Calibration of Vision Systems Operating in Separate Coordinate Systems

Damian Kacperski, Wojciech Sankowski, Michał Włodarczyk, and Kamil Grabowski

***Abstract*—Over the last years, the use of multiple cameras is becoming more and more popular in today's computer vision systems. Such approach is widely used in many applications, such as navigation of autonomous mobile robots, video surveillance, the movie industry, augmented reality or people tracking and identification systems. Surprisingly, little attention is paid in the literature to the practical calibration procedures that can be employed to map space between various vision systems. Therefore, in this paper a novel approach that allows to map space between cameras with different coordinate systems: Cartesian and polar is presented. The practical problems that occurs in such scenarios are analysed and thoroughly discussed. The authors present step-by-step description of the proposed calibration procedure. A series of experiments were conducted to confirm the correctness of the presented approach and to demonstrate how to apply the developed solution in practical applications. The proposed method does not require any additional equipment beyond the standard calibration chessboard. The achieved results indicate, that for evaluated cameras configuration, the maximum mapping error for the horizontal and vertical axes does not exceed $0.6°$. Obtained results are encouraging and useful for development of similar solutions.**

***Index Terms*—cameras space mapping, coordinate system transformation, stereo vision**

I. Introduction

OVER the last years, the use of multiple cameras is becoming more and more popular in today's computer vision systems. Such approach is widely used in many applications, such as navigation of autonomous mobile robots [1], video surveillance [2], [3], the movie industry [4], augmented reality [5] or people tracking and identification systems [6], [7]. In most of these applications, the high accuracy of cameras space mapping is required for further processing steps. For instance, in 3D video production systems, continuous monitoring of cameras alignment allows to automatically modify their positions in real time and save time during post production processes [8].

Surprisingly, little attention is paid in the literature to the practical calibration procedures that can be employed to map space between various vision systems. Most available descriptions, like [9] or [10], focus on sophisticated and dedicated for specific configurations solutions only. Therefore, in this paper a novel approach that allows to map space between cameras with different coordinate systems: Cartesian and polar is presented. The paper is addressed to practical problems that occur in such scenarios. Authors provide detailed description of the calibration and verification procedure. The proposed method does not require any additional equipment beyond the standard calibration chessboard. Presented method is dedicated for the non-cooperative biometric system where two wide-field of view (WFOV) and one narrow-field of view (NFOV) cameras are used. The WFOV cameras are employed to observe the entire scene and locate potential objects that are to be identified. They form a stereo pair. Once the system decides that the distance and pose of a tracked subject is sufficient to perform the recognition, the NFOV camera is directed to a specific point of the scene to acquire high quality images of selected traits, such as iris or periocular (small region surrounding eye).

The paper is organized as follows. Section II contains the overall description of the employed vision systems. Section III presents a calibration procedure proposed by the authors. Section IV shows performance evaluation and obtained results. Finally, Section VII concludes the paper.

D. Kacperski, W. Sankowski, M. Włodarczyk and K. Grabowski are with Department of Microelectronics and Computer Science, Lodz University of Technology, ul. Wolczanska 221/223, 90-924 Lodz, Poland (emails: {dkacperski, wsan, mwlodarczyk, kgrabowski}@dmcs.pl)

II. System Overview

The WFOV system consists of two monochromatic Imaging Source DKM33GR0134 cameras. Each camera is equipped with 1/3" optical sensor with a resolution of 1280 x 960 pixels and a pixel size of 3 um. They use global shutter so the acquired images are free of visual artefacts, like motion blur or distorted shapes. Cameras are equipped with H614-MQ lenses manufactured by Pentax. They are mounted on a c-profile handle with movable camera holders. It is presented in Fig. 1.

Fig. 1. Wide field of view vision system

ISSN 2080-8755

The NFOV system is realized with the use of a custom designed solution proposed by the authors. Its construction is protected by two patent applications: P30994PL00/MB and PCT/B2016/052779. Designed device is based on a specialized lens with advanced pupil and galvanometric motors equipped with mirrors. This allows to change the camera view position like in Pan-Tilt-Zoom (PTZ) cameras. However, in this case the motion is much faster and free of undesirable vibrations. This is somewhat similar to Okumura et al. solution [11]. The NFOV camera operating range is equal to $\pm 26^{\circ}$ and $\pm 30^{\circ}$ for pan and tilt directions respectively. It is equipped with specialized focusing mechanism that allows to capture sharp images within the distance range of 0.8 to 3.0 meters. Data acquisition is realized with the use of Genie TS-C3500 camera manufactured by Teledyne DALSA. It includes a CMOS color sensor with a resolution of 3520 x 2200 pixels and a pixel size of $6\mu m$. Described device is shown in Fig. 2. Its main advantage is the possibility to immediately switch between selected positions without spending additional time on stabilization. In authors solution, the only mechanical elements are galvanometers. They are extremely fast (mean angle step response of $330\mu s$) and offer a long operating lifetime. Therefore, the NFOV system can be efficiently used to target into moving objects in real-time.

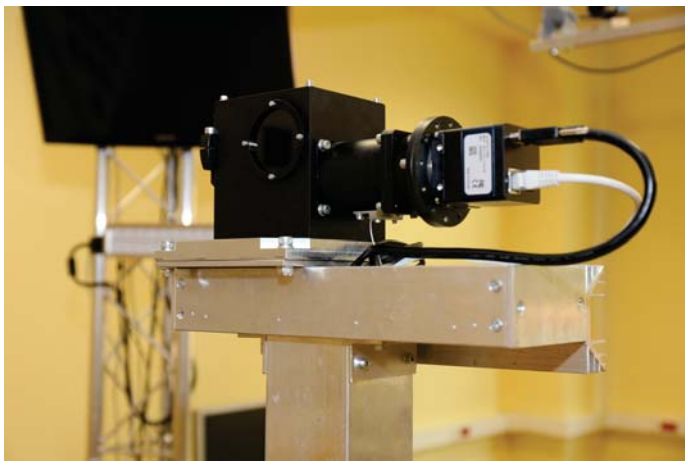

Fig. 2. Narrow field of view vision system

III. Proposed Calibration Procedure

In the described system, the WFOV stereo vision cameras return position of the observed object in 3D Cartesian coordinate system. A sample point is referred with the use of (x, y, z) coordinates expressed in metric units. The NFOV camera is described with the use of 3D polar coordinate system. This is because of the use of galvanometers and mirrors. A sample point is characterized with two angles: α and β which represents the deflection of horizontal and vertical galvanometers mirrors for which the desired part of the scene is observed. Additionally, the depth parameter is used to indicate the distance between the lens and the observed object.

These two different coordinates systems must be calibrated together. To perform the calibration procedure it is necessary to choose common key points in space and find their 3D locations in both coordinate systems. At least three points are required to perform calibration. These key points should be defined in such a way that its detection can be relatively easy. The more precisely these points are located by each camera, the more accurate results can be obtained. It was decided that the best solution would be to use the chessboard pattern for calibration. Authors used the board that contains 10×7 fields. Each field size is 120.2 mm $\times$ 120.0 mm. The board was mounted on a regulated stand. Such construction allows to easily move it around the scene and adjust its height. The same pattern is also used to calibrate the stereo vision system. The board is set in such a way that it is visible by both WFOV cameras. At the same time, it is also located at the smallest possible distance from the NFOV camera. This means that the calibration is performed for the area in which the NFOV camera depth is the smallest so the calibration accuracy should be the highest. Finally, the authors defined the points that are used for calibration and verification purposes. Their exact location is presented in Fig. 3. They are described as follows: Top Center (TC), Bottom Center (BC), Bottom Left (BL) and Bottom Right (BR).

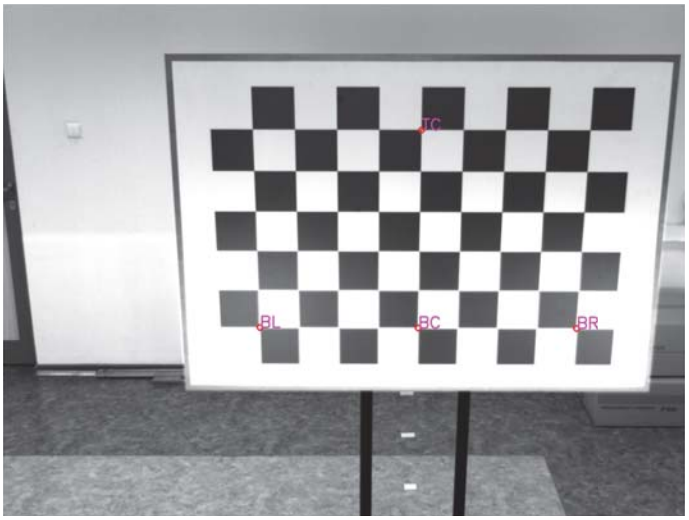

Fig. 3. WFOV keypoints collection

A. WFOV points detection

The first calibration step consists of detecting key points location in each WFOV camera image. For this purpose, the chessboard detector from the OpenCV computer vision library [12] is used. This method is used to find intersection points between chessboard fields. Next, with the use of the stereo vision, the (x, y, z) coordinates of each point are calculated.

B. NFOV points detection

The next calibration phase is to detect the location of the key points in relation to the NFOV camera. This stage is divided into two steps. Firstly, the galvanometer mirrors angle position is obtained for horizontal and vertical axis, defined as α and β respectively. This operation is repeated for each key point. Secondly, the distance between each point and the NFOV camera is calculated.

1) Intersection point detector: To obtain the position of galvanometer mirrors for each key point, it was necessary to precisely detect their location on the calibration board. Then, for each key point the mirrors position is set manually so that the key point is in the center of NFOV image. Because of the limited field of view of the NFOV camera, it was impossible to acquire the image of the entire calibration board. For this reason, it was necessary to divide the intersection points detection into two steps. Firstly, the NFOV camera is manually directed into selected key point. Then, with the use of automatic detector its position is adjusted automatically. The first step is realized manually by the operator. Basing on the live preview from NFOV camera, one can direct the camera into specific area with the use of keyboard. Then, in the second step, the points intersection detector is used to automatically adjust mirrors positions.

Unfortunately, the intersection points detector used for WFOV images can not be applied for the NFOV images. As it was mentioned before, the NFOV images do not contain the entire calibration board, but only its small parts. For this reason, the detector can not match pattern of the entire board so the intersection points are not found. Authors had to design dedicated intersection points detector that can work on a small area of the calibration board.

In the first step, the threshold operation is applied. Relatively high difference between the brightness of the top and bottom parts of the image implies that constant value threshold operation is not enough for this purpose. For this reason, authors used an adaptive threshold algorithm where the threshold value is the mean of neighbourhood areas. Result of this step is shown in Fig. 4.

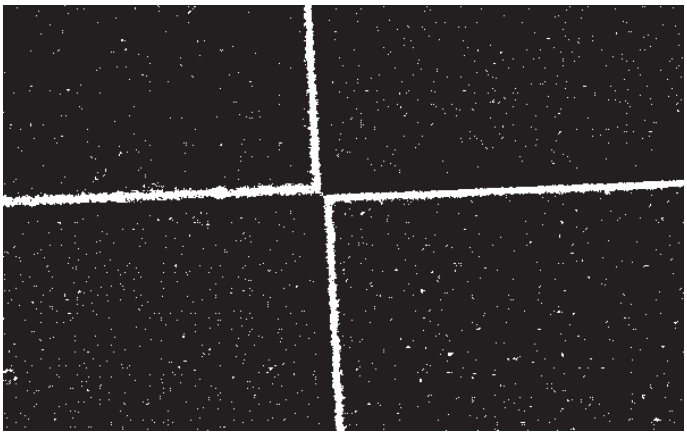

Fig. 4. Crossing points detection - adaptive threshold

The next step involves chessboard fields boundaries detection. For this purpose, the Hough Lines transformation is used. The minimum line length is set to the half size of the chessboard field. The results obtained with this operation are presented in Fig. 5.

In the next steps, the obtained lines are extended and the intersection points are detected. Additionally, the algorithm that filters the non-existing connections is introduced. For this purpose, the intersection between found lines are calculated and ones that deviate from 90° are removed. Moreover, the

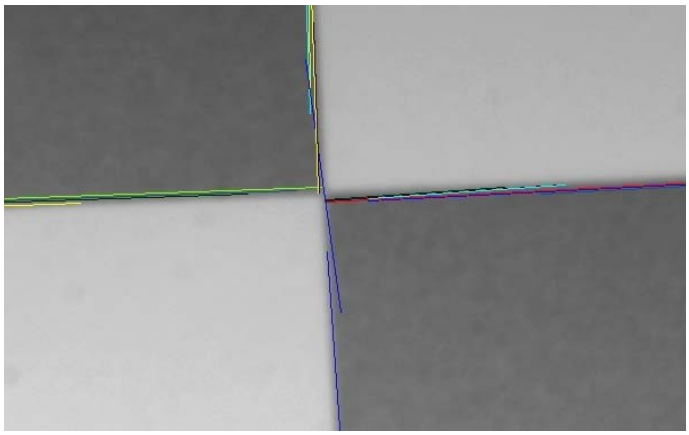

Fig. 5. Crossing points detection - Hough fitted lines

bounding boxes [13] method is used to assess the relative position between lines and filter out lines that do not have chance to intersect in the selected area. Finally, obtained intersection points are grouped into neighbourhood subsets. Their position is then averaged and the final result is calculated. The results of this operation is presented in Fig. 6.

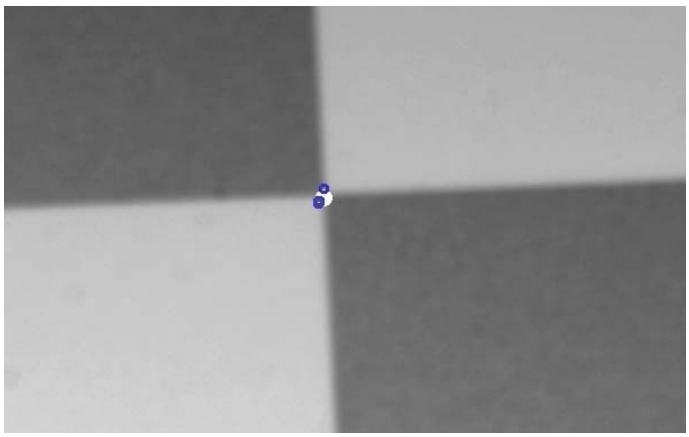

Fig. 6. NFOV keypoints collection

2) Calculating points depth: At this stage, the location of the intersection point visible by NFOV camera is already defined. However, the distance between this point and the camera is still missing. In order to acquire this information, the laser distance meter Leica DISTO D3a [14] is used. It operates in the range of 0.05 to 100 meters with measuring accuracy equals to ± 1.0 mm. The best performance is obtained when the angle between the optical path and the object is similar to 90°. For this reason, to reduce the possible errors, authors decided to measure the distance between the BC key point as it is located at angle similar to 90° (horizontally) in relation to the lens. The distance between the remaining points is calculated manually with the use of the Carnot theory. This operation is explained in Fig. 7. Applied equation is presented in (1), where L is the distance between the lens and BC point, d is the distance between BC and BR points and γ is the difference between galvanometers angles position in the direction of the Ox axis.

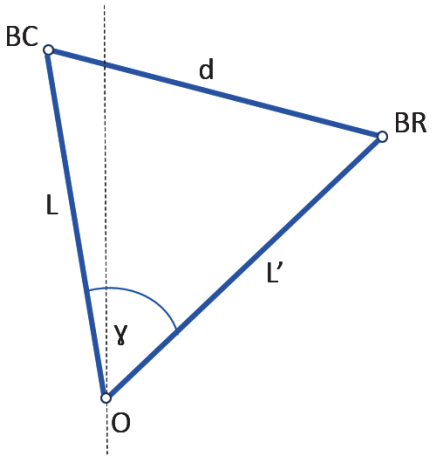

Fig. 7. NFOV points distance calculation

$$L^{'} = L \times \cos\gamma + \sqrt{d^2 - L^2 \times \sin^2\gamma} \tag{1}$$

3) Converting points: The obtained points location can be converted from the polar coordinate system to the corresponding values in the Cartesian coordinate system with the use of (2), (3) and (4). L is the distance between the point and the lens, α and β are the inclination angles for the $0x$ and $0y$ axis respectively. Presented equations are designated with the use of basic trigonometric transformations.

$$x = L \times \tan\alpha \tag{2}$$

$$y = L \times \tan\beta \tag{3}$$

$$z = \frac{L}{\sqrt{1 + \tan^2\alpha + \tan^2\beta}} \tag{4}$$

C. Transformation matrix calculation

The last stage of the calibration process is the transformation matrix generation. With the use of the detected points, two vectors $\vec{V}_{LR}$ and $\vec{V}_{TB}$ are defined. $\vec{V}_{LR}$ is defined with initial point BL and terminal point BR, while $\vec{V}_{TB}$ is described with initial point TC and terminal point BC. It cannot be assumed that they are perpendicular to each other. For this reason, the versors $\vec{V}_{OX}$, $\vec{V}_{OY}$ and $\vec{V}_{OZ}$ of the coordinate system are calculated using (5) - (7), similarly as in [15]. The versors are shown in Fig. 8. The origin of the coordinate system is set in the BC key point.

$$\vec{V}_{OX} = \frac{\vec{V}_{LR}}{|\,\vec{V}_{LR}\,|} \tag{5}$$

$$\vec{V}_{OZ} = \frac{\vec{V}_{LR} \times \vec{V}_{TB}}{|\,\vec{V}_{LR} \times \vec{V}_{TB}\,|} \tag{6}$$

$$\vec{V}_{OY} = \vec{V}_{OZ} \times \vec{V}_{OX} \tag{7}$$

The origin and the versors $\vec{V}_{OX}$ and $\vec{V}_{OY}$ provide a complete description of the coordinate system. The presented procedure is performed for both the WFOV and NFOV systems.

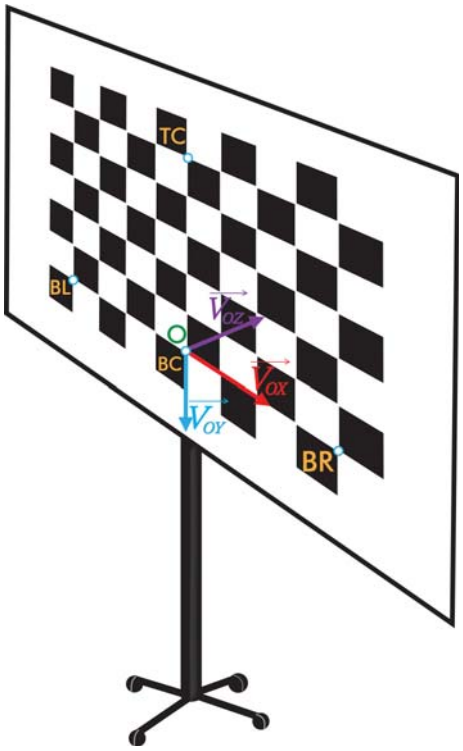

Fig. 8. Coordinate system determined by calibration chessboard

Then the transformation matrix is generated with the use of the Point Cloud Library (PCL) [16]. It is a 4x4 matrix which defines the rotation and translation operations. This data is saved in a file and is subsequently used for transformation between the coordinate systems in client applications.

IV. Verification Results

Verification procedure for proposed calibration method consists of two stages. Firstly, the targeting correctness test is used to calculate the error along X, Y axes. Secondly, the uncertainty along Z axis is verified with the depth correctness test. The following sections provide a detailed description of proposed verification procedures.

A. Targeting correctness verification

Detailed steps of the targeting correctness verification test are as follows:

- Calibration board is set to a fixed distance
- Verification point is selected from chessboard intersection points
- NFOV camera is targeted into selected verification point. Galvanometers angular deviation for horizontal and vertical axes is reported
- NFOV camera position is manually corrected with the use of chessboard key points detector. Reference galvanometers angular deviation for horizontal and vertical axes is reported
- Error is calculated as a difference between obtained galvanometers positions
- Above-mentioned operation can be repeated for all chessboard intersection points

Authors carried out the experiments for 4 different distances: 1.5, 2.0, 2.5 and 3.0 meters. The minimum distance of 1.5 meters stems from the minimum distance for which the calibration board is fully visible by WFOV cameras. The maximum distance of 3.0 meters results from the operation range of the NFOV camera, which is not able to acquire sharp images for further distances. For each distance, the accuracy of (α, β) coordinates localization for the four defined key points is calculated. As a result, the authors obtained 16 measurement points. The error for each axis was analysed separately. The experiments are summarized in Fig. 9 and Fig. 10. The results indicate, that the maximum mapping error for the horizontal and vertical axes does not exceed $0.6°$.

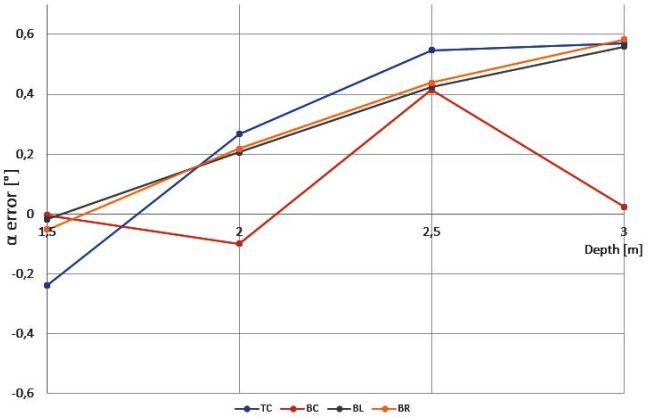

Fig. 9. Error of the α coordinate

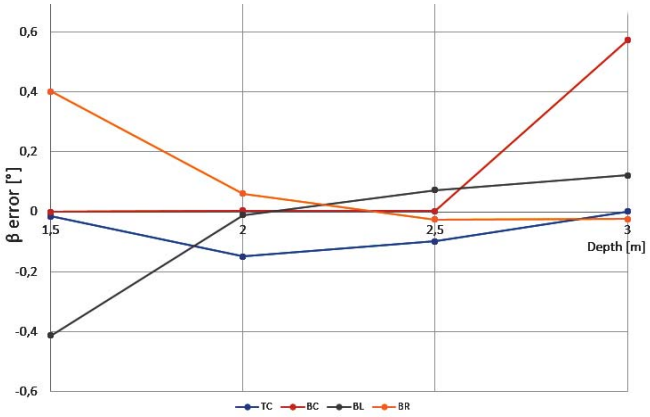

Fig. 10. Error of the $beta$ coordinate

B. Depth correctness verification

Depth correctness verification procedure consists of the following steps:

- Stereo vision calibration board is set to a fixed distance
- NFOV camera is targeted into TC intersection point which is in front of the lens. The calculated depth is reported
- The Leica DISTO D3a BT laser distance meter [14] is used to get the real distance between the top center point and the camera
- Error is calculated as a difference between calculated and real obtained with laser distance meter value
- Above-mentioned operation can be repeated for various distances

Depth correctness verification experiment was conducted for the same distance range as the targeting correctness verification test. The distance interval was 10 centimeters. As a results, authors gathered 17 measurement points. Obtained results are presented in Fig. 11.

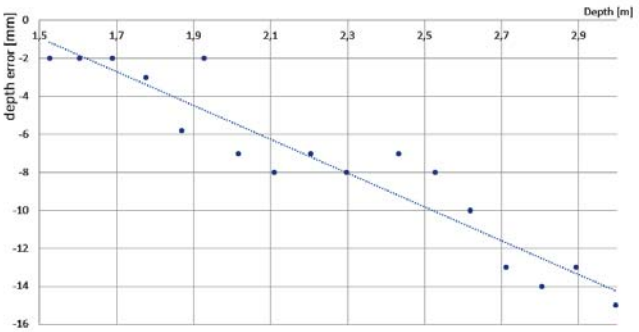

Fig. 11. Error of the $depth$

V. Conclusions

This paper describes a calibration method that can be employed to map space between cameras with different coordinate systems: Cartesian and polar. The authors present step-by-step description of the proposed procedure. A series of experiments were conducted to confirm the correctness of the presented approach and to demonstrate how to apply the developed solution in practical applications.

The proposed method does not require any additional equipment beyond the standard calibration chessboard. The achieved results indicate, that for evaluated cameras configuration, the maximum mapping error for the horizontal and vertical axes does not exceed $0.6°$. As for the depth, the obtained error consists of the random component and the large systematic component which can be approximated with the use of linear function. It can be noticed that the overall mapping accuracy degrades when the distance between the board and the cameras increases. However, this does not have a negative effect on the system performance because with a distance increasing also depth of field expands.

The obtained results are also very encouraging for the non-cooperative biometric system that the authors are developing. The obtained mapping accuracy is greater by an order of magnitude than the required value. Thereby, proposed space mapping calibration procedure can be successfully applied in the author's solution.

Acknowledgment

This presented research was funded by Polish National Centre for Research and Development in the frame of the project LIDER/027/591/L-4/12/NCBR/2013, entitled: "Non-COoperative bioMetric system for Positive AuthentiCaTion" (COMPACT)

REFERENCES

[1] M. Aranda, G. López-Nicolás, C. Sagüés, and Y. Mezouar, "Formation control of mobile robots using multiple aerial cameras," *IEEE Transactions on Robotics*, vol. 31, no. 4, pp. 1064–1071, Aug 2015.

[2] C. M. Huang and L. C. Fu, "Multitarget visual tracking based effective surveillance with cooperation of multiple active cameras," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, no. 1, pp. 234–247, Feb 2011.

[3] K. S. Kumar, S. Prasad, P. K. Saroj, and R. C. Tripathi, "Multiple cameras using real time object tracking for surveillance and security system," pp. 213–218, Nov 2010.

[4] H. Kim, R. Sakamoto, K. Kogure, and I. Kitahara, "Cinematized reality: Cinematographic 3d video system for daily life using multiple outer/inner cameras," pp. 168–168, June 2006.

[5] H. Saito, N. Inamoto, and S. Iwase, "Sports scene analysis and visualization from multiple-view video," vol. 2, pp. 1395–1398 Vol.2, June 2004.

[6] I. Everts, N. Sebe, and G. A. Jones, "Cooperative object tracking with multiple ptz cameras," pp. 323–330, Sept 2007.

[7] H. U. Chae, S.-J. Kang, and K. H. Jo, "Identification of a human using accorded blobs on the varied region from image sequence by multiple cameras," pp. 1887–1891, Oct 2008.

[8] P. Perek, D. Makowski, A. Mielczarek, A. Napieralski, and P. Sztoch, "Towards automatic calibration of stereoscopic video systems," pp. 134–137, June 2015.

[9] J. Li, P. Duan, and J. Wang, "Binocular stereo vision calibration experiment based on essential matrix," *Computer and Communications (ICCC), 2015 IEEE International Conference on*, pp. 250–254, Oct 2015.

[10] I. H. Chen and S. J. Wang, "An efficient approach for the calibration of multiple ptz cameras," *IEEE Transactions on Automation Science and Engineering*, vol. 4, no. 2, pp. 286–293, April 2007.

[11] K. Okumura, H. Oku, and M. Ishikawa, "High-speed gaze controller for millisecond-order pan/tilt camera," pp. 6186–6191, May 2011.

[12] G. Bradski, *Dr. Dobb's Journal of Software Tools*, 2000.

[13] T. M., "How to check if two line segments intersect - bounding boxes," 2013.

[14] "Leica disto d3a bt user manual," 2015.

[15] P. Nowak, W. Sankowski, and P. Krotewicz, "3d face and hand scans acquisition system dedicated for multimodal biometric identification," June 2016.

[16] R. B. Rusu and S. Cousins, "3d is here: Point cloud library (pcl)," May 9-13 2011.

Damian Kacperski is a Ph.D. student and teaching assistant at Lodz University of Technology. He received B.Sc. degree at Electronics and Telecommunication in 2012 and M.Sc. degree at Computer Science in 2013. His research is mainly focused on real-time, high performance vision systems and image processing algorithms.

Wojciech Sankowski received the MSc degree in Electronics and Telecommunication from Lodz University of Technology (Poland) in 2004. In 2009 he received the PhD degree in Computer Science from the same University. His PhD thesis concerned biometrics and was entitled "Eye image segmentation algorithms for iris pattern recognition systems". From 2009 till 2011 he worked in the software industry as a senior software developer in the TomTom Content Production Unit. Since 2011 he works as an assistant professor in the Department of Microelectronics and Computer Science, Lodz University of Technology. His research interests include image processing algorithms dedicated for biometrics and fusion of multiple biometric traits.

Michał Włodarczyk received the B.Sc. degree at Electronics in 2012 and M.Sc. degree in 2013 at Computer Science from the Lodz University of Technology, Poland. He currently serves as a PhD Student and teaching assistant at the Department of Microelectronics and Computer Science. His main reserach intrests are focused on a computer vision, pattern recognition and non-cooperative biometric recognition algorithms.

Kamil Grabowski received the Ph.D. degree in electronics from the Technical University of Lodz in 2010. His research interests are mainly focused on the field-programmable gate array and digital signal processor hardware systems for advanced image processing with an emphasis on the biometric applications, namely the study of iris-based authentication. Currently, he serves as Assistant Professor at the Department of Microelectronics and Computer Science of the Technical University of Lodz Poland, and is with the "DMCS Biometric Lab." research group. He is an author and coauthor of over 20 publications in International Journals and Conferences.