https://doi.org/10.2352/ISSN.2470-1173.2020.6.IRIACV-015

Creation of a fusion image obtained in different electromagnetic ranges used in industrial robotic systems

Evgenii Semenishchev; Moscow State Technology University (MSTU "STANKIN"); Leading Researcher; Moscow, Russian Federation
Voronin Viacheslav; Moscow State Technology University (MSTU "STANKIN"); Leading Researcher; Moscow, Russian Federation
Zelensky Aleksandr; Moscow State Technology University (MSTU "STANKIN"); Vice-rector for scientific work; Moscow, Russian Federation
Irina Tolstova; Don State Technical University (DSTU); postgraduate student; Rostov-on-Don, Russian Federation

Abstract

Modern robotic systems allow you to automate processes and increase employee productivity. To create such systems, finite state machines (sensor systems) and machine vision systems are used. The scope of their application may be the development of a robotic system within the framework of the INDUSTRY 4.0 project. The article proposes an approach based on combining data obtained by fusion from cameras working in various electromagnetic ranges. An approach is proposed that is based on a fusion of data on the shape, boundaries, and parameters of objects. The search for the boundaries and shape of objects is based on the layer-by-layer simplification of images with the allocation of local features at each level. The search for local features is based on the allocation of local stationary sections, the allocation of the boundaries of objects, the determination of their parameters and the combination of information in a single information space. The search for boundaries is based on the use of the method of combined image analysis with a joint analysis of the L2 norm criterion. As a measure of a discrepancy, the square of the difference of the input value and the resulting estimate is used. As an example, the results of fusing images based on the combination of infrared data, RGB data, and IR cameras are presented.

1. Introduction

Modern monitoring and control systems use, in addition to conventional sensors, information received from the outputs of photosensitive matrices. Analysis of the data obtained by cameras operating in the optical range helps the operator of the robotic complex to increase the efficiency of various operations. In automated systems based on the analysis of data obtained in the form of images, it is possible to completely or partially exclude human participation in conventional and monotonous operations [1-2]. As additional data, information obtained in ranges other than visible to a person can be used. Such data may be the ultraviolet range, allowing you to see static discharges and specialized labels. Near-infrared, allowing you to see in fog and lack of a light source. X-ray range allowing penetration into an object. The microwave range, which allows you to see the dispersion from the surface layer of objects with a dense structure. The far infrared spectrum, which allows you to see the temperature gradient of objects in the frame [3].

In this paper, we consider the method of fusion images. The resulting image combines data on the shape and characteristics of the objects. We proposed the methodology of fusion data in various ranges, for example, obtained in the IR spectrum, 3D and/or the visible range. The data obtained allow us to determine additional indicators that are not available from data in the one range. When creating content, the following problems are solved:

- Preprocessing. The components (noise, blur, etc.) are superimposed on images obtained in different electromagnetic ranges.
- The images obtained in the infrared spectrum can be very different from the data obtained in the visible spectrum.
- The boundaries, size, and shape of objects obtained in different electromagnetic ranges can vary greatly.

2. Model image and robotic system

The created machine vision system is based on a fusion of video data obtained by various sensor systems. The camera system is located on the manipulator of the universal robotic complex. The cameras are not stationary and can be moved. It is possible to divide the machine vision system into optics visual sensors, 3D, or IR.

2.1 Model robotic system

The model robotic system shown in Figure 1 allows you to organize the planning of the trajectory of the moving tool of the manipulator robot, for the analysis of processes and characteristics of the part, based on machine vision systems. The presented scheme uses serial data transmission. The exit from the technological cycle is carried out by the circumvention of the complex around the analyzed object. The movement is carried out by changing the coordinates of the robotic complex $O_tX_tY_tZ_t$ and the machine vision unit $O_bX_bY_bZ_b$ to the coordinate system of the part $O_pX_pY_pZ_p$. The beginning of the movement for the coordinate systems X and Y select the center or one of the corners of the part, and along the Z axis - the upper surface of the part. The conversion is as follows:

$$b = Fp + T, \quad (1)$$

when p – coordinates of the point in the part coordinate system, b – coordinates of the same point in the base coordinate system, F- rotation matrix size 3×3, T- alteration vector. To solve the problem of automated finding F and T requires the construction of a three-dimensional model of the workpiece.

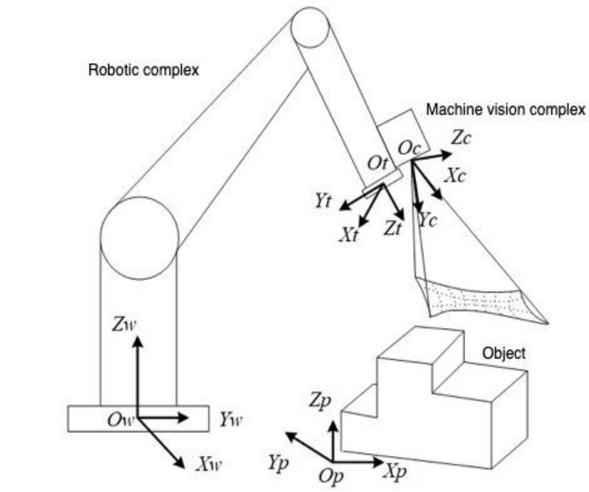


Figure 1. Model robotic complex

In order to simulate the complex optical properties of stage surfaces, we will consider the input image as the sum of the Lambert and mirror components. To do this, we introduce the equation of intrinsic decomposition based on the rendering equation:

$$S = \int_{\Omega+} \rho(\omega_i, \omega_0)(N \cdot w_i) L(w_i) d\omega_i \ . \qquad (2)$$

We use the following notation: ω_0 - shooting direction, ω_i - the direction of the upper hemisphere $\Omega+$, and N – the direction of the normal to the surface of the object. In the expression, each pixel of the input image S is represented as the integral of the product of the functions of the incident illumination L and the reflectivity of the surface ρ.

In general, surface reflectivity is a four-dimensional function, usually defined as a bidirectional reflectance distribution function (BRDF) [4]. The proposed BRDF models have a similar structure with the diffuse ρ_d, mirror term ρ_s and the corresponding coefficients α_d, α_s:

$$\rho = \alpha_d \cdot \rho_d(w_i, w_0) + \alpha_s \cdot \rho_s(w_i, w_0). \qquad (3)$$

The mirror component $\alpha_s S_s$ has characteristics that differ from the diffuse component $\alpha_d S_d$. Mirror albedo and shading have high-frequency spatial color changes, which makes decomposition ambiguous. Modeling the specular reflectivity leads to non-Lambertian expansion: $S = K \times V + M$, where the input image S is decomposed into the diffuse albedo K, diffuse shading V and mirror residual term M, respectively.

After the coordinate systems of the scanning object and the universal robotic complex are connected, we are using the operation of fusing images.

2.2 Model of images of machine vision systems

An image from a group of cameras is used as input data to a machine vision system. In [5] the example of combining into a single image. Figure 2 shows a scheme for fixing data obtained by different systems.

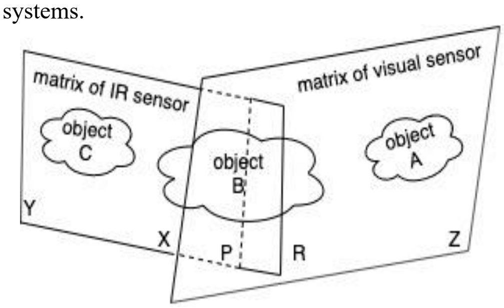


Figure 2. Scheme fixing images obtained different video systems

A simplified mathematical model of the combined image is represented as: $S_{i+k+l,n+j+m} = Y_{i+k,j+n} + Z_{k+l,j+m}$, where: $X_{i+k,j+n}$ - first image; $R_{l+k,m+j}$ - second image; $P_{k,j+n+m}$ - region that stitch the images together; $Y_{i+k,j+n} = X_{i+k,j+n} + P_{k,j+n}$ - the first image with the region that stitch the images together; $Z_{k+l,j+m} = P_{k,j+m} + R_{l+k,m+j}$ - the second image with the region that stitch the images together.

Substituting the expressions for $Y_{i+k,j+n}$ and $Z_{k+l,j+m}$ in and introducing the stitching parameter, we obtain the following form of the mathematical model: $S_{i+k+l,n+j+m} = X_{i+k,j+n} + \alpha P_{k,j+n+m} + \alpha R_{l+k,m+j}$, where: α – coefficient of the area transformations [5].

On figure 2 shows an example of combining images from objects located at various fixation zones: near (A), average (B) and distal (C). Denote $A(Z)_{lk_1,mj_1} + A(Y)_{k_1,j_1} = A$, $B(Z)_{k_2,j_2} + B(Y)_{k_2,j_2} = B$ and $C(Z)_{k_3,j_3} + C(Y)_{ik_3,nj_3} = C$, where $A \in S_{i+k+l,n+j+m}$, $B \in S_{i+k+l,n+j+m}$, $C \in S_{i+k+l,n+j+m}$, t, v – the width and height of the object of fixing. According to expression mathematical model stitching images, taking into account the semantic features will be: $S_{i+k+l,n+j+m} = X_{i+k,j+n} \cup \beta \cdot C + \alpha \cdot P_{k,j+n+m} \cup \beta \cdot B + R_{l+k,m+j} \cup \beta \cdot A$, where: β – coefficient of the type of distortion compensation.

3. Algorithm search local fiches

To create a fusion image with high additional content, the following tasks are required:

1. Image pre-processing.
2. Search local features and for correlation points between a pair of images.
3. Transformation images and fusion.

3.1 Algorithm preprocessing imaging

On figure 3 shows the primary processing image algorithm based on the application of the multicriteria method. To handle the two-dimensional digital signals (images) we propose a method based on simultaneous minimization L2 norm $\sum_i \sum_j \left(\bar{s}_{i,j} - s_{i,j}\right)^2$

[6]. Here is used measurement results $s_{i,j}$ of the assessments $\bar{s}_{i,j}$ and the first order square difference sequence of estimates row $\sum_i \sum_j (\bar{s}_{i,j} - \bar{s}_{i,j-1})^2$ and columns $\sum_i \sum_j (\bar{s}_{i,j} - \bar{s}_{i-1,j})^2$. The algorithm allows you to perform various operations such as suppressing noise, restoring blur along the edges, improving the boundaries of objects. The algorithm shown in Figure 3 is executed in a local window. The size of the window depends on the dimension of the local stationary regions. Window size is determined automatically based on the received function ratings.

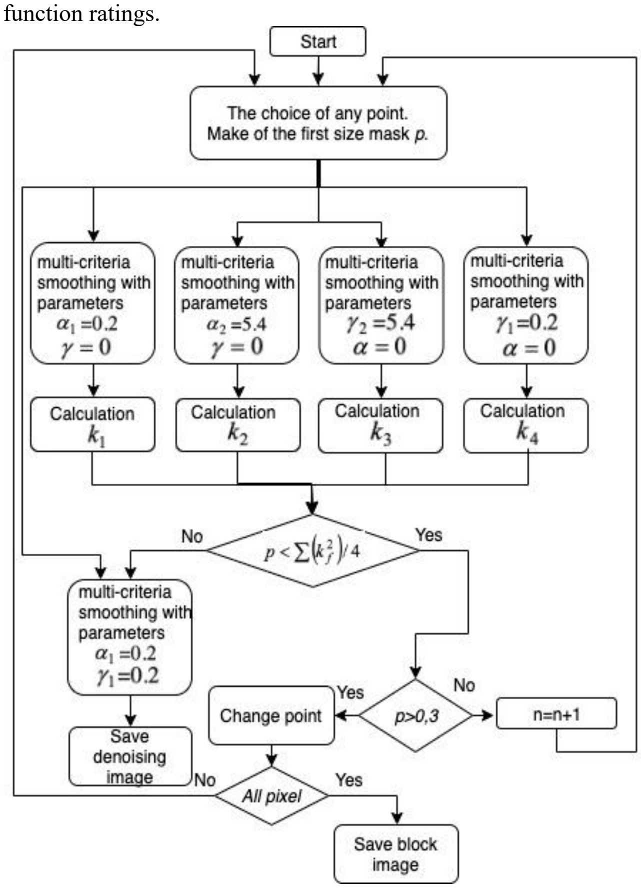


Figure 3. Algorithm preprocessing on multicriteria method

The algorithm presented in Figure 3 is implemented as follows:

1. The loading image and parameters camera.

2. In the second step, select any point on the analyzed image. This point is the center and creates a processing block with a size of at least 3x3 pixels (block size r). For the case of an image boundary, some of the directions are taken to be zero and are not taken into account when calculating the parameters k and p.

3. The image the processing in a block of size r. For this, a multicriteria method is used, based on minimizing the objective function.

$$\varphi(\bar{s}_{i,j}) = \sum_{i=0}^{m}\sum_{j=0}^{n}(\bar{s}_{i,j} - s_{i,j})^2 + \alpha\sum_{i=1}^{m}\sum_{j=0}^{n}(\bar{s}_{i,j} - \bar{s}_{i-1,j})^2 + \gamma\sum_{i=0}^{m}\sum_{j=1}^{n}(\bar{s}_{i,j} - s_{i,j-1})^2 \qquad (4)$$

where λ, μ are fixed positive constants, $s_{i,j}$ are measurement results, $\bar{s}_{i,j}$ is assessment of result.

The technique reduces the L2 norm and the quadratic deviation of the difference between the signal and the row and column estimate. We take the expression $\bar{s}^k = (\bar{s}^k_{i,j})_{i=0}^{m}{}_{j=0}^{n}$, $k = 0,1,2,...$, iterative formula for finding the estimates $\bar{s}^* = (\bar{s}^*_{i,j})_{i=0}^{m}{}_{j=0}^{n}$, $k = 0,1,2,...$, records of the signal:

$$\bar{s}^{k+1}_{i,j} = (1 - 2\cdot\alpha\cdot c_{i,j})\bar{s}^k_{i,j} + 2\cdot\alpha\cdot(s_{i,j} + \lambda(\bar{s}^k_{i-1,j} + \bar{s}^k_{i+1,j}) + \mu(\bar{s}^k_{i,j-1} + \bar{s}^k_{i,j+1})), \quad i = 0,1,...,m; \quad j = 0,1,..,n; \quad k = 0,1,2,...$$

The resulting solution allows you to minimize the objective function and perform the assessment in a given window [7].

The calculation is performed in four directions, along the rows and columns of the image. The minimization is performed with the parameters of the: α_1=0.2, $\gamma = 0$; α_2=5.4, $\gamma = 0$; γ_1=0.2, $\alpha = 0$; γ_2=5.4, $\alpha = 0$; on this method.

4. Determination of the parameter of reaching the end of the stationary border. To determine it, we calculate the average module of the elements of the studied direction.

5. Checking the condition of exceeding the threshold:

$$p < \sum (k_f^2)/4$$

If the threshold value is not exceeded, an operation is performed to reduce the effect of noise with the parameters α_1=0.2 and γ_1=0.2.

6. In the case of a positive decision step 5, the check of exceeding the threshold value established to determine the stationary border of the studied area. If the threshold is not exceeded, the processing window is resized. If the threshold value is exceeded, the image research point changes.

7. At the final stage, the conditions for the study of all pixels in the image are checked. If the study of the entire image is completed, data on the boundaries of objects and on stationary areas are saved.

3.2 The image simplification algorithm

The operation of image simplification is applied used only for IR images and based on averaging the values of local regions. The operation is based on reducing the bit depth of the image to the level of 3 bits. When changing the bit depth, the regions are increased. When identifying the internal structures of less than 3% of the image size, they are combined with larger areas. The algorithm is presented in the work [8].

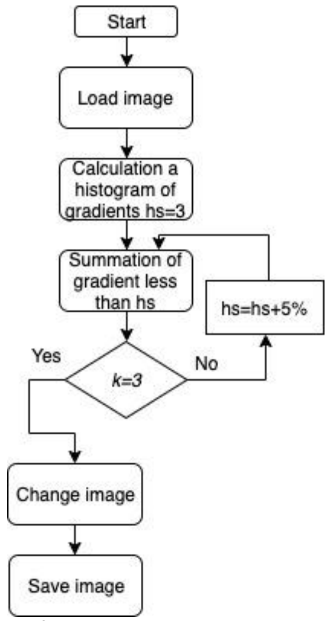
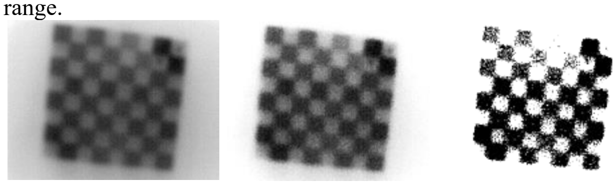


Figure 4. The image simplification algorithm

The algorithm presented in Figure 4 is implemented as follows:

1. The loading IR image and parameters camera.
2. We calculate the histogram of color gradients and accept threshold $hs=3\%$ equivalent to the max size.
3. The summation of the values of the ranges exceeding the threshold limit.
4. Checking the condition of the number of gradient histogram ranges. If the value of the ranges is not equal to 3, we increase the threshold value $hs=hs+5\%$. Otherwise, we combine the ranges and form the final image.
5. Save result of image.

Figure 5 shows an example of processing a test image of a chess field. The black squares of the chessboard are made of metal. White cells are made of foam, which is a heat insulator. With a sharp change in ambient temperature, the cell's visible in the IR range.



a) b) c)

Figure 5. example processing test images

Figure 5 shows an example of processing the input image of a chessboard (5a). Figure 5b shows the result of reducing the action of the noise component in the image. Figure 5c shows the result simplification image. The final step is to unionize the parameters of the cells of the chess field with the resulting image.

4. Algorithm fusion images

On Figure 6 shows the algorithm for fusion images obtained in various electromagnetic ranges. The input data can be high-resolution images obtained in the visible range, 3D, ultraviolet or IR images.

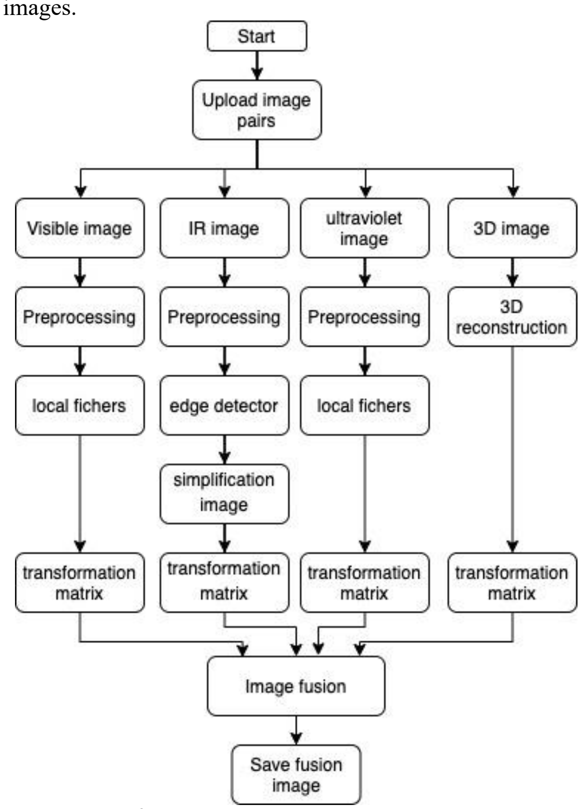


Figure 6. Algorithm fusion images

The algorithm presented in Figure 6 is implemented as follows:

1. The load image pairs and parameter cameras.
2. In the next step, an image pre-processing operation is performed. The step may include operations: reducing the effect of the noise component; image simplification; restore blur; gamma correction or others.
3. For 3D images, the operation of restoring the boundaries and edges of objects is performed. The use of this operation allows you to get rid of uncertain areas. For IR images, the operation of searching the boundaries of objects is performed. The following step simplifies thermal imaging images. This operation allows you to minimize the dimension of objects and emphasize the brightness.
4. For images of the visible and ultraviolet spectrum, the operation of searching for local features. This operation allows you to detect a group of persistent features that can be used as key points for the subsequent union.
5. The operation of image translation. This step includes zooming and searching for the transformation matrix.
6. The final step is the operation of mixing the color spaces of the images. The result of the fusion image is saved.

5. The results of the algorithm of fusion image

As test data, sets of images obtained by the machine vision system of a universal robotic complex are used. The frames were obtained in the infrared range using a FLIR c2 or SEAK thermal camera, as well as the visible spectrum on a Samsung camera. The algorithm shown in Figure 7 is used to merge the images. The camera shift was more than 0.3 meters. The distance from the camera to the object was 1.3 meters, the ambient temperature was + 22C (+ 71F).

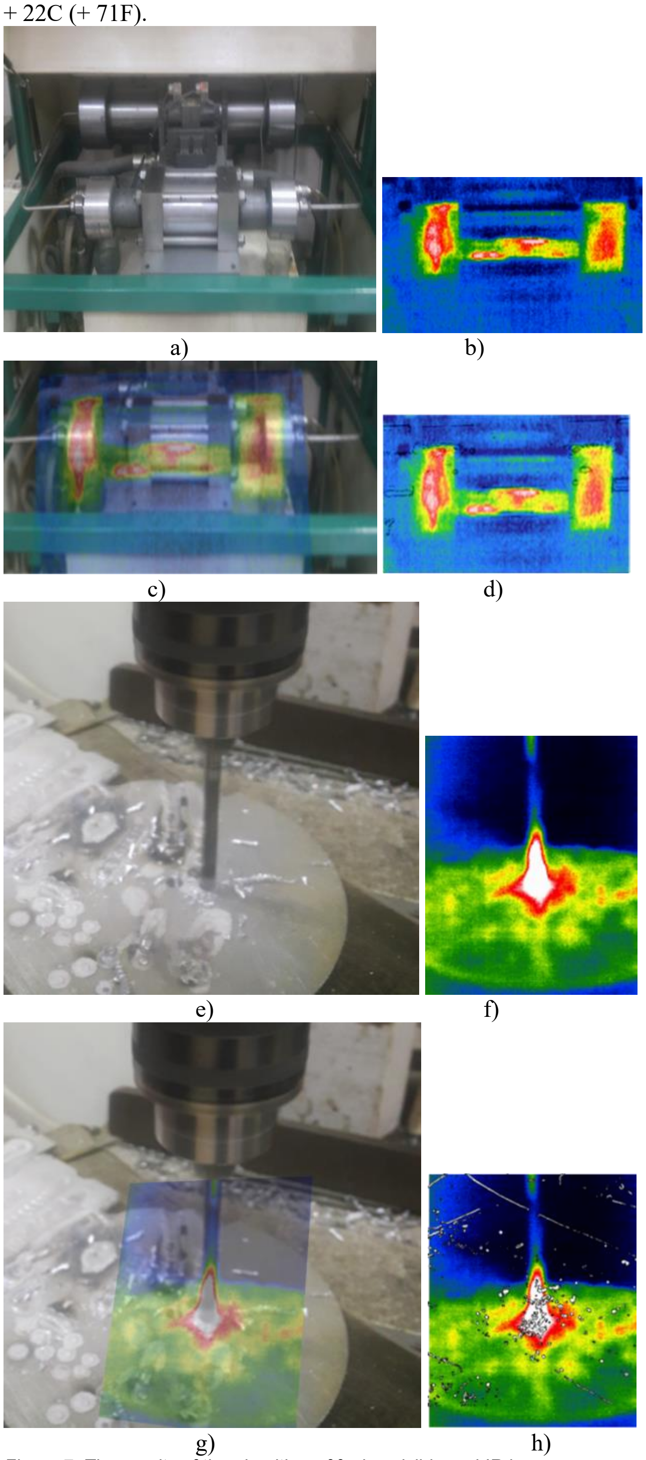


Figure 7. The results of the algorithm of fusion visible and IR image

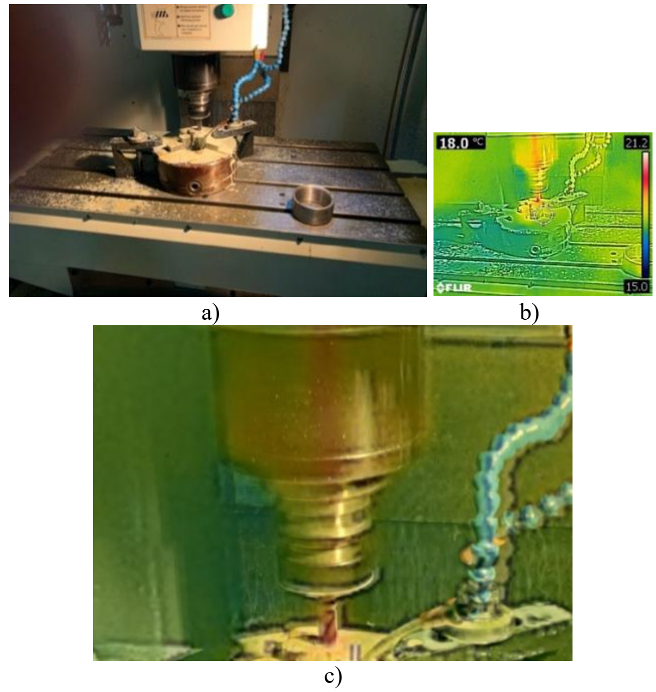


Figure 8. The results of the algorithm of fusion IR and visible image

As a result of applying the algorithm, it is possible to realize image merging, both visible (example in Figure 7c, 7g) and an IR image (example in Figure 7d, 7h or Figure 8c) can be used as a basis.

As a result of applying the algorithm (figure 7), the merging of visible and thermal images obtained from different fixation points and at different times was performed. The size of the studied images is (60 by 40 pixels - FLER C2) and (320 by 240 pixels - SEAK). The image obtained in the visible range has a resolution of 1024x768, with 8-bit color. Using standard approaches to search for local signs does not allow finding stable signs.

6. Conclusion

As a result of this work, it was developed by the method fusion images obtained in visible and other ranges. The algorithm is based on the search for local features, and the analysis of stationary edges. An example of a fusion image is shown on a set of test images captured by thermal and optic cameras

Acknowledgment

This work was supported by Russian Ministry of Education and Science in accordance to the Government Decree No 218 from April 9, 2010 (project number No 074-11-2018-013 from May 31, 2018 (03.G25.31.0284)).

References

[1] Fukuda T., Nakagawa S. “A dynamically reconfigurable robotic system,” IECON'87: Industrial Applications of Robotics & Machine Vision, International Society for Optics and Photonics, Vol. 856, pp. 588-595, 1987.

[2] Golnabi H., Asadpour A. “Design and application of industrial machine vision systems,” Robotics and Computer-Integrated Manufacturing, vol.23(6), pp. 630-637, 2007.

[3] Zhao, W., Hasser, C. J., Nowlin, W. C., & Hoffman, B. D. (2012). *U.S. Patent No. 8,108,072*. Washington, DC: U.S. Patent and Trademark Office.

[4] Shi J. et al. "Learning non-lambertian object intrinsics across shapenet categories," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1685-1694, 2017.

[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, vol. 9404, p. 94040R, 2015.

[6] Semenishchev E. et al. "Image denoising using a combined criterion," SPIE Commercial+ Scientific Sensing and Imaging. – International Society for Optics and Photonics, P. 98690E, 98690E-7, 2016.

[7] Semenishchev E.A., Voronin V. V., Marchuk V. I. "Image denoising and deblurring using multispectral data," Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIII. International Society for Optics and Photonics, vol.10198, 101981G, 2017.

[8] Semenishchev E. et al. "Solving the problem transform of size images on series base points found on frames in the problem of stitching images," IOP Conference Series: Materials Science and Engineering, IOP Publishing, Vol. 680, No. 1, p. 012045, 2019.

Author Biography

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 Federal University (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.

Zelensky Aleksandr 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 (2005), his MS in radio engineering from the South-Russian State University of Economics and Service (2007) and his PhD in technics from Novocherkassk Polytechnic Institute (2010). Zelensky A. is Vice-rector for scientific work on Moscow State Technology University (MSTU "STANKIN") (2017).

Irina Tolstova received her BS in «Radioelectronics system» from Don State Technical University (DSTU), Shakhty, Russia; received her MS in «Radioelectronics system» from Don State Technical University (DSTU), Shakhty; now is postgraduate student in Don State Technical University (DSTU), Rostov-on-Don, Russian Federation.