https://doi.org/10.2352/ISSN.2470-1173.2017.13.IPAS-087

Search the optimal border for combination of image pairs using neural networks

Semenishchev Evgeny Aleksandrovich, Dept. of Radio-Electronics Systems, Don State Technical University, Shevchenko 147, Shakhty, Russian Federation 346500
Voronin Viacheslav Vladimirovich, Dept. of Radio-Electronics Systems, Don State Technical University, Shevchenko 147, Shakhty, Russian Federation 346500
Marchuk Vladimir Ivanovich, Dept. of Radio-Electronics Systems, Don State Technical University, Shevchenko 147, Shakhty, Russian Federation 346500
Tolstova Irina Vladimirovna, Dept. of Radio-Electronics Systems, Don State Technical University, Shevchenko 147, Shakhty, Russian Federation 346500

Abstract

The modern world is filled with plenty of photo and video cameras, adapted to address a huge range of tasks. Using multiple fixation devices can extend the region surveillance and build a three-dimensional model of the observation. A further increase in the number of cameras gives a chance improve the sharp images or single objects on the image. Modern automated systems built by combining photos and video streams require an integrated approach to process and analyze the data. Of particular importance for the analysis of visual information have a mosaic image, allowing to observe a continuous scene entirely, instead of viewing parts. The problem of obtaining united image is relevant, since is need to: in security systems, with the analysis of the overall control of the zone; in medicine, for X-ray; the construction of cartographic images received from the satellite; in solving problems of photogrammetry ; in the preparation of 3D images used in construction; in microbiology, while creating images of biological objects of small dimensions taken with the microscope; security, about combining data obtained fingerprint reader; in genetics, about the creation of a single snapshot of nucleic acids; in industrial processes, for example in the production of films and glass to detect inclusions and irregularities in the casting or stretching products etc.

The paper presents a mathematical model of the image stitching process based on the use of linear algebra and the foundations of optics. This mathematical model takes into account the importance of objects on the image. The paper proposes the use of an algorithm based on the following features: Find objects on the images (using salience map). Selecting base points in the frames this data. Search for correspondences between base points is performed using the analysis of distances of the mutual arrangement the data points and the correlation analysis. Changing the image produced using projective transformations, as a criterion serves the boundaries of divergence. For an optimal combination boundaries will apply neural networks, with a deep learning. The use of this type of networks to minimize the difference and reduce or eliminate the visual distinction of the merger field. Using a unified color palette is based on the analysis of previously important areas and finding generalized correction factors. To eliminate double contours we used several approaches: the first is based on a combination of background gradient fields, the second on the analysis generic objects for closed contour in accordance with the weight of the image.

On the set of test images will be shown the effectiveness of the proposed algorithm. As test images used pair of medical images, satellite images and other cameras.

Introduction

Automated data analysis systems are used in many fields. The use of these systems is possible in medicine, security, space, automated control systems, automatic driving systems and others. The use of data obtained in different ranges allows expanding information field. Of particular importance for the analysis of visual information have a mosaic image, allowing to observe a continuous scene entirely, instead of viewing parts. The use of the camera which analyzes the temperature gradient allows you to provide additional information on the internal composition of objects. This makes it possible to introduce additional information parameter in the system. For systems using streaming video, is important the possibility of use the combined image. It is desirable to obtain minimally perceptible transition in the region the viewer attention. For this effect to be entered in algorithm the block which analyzes objects located on the images. Integration of information of different ranges can afford to reduce the divergence of the visual boundaries when combining images. The problem of obtaining united image is relevant, since is need to: in security systems, with the analysis of the overall control of the zone; in medicine, for X-ray; the construction of cartographic images received from the satellite; in solving problems of photogrammetry ; in the preparation of 3D images used in construction; in microbiology, while creating images of biological objects of small dimensions taken with the microscope; security, about combining data obtained fingerprint reader; in genetics, about the creation of a single snapshot of nucleic acids; in industrial processes, for example in the production of films and glass to detect inclusions and irregularities in the casting or stretching products etc.

In formulating the problem of obtaining a united image, there are problems associated with the location of objects in the image, in the case of several objects of their relationship, the structure of the background and the overlap area. Particularly acute, this problem occurs in the case of fixing frames received from different points, using different cameras and differences in the settings of internal parameters of cameras. Also, if not used device the camera stabilization, photos may turn out blurry or rotated. In some cases, a group of objects, fixed with different points can be seen as several object or single. In the case of combining the images recorded by a digital microscope, the result is strongly influenced by selected focal length. Performing operations on the images of all sizes, as well as obtained from different points, imposes restrictions on the use of the algorithms, and in some cases makes it impossible to use them. In this paper we consider an algorithm of image stitching.

In the work [1], presented stitching algorithm. This algorithm allows to receive a solution of multi-image matching problem. It is use invariant local features to find matches between all images. This method is insensitive to the ordering, orientation, scale and illumination of the input images. It is also insensitive to noise images that are not part of a panorama, and can recognize multiple panoramas in an unordered image dataset.

The stitching quality is measured visually by the similarity of the stitched image to each of the input images, and by the visibility of the seam between the stitched images [2]. In order to define and get the best possible stitching, introduce several formal cost functions for the evaluation of the stitching quality. In these cost functions the similarity to the input images and the visibility of the seam are defined in the gradient domain, minimizing the disturbing edges along the seam. A good image stitching will optimize these cost functions, overcoming both photometric inconsistencies and geometric misalignments between the stitched images. In this work we study the cost functions and compare their performance for different scenarios both theoretically and practically.

In the work [3] is to achieve seamless image stitching without producing visual artifact caused by severe intensity discrepancy and structure misalignment, given that the input images are roughly aligned or globally registered. This is approach is based on structure deformation and propagation for achieving the overall consistency in image structure and intensity. Depending on the compatibility and distinctiveness of the 2-D features detected in the image plane, single or double optimal partitions are computed subject to the constraints of intensity coherence and structure continuity. Afterwards, specific 1-D features are detected along the computed optimal partitions from which a set of sparse deformation vectors is derived to encode 1-D feature matching between the partitions. These sparse deformation cues are robustly propagated into the input images by solving the associated minimization problem in gradient domain, thus providing a uniform framework for the simultaneous alignment of image structure and intensity.

Part of the work is devoted to assessing the quality of stitching images. In the selected work [4], the get the best possible stitching, is introduce several formal cost functions for the evaluation of the quality of stitching. In these cost functions, the similarity to the input images and the visibility of the seam are defined in the gradient domain, minimizing the disturbing edges along the seam. A good image stitching will optimize these cost functions, overcoming both photometric inconsistencies and geometric misalignments between the stitched images.

To solve the problem of combining pairs of images in a single composition in the work will be used an approach based on the simultaneous analysis of detailed areas, as well as criteria for analyzing boundaries of objects with minimal divergence gradient.

I. The image model

The combining images from different sensors associated single process is a complex task. Fusing of data from different sources allows to obtain additional information that cannot be determined in a separate method of visualization.

We will allocate four basic levels of data fusion:

I. Data fusion of at regional level (blocks or regions). The preliminary decision to multimodal objects are taken in each of the image transmission channels. The final decision is carried in the data channels. Post-processing is based on the selection the features and fusion data the structures of objects.

II. The data fusion of formation at the level of features of objects. In this type of fusion in each of the channels formed vector feature objects. At this stage results of features is combined into a single big feature vector. This vector is used to decide to combine segmental data into a single array.

III. Complexation of signals at the elementary level, this data is received directly from sensors. Formation of identification features is made using a vector rather than a scalar, as in scheme II.

IV. Complexation of level elementary signals conducted by based on the analysis data multichannel optoelectronic system. This approach provides a single fusion image.

In this paper we use the approach used the spaced cameras. The cameras are not bound a single optical system. One possible application is to use approaches based on: the principle data fusion of at regional level (I), on the analysis of data fusion of formation at the level of features of objects (II) or on of level elementary signals conducted by based on the analysis data multichannel optoelectronic system (IV). The scheme combining the data suggests the presence of common computing module the descriptor. The approach requires the use of analysis of information obtained in each of the channels. For schemes (I) and (II), creating a single feature vector immediately after the sensor is impossible. We will be apply method (IV) because the approach will be used the analysis data of multi-channel optoelectronic system.

To form the combination of images we will assume that a pair of images obtained from different points. The image data are separated in time and have different the camera settings. However, between the pairs of images have general area. This area is part of the real world. On the figure 1 is a graphical representation of the image stitching tasks into a single composition. On image: $Y_{i+k,j+n} = X_{i+k,j+n} + P_{k,j+n}$ is the first image with the area of the stitching; $Z_{k+l,j+m} = P_{k,j+m} + R_{l+k,m+j}$ is the second image with the area of the stitching; $X_{i+k,j+n}$ is the first image; $R_{l+k,m+j}$ is the second image; $P_{k,j+n+m}$ is the stitching area; $A = A(Z)_{lk_1,mj_1} + A(Y)_{k_1,j_1}$ is the object on image [8].

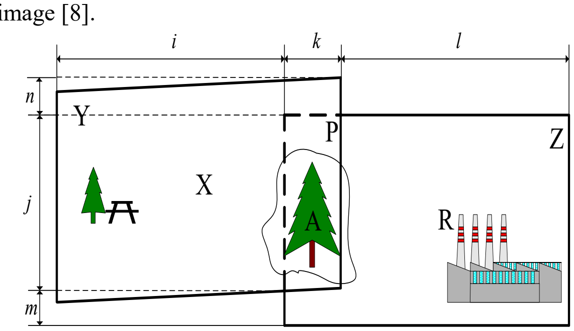

Figure 1. The image stitching model.

A simplified mathematical model the result of stitching images in is represented as:

$$S_{i+k+l,n+j+m} = A \cup X_{i+k,j+n} + \\ + A \cup \alpha \cdot P_{k,j+n+m} + A \cup R_{l+k,m+j},$$

where: α is the coefficient compensating type of distortion.

The type of distortion compensation coefficient can be both simple and composite, allowing the suppression of the main types of distortion caused by: the difference object distance shooting sensitive matrix; used the properties of lenses and the inaccuracy of their manufacture; dimensions and shapes of the objects, and others.

2. Image stitching

We present a flowchart for stitching images into a single composition in figure 2.

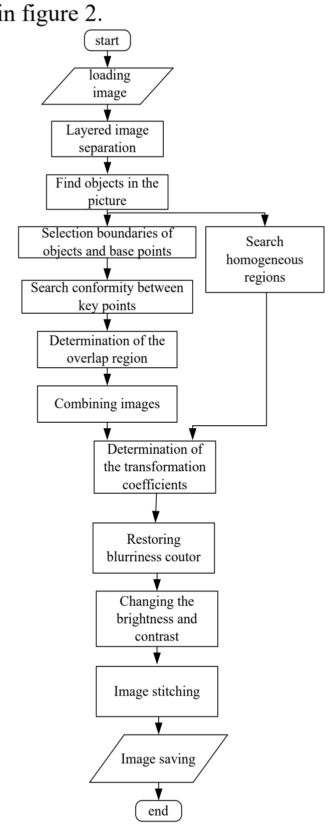

Figure 2. The algorithm for image stitching.

The algorithm is implemented as follows [5]:
1) The first step is loading a pair of images.
2) On the next step, in the case of operator involvement we selects the points which should use in stitching. In the case of automatic consolidation:
3.1) We make an exception the background image from consideration. As the methods for determining the background can be used method of one of the approaches based on the analysis of the boundaries of objects and their mutual arrangement [6]. It is also possible to use trained neural network, a threshold method and methods based on saliency map.
3.2) The next step is to search the boundaries of objects in images and borders in the internal structure of these objects. To do this, we use stepwise process. Tiered exception of analysis the found in the previous step the corners of objects and areas of the border. At each step, we define small (3-7 pixels) individual boundaries and their distance to the corners. Found for at this stage of the algorithm the data allow to construct a graph tree.
3.3) We find the scalable small areas, for which the distances are constant and the same for the pair of images. The search is conducted on the analysis of the resulting tree graph. According to analysis of the elements and structure of the objects, we will highlight the field with the close proximity of the borders or corners. These areas will be taken as the base. Later, when we use stitching images, elements from these areas will have a maximum weight.
3.4) To increase the accuracy in the choice of base points we use the correlation analysis and make search of correspondences between the elements of the base area. As a point of analysis will use the center of mass within the base area for each of the respective pairs. Stable shall be considered the area with the maximum correlation.
4) Above the pair of images we use the projective transformations. The main condition, the maximum coincidence of reference points in the overlap area.
5) In the case of fuzzy boundaries in enclosed object, we use search algorithm centerline.
6) To eliminate the visible transition between the images in the case of differences in illumination of objects or the difference in exposure, apply a gradient approach to stitching images. As a curve the transition will use natural boundaries of objects arranged on a pair of frames. We define the criteria for a given curve. This criteria Include the condition of maximum direction of the curve to the center of stitching region and passing along the borders of the objects.
7) In the final step, we make color correction on one of the selected images. We will limit the area of study of color. We choose a pair of identical square areas around the reference points by which to combine. We make the definition of the brightness histogram for them. Example for color images is shown in figure 6. We define the position of the maxima of the histogram data and carry out the correction of the pair of output images. The correction will be made by multiplying at a certain constant for each of the channels. For an area without objects apply a gradient transition.
8) On the final stage they are producing stitching of the images into a single content.

3. Analysis of the areas on the image

We use images combination operation [6]. The proposed method allows the selection of areas with high detail. The method is based on analysis of the boundaries of objects in an image. When looking for detalisation objects by "density" on the first step considered the general factor of detail on the whole image, which is defined by the formula [6]

$$\frac{\sum X(i,j)}{\tilde{i}\cdot\tilde{j}} = \tilde{P}_{base},$$

where: $X(i,j)$ is the value of the pixel with coordinates i and j; $\tilde{i}$ is the rows; $\tilde{j}$ is the columns; $\tilde{P}_{base}$ is the base coefficient of detalisation.

On the next step, similarly to expression produced calculation the density in each sliding window [6].

$$\frac{\sum X(i,j)}{0.1\cdot\tilde{i}\cdot\tilde{j}} = \tilde{P}_{windiws}$$

where: $\tilde{P}_{windiws}$ is the general coefficient of detalisation; 0,1 is the averaging factor associated with the automatic selection of the window size equal to 10% of the total image.

Next, $\tilde{P}_{base}$ and $\tilde{P}_{windiws}$ is compared and a decision is made about of detail in this window. In the next step, the window shifts and is made similar calculations for him. The proposed approach has allowed to exclude areas with by the uniform background from consideration. Account shall be taken only data on areas with a large number of edges.

Concurrently, we will use two approaches. The first is based on the analysis of areas of detail. The second is based on an analysis of the internal structure of an object. For the implementation of the second method we will use the simple linear iterative clustering (SLIC) proposed in the work [7]. Figure 3 shows an example of processing of the left image by this method.

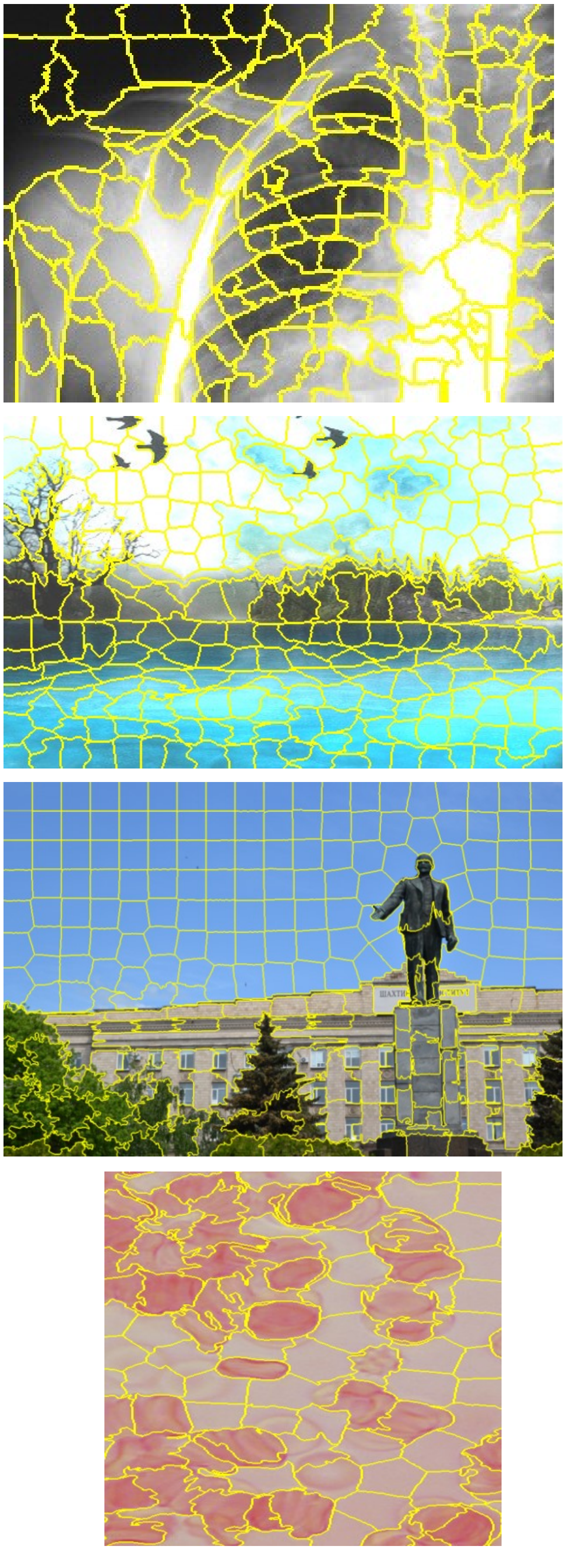

Figure 3. The example of processing of the left image by SLIC method.

This superpixel algorithm is simple linear iterative clustering (SLIC), which adapts a k-means clustering approach to efficiently generate superpixels. Despite its simplicity, SLIC adheres to boundaries. This algorithm it is fast and memory efficient, have improves segmentation performance, and is straightforward to extend to supervoxel generation.

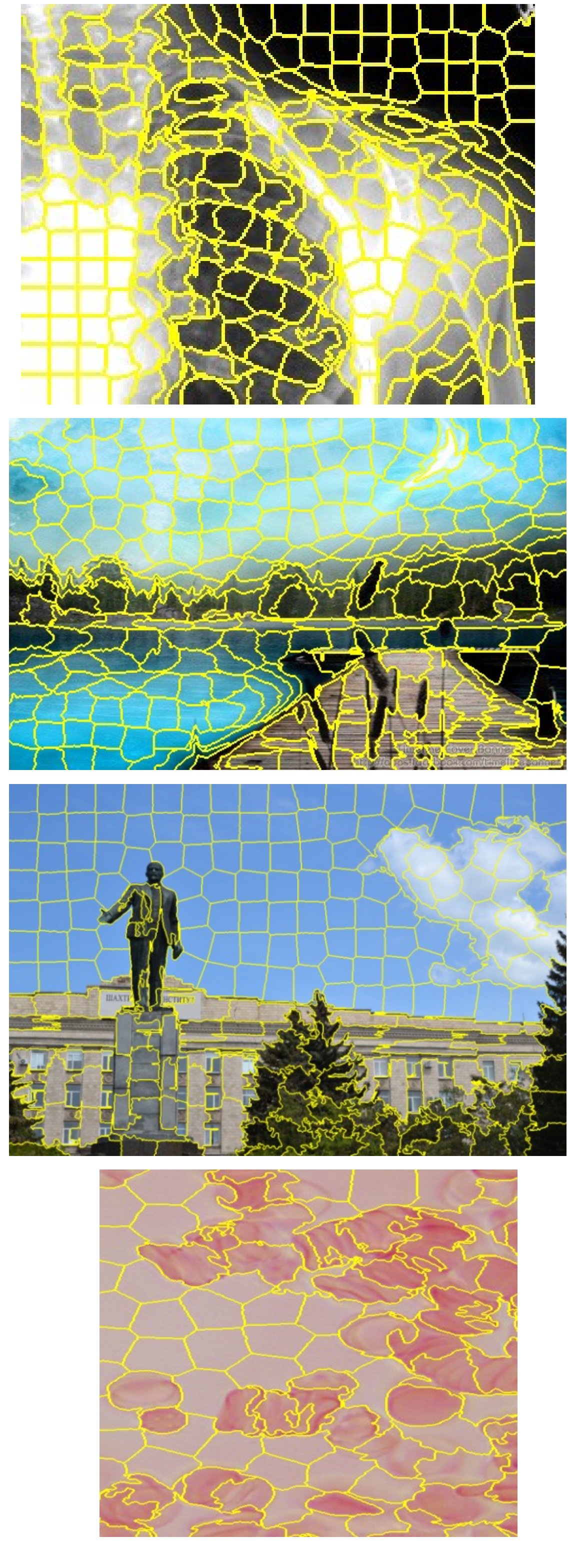

Figure 4. The example of processing of the right image by SLIC method.

We used the standard algorithm submitted by the sponsors. As an selected parameter used the medium value. On figure 4 shows an example of processing of the right image by this method.

Search stitching border was carried out using conditions, the priority have boundaries located close to the center of the combination area. The main steps are: 1. In the case coincidence two boundaries, of the obtained of the algorithm SLIC on both images. Selected as the stitching edge is as resulting use algorithm curved border. 2. In case of discrepancy boundaries of obtained SLIC. Selection the boundary combining is performed based on of the neural network. The neural network was trained on the basis of standard images. The regulation of the selection borders - minimally noticeable transition in the field of combining.

In step search transformation matrix image, have the highest weight the area which selected of algorithm "density". The examples of the final stitching are shown in Figure 5.

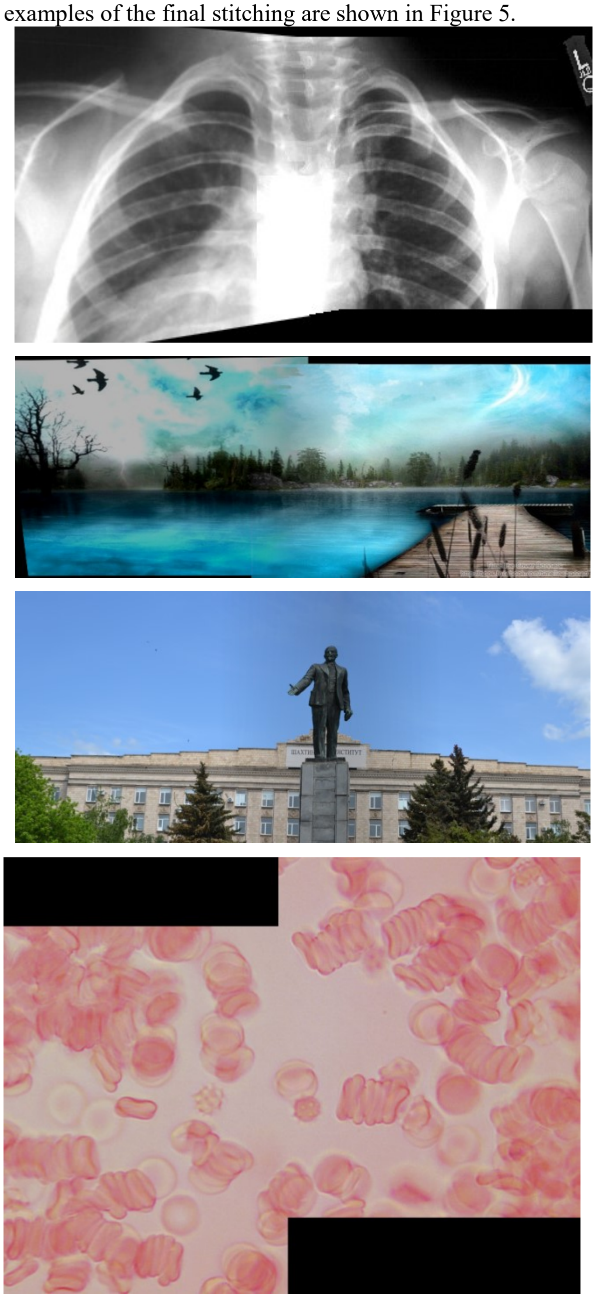

Figure 5. The examples of the image stitching.

On the figure 6 are presented examples of the image stitching. As can be seen from the example stitching area is visually imperceptible.

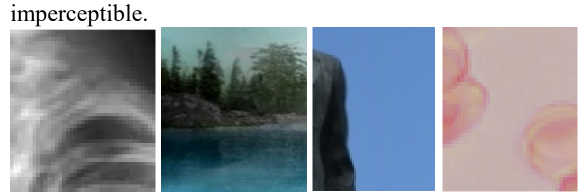

Figure 6. The examples of the image stitching in the field of combining.

Conclusion

In this paper, we have improved the method and image stitching algorithm. We used the previously obtained image stitching algorithm. The selection condition combining border based on combining results two algorithm: "density" and generate superpixels SLIC. Selection the boundary combining is performed based on of the neural network. The proposed algorithm allows receiving image stitching area as visually imperceptible. We show the effectiveness of the proposed approach on the different test images.

Acknowledgments

This work was supported by Russian Ministry of Education and Science in frame of the Federal Program "Research and development on priority directions of scientific-technological complex of Russia Federation in 2014-2020" (contract № 14.577.21.0212 (RFMEFI57716X0212)).

References

[1] Brown, Matthew, and David G. Lowe. "Automatic panoramic image stitching using invariant features." International journal of computer vision 74.1 (2007): 59-73.

[2] Zomet, Assaf, et al. "Seamless image stitching by minimizing false edges." IEEE Transactions on Image Processing 15.4 (2006): 969-977.

[3] Jia, Jiaya, and Chi-Keung Tang. "Image stitching using structure deformation." IEEE Transactions on Pattern Analysis and Machine Intelligence 30.4 (2008): 617-631.

[4] Levin, Anat, et al. "Seamless image stitching in the gradient domain." European Conference on Computer Vision. Springer Berlin Heidelberg, 2004.

[5] Semenishchev, Evgeny A., et al. "Stitching algorithm of the images acquired from different points of fixation." SPIE/IS&T Electronic Imaging. International Society for Optics and Photonics, 2015.

[6] Semenishchev, E. A., et al. "Investigation of methods to search for the boundaries on the image and their use on lung hardware of methods finding saliency map." SPIE Sensing Technology+ Applications. International Society for Optics and Photonics, 2015.

[7] Achanta, Radhakrishna, et al. "SLIC superpixels compared to state-of-the-art superpixel methods." IEEE transactions on pattern analysis and machine intelligence 34.11 (2012): 2274-2282.

Semenishchev Evgeny Alexandrovich received her BS in «Radioelectronics system» from South-Russian State University of Economics and Service, Shakhty, Russia (2005), received her MS in «Radioelectronics system» from South-Russian State University of

Economics and Service, Shakhty, Russia (2007), PhD in applied radiotechnics from Southern University fereralny (2009). Associate Professor of Dept. "Radio-Electronics Systems" Shakhty, Russia.

Voronin Viacheslav was born in Rostov (Russian Federation) in 1985. He received his BS in radio engineering from the South-Russian State University of Economics and Service (2006), his MS in radio engineering from the South-Russian State University of Economics and Service (2008) and his PhD in technics from Southern Federal University (2009). Voronin V. is member of Program Committee of conference SPIE. His research interests include image processing, inpainting and computer vision.