Joanna KULAWIK

Częstochowa University of Technology

doi:10.15199/48.2018.01.35

Comparison of selected methods of characteristic point detection in satellite images

***Abstract**. The article presents two methods of detecting objects in images of the surface of the earth from the air. The search was performed using local characteristic features, i.e. key points. In the first method, the corner detection was supplied using the Harris & Stephens algorithm. The descriptors were built for detection key points by the FREAK algorithm. In the second method the blobs were provided by the SURF algorithm. The descriptors were built by the SURF algorithm. After the usage of the above methods, a comparison was made. The obtained results were shown on the example images.*

Streszczenie.** W artykule przedstawiono dwa przykłady detekcji obiektów w zdjęciach powierzchni ziemi z powietrza. Wyszukiwanie wykonano przy użyciu cech charakterystycznych. W pierwszym przykładzie dokonano detekcji narożników przy użyciu algorytmu Harris & Stephens. Następnie zbudowano deskryptory do znalezionych punktów kluczowych w oparciu o algorytm FREAK. W drugim przykładzie zastosowano metodę SURF do odnalezienia plamek i zbudowania ich deskryptorów. Po użyciu powyższych metod dokonano porównania. Uzyskane wyniki zaprezentowano na przykładowych zdjęciach. –* ***Porównanie wybranych metod wykrywania charakterystycznych punktów w obrazach satelitarnych.

Keywords: corner detection, object detection, feature detection, descriptors.
Słowa kluczowe: wykrywanie narożników, wykrywanie obiektów, wykrywanie cech, deskryptory.

Introduction

Today's reality is strongly supported by digital systems. Nowadays, everybody has a video camera and digital camera. Such devices are light, small, precise, and they create good quality images. It also allowed acquiring images of the surface of the earth from the air. Video streams are used to observe, follow or detect a difference. Each video stream is thought to be a sequence of frames recorded at different time intervals. Therefore, the detection is reduced to the comparison of two images. For the computer to find objects of interest to us must be prepared for him a pattern, which will be looking for.

Working with the image is known to always start with a preliminary analysis [1]. The detection of characteristic key points is popular method of image analysis and processing [2]. The are many methods to choose from depending on the type of key points. One of the oldest methods is detecting corners and edges, as Chris Harris & Mike Stephens described in 1988 [3]. The SIFT algorithm proposed by Lowe in 1999 [4] is very popular, although it is computationally complex. The FAST corner detector [5] is much faster, but it is inaccurate. In 2006 the SURF method was proposed [6]. Over the last decade, a number of different algorithms have been developed to build the characteristic key points. One of the newer descriptors is FREAK [7], which can describe different key points. The above methods are used to study satellite imagery [8, 9, 10]. Work is continuing on problems exclusion with this kind of photos [11, 12].

An alternative solution is the detection with two-dimensional hidden Markov models [13]. Currently, the above mentioned methods are used for 3D image analysis [14]. Nowadays there is much research conducted on human identification. Besides fingerprint as a standard mean to determine one's identity, there is research on the shape of the ear [15, 16] or palm vein patterns [17, 18]. There on detection problems, for example, in difficult lighting conditions, is being conducted [19]. There is also research on medical images in which irregularity and roughness are important [20]. There are papers describing the problem of detection and classification of stamp instances in scanned documents such as stamps, logos, printed text blocks, signatures and tables [21, 22]. Another research directions is autonomy robot, which is based on Simultaneous localization and mapping (SLAM) [23, 24]. These research have vast application in many areas.

This article will present methods on the creation of patterns and the algorithms by which object detection and analysis can be made on satellite images. The purpose of presented analyzes is to compare selected methods. The results obtained will make it possible to choose a better method for further research. The Częstochowa area and its surroundings were depicted on photos which were used in research.

Fig.1. Example of the image subjected to detection characteristics of size 1536 x 768 saved in JPG format in RGB scale

Preliminary operation

Analyzed images has been downloaded from https://www.google.pl/maps and Google Earth Pro program. Many different zoom scales has been applied to obtain images having same elements but different sizes and placed differently. Images are from the same area. Images have been cropped to unified size of 1563 x 768. At this stage of preparation there has been no transformation on images. There has been assumed that exemplary image and the analyzed one have certain elements in common, what could lead to check if the analysis gives effect without detection. The image base has been created in above-mentioned way using MATLAB 2009.

The analysed images were introduced to the program as colour images in RGB format. To reduce the computational complexity in the first step, they were converted into monochrome images with a given grayscale range 0–255.

In order to suppress the unwanted image noise and improve the potential of poor technical quality of the image, filtration was made. Convolution was carried out for the two-dimensional discrete image (matrix) with the mask of the filter [1]. This operation is presented as the following formula:

$$I_2(x,y) = \sum_{p,q \in MF} I_1(x-p, y-q) w(p,q) \quad (1)$$

where: I_1 – input image (matrix), I_2 – image (matrix) obtained after filtration, *w(p,q)* – mask of the filter (presented in Table 1).

Table 1. Mask of the filter *w(p,q)*

		Column index (q)		
		+ 1	0	- 1
Row index (p)	+ 1	0.0113	0.0838	0.0113
	0	0.0838	0.6193	0.0838
	– 1	0.0113	0.0838	0.0113

As a result of filtration, values were obtained outside the range 0–255, which was the cause for the need for standardization. The image saturation values, which were received from the previous calculation, were scaled to the required range. Normalization was performed by applying the following equation:

$$I_3(x,y) = \frac{I_2(x,y) - \min(I_2(x,y))}{\max(I_2(x,y)) - \min(I_2(x,y))} 2^B \quad (2)$$

where: I_3 – image matrix obtained after normalization, *B* – is the number of bits representing the pixel

Image analysis

It is known that the compared reference image with the frame of the video stream, pixel by pixel, only returns a positive result in the laboratory. That is, the compared images were registered in the same circumstances while maintaining lighting conditions, the same or at least similar parameters hardware, maintaining the same size and resolution, and the identical set in space photographed objects. It is not possible to meet the above mentioned requirements in the real terms with images taken from the air. Moreover, such comparisons are thought to be very time consuming. It was therefore decided to analyze them based on the characteristic features. They are the specific configurations of pixels and were stacked in specific structures. Examples of such characteristic points [2] are points, corners, blobs, lines, regions and many others. To be able to compare two images, it is necessary to establish which characteristics will be analyzed.

In the first method, corners were detected. For this purpose the Harris & Stephens corner detection algorithm [3] was used, which is described by the formula:

$$E(u,v) = \sum_{x,y} w(x,y)[I(x+u, y+v) - I(x,y)]^2 \quad (3)$$

where: *w(x,y)* - mask of the filter, *I* - 2- dimensional discrete image in grayscale.

The convolution of matrix of the mask with matrix of the image was carried out. The mask of the filter was entwined consecutively with all points of 2-dimensional, discrete image. To determine the maximum corners metric value in the image Adopted minimum accepted quality of corners equal 0.01. Example output is shown in Figure 2.

Fig.2. The detection of the key points by Harris & Stephens

In the second method, the SPEEDED-UP ROBUST FEATURES (SURF) detector was used to detect the characteristic blobs. It was presented for the first time by Herbert Bay [6]. SURF is based on the sums response of two-dimensional Haar wavelets and effective use Integral images. The matrices Hesse is used for the location of key points. The determinant of this matrix reaches the maximum value for the characteristic points. Hessian matrix *H (p, σ)* at the point *P = (x, y)* of the scale *σ* is defined as follows:

$$H(p,\sigma) = \begin{pmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{xy}(p,\sigma) & L_{yy}(p,\sigma) \end{pmatrix} \quad (4)$$

where $L_{xx}(p,\sigma)$ is a convolution of the second derivative Gaussian $\frac{\partial^2}{\partial x^2} g(\sigma)$ of the input image *I* in point *p* (analogous $L_{xy}(p,\sigma)$, $L_{yy}(p,\sigma)$). Thanks to the approximation of the second derivative Gaussian function, corresponding mask filters are created that allow the use of images of the total to an efficient convolution calculation:

$$I_\Sigma(x,y) = \sum_{i=0}^{i \le x} \sum_{j=0}^{j \le y} I(i,j) \quad (5)$$

where $I_\Sigma(x,y)$ the value of total image at a point (x,y) on the input image *I*.

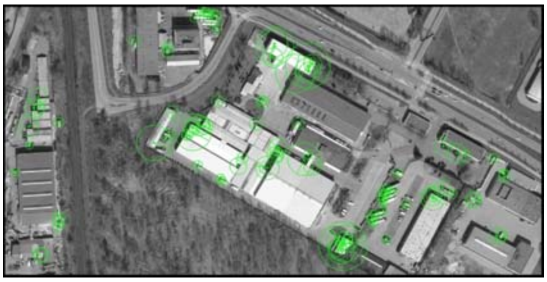

Fig.3. The detection of the key points by Speeded-Up Robust Features (SURF).

After applying above mentioned methods the key points was obtained in both the image and the reference image compared to. It should be noted, however, that although there was a reduction in the number of pixels being compared, it still can obtain unsatisfying results. Indeed suffice divergence in the resolution of images changes the lighting, movement, rotation or may change the orientation. The human eye can perceived the similarity, but the computer cannot. Therefore in the next step, descriptors will built for first method and second method.

Construction of descriptors

Having key points, it was possible to make a description of their features to the immediate surroundings of these points. It was necessary to choose such features that will be resistant to geometric transformations, scale change, and change of the lighting, while a distinctive characteristics of these points will be maintained. This will allow a significant reduction in the number of calculations and will make it possible to save time.

Corner descriptor

In the first method. To build the descriptors for the corners obtained by using Harris & Stephens algorithm the FREAK algorithm was applied (Fast Retina Key point) [7]. This manner is determined by the weighted average Gaussian for the surroundings of the characteristic point. The pattern formed by the average Gaussian shape is inspired by the human eye. Pixels are less diverse and more centred around the characteristic point. We create a matrix of nearly extracted key points. Each row corresponds to a key point represented with its large descriptor made of all possible pairs in the retina sampling pattern. We compute the mean of each column. In order to have a discriminant feature, high variance is desired. A mean of 0.5 leads to the highest variance of a binary distribution. We order the columns with respect to the highest variance. We keep the best column (mean of 0.5) and iteratively add remaining columns having low correlation with the selected columns. The algorithm sets the Orientation property of the key points output object to the orientation of the extracted features, in radians.

SURF descriptor

In the second method, the SURF algorithm contains a manner for constructing a descriptor. The first step to describe the key point is to determine its orientation [6]. Therefore, a response is calculated in Haar wavelets in a horizontal and vertical direction. We can use integral images for fast filtering. The wavelet responses are calculated and weighted with a Gaussian centred at the interest point. All calculated responses are presented in the form of points in a coordinate system. A local orientation vector is obtained by summation of all responses within a sliding window the size of 60°. Haring the point orientation the square area is created around the very point in comparison to it is orientation. In order to capture spatial information was divided into 16 smaller square subdivisions. For each of them the response of Haar wavelets is calculated. Replies d_x, d_y were weighed Gauss centred on a key point. The structure of the intensity of each subarea is described descriptor $v = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|\right)$. Four values calculated for each subarea together form a 64-dimensional descriptor key point [25].

Final detection

In both presented methods the search for characteristic features in the reference images were carried out. Then, for each of them, descriptors were built separately. Next were carried out analogously the same operations, in the analyzed images. The key points were extracted, then so do characteristic features and descriptors was made. After this the final part of the research become possible to carry out – Comparison of the images. Below there can be seen examples of images to which the described earlier methods were applied.

Fig.4. Examples of images subjected to detection of size 1536 x 768 pixels.

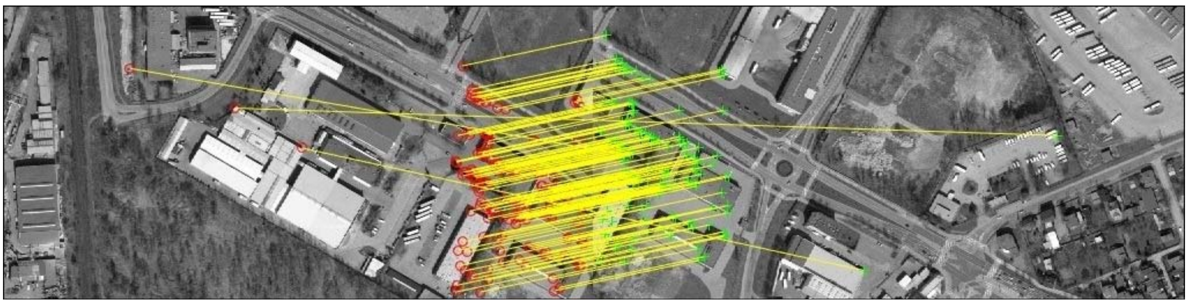

Fig.5. Example of detection corners by Harris & Stephens algorithm and FREAK description.

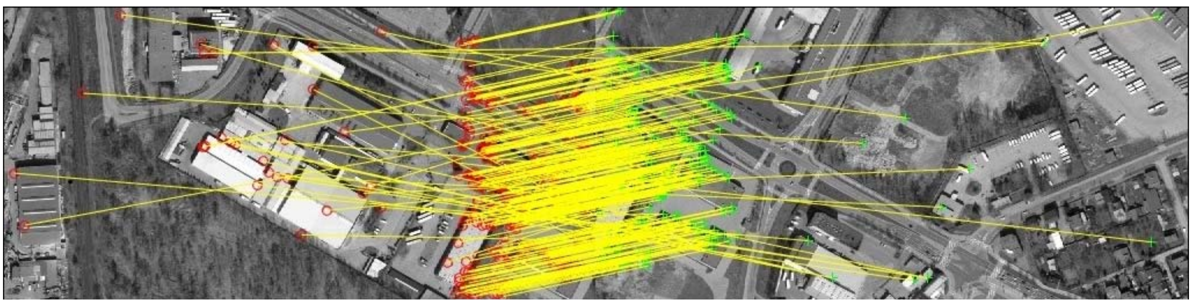

Fig.6. Example of detection objects by Speeded-Up Robust Features (SURF).

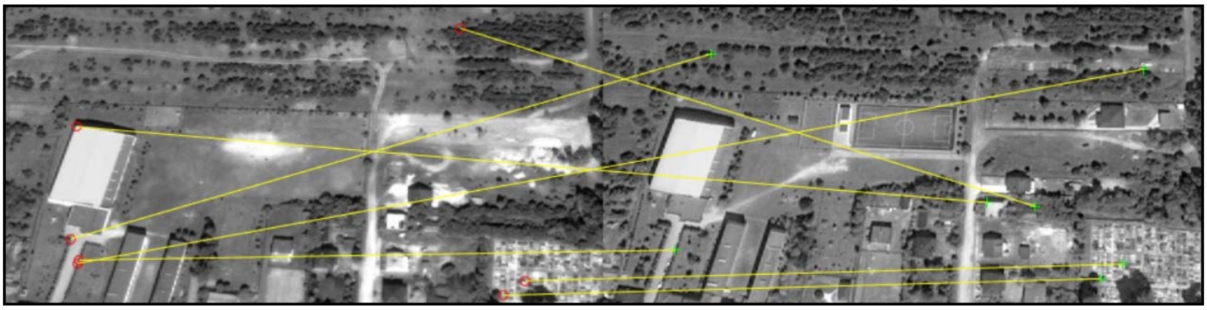

Fig.7. Example of corners detection by Harris & Stephens algorithm and FREAK description for two images of one size: the left image from 2011year; to the right image from 2015 year.

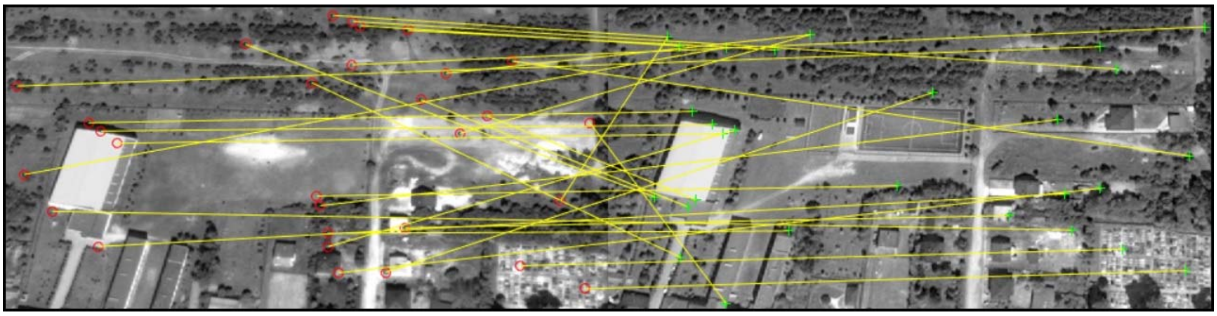

Fig.8. E Example of detection objects by Speeded-Up Robust Features (SURF) for two images of one size: the left image from 2011 year; to the right image from 2015 year.

Experiments results

The two considered methods were evaluated. In each method, four different analyses were carried out according to the types of pairs of photographs. In the first method, Harris & Stephens' key points were detected. Then FREAK descriptors were built for them. In the second method, key points were detected, and descriptors were built for them by the SURF algorithm . To obtain reliable results, the whole process for each analysis was performed for thirty pairs of comparable images. The study consisted of comparing the previously detected key points along with their descriptors from the master image in the set of key points and their descriptors of the analyzed image. Using the results obtained, the following metrics of completeness, correctness, and quality were computed [26].

"Completeness" is the percentage of key points that have been correctly identified with respect to the reference image:

$$\frac{\mathbf{TP}}{\mathbf{TP}+\mathbf{FN}} \quad (6)$$

where: *TP* - key points extracted in the analyzed image, *FN* - key points of the reference image.

"Correctness" defines the percentage of correctly identified key points in the analyzed image:

$$\frac{\mathbf{TP}}{\mathbf{TP}+\mathbf{FP}} \quad (7)$$

where: *FP* - key points obtained after the final detection:

"Quality" is a metric combining the value of "Completeness" and "Correctness":

$$\frac{\mathbf{TP}}{\mathbf{TP}+\mathbf{FN}+\mathbf{FP}} \quad (8)$$

The results obtained were averaged and presented in Table 2.

Table 2. The results of the assessment of the accuracy of detection of key points

	Completeness	Correctness	Quality
Analysis No.1	0.4533	0.7584	0.3932
Analysis No.2	0.5310	0.6188	0.4001
Analysis No.3	0.5568	0.9856	0.5536
Analysis No.4	0.4761	0.9976	0.4756
Analysis No.5	0.6192	0.9447	0.6516
Analysis No.6	0.5501	0.5841	0.3954
Analysis No.7	0.5389	0.9309	0.5179
Analysis No.8	0.4991	0.9836	0.4951

where:

- Analysis No.1 – Harris & Stephens algorithm and a FREAK descriptor for two photos of one size with a common part.
- Analysis No.2 – Harris & Stephens algorithm and a FREAK descriptor for two images where the analyzed image is part of the reference image.
- Analysis No.3 – Harris & Stephens algorithm and a FREAK descriptor for two photos of one size with a common part, and the analysed image is rotate.
- Analysis No.4 – Harris & Stephens algorithm and a FREAK descriptor for two images of one size, and the reference image comes from an older time.
- Analysis No.5 – SURF algorithm for two photos of one size with a common part.
- Analysis No.6 – SURF algorithm for two images where the analyzed image is part of the reference image.
- Analysis No.7 – SURF algorithm for two photos of one size with a common part, and the analysed image is rotate.
- Analysis No.8 – SURF algorithm for two images of one size, and the reference image comes from an older time.

Conclusions

An investigation has been conducted for the formulated first and second method. The first three analysed for each of them involved two photographs from a single time period. In contrast, in Analysis No. 4 and Analysis No. 8, two images of one area were compared, and the analysed image was from an older time. In order to be able to obtain reliable results, studies of 30 different pairs of images were performed for all eight analyses. The results obtained were averaged and presented in Table 2.

As the data were shown in Analysis No. 1, Analysis No. 2 and Analysis No. 6, very good results were obtained. In the case of Analysis No. 5 and Analysis No. 7, the results were also correct but much less accurate. Only in Analysis No. 3 from 30 tests the very few or none matching was returned. In addition, Analysis No. 4 and Analysis No. 8 have shown that image quality and various lighting conditions are a major problem in object detection. Therefore, further research will need to provide good image quality and adequate lighting. We noticed that the first method did not perform well for rotated images, and so is inapplicable in vehicle autonomy, where image rotation is important. The resulting calculation times are similar, but more favourable for the second method.

Detection carried out on individual frames were given fast and correct results, but unfortunately, there were difficulties processing them in real time. The environment in which the algorithms were prepared, required high computing power, which considerable delays were resulted. It is planned to optimize the construction of the descriptor to accelerate real-time processing.

For further research, an unmanned aerial vehicle (UAV) project, commonly known as a drone, is planned. This will be a fully autonomous vehicle. It is planned to use the simultaneous localization and mapping (SLAM) method to navigate and detect obstacles. The research described in this article is an introduction to the project being planned. Conducting a comparison of the presented methods allowed us to choose of the method we plan to use in further work.

***Authors**: mgr inż. Joanna Kulawik, Czestochowa University of Technology, Institute of Computer and Information Sciences, ul. Dąbrowskiego 73, 42-201 Czestochowa, Poland, e-mail: joanna.kulawik@icis.pcz.pl.*

REFERENCES

[1] R. Tadeusiewicz, P. Korohoda, Komputerowa analiza i przetwarzanie obrazów. Wydawnictwo Fundacji Postępu Telekomunikacji Kraków, (1997)
[2] S. Gauglitz, T. H¨ollerer, and M. Turk. Evaluation of interest point detectors and feature descriptors for visual tracking. International Journal of Computer Vision, (2011), pages 1–26
[3] C. Harris and M. Stephens, A combined corner and edge detector. Proceedings of the 4th Alvey Vision Conference, (1988), pp. 147–151
[4] D. Lowe. Object recognition from local scale-invariant features. In Computer Vision. The Proceedings of the Seventh IEEE International Conference on, vol. 2 (1999) Ieee, pp. 1150–1157
[5] E. Rosten and T. Drummond. Machine learning for high speed corner detection. Computer Vision–ECCV 2006, pp. 430–443
[6] Bay H., Tuytelaars T., Van Gool L., SURF: Speeded-up robust features, [w:] Lecture Notes in Computer Science, t. 3951 (2006), pp. 404-417
[7] A. Alahi, R. Ortiz, and P. Vandergheynst. Freak: Fast retina key point. In CVPR, 2012. 2
[8] Chaitanya Malladi, Detection of Objects in Satellite images using Supervised and Unsupervised Learning Methods, URL: http://www.divaportal.se/smash/get/diva2:1119048/FULLTEXT01.pdf, (2017)
[9] Gong Cheng, Junwei Han, A Survey on Object Detection in Optical Remote Sensing Images, DOI: 10.1016/j.isprsjprs.2016.03.014, arXiv:1603.06201, (2016)
[10] Gong Cheng, Junwei Han, Xiaoqiang Lu, Remote Sensing Image Scene Classification: Benchmark and State of the Art, DOI: 10.1109/JPROC.2017.2675998, arXiv:1703.00121, (2017)
[11] Qizhi Xu, Yun Zhang & Bo Li, Improved SIFT match for optical satellite images registration by size classification of blob-like structures, Remote Sensing Letters Vol. 5 , Iss. 5, (2014)
[12] Mohamed Tahoun, El Rahman Shabayek, Ralf Reulke, Aboul Ella Hassanien, Co-registration of Satellite Images Based on Invariant Local Features, Advances in Intelligent Systems and Computing book series (AISC, volume 323), Intelligent Systems'2014, pp 653-660
[13] J. Bobulski, H. Piech, Road signs recognition with two-dimensional hidden Markov models, Przegląd Elektrotechniczny, 91, 2/2015, 123-126
[14] V. V. Hnatushenko, O. O. Kavats, M. Kubanek, Y. O. Kibukevych, Conditions and limitations of digital satellite image pre-processing for the further 3D modeling. Journal of Applied Mathematics and Computational Mechanics, Volume 15 (2016), Issue 3, pp 57-65
[15] Choraś Michał, Image Processing Methods in Person Identification Applications - Ear Biometrics, Electrotechnical Review (Przegląd Elektrotechniczny) no. 4/81 (2005), 38-41
[16] Choraś Michał, Ear biometrics: feature extraction methods based on geometrical parameters, Przegląd Elektrotechniczny (Electrotechnical Review) no. 12/2006, pp. 5-10, (written in Polish)
[17] M. Kubanek, D. Smorawa, Verification of Identity Based on Palm Vein and Palm-Print. Advances in Intelligent Systems and Computing, Soft Computing in Computer and Information Science, Volume 342 (2015), pp 139-146, Springer
[18] D. Smorawa, M. Kubanek, Biometric Systems Based of Palm Vein Patterns. Journal of Telecommunications and Information Technology, Volume 2 (2015), pp 18-22
[19] Forczmański Paweł, Kukharev Georgy, Shchegoleva Nadezdha, 2012, An algorithm of face recognition under difficult lighting conditions, Przegląd Elektrotechniczny (Electrical Review), r. 88 No. 10b/2012, pp. 201-204.
[20] M. Borowska, A. Kitlas, E. Oczeretko, A. Radwański, I. Szarmach, J. Szarmach, Fractal Analysis of Dental Radiographic Images in the Irregular Regions of Interest, Information Technologies in Biomedicine, vol. 2, 2nd International Conference, Kamień Śląski, 2010, June 7-9, (E. Piętka, J. Kawa Ed(s).), Adv. Intelligent Soft Comput. vol. 69, (publ. by) Springer, (2010), pp. 191-199
[21] Forczmański P., Dziurzański P., System-level hardware implementation of simplified Low-Level color image descriptor, Proceedings 8th International Conference On Computer Recognition Systems CORES 2013, Advances in Intelligent Systems and Computing, vol. 226 (2013), pp. 461-468
[22] Paweł Forczmański, Andrzej Markiewicz, Two-stage approach to extracting visual objects from paper documents, Machine Vision and Applications, Volume 27 (2016), Issue 8, pp 1243–1257
[23] Chen-Chien Hsu, Cheng-Kai Yang, Yi-Hsing Chien, Yin-Tien Wang, Wei-Yen Wang, Chiang-Heng Chien, (2017) Computationally efficient algorithm for vision-based simultaneous localization and mapping of mobile robots, Engineering Computations, Vol. 34 Issue: 4
[24] Louis-Kenzo Cahier, Tetsuya Ogata, Hiroshi G. Okuno, Incremental probabilistic geometry estimation for robot scene understanding, Robotics and Automation (ICRA) 2012 IEEE International Conference on, pp. 3625-3630, (2012), ISSN 1050-4729
[25] Bay H., Ess A., Tuytelaars T., Van Gool L., SURF: Speeded-Up Robust Features, [w:] Computer Vision and Image Understanding (CVIU), t. 110 (2008), pp. 346-359
[26] Zeng C., Wang J., Automated Building Information Extraction and Evaluation from Highresolution Remotely Sensed Data, The University of Western Ontario, Electronic Thesis and Dissertation Repository, (2014)