Full-Focused Image Fusion in the Presence of Noise

Andrey Noskov, Vladimir Volokhov, Andrey Priorov, Vladimir Khryashchev
Yaroslavl Demidov State Univercity
Yaroslavl, Russia
noskoff.andrey@gmail.com, volokhov@piclab.ru, andcat@yandex.ru, vladimir@piclab.ru

***Abstract*—The implementation and analysis of the algorithm for the full-focused image fusion in the presence of noise are presented. Three methods of combining noisy images are considered: without pre-processing and post-processing, using prefiltration of original images, using post-filtering of the fused image. The database of test scenes created by the authors was used for testing the proposed algorithm for full-focused image fusion. Additive white Gaussian noise was considered as an noise model. Two-stage digital image processing scheme, based on principal components analysis was used as a filtering algorithm. Quantitative and visual results are shown and demonstrate the main features of the proposed algorithm.**

I. Introduction

The transition from analog to digital photography is a big step forward. Digital photography [1] has opened many new features, which include instant preview of captured images, quick editing, the ability to easily record video sequences, etc. Nowadays, digital cameras with megapixel resolution allows to create high-quality images for a wide range of consumers and professional applications. Some researchers consider that the next big step forward in the field of the formation and processing of visual information is computational photography [2], which extends the boundaries of traditional digital photography. This is provided by the ability to record much more information about the captured scene, as well as better processing of this information afterwards.

One of the important problems considered in the field of computational photography is the task of full-focused image fusion [3], [4], [5], [6], [7]. The main issue of the task is to create an algorithm that allows to combine several images of a fixed scene, which have a limited depth of field and are formed with different focal length, into one completely focused. It should be noted that this problem is considered in isolation from the problem of noise generation appearing in the process of forming a digital image and having a negative impact on the procedure for constructing fully focused images. Therefore, in this paper an approach is considered that allows to create a full-focussed image with noise impact. The essence of the approach is the use of prefiltration (filtration before fusing noisy images) or postfiltration (filtration after fusing noisy images) algorithms in order to form a quality full-focused image.

II. Digital image filtration

Nowadays, digital image filtration algorithms [8] are widely used in the field of modern science and technology and have many practical applications. An interesting approach to the problem of image filtration is the use of machine learning methods [9], [10]. Two-stage image filtering scheme based on the principal components analysis (PCA) [11], [12] was used in the present work as a solution of the noise reduction problem. The article assumes that each of the original digital images x was distorted by additive white Gaussian noise (AWGN) with zero mathematical estimation and the standard deviation σ.

A. First stage of processing

Step 1. Propose that the standard deviation σ of noise on the input noisy image $y = x + n$ is known.

Step 2. Split the input noisy image into a collection of overlapping blocks. Within each of them, next types of region can be select: the training region, the filtering region and the region where the blocks are overlapping. The size of the regions can vary.

Step 3. Inside the training region, select all possible square blocks (training vectors). They are represented as vector-columns and can be combined to sample matrix.

Step 4. Create a covariation matrix based on centered sample matrix from step 3. Find eigenvalues and eigenvectors of co-variation matrix.

Step 5. Find the projections (transformation coefficients) of the set of vectors from the centered sample matrix to the set of eigenvectors found in step 4.

Step 6. Process the obtained set of projections using the linear mean-square estimator [10].

Step 7. Based on the set of processed data, the estimate of the sampled matrix of noiseless data can be restored. This estimate is used for processing of separated regions on image. In this case, firstly, the training region is restored by investing training vectors in the corresponding spatial positions in it. Training vectors, which are inside the evaluation of the sample matrix of noiseless data, were transformed back into square blocks before investing in the training region. It should to be noted that the overlap region of training vectors is averaged by using arithmetic averaging. Secondly, after the restoration of the training region, a filtration region of smaller size is allocated from it. By the repeating of the same operation for the remaining filtering regions, taking into account their overlap, the primary "rough" estimate $\hat{x}^I$ of the undistorted image x can be get. In this case, the processed filtering regions are embedded in the corresponding spatial positions of the image, their overlapping region is arithmetically averaged.

B. Second stage of processing

Step 1. Repeat steps 2-5, considered in the first stage of processing using a noisy image. Other sizes of training regions, filtering regions and overlapping regions, as well as training vectors are set to different size.

ISSN 2305-7254

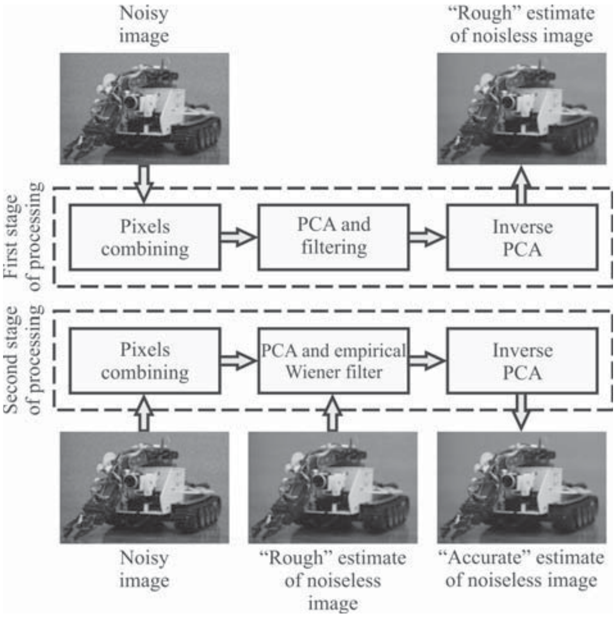


Fig. 1. A block diagram of digital image processing using a two-stage image filtering scheme based on the principal components analysis [11], [12]

Step 2. Process the resulting set of projections by using the empirical Wiener filter, which is presented in the field of principal components and calculated based on the primary estimate $\hat{x}^{I}$.

Step 3. Repeating step 7 of the first stage of processing, get the second, "accurate" estimate $\hat{x}^{II}$ of the undistorted image x.

Block diagram of the described algorithm for filtering digital images is shown in Fig. 1. It should to be noted that RGB images are processed separately per-channel.

Fig. 2 shows an example of filtered digital images from a database reviewed in the article, which will be briefly described below. Each of the images of the "Robot" scene, formed with a different focal length, and noisy by AWGN with the standard deviation $\sigma = 35$. In addition, Fig. 2 shows the numerical estimates of the PSNR, dB (the peak signal-to-noise ratio) [13] and the SSIM (structural similarity index) [14] for noisy and restored images.

III. Full-focused image fusion

It was noted above that the full-focused images fusion is the process of obtaining single image from several source images. The resulting image contains in itself more information about the scene than each of the original ones individually. Such an image can be more convenient for further work by human or an automatic digital image processing. In this work, the algorithm [7] based on cellular automata [15], [16] and image pyramids [17], was used to solve the problem of the full-focused image fusion. Assuming that the data processed by the algorithm are contained in several images of a fixed scene that have a limited depth of field and are formed with a different focal length, let's briefly describe the main steps of this algorithm. A block diagram of the described algorithm is shown in Fig. 3.

An examples of source images for "Toys" series are shown in Fig. 4.

A. First stage of processing

Step 1. Calculate the focus measure (MS) for each pixel in each of the N original images x_k of the fixed scene. Form N matrices containing the calculated measures. The index k here is the image number, which varies from 1 to N.

Step 2. Calculate the maximum value of the focus measure for each of the N matrices obtained in step 1.

Step 3. Perform the threshold binarization of the matrices calculated in step 1 with the thresholds based on the maximum values of the focus measure found in step 2 and the value α, which is the parameter of the algorithm.

Step 4. Create a label matrix containing information about which pairs of coordinates (i, j) of pixels from N images of the fixed scene should participate in the formation of the final full-focused image. The label matrix is formed using data of the threshold binarization output (step 3).

Step 5. Correction of the matrix of labels using a cellular automaton.

B. Second stage of processing

Step 1. Form binary masks m_k based on the label matrix obtained in step 5 of the first stage of processing. The number of binary masks is N, and their dimension coincides with the resolution of the initial images of the fixed scene. Binary masks are needed to implement the subsequent steps of combining the source images of a fixed scene into one full-focused. The fusion is perform using pyramids of Gaussians and Laplacians [17].

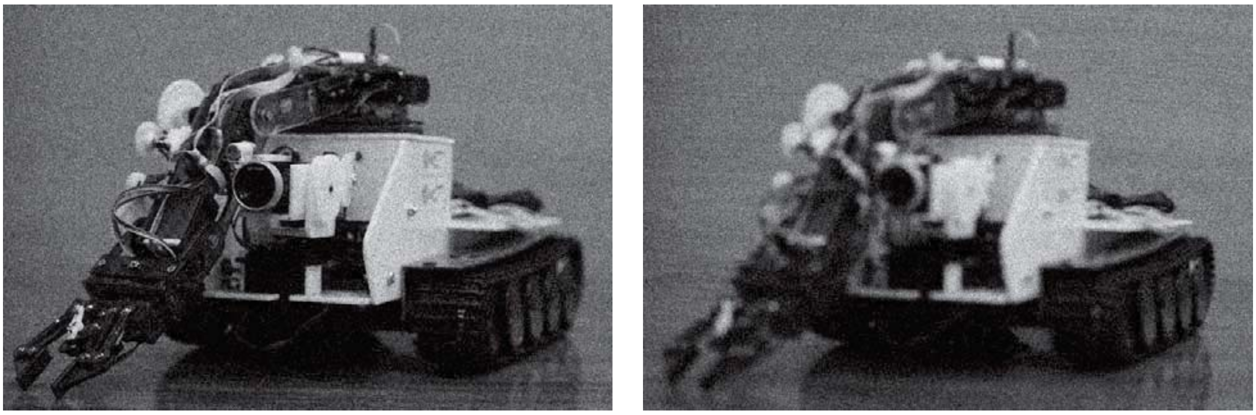
Step 2. Expand the source images of the fixed scene and the corresponding binary masks to the nearest 2^n in both demension. This procedure is necessary to perform an integer division by two during decimation of images in the process of construction of Gaussian and Laplacian pyramids.

Step 3. Create a Laplacian pyramid LP_k for each source image of a fixed scene, specifying the required number of decomposition levels. For color RGB images, the formation of the Laplacian pyramid is performed per-channel.

Step 4. Iteratively perform the fusion of Laplacian pyramids LP_k to form a hybrid Laplacian pyramid GLP, the inverse transformation of which will allow the formation of a fully focused image. To perform this procedure, need to follow next steps:

Step 4.1. Introduce the integral mask im, which has the dimension of binary masks m_k. Initialize the values of the integral mask and the hybrid Laplacian pyramid with values m_1 and LP_1, respectively, calculated for the first image of the fixed scene.

Step 4.2. Create a pyramid of Gaussians GP^{i-1} for the current integral mask im^{i-1}. Here, the index i describes the iteration number on which the hybrid Laplacian pyramid is formed.



a) PSNR = 17.86, dB / SSIM = 0.252 b) PSNR = 7.77, dB / SSIM = 0.197

c) PSNR = 29.34, dB / SSIM = 0.845 d) PSNR = 31.92, dB / SSIM = 0.871

Fig. 2. Examples of images for the ”Robot” scene with PSNR and SSIM: a) and b) noisy images ($\sigma = 35$); c) and d) the corresponding restored images

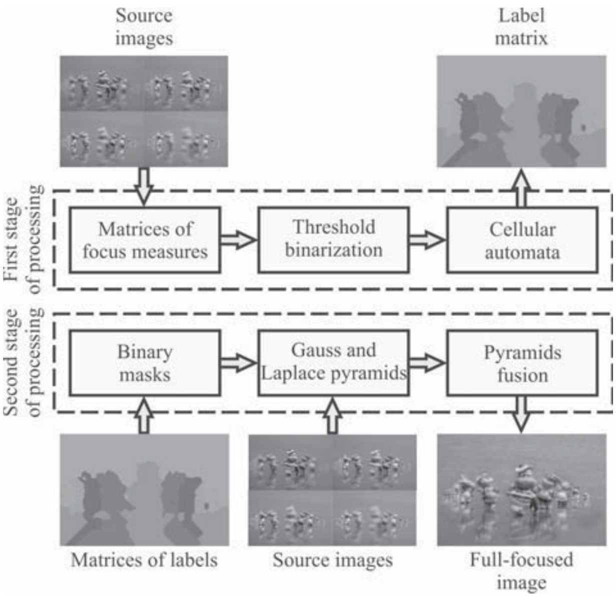


Fig. 3. Block diagram of the algorithm for full-focused image fusion based on cellular automata and image pyramids

Step 4.3. Create a hybrid pyramid of Laplacians according to the eq. 1.

$$GLP^i = GLP^{i-1} * GP^{i-1} + LP_i * \left(1 - GP^{i-1}\right). \quad (1)$$

Step 4.4. Update the current integral mask using the eq. 2.

$$im^i = im^{i-1} + m_i. \quad (2)$$

Step 4.5. Repeat steps 4.2-4.4 until the Laplacian pyramids LP_k are combined for all the original images of the fixed scene. Get a hybrid pyramid of Laplacians.

Step 5. Create the final full-focused image by reversing the hybrid Laplacian pyramid.

Fig. 5 presents examples of full-focused images fused using the algorithm [7] for groups of images of the four scenes considered in the work.

IV. Modeling results

To perform the modeling procedure, four groups of color RGB images of fixed scenes were used:

1) ”Numerical cards”, 7 images, resolution 892×592.
2) ”Robot”, 8 images, resolution 786×523.
3) ”Soldiers”, 8 images, permission 786×523.
4) ”Toys”, 5 images, resolution 786×523.

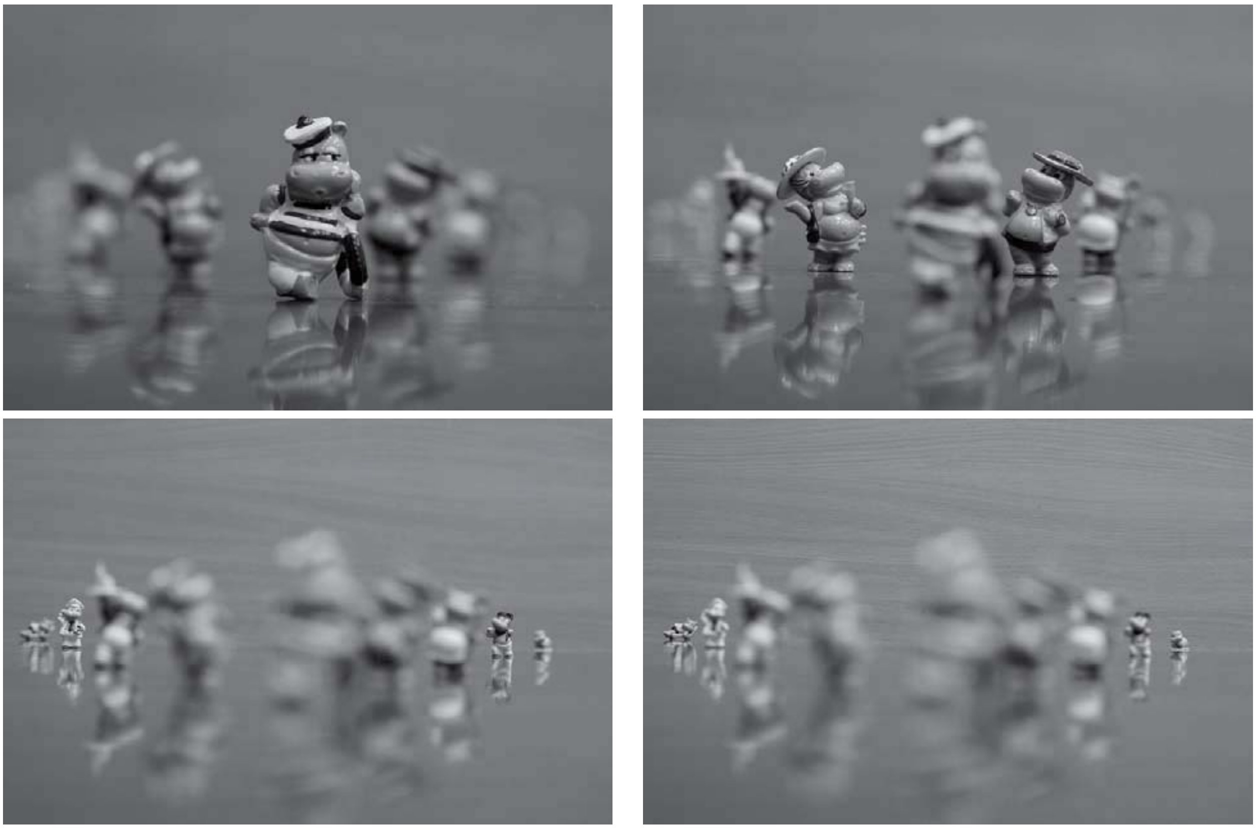


Fig. 4. Examples of the source images for the "Toys" scene

The test images were obtained by the authors using a digital camera fixed on a tripod, which has the possibility of changing the focal length. The formed groups of images allowed using the algorithm described at [7]. Images created by algorithm are taken as reference for the quality assessment of the algorithm in the presence of noise.

Three approaches of combining images are considered in the article.

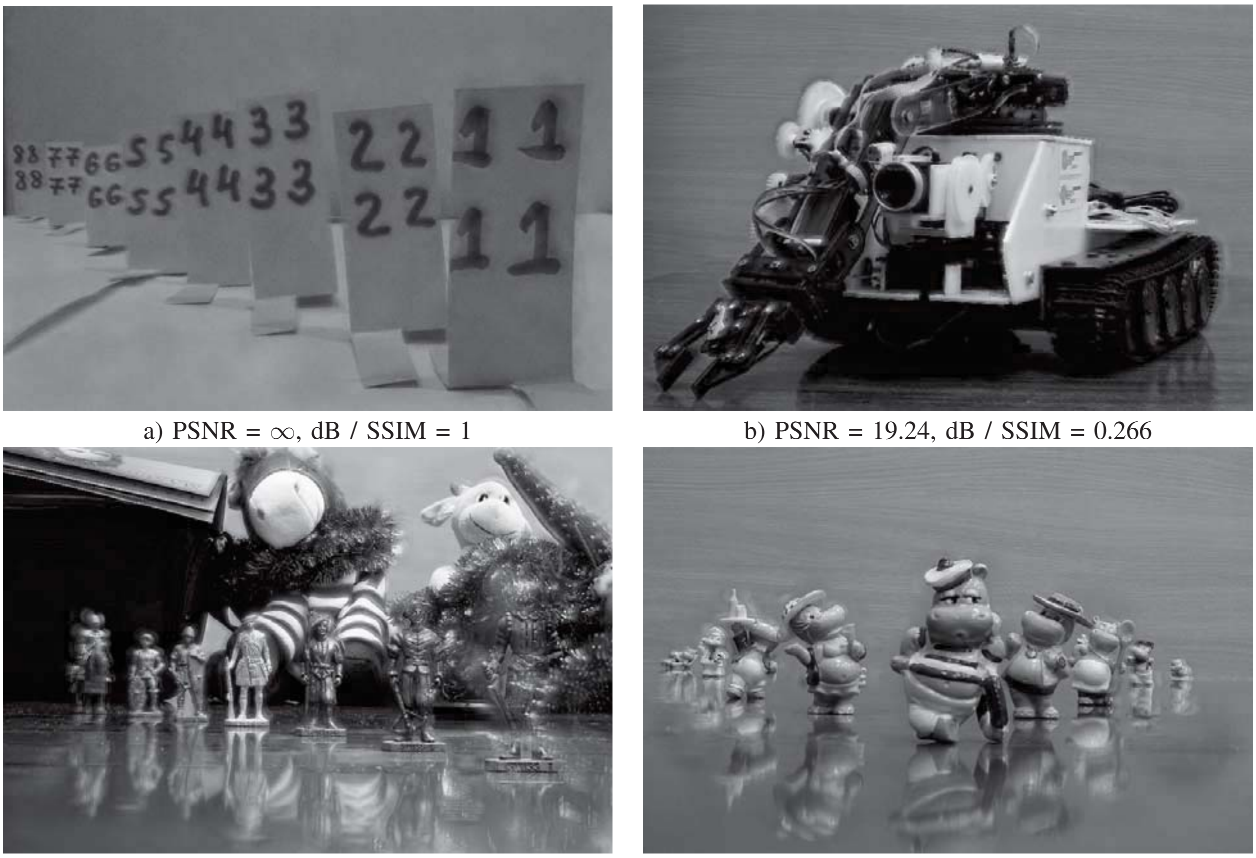
1) *Approach 1*. The source images of the scenes were noisy by AWGN with a fixed σ and then combined using the algorithm [7].
2) *Approach 2*. The source images of the scenes were noisy by AWGN with a fixed σ, they were filtered using a two-stage image processing scheme based on the principal component analysis [11], [12], and then combined using the algorithm [7].
3) *Approach 3*. The original images of the scenes were noisy by AWGN with fixed σ, combined using algorithm [7], and then the resulting fused image was filtered using a two-stage image processing scheme based on the principal components analysis [11], [12]. In this case, it was assumed that the noise model and its parameters in the fused image are identical to those of the source images before fusion.

Table I presents a numerical comparison of different approaches of fusion several images of a fixed scene for different σ. The best results are shown in bold. In addition, at Fig. 6 presents the visual results of the construction of full-focused images. For this research, standard metrics for quality assessment were used — PSRN and SSIM. These metrics were applied to two images. The first image, obtained by using the algorithm [7] in the absence of noise, was taken as the reference one. The second is obtained using one of the three approaches.

Analysis of the results shows that an acceptable quality of fusion of noisy images using the algorithm [7] is possible (Fig. 6c), if the original noisy data will be pre-processed using a filtration algorithm, for example [11], [12]. Combining images in the absence of this scheme does not allow qualitatively identifying the pixels that should make a significant contribution to the construction of a full-focused image. As a consequence, the combined image in this case will contain noise, as well as a significant blurring of the objects of interest present in the scene (Fig. 6b). Therefore, filtration after fusion does not give a qualitative result of the processing (Fig. 6d). However, it should to be noted that filtering before images fusion leads to the repeated use of the denoising algorithm. In this case, every noisy image of the fixed scene is filtered, and this is greatly increases the computational cost of approach 2. In the case of approach 3, the filtering is performed once for the combined image.

V. CONCLUSION

Three methods of noisy images fusion are considered: without pre-processing and post-processing, using prefiltration



a) PSNR = ∞, dB / SSIM = 1

b) PSNR = 19.24, dB / SSIM = 0.266

c) PSNR = 28.49, dB / SSIM = 0.858

d) PSNR = 24.32, dB / SSIM = 0.792

Fig. 5. Examples of the fused images for different test scenes with PSRN and SSIM: a) ”Numerical cards”; b) ”Robot”; c) ”Soldiers”; d) ”Toys”

TABLE I. PSNR, DB / SSIM OF FULL-FOCUSED IMAGES IN THE PRESENCE OF AWGN

σ	Approach 1	Approach 2	Approach 3
Scene ”Numerical cards”			
5	31.28 / 0.749	32.51 / 0.876	**34.83 / 0.879**
15	24.21 / 0.416	**31.99 / 0.862**	31.83 / 0.845
20	22.03 / 0.308	**31.82 / 0.857**	31.25 / 0.838
25	20.26 / 0.230	**31.80 / 0.852**	30.84 / 0.833
35	17.52 / 0.140	**31.46 / 0.843**	30.17 / 0.825
Scene ”Robot”			
5	32.21 / 0.881	**35.60 / 0.938**	34.63 / 0.933
15	24.28 / 0.510	**31.04 / 0.847**	29.78 / 0.805
20	22.23 / 0.415	**29.84 / 0.816**	28.73 / 0.776
25	20.55 / 0.340	**28.94 / 0.791**	27.86 / 0.753
35	17.99 / 0.243	**27.39 / 0.748**	26.49 / 0.717
Scene ”Soldiers”			
5	31.87 / 0.907	**33.85 / 0.954**	33.63 / 0.953
15	23.32 / 0.584	**28.39 / 0.869**	26.42 / 0.809
20	21.34 / 0.488	**27.06 / 0.834**	24.94 / 0.752
25	19.79 / 0.412	**26.04 / 0.801**	23.92 / 0.709
35	17.40 / 0.298	**24.53 / 0.742**	22.55 / 0.646
Scene ”Toys”			
5	32.67 / 0.887	33.92 /0.942	**35.96 / 0.951**
15	22.66 / 0.450	**29.80 / 0.885**	26.19 / 0.835
20	20.72 / 0.342	**29.06 / 0.870**	24.95 / 0.809
25	19.24 / 0.266	**28.49 / 0.858**	24.32 / 0.792
35	16.84 / 0.167	**27.58 / 0.835**	23.38 / 0.766

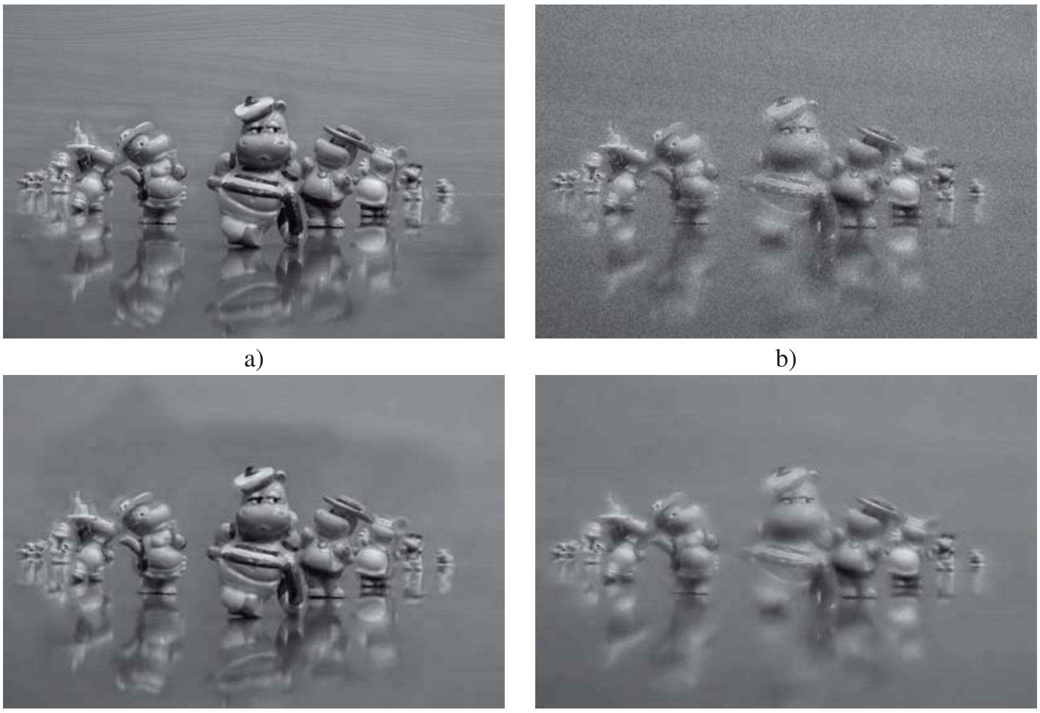


Fig. 6. Examples of the fused full-focused images for different test scenes: a) ”Numeric cards”; b) ”Robot”; c) ”Soldiers”; d) ”Toys”

of source images, using post-filtering of the fused image. Analysis of the results shows that combining noisy images with prefiltration allows better identification of pixels that should make a significant contribution to the construction of a full-focused image. However, it should be noted that filtering before the fusion leads to a multiple uses of the denoising algorithm, which significantly increases the computational cost of processing.

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

Acknowledgment

We thank Ilya Trapeznikov who provided insight and expertise that greatly assisted the research, although he 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 for comments that greatly improved the manuscript.

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

References

[1] R.Gonzalez, R.Woods, *Digital image processing*, Pearson, 2005.

[2] R.Szeliski, *Computer vision: algorithms and applications*, Springer, 2010.

[3] M.Subbarao, T.Choi, A.Nikzad, “Focusing techniques”, *Optical Engineering*, 1993, vol.32, 11. pp. 2824–2836.

[4] C.-H.Shen, H.H.Chen, “Robust focus measure for low-contrast images”, *Proc. IEEE Int. Conf. Consumer Electronics*, 2006, pp. 69–70.

[5] H.Li, Y.Chai, H.Yin, G.Liu, “Multifocus image fusion and denoising scheme based on homogeneity similarity”, *Optics Communications*, 2012, vol.285, pp. 91–100.

[6] V.P.S.Naidu, J.R.Raol, “Pixel-level image fusion using wavelets and principal component analysis”, *Defense Science*, 2008, vol.58, pp. 338–352.

[7] A.Noskov, E.Aminova, A.Priorov, “Full-focused image fusion based on cellular automata and images pyramid”, *Telecommunications and Radio Engineering*, 2016, pp. 39–46.

[8] V.Katkovnik, A.Foi, K.Egiazarian, J.Astola, “From local kernel to nonlocal multiple-model image denoising”, *Int. J. Computer Vision.*, 2010, vol. 86, pp. 1–32.

[9] S.Marsland, “Machine learning: an algorithmic perspective”, *Chapman and Hall*, 2009.

[10] D.D. Muresan, T.W.Parks, “Adaptive principal components and image denoising”, *Proc. IEEE Int. Conf. Image Processing*, 2003, vol.1, pp. 101–104.

[11] V.A.Volokhov, E.V.Sergeev, A.L.Priorov, “Two-staged digital images filtration procedure based on the principal components analysis”, *Design and technology of electronic resources*, 2011, pp. 40–44.

[12] A.Priorov, K.Tumanov, V.Volokhov, “Efficient denoising algorithms for intelligent recognition systems”, *In: Favorskaya M., Jain L.C. (eds.) Computer Vision in Control Systems-2*, Springer International Publishing. 2015. vol. 75, pp. 251–276.

[13] D.Salomon, *A Guide to Data Compression Methods*, Springer-Verlag, New York, 2002.

[14] Z.Wang, A.C.Bovik, H.R.Sheikh, E.P.Simoncelli, “Image quality assessment: from error visibility to structural similarity”, *IEEE Trans. Image Processing*, 2004, vol. 13, pp. 600–612.

[15] T.Toffoly, N.Morgolus, *Cellular automata machines*, Mir, Moscow, 1991.

[16] A.G.Hoekstra, J.Kroc, P.M.A.Sloot, *Simulating complex systems by cellular automata*, Springer, 2010.

[17] P.J.Burt, E.H.Adelson, “The Laplacian pyramid as a compact image code”, *IEEE Trans. Communications*, 1983, vol. COM-31, pp. 532–540.