1. Karina G. SELIVANOVA[1], 2. Oleg G. AVRUNIN[1], 3. Maksym Y. TYMKOVYCH[1], 4. Tamila V. MANHORA[2], 5. Oleh S. BEZVERKHYI[3], 6. Zbigniew OMIOTEK[4], 7. Aliya KALIZHANOVA[5,6], 8. Ainur KOZBAKOVA[5,7]

Kharkiv National University of Radio Electronics (1), Vinnytsia National Agricultural University (2), Vinnitsa State Pedagogical University named after M. Kotsiubynskyi (3), Lublin University of Technology, Poland (4), Institute of Information and Computational Technologies CS MES RK (5), University of Power Engineering and Telecommunications (6), University of Power Engineering and Telecommunications (7)
ORCID: 1. 0000-0003-1002-0761; 2. 0000-0002-6312-687X; 3. 0000-0001-5613-1104; 4. 0000-0002-7010-8768; 5. 0000-0001-9322-2646; 6. 0000-0002-6614-7799; 7. 0000-0002-5979-9756; 8. 0000-0002-5213-4882

doi:10.15199/48.2021.09.06

3D visualization of human body internal structures surface during stereo-endoscopic operations using computer vision techniques

***Abstract**. The paper describes visualization steps of the surface of internal structures of the human body during stereo-endoscopic and laparoscopic operations using modern computer vision techniques. The presented stages make it possible to obtain three-dimensional representation (more useful for representation and analysis), which is especially important for assessing the state of the examined area and for training health care specialists. The direction of further research is the development of training tools using the proposed approaches.*

Streszczenie.** W pracy opisano etapy wizualizacji powierzchni struktur wewnętrznych ciała ludzkiego podczas operacji stereo-endoskopowych i laparoskopowych z wykorzystaniem nowoczesnych technik widzenia komputerowego. Przedstawione etapy pozwalają na uzyskanie trójwymiarowej reprezentacji (bardziej przydatnej do reprezentacji i analizy), co jest szczególnie istotne dla oceny stanu badanego obszaru oraz dla szkolenia specjalistów ochrony zdrowia. Kierunkiem dalszych badań jest opracowanie narzędzi szkoleniowych wykorzystujących proponowane podejście.* ***(Etapy wizualizacji powierzchni struktur wewnętrznych ciała ludzkiego podczas operacji stereo-endoskopowych i laparoskopowych)

Keywords: computer vision, endoscopy, stereoscopic imaging, visualization.
Słowa kluczowe: widzenie komputerowe, endoskopia, obrazowanie stereoskopowe, wizualizacja.

Introduction

Modern surgery is a highly intellectual field of medicine that relies heavily on numerous technical means. Such technical means are aimed at improving or simplifying the performance of surgical procedures. This is especially important when you consider the fact that every small mistake can lead to irreparable outcomes such as the death of the patient or his injury. Endoscopy [1] and laparoscopy [2] maintain an important place among such procedures. The main purpose of which is to provide doctors with the information necessary for diagnosis on the visual state of internal tissues, organs, etc. The construction of such means is possible by using various types of optical effects [3. 4].

An important stage in the development of medical endoscopy was the shift from a two-dimensional monocular image to a three-dimensional stereoscopic image [5]. The development of modern endoscopic systems consists of two key components: hardware - an optical system and modules for converting the optical signal to a digital form; as well as software - software that provides a wide range of basic and specialized functionality.

As noted earlier, software by processing information allows you to select significant information and provide it in the required form. Such examples are both specialized medical diagnostic systems that perform processing in a pure software environment [6, 7, 8] and medical systems using special hardware signal processing .

In the case of endoscopic imaging, such systems can be evaluated based on the following technical parameters: frames per second; delay; field of view; number of optical channels; etc. [9].

Ensuring comfortable work of medical personnel during a diagnostic study largely depends on the before mentioned parameters of time consumption. Thus, the software subsystem should not impose additional significant latency and it should have high enough FPS during medical examination.

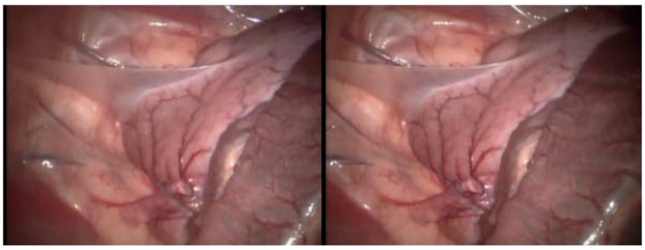
In general, such software subsystems are software image processing tools. Currently, depending on the tasks to be solved, various methods of image processing and analysis are distinguished for their purpose.

In recent years in the field of medical endoscopy, there have been trends in the use of both several optical modalities (stereo-endoscopy) [5,9], the use of such specialized modalities as polarization images, time-of-flight, a combination of different modalities, and the use of various methods of image processing and analysis.

Thus, the development and improvement of software for visualization of the surface of the internal structures of the human body based on stereo-endoscopic and laparoscopic images, taking into account the limitations on the processing and visualization time is the purpose of this study.

Materials and methods

As input data, we used the data of stereo-laparoscopic studies presented in the form of two RGB video streams, with a resolution of 640x480 px, with a frame rate of 30 frames per second (Fig. 1). The study and the corresponding data were produced and published by a group of researchers in the relevant works.



A B

Fig.1. An example of an endoscopic stereo-pair: A – left image, B – right image

The internal parameters of each of each cameras are presented in the form of a matrix K (1):

$$K = \begin{bmatrix} fk_x & 0 & p_x \\ 0 & fk_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where: K – intrinsic camera matrix, fk_x – scaled focal length in x axis, fk_y – scaled focal length in y axis, p_x – center pixel in x axis, p_y – center pixel in y axis.

Thus, for instance, for left camera in a given dataset (Fig. 1), the matrix *K* is as follows (2):

$$K = \begin{bmatrix} 751,7 & 0 & 338,7 \\ 0 & 766,3 & 258,0 \\ 0 & 0 & 1,0 \end{bmatrix} \tag{2}$$

Besides, to describe nonlinear distortions, the following distortion model is used (3) [23,21,22]:

$$\begin{aligned} x'' &= x' \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^{62}} + \\ &\quad + 2 p_1 x' y' + p_2 (r^2 + 2x'^2); \\ y'' &= y' \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^2} + \\ &\quad + p_1 (r^2 + 2y'^2) + 2 p_2 x' y', \end{aligned} \tag{3}$$

where: x', y' – projected pixel coordinates without distortion in x and y axis, respectively, x'', y'' – projected pixel coordinates after distortion in x and y axis, respectively, k_1, k_2, k_3, k_4, k_5, k_6, – radial distortion coefficients, p_1, p_2,– tangential distortion coefficients, $r^2 = x'^2 + y'^2$.

For the dataset (Fig. 1) of left camera, nonlinear distortion coefficients are presented in Table. 1.

Table 1. Used nonlinear distortion coefficients

Coefficient	Value
k_1	-0,328424
k_2	0,856059
k_3	0
k_4	0
k_5	0
k_6	0
p_1	0,003430
p_1	0,000248

The schamatical representation of stereo-endoscope presented bellow (Fig. 2).

Obtaining the distance (Z) to the object (the rendered surface point of internal structures) can be expressed based on the expression for the projection of the point (*X*) to the left and right cameras (4):

$$\begin{aligned} x_L &= f \frac{X}{Z}; \\ y_L &= f \frac{Y}{Z}; \\ x_R &= f \frac{X - B}{Z}; \\ y_R &= f \frac{Y}{Z}. \end{aligned} \tag{4}$$

Based on this, the parameter disparity (d) can be expressed (5) as follow:

$$d = x_L - x_R = f \frac{X}{Z} - \left(f \frac{X}{Z} - f \frac{B}{Z} \right) = \frac{fB}{Z}. \tag{5}$$

And accordingly, the distance (Z) itself is calculated based on the expression (6):

$$Z = \frac{fB}{d}. \tag{6}$$

Therefore, to calculate the distance Z, the parameter d is required, which is calculated by subtracting the corresponding projections in the x-axis on pairs of images (5).

As seen from Fig. 2, the main components are two cameras (left – *L* and right – *R*). They register optical radiation in the visible range, registering color in the form of an additive RGB-model. Also, one should take into account the presence of an illumination source that ensures the presence of light in the examined cavities.

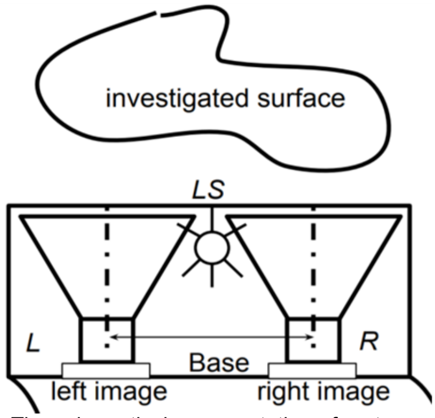


Fig.2. The schematical representation of a stereo-endoscope: *L*, *R* – left and right cameras respectively; *LS* – light source

The optical system of a stereo-endoscope is presented in the traditional form for stereovision systems and can be shown in the following form (Fig. 3).

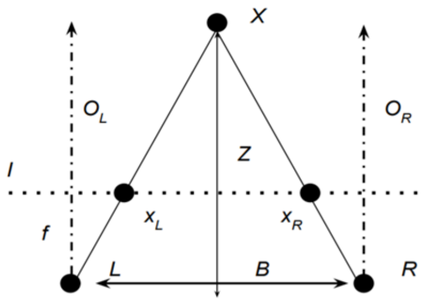


Fig.3. The optical system of a stereo-endoscope: *L*, *R* – left and right cameras, *X* – point in 3d space, *Z* – distance to the point, *I* – image plane. *f* – focal length of cameras, O_L, O_R – Optical axis for left and right cameras respectively, x_L, x_R – projections of point X into image plane of left and right cameras respectively, B – baseline

Consequently, the development of such a software solution should solve the problem of finding correspondences between the stereo pairs (Fig. 4).

An example of keypoint matching on a stereo-endoscopic pair of images is shown in Fig. 5.

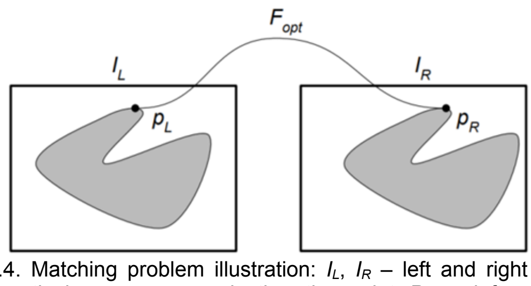


Fig.4. Matching problem illustration: I_L, I_R – left and right images respectively, p_L, p_R – projection the point *P* on left and right cameras respectively, F_{opt} – matching function

A huge number of articles are devoted to solving this problem. One approach is to find specialized image elements - keypoints.

These points must have certain properties that would allow their unambiguous selection from the image. Also, for each such key point, a descriptor is calculated, which is a specialized vector of features that allows you to determine the measure of the similarity of one keypoint with another keypoint. Certain requirements are also imposed on descriptors, namely, the descriptor must be invariant to: scale; rotation; illumination change; noise; etc.

In the field of image processing and analysis, researchers have developed a different set of methods for extracting keypoints, as well as their descriptors. Each of the techniques has its advantages and disadvantages.

The most widely used methods are SIFT (Scale Invariant Feature Transform), SURF (Speed Up Robust Features), ORB or oriented FAST (Features from accelerated segment test) rotated BRIEF (Binary Robust Independent Elementary Features).

An example of keypoint matching on a stereo-endoscopic pair of images is shown in Fig. 5.

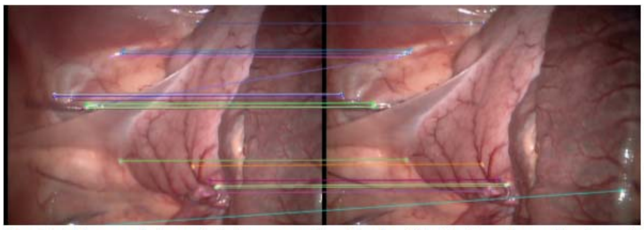


Fig.5. Key point matching results (only 25 first pairs displayed) using the ORB approach

As seen from the figure, correspondences were determined on stereopairs of endoscopic images. At the same time, there is a possibility of false positive connections. In this case, false connections appeared in the regions of glare (areas with high reflectivity; the intensity of the reflected optical radiation is quite high).

Also important is the fact that this type of matching also is sparse matching (the pairs are determined per keypoint). This approach is insufficient for a three-dimensional reconstruction of dense data. Besides, additional restrictions should be imposed on the position of the matching pairs, thereby providing a more correct result.

This result can be accomplished using specialized approaches that are aimed at getting the most optimal areas of pairs of image areas, taking into account the relative spatial position.

Conventional surface rendering techniques only display texture information, namely color. At the same time, to ensure the clarity of information, as well as the contrast of individual surface elements, it is advisable to carry out an additional stage of segmentation.

Thus, in automatic or semiautomatic mode, the selection of texture features that have certain color characteristics is carried out. This is especially important when undertaking time-limited or stressful (by nature) research. This allows the researcher to signal the presence of some kind of anomaly, and, accordingly, to attract his influence. In addition, it is possible to build specialized 3D maps. For example, 3D maps of blood vessels.

In general, the segmentation of such images, for the subsequent display of the investigated surface, can be carried out in different ways [12, 13]. In the field of medical image processing, widely used classical methods of thresholding, which, based on the intensity of a particular color component, or their combination in a particular color-space, allows separate pixels that correspond to one type of anatomical structure from others. And accordingly, obtain a three-dimensional distribution of such structures in space. So also more complex ones, which use approaches using deep neural networks, such as U-Net [13].

For this type of image, it is advisable to perform segmentation using thresholds. For this, histogram measurements of the brightness levels of pixels were carried out, which correspond to different types of tissues. In fig. 6 shows the example image histogram.

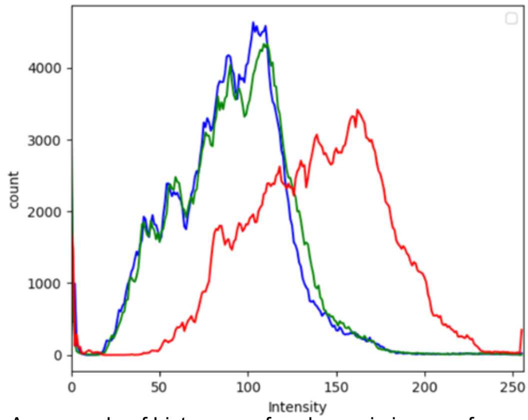


Fig.6. An example of histogram of endoscopic image of one view of stereo-pair

For this, in addition to threshold segmentation, the Sobel filter was also used, which allows detecting changes in intensity, and for this type of image corresponds to vessels. The segmentation result is shown in Fig. 7.

As seen from Fig. 7 some areas are lonely pixels, falsely segmented.

To combat this problem, it is advisable to apply morphological operations such as erosion and dilation.

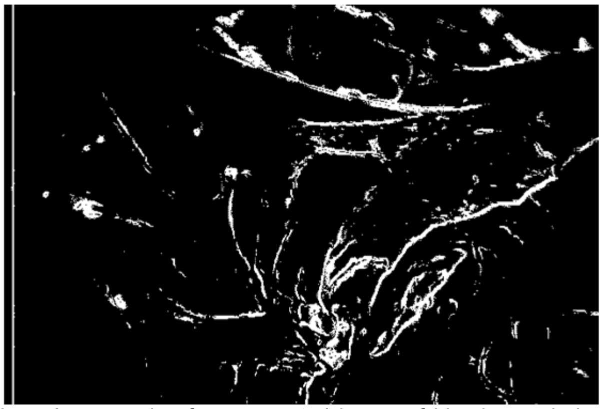


Fig.7. An example of a segmented image of blood vessels in an endoscopic image

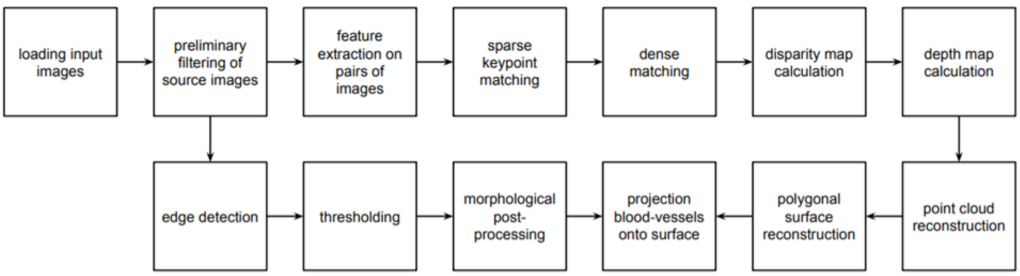


Fig.8. Generalized scheme of 3D visualization of surface of the human body's internal structures during stereo-endoscopic/laparoscopic procedures

Based on the above, the generalized scheme of the proposed pipeline consists of the following stages (Fig. 8):
– loading input images;
– preliminary filtering of source images;
– feature extraction on pairs of images;
– procedure of sparse match on key points on pairs of images;
– implementation of matching of all image points;
– disparity map calculation;
– depth map calculation;
– reconstruction of point cloud;
– polygonal surface reconstruction;
– edge detection;
– thresholding;
– morphological postprocessing;
– projection of blood vessels onto the reconstructed surface.

Results and discussion

Thus, as a result of the work of the developed software solution, at the stage of obtaining the map of disparity, the corresponding result was obtained (Fig. 9).

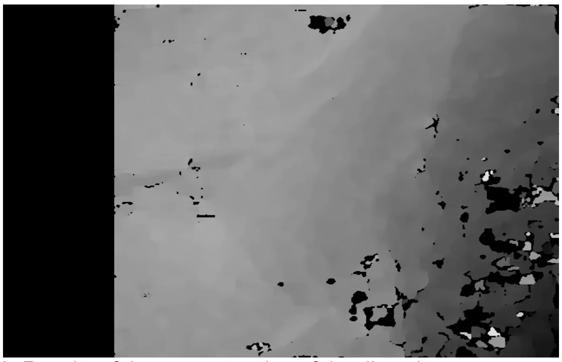


Fig.9. Results of the reconstruction of the disparity map

As seen from Fig. 9, one part of the image consists of zero values, which correspond to pixels for which disparity cannot be calculated.

Using the described sequence of actions, a three-dimensional reconstruction of the inner surface under investigation was carried out using pairs of stereoscopic endoscopic images.

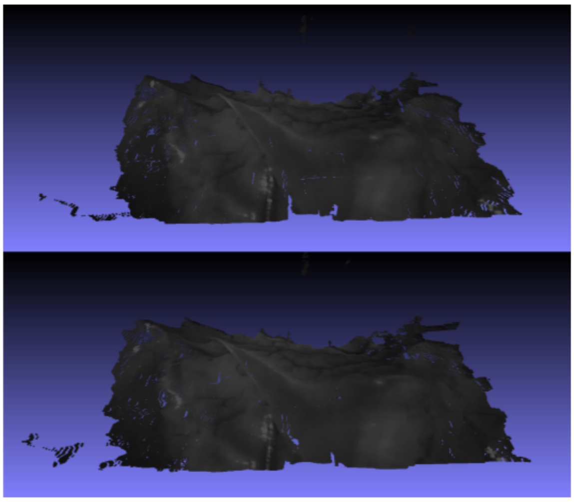


Fig.10. Results of three-dimensional reconstruction of the surface of internal structures according to the data of a stereo-endoscopic / laparoscopic examination for two frames

The results obtained reflect the surface of interest. At the same time, the generated polygonal surfaces contain inaccuracies, which are associated with both the inaccuracy of the search for matches and the presence of noise.

Conclusions

In the paper, research and development of specialized tools for visualizing the three-dimensional surface of a stereo-endoscopic / laparoscopic study was carried out. For this, an approach is used in the search for pairs of elements on stereoscopic images with subsequent calculation of disparity. It is also proposed to use additional segmentation methods to obtain visual three-dimensional objects (for example, blood vessels). The proposed pipeline can be expanded through the use of more advanced segmentation methods, which are based on the use of deep neural networks.

A promising area of the further research is the inclusion of this approach, in the development of various kinds of simulators, as a visualization tool.

***Authors**: dr inż. Karina G. Selivanova, Kharkiv National University of Radio Electronics, Nauky Ave 14, 61166, Kharkiv, Ukraine, E-mail: karina.selivanova@nure.ua; prof. dr hab. inż. Oleg G. Avrunin, Kharkiv National University of Radio Electronics, Nauky Ave 14, 61166, Kharkiv, Ukraine, E-mail: oleh.avrunin@nure.ua; dr inż. Maksym Y. Tymkovych, Kharkiv National University of Radio Electronics, Nauky Ave 14, 61166, Kharkiv, Ukraine, E-mail: maksym.tymkovych@nure.ua, Ph.D., senior lector Tamila V. Manhora, Vinnytsia National Agricultural University, e-mail: tmangora@gmail.com, Ph.D., senior lector Oleh S. Bezverkhyi, Vinnitsa State Pedagogical University named after M. Kotsiubynskyi, e-mail: bezoleg2016@gmail.com; Zbigniew Omiotek, Lublin University of Technology, Lublin, Poland, e-mail: z.omiotek@pollub.pl; Aliya Kalizhanova, Institute of Information and Computational Technologies CS MES RK, Almaty, Kazakhstan; University of Power Engineering and Telecommunications, Almaty, Kazakhstan, e-mail: kalizhanova_aliya@mail.ru; Ainur Kozbakova, Institute of Information and Computational Technologies CS MES RK, Almaty, Kazakhstan;, Almaty University of Power Engineering and Telecommunications, Kazakhstane-mail: ainur79@mail.ru*

REFERENCES

[1] Avrunin O., Tymkovych M., Drauil J. Automatized technique for three-dimensional reconstruction of cranial implant based on symmetry, *ITIB 2015 – Proceedings*, (2015), 39-42. doi:10.15199/48.2017.05.2
[2] Selivanova, K. G., et al. Virtual training system for tremor prevention. Information *Technology in Medical Diagnostics II, (2019),* 9-14.
[3] Beg S., Parra-Blanco A., Ragunath K. Optimising the performance and interpretation of small bowel capsule endoscopy, *Frontline Gastroenterol*, 9, No. 4, (2018), 300-308.
[4] Zhang H., Zeng D., Zhang H., Wang J., Liang Z., Ma J.. Applications of nonlocal means algorithm in low-dose X-ray CT image processing and reconstruction: a review, Med Phys., 44, No. 3, (2017) 1168-1185.
[5] Tymkovych M., Gryshkov O., Selivanova K., Mutsenko V., Avrunin O., Glasmacher B.. Application of Artificial Neural Networks for Analysis of Ice Recrystallization Process for Cryopreservation, *IFMBE Proceedings*, 80 (2021), 102-111. doi:10.1007/978-3-030-64610-3_13.
[6] Maćkowiak S., Segmentation of football video broadcast, *International Journal of Electronics and Telecommunications*, 59 (2013), No. 1, 33-39.
[7] Shandro B.M., Emrith K., Slabaugh G., Poullis A., Smith M.L.. Optical imaging technology in colonoscopy: Is there a role for photometric stereo?, *World J Gastrointest Endosc.*, 12, No. 5, (2020) 138-148.
[8] Kvyetnyy R., Sofina O., Orlyk P., Utreras A. J., Smolarz A., Wójcik W., Orazalieva S., Improving the quality perception of digital images using modified method of the eye aberration correction, *Proc. SPIE* 10031, (2016), 1003113; https://doi.org/10.1117/12.2249164
[9] Rymarczyk, T., Analysis medical and stereoscopic images by E-Medicus system, *Informatyka, Automatyka, Pomiary W Gospodarce I Ochronie Środowiska* 8 (2018), vol. 2, 54-57. https://doi.org/10.5604/01.3001.0012.0707