Intrinsic Parameters Estimation for a Multiview System

Dawid Mieloch
Chair of Multimedia Telecommunications and Microelectronics
Poznan University of Technology
Poznań, Poland
dawid.mieloch@put.poznan.pl

***Abstract*—This paper presents a new method of estimation of camera parameters which uses spherical structures as calibration objects. The method allows estimation of intrinsic parameters and parameters of the optical distortion of a camera and is adapted to the multi-camera systems. The algorithms shown in the paper, as well as implementations of the optical center and lens distortion coefficients estimation use novel solutions developed for the purpose of this paper. The presented method is tested experimentally by the authors. The results of the experiment prove that the method provides the accuracy of the estimation of the camera parameters comparable to the widely used state-of-the-art methods, therefore can be successfully used with any multi-camera system.**

***Keywords*—intrinsic parameters, camera calibration, lens distortion**

I. Introduction

The calibration of a multiview system, i.e. the estimation of the parameters of used cameras, is the first step of the processing of the acquired views, e.g. in free-viewpoint television systems [14], [15], and augmented reality applications [16]. The calibration is necessary when the information about the relative distances between objects of a 3D scene has to be known [5].

The goal of this article is presentation of a new method of estimation of the camera intrinsic parameters (i.e. the focal length, the camera optical center, and lens distortion coefficients [3]) that is adapted to be used with multiview systems. The presented method is based on the calibration using a spherical object. As it is described in Section II, such object is adapted to the requirements of multi-camera systems, and, as it is presented in Section III, can be used to estimate all the required parameters of a camera.

In the work, a camera is assumed to have an Euclidean plane of an image, i.e. the skewness of the plane $o_x = 0$ and the focal length the same for the x and y planes of an image ($f_x = f_y$) [4]. It is a simplification that is assumed and used in the multiview systems with modern cameras, equipped with image sensors of the high quality.

II. Calibration Object

In order to estimate the intrinsic camera parameters, in the first step it is necessary to provide a position of some reference points from the acquired views. The camera parameters are calculated on the basis of the geometrical dependencies between these points.

There are many methods to acquire the reference points from an image acquired by a camera. Methods of the reference points searching are usually adapted to the used calibration object (e.g. a chessboard pattern [10]) and used method of the camera parameters estimation. The reference points can also be acquired manually, e.g. by a marking the lines that should be linear in an acquired view – such points can be used to eliminate the lens distortion. Nevertheless, other parameters of a camera must be calibrated as well, and in order to assure a higher convenience of the calibration, the automatic search for reference points is preferred.

Choosing a calibration object has a direct influence on the performance of the camera parameters estimation. Moreover, not all calibration objects can be successfully used for the estimation of all parameters, e.g. the lens distortion cannot be properly estimated with the very small object, such as small lamp or the laser pointer, because the influence of the distortion on the shape of such object is negligible and hard to evaluate. On the other hand, these objects are very good for estimating the position of cameras (extrinsic camera parameters [3]).

In order to choose a proper calibration object, it is necessary to describe the requirements for such an object. In this paper the method for estimating intrinsic parameters is proposed but the possibility of using the same object for the estimation of both intrinsic and extrinsic parameters increases usefulness of the presented method for multiview camera systems. Therefore, the calibration object should:

- have shape that can be easily used for the estimation of all parameters of a camera,
- be easily automatically detected in an acquired image,
- be visible simultaneously by all cameras of a system,
- be convenient in use.

One of the objects that meets all the requirements is a sphere. As it is presented in Section III, the spherical object can be used to estimate all required parameters of the camera, i.e. the focal length, the principal point and the lens distortion. Moreover, because of the specified shape of a sphere, which can be easily represented mathematically, it can be automatically detected. A sphere can also be used for more than one camera simultaneously – even if cameras of a system are placed at an

The presented work has been funded by the Polish Ministry of Science and Higher Education within the status activity task "Theory and algorithms of multidimensional signal processing" (DS) in 2018.

angle, a sphere is visible, what is not a case e.g. when a planar checkerboard is used. Using a sphere is easier than the abovementioned checkerboard, because of much smaller size of the used spherical object. There are other methods that use a sphere as the calibration object [6], [11], [12]. Unfortunately, these methods do not estimate the lens distortion ([6], [12]), or require to use a sphere with the grid [11].

III. Intrinsic Parameters Estimation

In this Section, the new method of the intrinsic parameters estimation is presented. The proposed method consists of few stages (as shown in Fig. 1) and uses both the linear and non-linear optimizations during the estimation process. As it is presented in [5], methods of camera parameters estimation that consist of few steps and simultaneously utilize linear transformations (calculated on the basis of the geometry of the calibration object) and non-linear optimizations provide the highest accuracy of the estimation. Such algorithms are much less dependent on the initial parameters used in the optimization process.

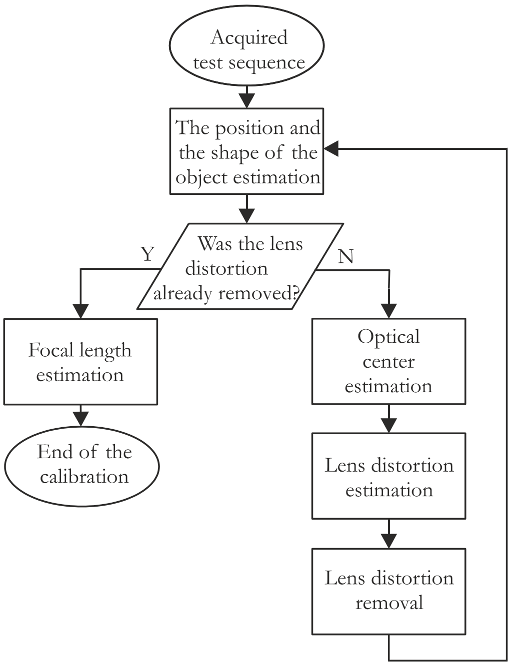

Fig. 1. The scheme of the proposed method of the intrinsic parameters estimation.

A. *The Estimation of the Position and the Shape of the Calibration Object*

The acquisition of a sequence used for the calibration of a camera should be performed immediately before or after the acquisition of the proper sequence. Often, some parameters of a camera (e.g. the focal length) are changed before the acquisition of the video, in order to adjust them to the conditions present during the recording (e.g. to the size of an acquired scene). Therefore, the estimation of the position of a calibration object has to be insensitive to the conditions of the acquired calibration sequence, i.e. to the background of the scene and to the lightning conditions too.

The estimation of the position and shape of the calibration object is based on the Hough transform for circles [2]. Firstly, the image of the calibration object is binarized in order to extract the white sphere from a scene. All objects left in the image after the binarization are parametrized – their size, contour, and geometric center are estimated. These parameters are compared to the results of the Hough transform (performed on the input image, without binarization). The transform usually finds many circles in the image. If any of the circles has similar size and center as the objects on the binarized image, then this circle is used in the calibration process.

B. *Lens Distortion Estimation*

The lens distortion is expressed as a set of n distortion coefficients $k = [k_1, k_2, \dots, k_n]$ [1]. In case of the estimation for two coefficients k_1, k_2, the linear equation that links the position of the point before and after the distortion is:

$$\begin{bmatrix} r^2 & r^4 \\ r^2 & r^4 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} x'/x - 1 \\ y'/y - 1 \end{bmatrix}, \tag{1}$$

where x, x' are real and distorted positions of the point in the horizontal plane, y, y' are real and distorted positions of the point in the vertical plane, r is the Euclidean distance from the optical center of the camera.

The use of only one position of the calibration object is sufficient for the estimation of k_1 and k_2 but the estimated model of distortions would be correct only in a small neighborhood of the object. Therefore, it is necessary to acquire more than one position of the calibration object, in order to cover as big as possible area of the camera. It results in a redundant linear equation for n positions of the calibration object:

$$\begin{bmatrix} r_1^2 & r_1^4 \\ r_1^2 & r_1^4 \\ r_2^2 & r_2^4 \\ r_2^2 & r_2^4 \\ \vdots & \\ r_n^2 & r_n^4 \\ r_n^2 & r_n^4 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} x_1'/x_1 - 1 \\ y_1'/y_1 - 1 \\ x_2'/x_2 - 1 \\ y_2'/y_2 - 1 \\ \vdots \\ x_n'/x_n - 1 \\ y_n'/y_n - 1 \end{bmatrix}. \tag{2}$$

As it can be seen, to estimate the distortion coefficients, besides the position of the calibration object on the distorted view, the theoretical position of the object on the undistorted view has to be also known.

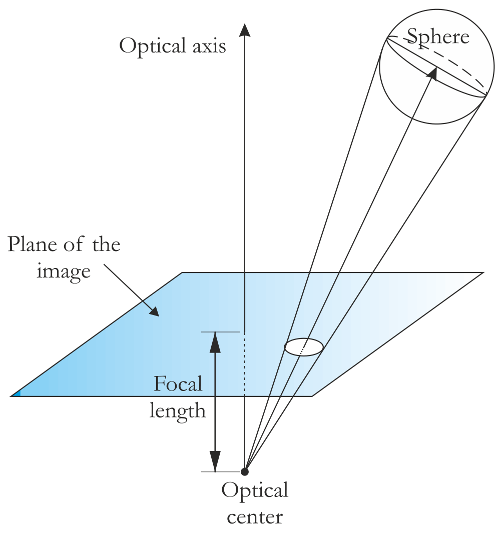

Fig. 2. A projection of a sphere on the plane of the image.

When a sphere is visible by the pinhole camera model, a cone from the optical center of the camera to the sphere can be drawn (Fig. 2). Let us assume that the camera has no lens distortion. If the middle of the sphere would be on the optical axis of the camera, the projection of the sphere on the plane of the image would be round. If the sphere is moved from the optical axis, the projection becomes an ellipse.

Fig. 3 presents the computer-generated set of 7 positions of the spherical calibration object. When the image has no lens distortion, there is an almost linear dependency between the difference of the smallest and the biggest radius of an ellipse and the distance from the optical center of the image (Fig. 4). Additionally, these radii are always perpendicular.

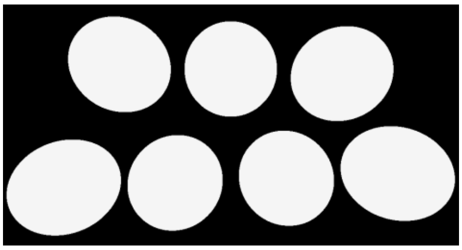

Fig. 3. A generated set of the 7 positions of the calibration object (without the lens distortion).

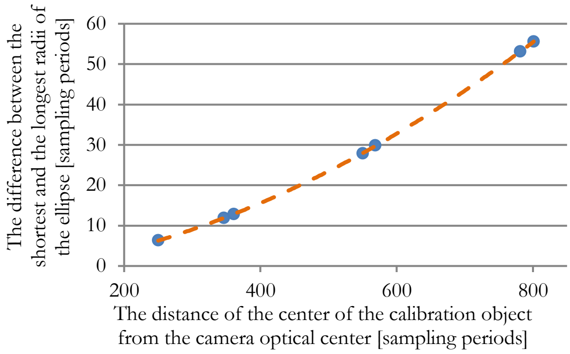

Fig. 4. The difference of the smallest and the biggest radii of an ellipse and the distance from the optical center of the image (without the lens distortion).

Fig. 5 shows a shape of the calibration object when the lens distortion is present (the distortion coefficients in the example were k_1=-0.2, k_2=0.1, a distortion similar to the Canon Xh G1 camera, used in the FTV system described in [13]).

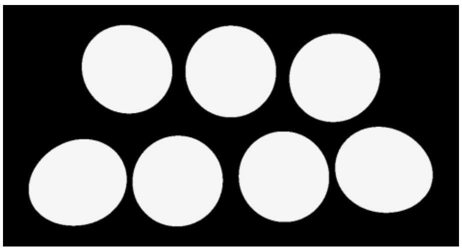

Fig. 5. A generated set of the 7 positions of the calibration object (with the lens distortion).

In Fig. 6 the difference of the smallest and the biggest radii of an ellipse and the distance from the optical center of the image are presented, but now for the distorted image. The dependency between these measures is now not linear and the angle between abovementioned radii is also no longer right.

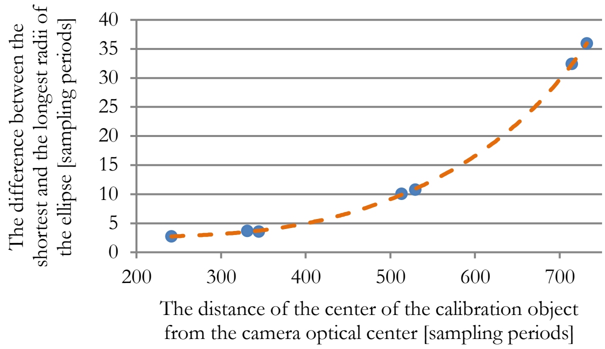

Fig. 6. The difference of the smallest and the biggest radii of an ellipse and the distance from the optical center of the image (with the lens distortion).

The resultant distortion coefficient K, used in the following calculations, can be defined as:

$$K = (1 + k_1 \cdot r^2 + k_2 \cdot r^4). \tag{3}$$

In order to estimate the real position of the calibration object (u, v) using the position on the distorted view (u', v'), we achieve:

$$\begin{aligned} u &= (u' - c_x)/K + c_x \,, \\ v &= \left(v' - c_y\right)/K + c_y \,. \end{aligned} \tag{4}$$

Therefore, it is necessary to calculate K to estimate the lens distortion. On the basis of the abovementioned dependencies, the authors propose such definition of K:

$$K = (100 - |(r_{max} - r_{min}) \cdot cos(\varphi)|)/100, \tag{5}$$

where r_{max} is the length of the longest radius of the ellipse, r_{min} is the length of the shortest radius of the ellipse, φ is the angle between these radii. When the angle is equal to 90 degrees, $cos(\varphi) = 0$, so the K is equal to 1, what indicates that there is no lens distortion in the image. Fig. 7 presents the comparison of the value of K that was calculated from the exemplary positions of the calibration object (Fig. 3 and 5) with the K estimated using proposed definition.

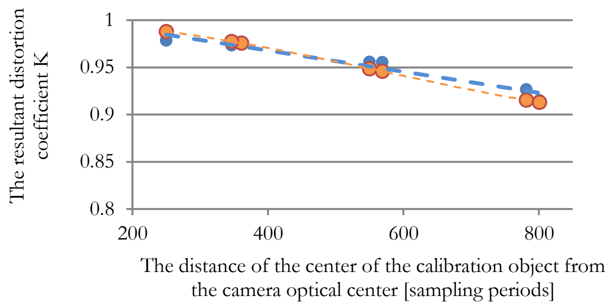

Fig. 7. The value of K calculated for an exemplary positions of the calibration object (orange line) and estimated values (blue line).

Fig. 8 shows the longest and the shortest radii found in exemplary positions of the distorted positions of the calibration object. These radii, and angle between them, are used to estimate K using (5), which is then used to estimate the real positions of the calibration object with (4). The real and distorted positions are then used in (2), which is solved for k_1 and k_2 using the least squares method.

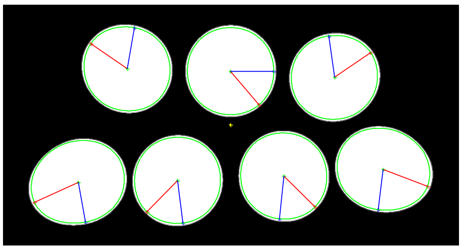

Fig. 8. The shortest (blue lines) and the longest (red lines) radii of the calibration object.

C. The Estimation of the Camera Optical Center

The camera optical center is estimated before the distortion coefficients but is also based on the properties of the projection of the sphere on the plane of the camera (shown earlier in Fig. 2). The longest radius of the projected ellipse lies on the line which passes through the optical center of the camera. Fig. 9 shows the longest radii of the 7 positions of the calibration object (red lines) and lines that contain these radii (orange lines). As it can be seen, the extended lines cross in the close neighborhood of the optical center of the camera (yellow dot).

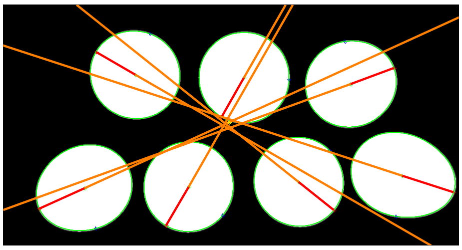

Fig. 9. The estimation of the optical center of the camera.

In order to estimate the optical center of the camera, an minimization of the sum of the distances from the lines that contain the longest radii of the acquired positions of the calibration object to the estimated position of the optical center is performed:

$$f_{min}(x, y) = \sum_{i=1}^{N} \frac{|A_i \cdot x + B_i \cdot y + C_i|}{\sqrt{A_i^2 + B_i^2}}, \tag{6}$$

where: A_i, B_i, C_i – coefficients of the line i that contain the longest radius of the i-th position of the calibration object, x, y – the estimated position of the optical center of the camera, N – the number of the acquired positions of the calibration object.

D. The Estimation of the Focal Length

The last step of the intrinsic parameters estimation is the calculation of the focal length of the camera. The solution is based on the method described in [6]. As it can be observed in Fig. 2, a change in the focal length has direct impact on the shape of the projection of the sphere on the plane of the camera. A cone that goes from the sphere to the optical center of the camera can be described with 6 coefficients (A, B, C, D, E, F). All points of the cone Q satisfy the equation:

$$Q(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0. \tag{7}$$

The coefficients of the cone are estimated for N positions of the calibration object using the minimization with the BFGS method [7]:

$$f_{min} = \sum_{i=1}^{N} Q^2(x_i, y_i). \tag{8}$$

As [6] shows, when the lens distortion is removed from the image, the focal length f_i for each position of the calibration object can be calculated using the equations:

$$f_i^2 = f_i^2 \cdot k \frac{-D \cdot E + A \cdot F - D \cdot c_y - E \cdot c_x}{B} - c_x^2 - c_y^2 - \frac{c_x \cdot c_y}{f_i^2 \cdot k \cdot B},$$
$$f_i^2 \cdot k = \frac{A + C - \sqrt{(A-C)^2 + 4 \cdot B^2}}{2 \cdot (A \cdot C - B^2)}. \tag{9}$$

On the end, resulting N focal lengths f_i are averaged in order to achieve the final focal length f of the camera.

IV. Experimental Results

Each of the parts of the proposed intrinsic parameters estimation method (i.e. the estimation of the optical center, the lens distortion, and the focal length) is tested in the performed experiments. The experiments are based on the generated calibration sequences

A. The Accuracy of the Optical Center Estimation

The calibration sequence for testing of the accuracy of the optical center estimation was generated for the fixed focal length equal to 1000 sampling periods, the lens distortion coefficients were $k_1 = -0.2$ and $k_1 = 0.1$. The resolution of the image that contained 3 positions of the calibration object was 2048x1080. The estimation was performed for a set of 20 sequences that varied in the position of the optical center of the camera. The results of the experiment are presented in Table I.

TABLE I. The Accuracy of the Optical Center Estimation

Test sequence number	*Ground truth values of the optical center [sampling periods]*		*Estimated values of the optical center [sampling periods]*		*Error of the estimation [sampling periods]*	
	c_x	c_y	$\hat{c_x}$	$\hat{c_y}$	err_{c_x}	err_{c_y}
1.	791	425	792.2	417.6	1.2	7.4
2.	800	374	792.5	362.8	7.5	11.2
3.	830	714	833.9	706.7	3.9	7.3
4.	909	449	904.0	447.6	5.0	1.4
5.	917	407	920.7	405.6	3.7	1.4
6.	940	399	939.7	399.4	0.3	0.4
7.	1013	646	1020.3	644.0	7.3	2.0
8.	1029	645	1021.0	643.5	8.0	1.5
9.	1034	408	1038.5	406.5	4.5	1.5
10.	1053	585	1047.5	576.0	5.5	9.0
11.	1058	473	1050.6	469.5	7.4	3.5
12.	1068	485	1060.5	481.4	7.5	3.6
13.	1116	392	1121.2	397.5	5.2	5.5
14.	1117	582	1112.4	581.1	4.6	0.9
15.	1198	528	1192.5	533.4	5.5	5.4
16.	1220	644	1213.3	647.7	6.7	3.7
17.	1225	543	1230.8	539.7	5.8	3.3
18.	1250	465	1257.2	461.3	7.2	3.7
19.	1255	620	1243.6	611.0	11.4	9.0
20.	1317	630	1312.1	626.0	5.0	4.0
				Mean errors:	***5.7***	***4.3***

B. The Accuracy of the Lens Distortion Estimation

In the next experiment the focal length was again equal to 1000 sampling periods. The resolution of the image that contained 3 positions of the calibration object was 2048x1080, the optical center was fixed in the center of the image (1024,540). The estimation was performed for a set of 20 sequences that varied in the lens distortion of the camera. The results of the experiment are presented in Table II.

TABLE II. THE ACCURACY OF THE LENS DISTORTION ESTIMATION

Test sequence number	*Ground truth values of the distortion coefficients*		*Estimated values of the distortion coefficients*		*Error of the estimation*	
	k_1	k_2	$\widehat{k_1}$	$\widehat{k_2}$	err_{k_1}	err_{k_2}
1.	-0.06	0.02	-0.06	0.06	0.00	0.04
2.	-0.10	0.05	-0.08	0.08	0.02	0.03
3.	-0.10	0.10	-0.11	0.15	0.01	0.05
4.	-0.11	0.15	-0.15	0.14	0.04	0.01
5.	-0.11	0.06	-0.10	0.10	0.01	0.04
6.	-0.12	0.12	-0.12	0.11	0.00	0.01
7.	-0.12	0.06	-0.09	0.09	0.03	0.03
8.	-0.14	0.13	-0.12	0.13	0.02	0.00
9.	-0.15	0.10	-0.17	0.03	0.02	0.07
10.	-0.16	0.12	-0.15	0.18	0.01	0.06
11.	-0.18	0.14	-0.15	0.17	0.03	0.03
12.	-0.20	0.10	-0.19	0.08	0.01	0.02
13.	-0.20	0.14	-0.15	0.14	0.05	0.00
14.	-0.21	0.13	-0.21	0.20	0.00	0.07
15.	-0.23	0.15	-0.19	0.15	0.04	0.01
16.	-0.24	0.15	-0.17	0.12	0.07	0.03
17.	-0.24	0.19	-0.24	0.22	0.00	0.03
18.	-0.26	0.15	-0.16	0.12	0.10	0.03
19.	-0.26	0.18	-0.22	0.20	0.04	0.02
20.	-0.27	0.11	-0.32	0.10	0.05	0.01
				Mean errors:	0.03	0.03

C. *The Accuracy of the Focal Length Estimation*

In the last experiment the accuracy of the estimation of the focal length was tested. The resolution of the image that contained 3 positions of the calibration object was 2048x1080, the optical center was fixed in the center of the image (1024,540). Because the estimation of the focal length is performed after the lens distortion is removed from the image, the lens distortion coefficients were set to 0. The estimation was performed for a set of 20 sequences that varied in the focal length of the camera. The results of the experiment are presented in Table III.

TABLE III. THE ACCURACY OF THE OPTICAL CENTER ESTIMATION

Test sequence number	*Ground truth values of the focal length f [sampling periods]*	*Estimated values of the focal length $\hat{f}$ [sampling periods]*	*Error of the estimation [sampling periods]*
1.	517	561.4	44.4
2.	538	506.3	31.7
3.	798	695.6	102.4
4.	802	708.4	93.6
5.	815	820.7	5.7
6.	1033	1115.5	82.5
7.	1043	967.1	75.9
8.	1096	1168.2	72.2
9.	1121	1078.4	42.6
10.	1164	1017.7	146.3
11.	1258	1274.3	16.3
12.	1460	1390.3	69.7
13.	1593	1445.4	147.6
14.	1605	1472.8	132.2
15.	1692	1567.7	124.4
16.	1727	1615.1	111.9
17.	1754	1767.0	13.0
18.	1795	1568.2	226.8
19.	1803	1852.1	49.1
20.	1880	1819.4	60.6
		Mean errors:	82.4

D. *Comparison with Other Methods*

In [17], the author proposed a method of testing of camera parameters estimation techniques, together with the accuracies of 2 methods: the Self-calibration method [9] and the calibration with the chessboard pattern [10]. The comparison of these methods with the proposed method of the camera calibration is presented in Table IV.

The proposed method achieved comparable accuracy of the estimation of the camera parameters as [10], but is better adjusted to use with the multiview systems. [10] requires to use the large chessboard pattern, which is problematic to use with cameras with many cameras (see Section II). The camera self-calibration [9] uses small laser pointer and simultaneously estimates the extrinsic parameters of cameras (their position) but unfortunately the accuracy of the estimation of intrinsic parameters is the lowest from all methods.

TABLE IV. THE ACCURACY OF THE CAMERA CALIBRATION METHODS

The estimated camera parameter		*Error of the estimation*		
		Proposed method	*Chessboard [10]*	*Self calibration [9]*
Focal length [sampling periods]	f	82.44	70.16	128.46
Optical center [sampling periods]	c_x	5.65	3.82	36.19
	c_y	4.28	6.32	58.72
Lens distortion coefficients	k_1	0.03	0.01	0.30
	k_2	0.03	0.02	-

V. CONCLUSIONS

In this paper, the new method of the intrinsic parameters estimation is presented. The proposed method is adapted to be used with the multiview systems because of using a spherical calibration object. Such object was found to meet the requirements of the multi-camera systems – easiness of use, simultaneous visibility by many cameras of the system, and a shape that makes it easy to find the position of the calibration object.

The method uses novel methods of the optical center and the lens distortion estimation. Performed experiments shown, that the method provides the accuracy of the estimated camera parameters comparable to other methods, but because of being adapted to use with multiview systems, is a good alternative to the state-of-the-art methods.

REFERENCES

[1] D.C. Brown, "Close-range camera calibration," Photogrammetric Engineering, vol. 8, vol. 37, 1971.

[2] R. Duda, P. Hart, "Use of the Hough transformation to detect lines and curves in picture," Magazine Communications of the ACM, vol. 15, pp.11-15, 1972.

[3] R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, 2nd ed., Cambridge University Press, 2003.

[4] A. Heyden, K. Åström, "Minimal conditions on intrinsic parameters for euclidean reconstruction," in Third Asian Conference on Computer Vision Hong Kong, China, Proceedings, pp. 169-176, 1998.

[5] J. Salvi, X. Armangue, and J. Batlle, "A comparative review of camera calibrating methods," Pattern Recognition, vol. 35, pp. 1617–1635, 2002.

[6] H. Teramoto, G. Xu, "Camera calibration by a single image of balls: From conics to the absolute conic," in Proc. of 5th Asian Conference on Computer Vision, s. 499-605, 2002.

[7] D. Yu-Hong, “A perfect example for the BFGS method,” Mathematical Programming, vol. 138, pp. 501-530.

[8] Z. Zhang, “Parameter-estimation techniques: A tutorial with application to conic fitting,” Image and Vision Computing, vol. 15, pp. 59-76, 1997.

[9] T. Svoboda, Daniel Martinec, T. Pajdla, “A convenient multi-camera self-calibration for virtual environments,” Presence, vol. 14, 2005.

[10] Z. Zhang, “A flexible new technique for camera calibration,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330-1334, Nov 2000.

[11] Z. Liu, Q. Wu, S. Wu, and X. Pan, “Flexible and accurate camera calibration using grid spherical images,” Opt. Express, vol. 25, pp. 15269-15285, 2017.

[12] H. Zhang, K. Y. K. Wong and G. Zhang, “Camera Calibration from Images of Spheres,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 3, pp. 499-502, March 2007.

[13] M. Domański, A. Dziembowski, D. Mieloch, A. Łuczak, O. Stankiewicz and K. Wegner, “A practical approach to acquisition and processing of free viewpoint video,” in 2015 Picture Coding Symposium (PCS), Cairns, QLD, 2015, pp. 10-14.

[14] O. Stankiewicz, M. Domanski, A. Dziembowski, A. Grzelka, D. Mieloch and J. Samelak, “A Free-viewpoint Television system for horizontal virtual navigation,” IEEE Transactions on Multimedia, early access, DOI 10.1109/TMM.2018.2790162, 2018.

[15] M. Tanimoto, M. Panahpour Tehrani, T. Fujii and T. Yendo, “FTV for 3-D spatial communication,” Proceedings of the IEEE, vol. 100, no. 4, pp. 905–917, April 2012.

[16] H. Ling, “Augmented Reality in reality,” IEEE MultiMedia, vol. 24, no. 3, pp. 10-15, 2017.

[17] D. Mieloch, “Analiza dokładności wyznaczania parametrów kamer,” [“Analysis of the camera parameters estimation accuracy”], B.Eng. thesis, Chair of Multimedia Telecommunications and Microelectronics, Poznan University of Technology, Poland, 2013.