2018 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)

Robust Coded Target Recognition in Adverse Light Conditions

Tushev S., Sukhovilov B., Sartasov E.
South Ural State University
Chelyabinsk, Russia
science@tushev.org, sukhovilovbm@susu.ru, sartasovem@susu.ru

***Abstract*—We propose a new algorithm for the robust identification of coded targets in adverse light conditions. The new algorithm relies on the recognition of circle retroreflective targets. The algorithm finalizes the process of development of multi-functional coded targets for high-precision photogrammetric systems. The main advantages of the new algorithm are tolerance to adverse light conditions and a high speed. This paper reviews existing architectures of coded targets and their identification methods. We propose and describe the architecture of a coded target for high-precision photogrammetric measurements. We also analyze the efficiency of new algorithms and the range of their possible applications.**

Keywords—photogrammetry; computer vision; coded targets; fiducial markers; artificial targets; camera pose estimation; pattern recognition

I. Introduction

Fiducial markers are widely used in computer vision. They allow simplifying the determination of camera and objects' position and orientation. They also can be used to perform various measurements [1]. Moreover, artificial markers are more preferable over the naturally occurring ones (such as feature points, textures etc.) due to the fact that artificial targets are easier to recognize (because of their distinctive appearance). Thus, most algorithms expect the appearance of the fiducial markers not to undergo significant changes. This imposes certain restrictions on lighting conditions and location of the markers.

We are currently developing a photogrammetric system [2-7] that performs high-precision measurements of 3D coordinates of control points on various objects. These control points are represented by retroreflective circle targets. The system also uses special multifunctional coded targets (fiducial markers) that allows to compute relative pose of cameras, to match control points across the image set as well as to compute the scale of the scene in case large-sized high-precision scale bars are not present in the scene.

Developing the system, we faced a problem of unsatisfactory recognition of coded targets in adverse light conditions. Our first prototypes of the system used ArUco [1] as fiducial markers. At later stages, those markers were modified considerably, and finally they evolved into our own architecture of coded targets named *PhisTag* [8].

Our photogrammetric system uses flash light along with retroreflective circle targets. The measurements are performed frequently in relatively dark premises; the images are taken from a considerable distance. The shots are taken with the narrow diaphragm (aperture stop) to ensure large depth of field. This all leads to quite dark images; and the photogrammetric system cannot always recognize coded targets in these images.

Therefore, there is a need in a robust algorithm for coded targets recognition that provides high-precision data for the relative position of the camera and works reliably in adverse light conditions.

In this paper, firstly we review related works in section II. Next, we describe the architecture of our *PhisTag* coded target *in brevi* in section III (the detailed description of the architecture is available in [8]). Section IV lays down the problem statement and Section V describes our newly-proposed algorithm for robust identification of coded targets (which is based on recognition of circle retroreflective targets). The experimental results are given in Section VI.

II. Related works

Choosing an optimal architecture and functionality of coded targets is one of the key problems in computer vision and photogrammetry. In most cases, the coded targets are used for identifying the images and determining the coordinate system (used to compute cameras' relative pose). Coded targets can be also used for point matching – either directly, as in case of circular code add-ons for circle targets, or indirectly, by means of the epipolar geometry [3, 9].

There are several primary types of coded targets. Generally, the coded target (CT) includes one or more *base points* (reference points), as well as an identification part. In certain cases, some systems use only artificial points, without any identification tags; however, such systems are limited in use.

Fig. 1 shows several common types of coded targets. Circular targets are quite popular in computer vision. They consist of the primary reference point and one or more circular-shaped identification tags. The main disadvantage of this type of targets is that it is impossible to compute camera position from a single coded target of this type. Thus, the camera must see at least 5-7 coded targets in the image to determine its relative position successfully [9, 10].

Coded targets with several reference points do not have these limitations. Among popular targets of this kind, we can name CyberCode [11], ReacTIVision [12], VisualCode [13], ArUco [1] and many others. As a general rule, these fiducial

The work was supported by Act 211 Government of the Russian Federation, contract № 02.A03.21.0011

markers are square-shaped, while the corners of the square are considered as base (reference) points. However, other types of markers are not necessarily square-shaped while keeping several reference point in their architecture.

The presence of several reference points with known relative positions makes it possible to compute camera position in the coordinate system of the coded target.

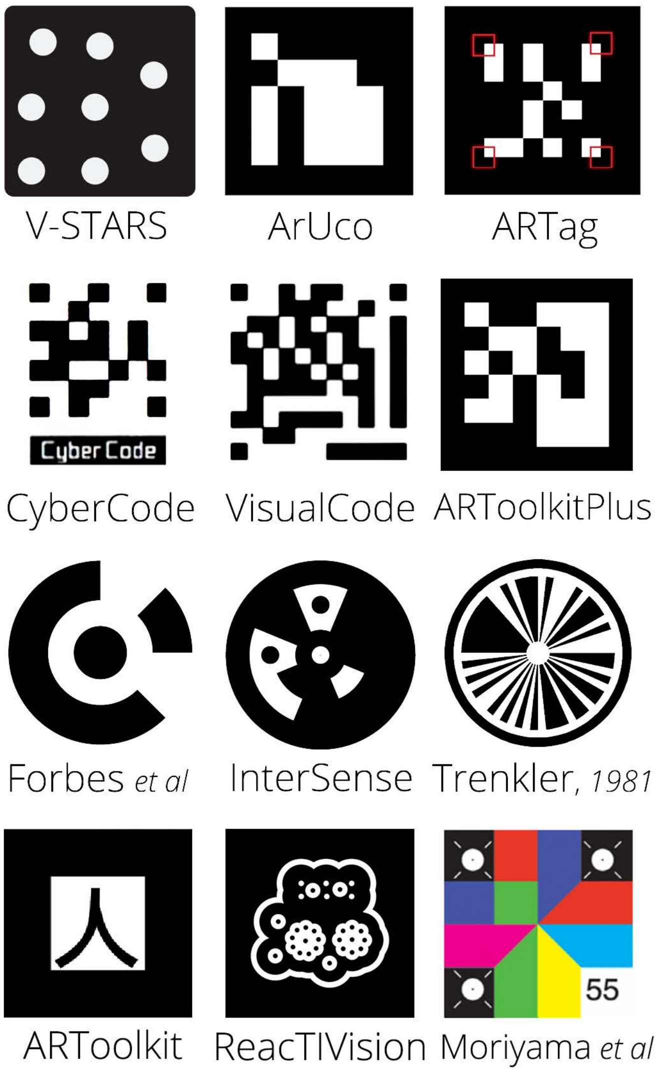

Fig. 1. Some types of coded targets in photogrammetry and computer vision

Another important problem in photogrammetry that is related to the coded targets is the robustness of the CTs' architecture to perspective transformation and various lighting conditions. To solve these issues some coded targets use binary code with error correction. For instance, Matrix [14] is based on redundant binary code. ARTag [15] relies on edge detection, as well as on error correction codes to improve detection tolerance to different lighting conditions. ARToolkitPlus [16] adjusts the system to the lighting conditions automatically (along with error correction codes) to gain even more tolerance to the illumination of the scene.

An alternative approach to recognition of fiducial markers in challenging conditions relies on machine learning. Paper [17] describes the use of support vector machines (SVM), while [18] analyzes the efficiency of neural networks in the tasks of detection of the coded targets.

The first prototype of our photogrammetric system used ArUco coded targets [1]. ArUco markers are invariant to projective transformations. They also feature error correction codes, as well as an algorithm for generation dictionaries of coded targets of the desired size. One target that is visible in the image makes it possible to compute camera's relative pose. At its heart, ArUco is based on detection of a square matrix that comprises the binary code. This code represents the ID of the coded target, in accordance with the dictionary. The system uses local adaptive thresholding, edge detection, and filtering. Then it locates and rectifies quadrangles in the images. The system again performs edge detection, as well as brightness analysis of the rectified rectangle, in an attempt to recognize identification part of the coded target, which is represented by the binary matrix [1].

Our further research in this field and development of the photogrammetric system resulted in discovery of several weak points in the ArUco system. In particular, ArUco markers cannot be reliably recognized in very low-light conditions. This leads to failures in the photogrammetric streamline.

The accuracy of finding the relative position of the camera in the coordinate system of the fiducial marker depends on the accuracy of establishing this very system. The coordinate system is anchored to the base points (reference points) of the coded target. Original ArUco CTs use the corners of the black square as the reference points; starting from ArUco v. 1.3, it supports tangency points of the corners of the main square and four auxiliary squares [19]. This allows using subpixel refinement in order to get more precise coordinates of the reference points.

Nevertheless, as we found out, even subpixel refinement does not provide the accuracy required for high-precision photogrammetric measurements. In order to resolve the aforenamed issues we modified the original ArUco architecture. Our modifications have finally evolved into our own architecture of the coded target, named *PhisTag*. Its architecture is described in the next section.

It is also worth noting that we regard the V-STARS photogrammetric system [20] of Geodetic Systems, Inc. as a point of reference in functionality for our own system under development. The V-STARS coded targets are also given in Fig. 1. However, they provide only one reference point and are primarily used for point matching. In general, the V-STARS remains a closed-source system. The details that are related to the key principles of its internal architecture are not publicly available and are of commercial interest. This is a common practice for other commercial photogrammetric systems [21–24]. Due to this, there is a limited number of available papers related to the field of high-precision photogrammetry.

III. Architecture of our PhisTag Coded Target

The primary objective of the coded target is to determine the position of the camera in the coordinate system of the coded target. Once the positions and orientations for all cameras have been computed, the system starts point matching procedure that is based on epipolar geometry [3].

Another key requirement from the coded targets is the possibility of its robust detection from any point of view (i.e. invariance to projective transformations) and in any lighting conditions. The first prototype of our system used ArUco coded targets. However, as experiments show, ArUco targets do not provide camera position data that are accurate enough for high-precision photogrammetric measurements. We introduced several modifications of the ArUco system in order to improve the accuracy of determination of the camera's relative pose and orientation. These modifications finally evolved into our own architecture of the coded target, which we have named *PhisTag*. Its appearance is given in Fig. 2.

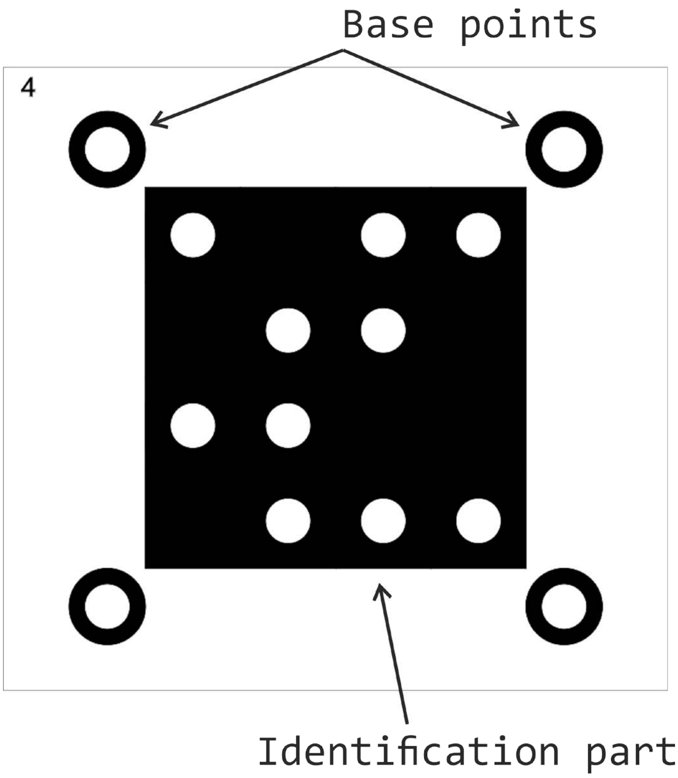

Fig. 2. Our new-proposed PhisTag coded target.

PhisTag uses circle retroreflective targets as base points (reference points). As experiments show [25], the coordinates of the center of the circle targets can be determined with more precision rather than coordinates of the corners of the square (or corner tangency points for ArUco v. ≥ 1.3 [19]). In order to achieve the required precision we use four so-called "base points" as reference points to compute the coordinate system of the coded targets [8, 11]. These points are located near the corners of the main square and are represented by the circle targets.

We have changed the coding elements of the identification part of the CT to the circle retroreflective targets as well for the purpose of universalization. These targets are identical to base points of the CT, as well as to the uncoded targets that represent the control points of the object that is being measured. The coding algorithm remains identical to the one used in ArUco.

Special emphasis should be made on the fact that we use only retroreflective coating on our circle targets. This increases target visibility in the image (due to the use of flash light), and makes it possible to recognize the coded targets from longer distances (more robustly). Once recognized, the target may be used for computation of the camera relative pose.

Aside from determining the camera position, it is possible to use PhisTag targets for several other purposes. For instance, the PhisTag CTs can be used as a scale reference (to take the function of the scale bar) in a range of photogrammetric tasks that do not require micron-level accuracy. Each coded target contains four preliminary known distances between the base points, which can be additionally calibrated to gain even higher measurement accuracy [8]. Nevertheless, using a distance, that is shorter than an average distance in the scene, as a scale reference, leads to proportional increase in measurement error. Several fields of application, such as crime analysis, road accident investigation, construction etc. do not require submillimeter precision. Such requirements make it possible to rely solely on the coded targets in these cases, thus eliminating the need in a high-cost precision scale bar. This can reduce the price of the system for several thousand Euros.

As well as, our system uses PhisTag targets to stitch the images of the lengthy objects into the single scene. This feature requires 2 CTs to be visible in at least one image, thus computing the cameras' positions in one global coordinate system.

IV. Problem Statement

As it was mentioned above, we are developing an industrial close-range photogrammetric system. The photogrammetric sessions are frequently taken in relatively dark premises or factory floors. The camera remains on the long distances from the targets frequently. The photos are taken with narrow f-stop, such as $f/11$, in order to provide large depth of field. All this leads to dark images with barely distinguishable objects even with the use of flash light (except for the circle retroreflective targets, which are almost always easily distinguishable from the background).

We noticed that in a range of scenes the existing algorithms fail to recognize the coded target. Investigation showed that this happens because those algorithms search for the black quadrangles as a principal feature of the coded target. However, in low-light conditions, the black squares may not be easily distinguished due to the low contrast with the background.

Thus, a new algorithm for the coded targets identification is in demand. The new algorithm should rely solely on the configuration of the circle retroreflective targets. We expect this algorithm to be robust to adverse light conditions, thus increasing the operating specifications of our photogrammetric system.

V. THE NEW ALGORITHM FOR ROBUST IDENTIFICATION OF THE CODED TARGETS

In this section, we present our new algorithm for robust recognition of PhisTag coded targets. The new algorithm relies on circle retroreflective targets for the recognition of the CT. It uses pixel coordinates of the circle targets as input data. This data are prepared on the preceding step. Besides, the algorithm uses the raw image data as well.

In the first step, the algorithm tries to pick the quadruples of points (circle targets) that could be potential base points of the coded target. In order to do this, the algorithm searches for quadruples of points that are located in close proximity. As the scale of the scene has not been calculated yet, the algorithm uses the diameter of the circle target as a reference for estimating the distance between the targets. Next, the algorithm checks whether the quadruple forms a parallelogram and arranges the points in a clockwise order. Then the algorithm counts the number of targets inside each quadruple that resemble the configuration of the base points of the coded target. These targets possibly belong to the identification part of the expected CT. If the number of targets inside the analyzed quadruple lies within the range of $[N_{min} - \beta; N_{max}]$, then the quadruple is likely to be the real coded target. (Here N_{min} and N_{max} are minimal and maximal number of points that can form an identification part of the coded target, according to the dictionary used; β is the maximal number of obscured points that can be handled by error correction code). In this case, the algorithm moves to the next step.

In the second step, the algorithm warps all image areas that potentially contain the coded target, in order to get the square image. Then the algorithm calculates the location of the internal area that contains the identification part of the CT. Finally, it performs image thresholding in order to get the binary image data.

According to the specifications of the coded target, the circle targets may be located strictly in the nodes of the grid. The algorithm consequently analyzes the small square-shaped image areas in the vicinity of each node. The binarized area should contain the white circle on the black background. This data from the raw image is processed along with the list of the circle targets obtained from the preCT-stage of image processing sequence. At the end of this step the algorithm produces the n-by-n binary matrix that corresponds to the configuration of the circle targets in the identification part of the CT (where n is the defined by the CTs dictionary in use).

In the final stage of the coded target recognition, the algorithm computes the four consequent clockwise rotations of the binary matrix in order to consider all possible observing directions of the CT. The coding scheme used in the identification part is invariant to rotations, meaning that any of the four possible rotations of the binary matrix will never match any other matrix (i.e. coded target) in the dictionary. These matrices are matched across the current dictionary of the coded targets. Once a match is found, the CT receives its ID, and all the circle targets that compose the CT are excluded from further analysis.

As algorithm output data, the software prepares the list of recognized coded targets, the list of cameras' positions in the coordinate systems of the CTs, as well as the list of targets that should be excluded from further processing as uncoded targets located in the control points of the measured objects.

VI. EXPERIMENT

In order to analyze the efficiency of the proposed algorithm, we conducted a photogrammetric session. The scene comprised two *PhisTag*–type coded targets with the distances between base points (base distance) of 120 and 240 mm. The diameter of the circle targets in the CTs was 12 and 24 mm correspondingly. The coded targets were located on the rear wall of the corridor. The camera was moving away from the CTs in the course of action. The distance from CTs was controlled using laser ranging device.

We took two images in each photography position. The first one used the built-in flash light of the Nikon D810 camera. The second one was taken with the external Speedlight SB-910 flash light set to full impulse, manual mode. The images were analyzed by out photogrammetric system by two algorithms for the coded target recognition: the old one, which is based on ArUco code and searches for the black image, and the new one, that relies solely on the retroreflective white circles. We did not use any brightness correction algorithms for this test.

The result data are given in Table 1. The data are shown in pairs (● / ●), where the first value represents the result of the old algorithms, while the second value represents the new one. The bullet stands for successful recognition, while minus stands for unavailability of the algorithm to recognize the coded target in the image.

TABLE I. COMPARATIVE EFFICIENCY OF CODED TARGET IDENTIFICATION ALGORITHMS (ARUCO / PHISTAG)

Shooting distance, m	**Built-in flash light (Nikon D810)**			**External flash light (Speedlight SB-910, M1/1)**		
	Average image brightness	***CT 120 mm***	***CT 240 mm***	***Average image brightness***	***CT 120 mm***	***CT 240 mm***
5	26.4567	● / ●	● / ●	124.314	● / ●	● / ●
7	47.9659	– / ●	● / ●	143.050	– / ●	● / ●
9	42.6354	– / –	● / ●	120.110	– / –	● / ●
11	42.4275	– / –	– / ●	137.684	– / –	● / ●
13	46.8087	– / –	– / ●	144.084	– / –	● / ●
15	39.2702	– / –	– / –	134.387	– / –	– / ●

Shooting distance, m	Built-in flash light (Nikon D810)			External flash light (Speedlight SB-910, M1/1)		
	Average image brightness	*CT 120 mm*	*CT 240 mm*	*Average image brightness*	*CT 120 mm*	*CT 240 mm*
17	37.1062	– / –	– / –	132.524	– / –	– / ●
19	38.7696	– / –	– / –	139.231	– / –	– / ●
21	36.9687	– / –	– / –	136.833	– / –	– / ●

The average brightness of the image was computed by the means of *identify* tool from the ImageMagick [26]. The brightness values belong to the 0÷255 range, where 0 stands for an absolutely black image, while 255 stands for a pure white image.

As we see from Table 1, the new algorithm surpasses the original ArUco CT recognition algorithm. This claim is supported by author's experience with many other photogrammetric sessions. We also analyzed the processing time of the session. The new algorithm is 1.91 times faster on average.

The examples of the images taken in adverse light conditions during the experimental session are given in Fig. 3 and 4. The images were taken at the distances of 5 and 19 meters from the coded targets with the built-in flash light. We used ISO 100 sensitivity mode, aperture of $f/11$ and shutter speed of 1/100 s.

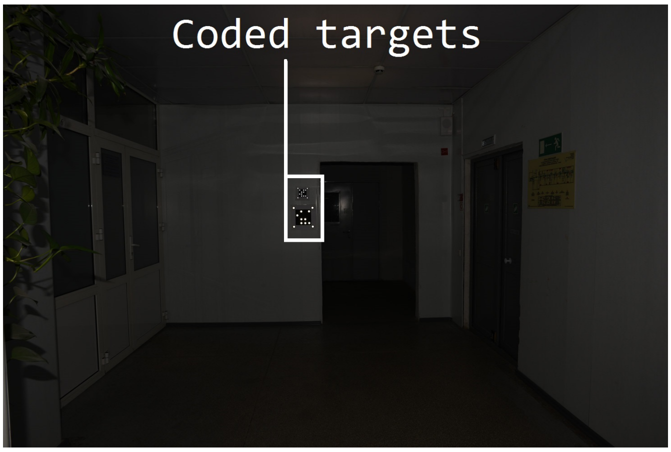

Fig. 3. An image taken from the distance of 5 m from the rear wall.

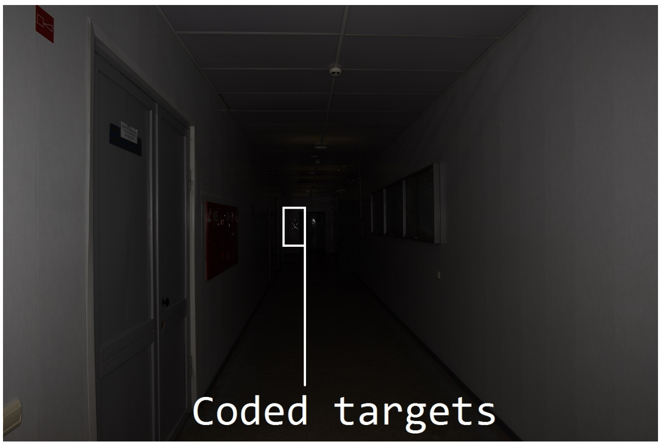

Fig. 4. An image taken from the distance of 19 m from the rear wall. The coded targets are barely visible.

VII. Conclusion and Future Work

We propose the new algorithm for robust recognition of the coded targets in adverse light conditions. The new algorithm relies solely on the use of circular retroreflective targets.

The algorithm finalizes the process of development of multi-functional coded targets for high-precision photogrammetric systems.

The main advantages of the new algorithm are robustness to adverse light conditions and high processing speed. It is also worth noting that the new algorithm is completely unified with the internal standards of the photogrammetric system.

The new algorithm (alongside with the proposed *PhisTag* architecture) makes it possible to determine camera's position in the coordinate system of the CT with high speed and accuracy. The coded targets are reliably recognized even in adverse light conditions due to the use of retroreflective circle targets.

However, the use of the new algorithm and the architecture of the coded targets is not limited to our photogrammetric system. *PhisTag* may be used for any computer vision problem that requires determining relative position of the camera, to estimate the scale of the scene. It also allows to stitch the series of lengthy objects into the single scene.

We consider the development of the publicly available software library for working with *PhisTag* as one of the directions of the future work.

References

[1] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas, and M.J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," Pattern Recognition, vol. 47, no. 6, pp. 2280-2292, 2014.

[2] B.M. Sukhovilov and E.A. Grigorova, "Development of a photogrammetry system of measuring spatial coordinates of construction elements of a low-floor tram frame," Science of SUSU, pp. 458-463, 2015.

[3] S.A. Tushev and B.M. Sukhovilov, "Parallel algorithms for effective correspondence problem solution in computer vision," Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering, vol. 6, no. 2, pp. 49-68, 2017. DOI: 10.14529/cmse170204

[4] S. Tushev and B. Sukhovilov, "Photogrammetric system accuracy estimation by simulation modelling," 2017 Int. Conf. on Industrial Engineering, Applications and Manufacturing (ICIEAM), pp. 1-6, 2017. DOI: 10.1109/ICIEAM.2017.8076464.

[5] S.A. Tushev and B.M. Sukhovilov, "Effective graph-based point matching algorithms," 2nd Int. Conf. on Industrial Engineering, Applications and Manufacturing (ICIEAM), pp. 1-5, 2017. DOI: 10.1109/ICIEAM.2016.7911628.

[6] S.A. Tushev and B.M. Suhovilov, “Some ways of improving the efficiency of automatic digital camera calibration,” Young researcher: materials of the 2nd scientific exhibition-conf. of scientific, technical and creative works of students, pp. 434-439, 2015.

[7] B.M. Sukhovilov, E.A Grigorova, E.M. Sartasov, and E.N. Gornyh, “Experimental analysis of photogrammetry system errors in measuring spatial coordinates,” Science of SUSU, pp. 458-463, 2016.

[8] S. Tushev, B. Sukhovilov and E. Sartasov, “Architecture of industrial close-range photogrammetric system with multi-functional coded targets,” 2017 2nd Int. Ural Conf. on Measurements (UralCon), = pp. 435-442, 2017. DOI: 10.1109/URALCON. 2017.8120748

[9] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision. Cambridge University Press, 2004.

[10] D. Forsyth and J. Ponce, Computer Vision: A Modern Approach. Pearson, 2011.

[11] J. Rekimoto and Y. Ayatsuka, “CyberCode: designing augmented reality environments with visual tags,” Proc. of DARE 2000 on Designing augmented reality environments, pp. 1-10, 2000.

[12] M. Kaltenbrunner and R. Bencina, “ReacTIVision: a computer-vision framework for table-based tangible interaction,” Proc. of the 1st int. conf. on Tangible and embedded interaction, pp. 69-74, 2007.

[13] M. Rohs and B. Gfeller, “Using Camera-Equipped Mobile Phones For Interacting with Real-World Objects,” Advances in Pervasive Computing, vol. 176, pp. 265-271, 2004.

[14] J. Rekimoto, “Matrix: a realtime object identification and registration method for augmented reality,” Proc. of the 3rd Asia Pacific Computer Human Interaction, pp. 63-68, 1998.

[15] M. Fiala, “Designing highly reliable fiducial markers,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 7, pp. 1317-1324, 2010.

[16] D. Wagner and D. Schmalstieg, “ARToolKitPlus for Pose Tracking on Mobile Devices,” Proc. of 12th Computer Vision Winter Workshop CVWW07, pp. 139-146, 2007.

[17] V.M. Mondéjar-Guerra, S. Garrido-Jurado, R. Muñoz-Salinas, M.J. Marín-Jiménez, and R. Medina-Carnicer, “Classification of fiducial markers in challenging conditions with SVM,” Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 10255 LNCS, pp. 344-352, 2017.

[18] V. Mondéjar-Guerra, S. Garrido-Jurado, R. Muñoz-Salinas, M.J. Marín-Jiménez, and R. Medina-Carnicer, “Robust identification of fiducial markers in challenging conditions,” Expert Systems with Applications, vol. 93, pp. 336-345, 2018.

[19] ARUCO: a minimal library for Augmented Reality applications based on OpenCV. [Online]. Available: https://www.uco.es/investiga/grupos/ava/node/26

[20] V-STARS. [Online]. Available: http://www.geodetic.com/products/systems/v-stars-n.aspx

[21] Zeiss Optotechnik. [Online]. Available: http://optotechnik.zeiss.com/en/products/3d-scanning/photogrammetry

[22] Linearis3D. [Online]. Available: http://www.linearis3d.com/

[23] TRITOP. [Online]. Available: http://www.capture3d.com/3d-metrology-solutions/photogrammetry/tritop

[24] Creaform MaxSHOT 3D. [Online]. Available: https://www.creaform3d.com/en/metrology-solutions/optical-measuring-systems-maxshot-3d

[25] B.M. Sukhovilov, E.M. Sartasov and E.A. Grigorova, “Improving the accuracy of determining the position of the code marks in the problems of constructing three-dimensional models of objects,” 2016 2nd Int. Conf. on Industrial Engineering, Applications and Manufacturing (ICIEAM), pp. 1-4, 2017. DOI: 10.1109/ICIEAM.2016.7911682.

[26] ImageMagick. [Online]. Available: https://www.imagemagick.org/script/index.php