identify the most appropriate algorithm for this propose. So as far as is known the most of algorithms suggest analysis of each pixel in the image, accordingly it is obviously that the most preferred methods based on convolution of original images with different forms of masks. However, several metrics that used statistical and other algorithms were included in this study for generality.

A. Absolute central moment

The first time when this metrics was considered is the work of [8]. This estimation is designed on statistical estimates and the image histogram H.

$$ACM = \sum_{k=1}^{L} |k - \mu| P_k$$

where μ — mean value of H, L — gray level range on the image, P_k — the relative frequency of k-th gray level range.

B. Sobel masks

The Tenengrad focus measure (Sobel mask) [9] is one of the most well-known image focus measurement, which is represent by brightness gradient of image estimation [10].

$$T(Sobel) = \sum_{(i,j)\in\Omega(x,y)} \left[G_x^2(i,j) + G_y^2(i,j)\right]$$

where Ω — image focus field of interest, G_x and G_y — the vertical and horizontal brightness gradient commutated via convolution of the original image with the corresponding Sobel masks:

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

C. Roberts masks

This is the same approach whereas used Roberts masks as the Tenengrad image focus estimation [11], [12]:

$$S_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} S_y = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

D. Prewitt masks

The other one Tenengrad image focus estimation is to apply Prewitt mask [13], [14]:

$$S_x = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} S_y = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

E. Gray scale level variance measurement

One of the most sought-after method for autofocus tasks and object of interest on image reconstruction is gray scale level variance measurement and described as [15], [16]:

$$GLV = \sum_{(i,j)\in\Omega(x,y)} (I(i,j) - \mu)^2$$

F. Wavelet transform coefficients ratio

This appraisal was proposed by [17] and suggests usage of high frequency to low frequency coefficient ratio of wavelet transform [18], [19]. The referred evaluation is described by following expression:

$$WCR = \frac{M_H^2}{M_L^2}$$

where

$$M_H^2 = \sum_k \sum_{(i,j)\in\Omega} W_{LHk}^2(i,j) + W_{HLk}^2(i,j) + W_{HHk}^2(i,j)$$

$$M_L^2 = \sum_k \sum_{(i,j)\in\Omega} W_{LLk}^2(i,j)$$

In this case $W_{LHlk}(i,j)$, $W_{HLk}(i,j)$ and $W_{HHk}(i,j)$ is the detailing wavelet transform coefficients, at the same time, $W_{LLk}(i,j)$ — the approximating wavelet transform coefficients and k — number of wavelet transform levels.

The suggested above estimation ways are selected for testing in the experiment due to their high efficiency and overall popularity in the computational photography applications and the relative computer vision and image processing.

III. Experiments for comparison of image focus methods

A special program of experimental research and original experimental stand were developed for the numerical evaluation of the properties of various polling focus metrics.

A. Test bench construction

Experimental setup consists of video-capture device (1), the reference set of key points containing image with high frequency details (2) and experimental images (3). The general scheme is shown in Fig. 4.

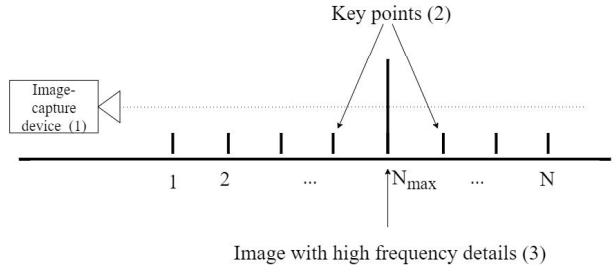


Fig. 4. General scheme of the test bench

As for video-capture device, Sony Alpha 37 with camera lens Sony 50mm 1.8/f was used. The reference points are needed to set the focal plane of photographic equipment at a certain position, so that the distance between its two neighboring positions is unaltered. The experimental image constitutes a pre-test pattern TIT-0249. The referred image is depicted in Fig. 5.

B. Experimental steps

- *Step 1*: Test sequences formation, composed of multiple images and each in accordance with the field of interest.
- *Step 2*: Usage each of considered above estimation on taken after the step 1 images.
- *Step 3*: Obtaining and evaluation of the complex quality assessment metric.

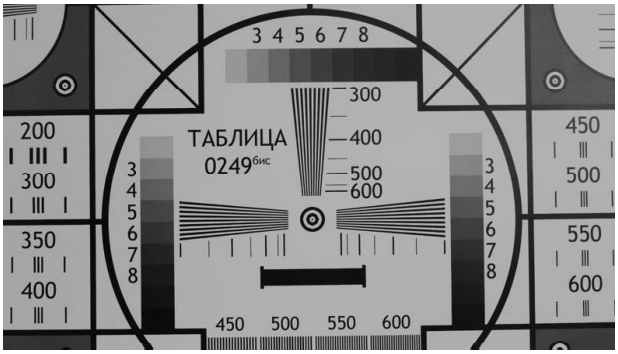


Fig. 5. Image for testing: pre-test television pattern TIT-0249

It should be noted that to carry out the first step, an expanded post-test sequence was formed, which consists of 16 images with different focal position. It was defined that the clearest image is located at a distance of 70 centimeters, and has the number 9. The change scale of the focal plane of the relieving device position is selected as 5 centimeters, starting with 30 centimeters indent from the capture device. As for test sequences, the various fields with varied content of surrounding scene were chosen in order to set different test sequences. The areas of interest and appropriate test patterns are shown in Fig. 6.

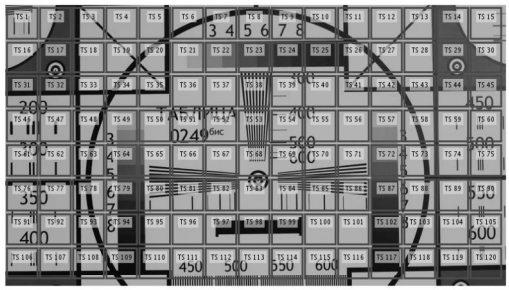
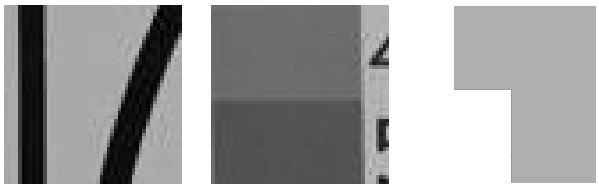


Fig. 6. Test sequences division on test sets

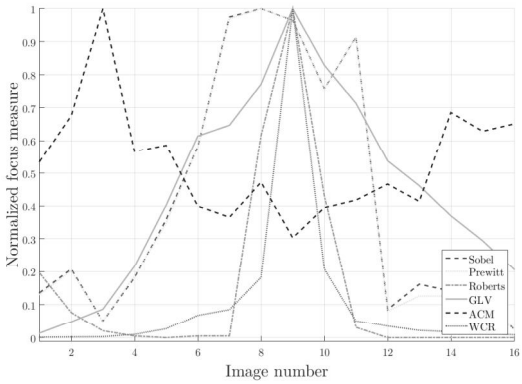
During the *step 2* of the experiment it is needed to weigh image focus estimation of selected areas by using analyzed algorithms. The calculated values of valuation metrics were normalized to the maximum value of the assessment in the current series of images for bringing to the range $[0; 1]$. Thereby, the values visualization of the obtained image focus approaches of the test series is depicted in Fig. 8.



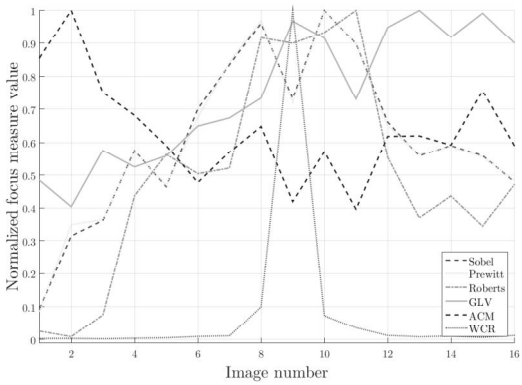
a) b) c)

Fig. 7. Image samples which are used in test sets with numbers: a) 48; b) 72; c) 96

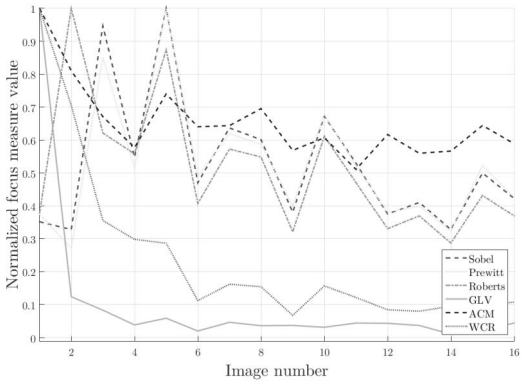
Notably the chart corresponding to the test area with the number 96. This test was assessed completely homogeneous area of the image. In this regard, the metrics do not work correctly. Some examples of image areas which was used in test are depicted in Fig. 7.



a)



b)



c)

Fig. 8. Image focus assessment value of the test set with numbers: a) 48; b) 72; c) 96

IV. Referenced image focus characteristics

For a numerical comparison of various image focus estimation algorithms in his thesis the value of the Pearson correlation coefficient between the taken X and a Y and referenced characteristic is encouraged to use. This coefficient

of correlation is calculated as follows:

$$r = \frac{\sum_n \left[(x_n - \bar{x}) (y_n - \bar{y}) \right]}{\sqrt{\sum_n (x_n - \bar{x}) \sum_n (y_n - \bar{y})}}$$

where $x_n = X(n)$, $y_n = Y(n)$, x and y — the average values of X and Y, respectively. Consequently, based on the common understanding of the described above system the referenced characteristics can be describe as:

1) A single pulse. The mention above referenced characteristic assumes that the unit-pulse usage and the assumption that the best metric should have a value of 1 in assessing the most focused image N_{max} and 0 for assessing the remaining images. Characteristics graph is depicted in Fig. 9.

$$FM(n) = \begin{cases} 1, & \text{if } n = N_{max} \\ 0, & \text{otherwise} \end{cases}$$

2) Thin lens formulas (linear relationship). The referenced characteristics via a thin lens guess assumes physical aspects of the observed dynamics changes of image focus. This is in accordance with the thin lens formula 1, in obedience to the radius of the spot where the blurred image point is directly proportional to the removal of the focal plane from the plane of the camera sensors [20]. This implies that if the value of the evaluation on a perfectly focused image is 1, it will decrease with increasing distance from the object plane by the rule $\frac{1}{x}$. Thence, characteristics graph is shown in Fig. 9.

$$FM(n) = \begin{cases} \frac{1}{-n+N_{max}+1}, & \text{if } n < N_{max} \\ \frac{1}{n-N_{max}+1}, & \text{otherwise} \end{cases}$$

3) Thin lens formulas (quadratic relationship). The referenced characteristic has the same statements as in a previous approach. However, square of the blurring spot is supported to use like a referenced parameter instead of radius. Therefore the general rule is in proportion $\frac{1}{x^2}$. This case is depicted in Fig. 9.

$$FM(n) = \begin{cases} \frac{1}{(-n+N_{max}+1)^2}, & \text{if } n < N_{max} \\ \frac{1}{(n-N_{max}+1)^2}, & \text{otherwise} \end{cases}$$

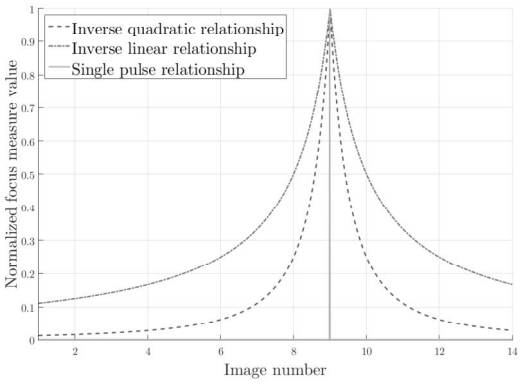


Fig. 9. Referenced characteristics

It should be mentioned that in the Fig. 10 histogram of the average coefficient correlation values for the proposal referenced characteristics and image focus assessments. The generalized correlation coefficient values were calculated using the following expression:

$$r = \sum_{k=1}^{N} w_k r_k$$

where N — the total number of test sequences, r_k — correlation coefficient of k-th test sequences, which is calculated as $w_k = n_k * NI$. In this case n_k — the number of images of k-th test sequences, NI — total number of images in all sequences.

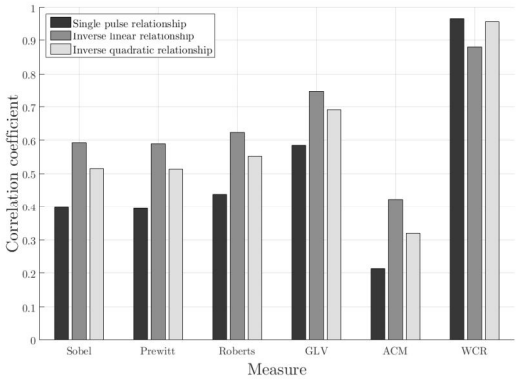


Fig. 10. Total correlation coefficient value for proposed assessments and referenced characteristics

Fig. 11 is a bar graph the average computation time focus metrics in the block size of 50 x 50.

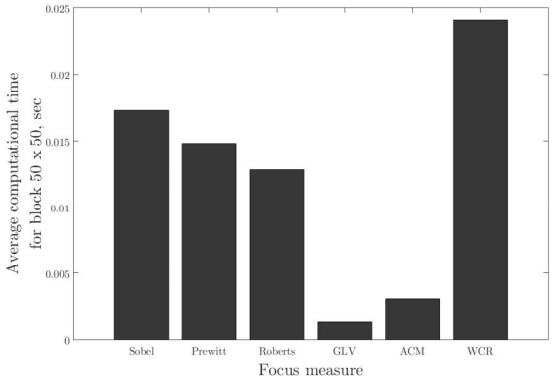


Fig. 11. The average computational time for all described estimations

V. Conclusion

To sum up, based on the results of the experiment, it is possible to make some conclusions, suggest the following recommendations on the use of image focus assessment methods.

1) From the analysis of correlation coefficients for different ways to assess the image focus value and ways to set the referenced characteristics can be concluded

that the choice of the most suitable referenced characteristic is not important due to all of them behave in a similar way. Indeed, if we calculate the correlation matrix for the obtained values of estimations, it would be as follows:

$$\begin{bmatrix} 1 & 0.9270 & 0.9828 \\ 0.9270 & 1 & 0.9821 \\ 0.9828 & 0.9821 & 1 \end{bmatrix}$$

This aspect underlines a strong positive relationship.

2) Analyzing the considered algorithms it ensued that most closely focused areas allocated on the basis of the algorithm calculating the ratio of the coefficients of the wavelet transform ($r = 0.8798$). Nevertheless, it cant be recommended for usage in real-time applications on the strength of high computational complexity.
3) Among the algorithms based on convolution operations, the best results are focused on the allocation of areas showing an algorithm for calculating the Tenengrad metric using masks Roberts ($r = 0.6241$). This method is approximately equal in computational complexity of the rest of the convolutional algorithm and can be used in real-time tasks.
4) The best results in terms of algorithm performance shown by calculating the variation of gray scale level. Additionally, among those represented approaches, this algorithm has the second highest quality selection of image focused areas.

In conclusion we provide an example of the execution result of the image fusion system based on research [1] and the results of current paper. Source images and resulted fused image are shown on Fig.12.

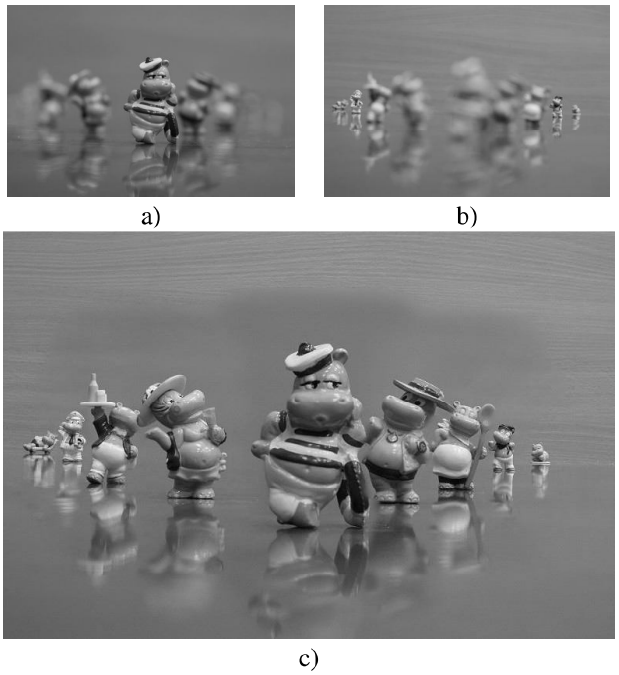


Fig. 12. Multi-focus image fusion: a), b) examples of source images; c) fused image

ACKNOWLEDGMENT

We thank Ilya Trapeznikov who provided insight and expertise that greatly assisted the research, although they may not agree with all of the interpretations/conclusions of this paper.

At the same time, we thank Ivan Mochalov for assistance with particular technique, methodology and Andrey Priorov for comments that greatly improved the manuscript.

We would also like to show our gratitude to the Yuriy Brjukhanov for sharing their pearls of wisdom with us during the course of this research, and we thank 3 anonymous reviewers for their so-called insights.

This research was supported supported by RFBR Grant 16-37-00301.

REFERENCES

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

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

[3] 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.

[4] 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.

[5] 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.

[6] S.Petruz, D.Puig, M.A.Garcia, "Analysis of focus measure operators for shape-from-focus", *Pattern Recognition*, vol. 46, Issue 5, May 2013, pp. 1415-1432.

[7] S. V. Voronov, "Development and modeling of pseudo-gradient procedures image attachment via informative criteria", Ulyanovsk, 2014, p.31

[8] M. Shirvaikar, "An optimal measure for camera focus and exposure", *Southeastern Symposium on System Theory*, 2004, pp. 472-475.

[9] Sobel, I., Feldman,G., "A 3x3 Isotropic Gradient Operator for Image Processing" *Stanford Artificial Project*, 1968.

[10] M. Subbarao, T. Choi, A. Nikzad, "Focusing techniques", *Journal of Optical Engineering*, vol. 32, 1993, pp. 2824-2836.

[11] Lawrence G. Roberts Machine "Perception Of Three-Dimensional Solids", 22 May 1963, Reissued May 1965

[12] A.S. Malik, T.S. Choi, "A novel algorithm for estimation of depth map using image focus for 3D shape recovery in the presence of noise", *Pattern Recognition*, vol. 41, 2008, pp. 2200-2225.

[13] J.M.S. Prewitt "Object Enhancement and Extraction", *Picture processing and Psychopictorics*, Academic Press, 1970.

[14] N.K. Chern, P.A. Neow, M.H. Ang, "Practical issues in pixel-based autofocusing for machine vision", *Proceedings of the International Conference on Robotics and Automation*, vol. 3, 2001, pp. 2791-2796.

[15] C.H. Shen, H.H. Chen, "Robust focus measure for low-contrast images", *Digest of Technical Papers of International Conference on Consumer Electronics*, 2006, pp. 69-70.

[16] C. Wee, R. Paramesran, "Measure of image sharpness using eigenvalues", *Information Sciences*, vol. 177, 2007, pp. 2533-2552.

[17] H. Xie, W. Rong , L. Sun, "Wavelet-based focus measure and 3-d surface reconstruction method for microscopy images", *RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 229-234.

[18] Dr. H. B. Kekre, Archana Athawale, Dipali Sadavarti, "Algorithm to Generate Kekres Wavelet Transform from Kekres Transform", *International Journal of Engineering Science and Technology*, Vol. 2(5), 2010, pp. 756–767.

[19] I. De, B. Chanda, "A simple and efficient algorithm for multifocus image fusion using morphological wavelets", *Signal Processing*, vol. 86, 2006, pp. 924-936.

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