Accuracy Improvements of Consumer-Grade 3D Sensors for Robotic Applications

Branko Karan
Institute of Technical Sciences
Serbian Academy of Sciences and Arts
Belgrade, Serbia
branko.karan@itn.sanu.ac.rs

Abstract—**Recent advances in development of low-cost 3D sensors, such as Microsoft Kinect, bring attractive opportunities to robot system integrators. The accuracy provided by such sensors is generally unsatisfactory for many robotic applications, but it may be improved through calibration. This paper presents a calibration case study that is based on the sensor calibration procedure involving only a use of a simple checkerboard. It is shown that the calibration enables improving sensor accuracy 3 to 5 times, depending on the anticipated use of the sensor. Additionally, results obtained using different levels of complexity of calibration models reveal that depth measurement correction is an important component of calibration as it may reduce by 50% the errors in sensor reading.**

Keywords—robot vision; Kinect; 3D sensing; depth camera; RGB-D camera; camera calibration

I. INTRODUCTION

An increased need for low cost sensors is one of determinants in contemporary robotics. Enriching the mobile and service robots with innovative sensing capabilities without jeopardizing the overall system cost is an important precondition for success on today's market. Thus, there is an understandable interest in development of consumer-targeted sensing devices.

Microsoft Kinect, with its high-tech 3D visual and audio sensing abilities, is one notable example. Initially intended as a sophisticated human interface device for computer games, the sensor has been receiving a growing attention from robotic researchers and practitioners. Thanks to its mass production and thus low price, Kinect is perceived as an attractive option in many robotic applications.

Since the initial release of Kinect XBOX 360 in 2010 [1], several variants of the sensor have also appeared. In 2012, Microsoft presented the enhanced Kinect for Windows [2]. At the same time, ASUSTeK and PrimeSense used the same technology to develop 3D vision sensors with the similar features [3, 4] so that it is possible to speak now about a generation of consumer-grade 3D sensors.

This work was supported in part by the Serbian Ministry of Education, Science and Technology Development under projects III-44008 and TR-35003.

Kinect and similar 3D sensors are low accuracy, low precision devices. Diverse studies, including this one, show that their accuracy is on the order of 2-3%. At a distance of 4m from the sensor, this corresponds to RMS error on the order of 10cm. This level of accuracy is quite satisfactory in e.g. human interaction applications. However, it may appear unsuitable in some specific robotic uses (e.g., indoor navigation or fine manipulation).

Accuracy can be improved by software correction of sensor outputs. The correction is based on a specific calibration model whose parameters are identified during the calibration process. The calibration procedure consists in essence in collecting sensor outputs and comparing them to reference data, and it assumes using a special calibration rig, i.e. an object of precisely known dimensions, and/or use of a high-precision measurement tool.

In this paper, a case study is presented that is based on a two-step calibration procedure recently proposed by this author [5]. The procedure is simple in the sense that it relies on a commonly adopted camera calibration tool (i.e., a checkerboard) and it does not require additional specific calibration objects or external measurement devices. The case study involved a calibration and analysis of accuracy of Kinect XBOX 360 sensor, but the results and the approach are applicable to other similar sensors.

The paper is organized as follows. A short review of several representative works on accuracy and calibration of Kinect-type sensors is given in the next section. Section 3 contains a description of employed sensor calibration model and model identification procedure. Section 4 presents analysis of accuracy that can be expected in different scenarios and with different levels of complexity of calibration models. Section 5 summarizes conclusions on the attained results.

II. RELATED WORK

Accuracy of Kinect-type sensors was a subject of investigation involving use of different measuring devices. Dutta [6] measured accuracy of Kinect sensor using a high-precision seven-camera Vicon 3D motion capture system and reported mean errors of up to 10cm (standard deviation on the order of 5cm) in the range covering the distances of up to 3m from the sensor. The sensor was previously calibrated and the calibration involved identification of parameters of the sensor camera. Gonzales-Jorge et al. [7] investigated accuracy of uncalibrated Kinect

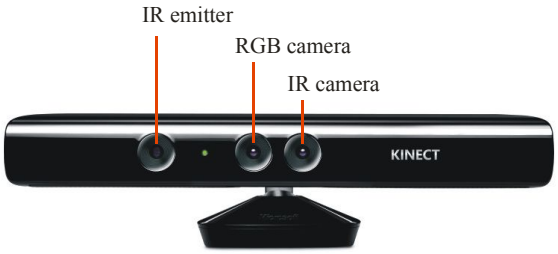


Figure 1. Kinect sensor

XBOX 360 and Asus Xtion using a specially designed measurement fixture. Their measurements confirm that both accuracy and precision deteriorate with the distance. These authors obtained similar accuracy/precision for both examined sensors and the values of RMS error were on the order of 10mm (standard deviation on the order of 8mm) at the distance of 2m.

Several works addressed improvement of sensor accuracy. In earlier works, e.g. Burrus [8] and Zhang and Zhang [9], the main concern was on procedures and methods for identification of intrinsic parameters of sensor cameras. Later, the focus has moved to calibration of sensor depth measurement model, with important works of Khoshelham and Elberink [10], Smisek et al. [11], and Herrera C. et al. [12]. These works addressed transformation of disparity maps provided by the sensor into depth maps. However, with the actual OpenNI [13] and Microsoft Kinect SDK [14] application programming interfaces, disparity data are already converted into depth using a nominal factory model. To account for this change, a reformulation of depth calibration model was proposed in [5], advocating to utilize a linear relationship between actual and sensor-provided inverse depths.

Calibration of Kinect-type 3D sensors is naturally split into two parts: identification of parameters of sensor cameras and identification of parameters of depth measurement model. It is naturally and simplest to identify camera parameters from raw camera data, but there were also attempts to calibrate sensor camera directly from depth maps [9] and perform a joint depth/RGB camera calibration [12]. Although advantageous in principle, this approach suffers from the problem of calibrating the camera using a low precision depth map: the low precision problem naturally results in a need for extremely large number of measurements. Additionally, the joint calibration, although having a potential of improving the optimal solution, may display an undesired interaction: for example, the data on joint calibration reported in [12] show that a refinement of depth model can paradoxically lead to enlargement of reprojection errors of RGB camera.

An important issue in depth model calibration is the measurement of depth as it normally requires either a special 3D calibration rig or an external measuring tool. In [10], depth was measured using a simple measuring tape. In [12], correspondence between depth map and RGB camera image was established using external corners of calibration table. A similar approach was proposed by Draelos et al. [15]. Geiger et al. [16] proposed more complex calibration object, consisting of multiple checkerboards. Shibo and Qing [17] designed a specific calibration board with regularly-spaced drilled holes allowing their easy identification in both RGB images and depth maps.

In this work, depth calibration follows a procedure proposed in [5], where the sensor RGB camera, once calibrated, is used as a depth measuring device for subsequent calibration of the depth model. Details of both the model and the calibration procedure are given in the next section.

III. Sensor Calibration

Kinect-type 3D sensors considered in this work operate as structured light sensors. A sensor (Fig. 1) incorporates a laser IR diode for emitting a dotted light pattern and an IR camera for capturing reflected patterns. Depth is calculated by sensor software on the basis of disparity of reflected patterns with the respect to the reference patterns obtained for a plane placed at a known distance from the sensor. A supplementary RGB camera is added to provide additional information on color and texture of the surface. Thus, sensor output consists of three data flows: images from RGB camera, raw images from IR camera, and depth maps calculated by sensor firmware. Sensor calibration can be viewed as a refinement of correspondences between 3D object coordinates and coordinates in RGB, IR, and depth images.

The proposed calibration procedure consists of two steps: the first step comprises calibration of sensor's RGB/IR cameras, whereas the calibration of depth model is performed within the second step.

A. Camera Calibration

Camera calibration assumes identification of parameters of functions modeling transformation of 3D coordinates of external objects into coordinates in image plane. For a point with 3D homogeneous coordinates $\mathbf{X}_e = \left[X_e, Y_e, Z_e, 1\right]^T$ in some external world coordinate frame, transformation of coordinates involves (a) transformation $\mathbf{X} = \mathbf{T}_e \mathbf{X}_e$ into coordinates $\mathbf{X} = \left[X, Y, Z, 1\right]^T$ in camera frame, (b) projection $Z \cdot \mathbf{x} = \left[\mathbf{I} \quad \mathbf{0}\right]\mathbf{X}$ into the point $\mathbf{x} = \left[u, v, 1\right]^T$ in normalized image plane, (c) distortion $\mathbf{x} = f^{(d)}(\mathbf{x})$, yielding distorted normalized coordinates $\mathbf{x}^{(d)} = \left[u^{(d)}, v^{(d)}, 1\right]^T$, and finally (d) transformation into pixel coordinates, using transformation of the form $\mathbf{x}^{(p)} = \mathbf{K} \cdot \mathbf{x}^{(d)}$, where $\mathbf{x}^{(p)} = \left[c, r\right]^T$ and c, r are column/row indices of image pixels. Thus, calibration consists in identification of intrinsic parameters — elements of camera matrix $\mathbf{K}$ and parameters of distortion function $f^{(d)}(\cdot)$, and extrinsic parameters — elements of transformation matrix $\mathbf{T}_e$.

For a stereo pair of Kinect cameras, calibration encompasses identification of camera matrices $\mathbf{K}_{RGB}$, $\mathbf{K}_{IR}$, distortion functions $f_{RGB}^{(d)}(\cdot)$, $f_{IR}^{(d)}(\cdot)$, and homogeneous transformation matrix ${}^{IR}\mathbf{T}_{RGB}$ for transformation of homogeneous 3D coordinates from the coordinate frame of RGB camera to the coordinate frame of IR camera.

In this work, the following functional forms of distortion function and camera matrix were considered:

$$f^{(d)}(\mathbf{x}) = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \cdot \mathbf{x},\ r^2 = \sqrt{u^2 + v^2} \quad (1)$$

$$\mathbf{K} = \begin{bmatrix} f_x & \alpha \cdot f_x & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Early tests with Kinect sensor revealed that tangential distortion of its RGB camera was below the achievable level of precision. Besides, distortion found in IR camera was extremely low and neglecting the tangential distortion did not change much the total reprojection error. Thus, the adopted functional form (1) involved only a radial distortion specified by parameters k_1, k_2, k_3 .

Identification of RGB/IR camera parameters was conducted in Matlab environment, using Bouguet's camera calibration toolbox [18]. A 9×8 checkerboard with 30mm square fields was employed as a calibration rig. Then, a set of ten pairs of close-up RGB/IR images of the checkerboard placed in different orientations were collected and submitted to calibration. To achieve appropriate light conditions for calibrating the IR camera, the IR emitter was disabled during imaging (an ordinary stick tape was used to cover the projector; the newer Kinect for Windows model allows programmable control over the IR emitter). While acquiring images, both cameras were set to their maximum resolutions, which were 1280×960 for the RGB camera and 640×480 for the IR camera.

The results are summarized in Table I. The most important difference between nominal and identified parameters is in focal length which differs about 1.8% for the RGB camera and 2.4% for the IR camera.

Once the parameters of transformations are known, depth maps provided by the sensor are easily converted into 3D maps. Assuming that a point in the depth map is available in of the form $\left(\mathbf{x}_{IR}^{(p)}, z_{IR}\right)$, where $\mathbf{x}_{IR}^{(p)}$ are its (column, row) coordinates in IR image coordinate frame and z_{IR} is the distance from the sensor, the transformation is given by:

TABLE I.
CAMERA PARAMETERS

	RGB camera intrinsic parameters		IR camera intrinsic parameters	
	Nominal	Identified	Nominal	Identified
n_x	1280		640	
n_y	960		480	
f_x	1062.3	1043.2	571.26	585.5
f_y	1062.3	1044.0	571.26	586.5
c_x	639.5	650.8	319.5	327.9
c_y	479.5	510.4	239.5	246.2
α		0.00122		0.00130
k_1		0.224		-0.125
k_2		-0.715		0.438
k_3		0.751		-0.556
Rotation vector RGB→IR:	[0.00260, -0.00600, 0.00175]			
Translation vector RGB→IR:	[-25.07450, 0.28102, 0.79722]			

$$\mathbf{x}_{IR} = \left[f_{IR}^{(d)}\right]^{-1} \left(\mathbf{K}_{IR}^{-1} \cdot \mathbf{x}_{IR}^{(p)}\right) \quad (3)$$

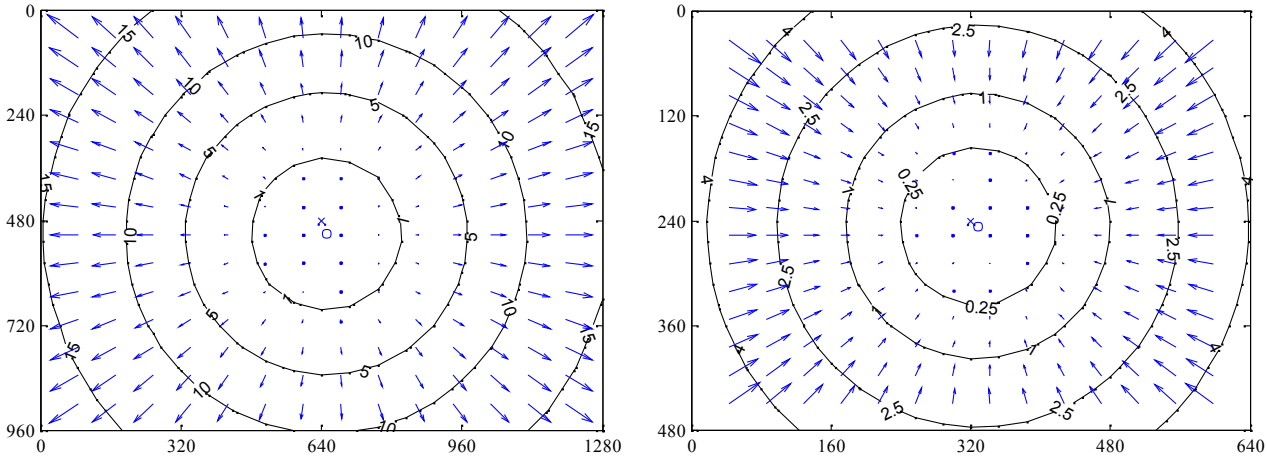
$$\mathbf{X}_{IR} = \begin{bmatrix} z_{IR}\mathbf{x}_{IR} \\ 1 \end{bmatrix} \quad (4)$$

Besides, the depth map can be extended with texture and color information from RGB image using the coordinate transformations:

$$\mathbf{X}_{RGB} = {}^{IR}\mathbf{T}_{RGB}^{-1} \cdot \mathbf{X}_{IR} \quad (5)$$

$$\mathbf{x}_{RGB}^{(p)} = \mathbf{K}_{RGB} \cdot f_{RGB}^{(d)} \left\{ \left[X_{RGB}/Z_{RGB}, Y_{RGB}/Z_{RGB}, 1 \right]^T \right\} \quad (6)$$

Evaluation of expressions (3-6) involves computation of distortions and therefore it is of interest to explore whether a simpler approximation of distortions is possible and whether the distortions could be perhaps completely neglected. Additional insight into the actual level of distortions introduced by camera optics provides Fig. 2, where the amount of distortions, expressed in pixels, is shown on contour lines and the direction of distortions is shown by blue arrows. It is seen that the direction of distortions is opposite for the RGB and IR cameras, so that they effec-



a. RGB camera b. IR camera

Figure 2. Radial distortions

tively affect the net deviation in the same direction. Distortions increase from the center to periphery of images: if the object of interest is kept within the central circles on Fig. 2, the distortions introduced by cameras could be as low as 1 pixel for the RGB camera and 0.25 pixels for the IR camera. Since the distortions are multiplied to 3D deviations with the factor of z/f, these pixel distortions correspond to 3D deviations on the order of, respectively, 3.8/1.7mm for the RGB/IR camera at the distance of 4m from the sensor. Thus, the deviations could be neglected for such central objects. On the other hand, by approaching peripheral image area, the deviations enlarge and a more complex model of deviations becomes necessary.

B. *Depth Measurement Calibration*

Sensor reading of Kinect-type sensors is based on internal computation of depth using the detected disparity between images of light beans obtained after reflection from the measurement (object) surface and the reference surface. The value of such inferred depth depends on sensor geometry and optical characteristics that are subject to manufacturing variations. Therefore, the actual output Z_S from the sensor is really an approximation based on the nominal model. It was shown in [5] that the relationship between the actual Z and measured Z_S distances can be modeled using the model:

$$\frac{1}{Z} = a_Z \cdot \frac{1}{Z_S} + b_Z \tag{7}$$

where a_Z, b_Z are the values that are characteristics of particular sensor. Ideally, $a_Z = 1$ and $b_Z = 0$. However, for a particular sensor, depth model parameters a_Z, b_Z may differ from ideal values and it leads to systematic errors in depth measurement. Therefore, appropriate tuning of depth model parameters may improve the accuracy.

In this work, identification of parameters a_Z, b_Z is conducted by following the procedure proposed in [5]. First, a set of pairs of RGB/depth images of the same checkerboard that was used in camera calibration is collected (note the difference to the camera calibration case, where the pairs of RGB/IR images were acquired). Each of RGB images is converted into grayscale and corner coordinates are extracted for inner 10×9 corners. The extracted pixel coordinates $\mathbf{x}_{RGB}^{(p)}(i,j,k)$, $i = 1,\ldots,10$, $j = 1,\ldots,9$, are then paired to known external coordinates $\mathbf{X}_C(i,j) = \left[w(i-1), h(j-1), 0, 1\right]^T$, where w, h denote the width/height of checkerboard fields, to infer the position and orientation of the checkerboard ${}^{RGB}\mathbf{T}_C(k)$ for k^{th} view. Using the known transformation between camera frames, corner coordinates are expressed in IR camera frame as:

$$\mathbf{X}_{IR}(i,j,k) = {}^{IR}\mathbf{T}_{RGB} \cdot {}^{RGB}\mathbf{T}_C(k) \cdot \mathbf{X}_C(i,j) \tag{8}$$

The z-component $Z(i,j,k)$ of such obtained position $\mathbf{X}_{IR}(i,j,k)$ is afterward compared to sensor reading. The sensor value is determined by converting $\mathbf{X}_{IR}(i,j,k)$ into pixel coordinates of IR camera and by searching for the nearest neighbor in sensor depth map m_k:

$$\mathbf{x}_{IR}^{(p)}(i,j,k) = \mathbf{K}_{IR} \cdot f_{IR}^{(d)}(\mathbf{x}_{IR}(i,j,k)) \tag{9}$$

$$Z_S(i,j,k) = m_k(\text{round}(\mathbf{x}_{IR}^{(p)}(i,j,k))) \tag{10}$$

Finally, the obtained set of pairs $Z(i,j,k)$ and $Z_S(i,j,k)$ is employed to fit the parameters a_Z, b_Z of depth measurement model (7) using the least squares fit.

Calibration was performed using three central views displayed in Fig. 4 (central front-side checkerboard views, one from each row in Fig. 4) and the resulting values of parameters were $a_Z = 0.9969$, $b_Z = 4.2881 \cdot 10^{-6}$.

Fig. 3 illustrates the depth correction curve obtained for these values (the curve shown in black). It is seen that the correction increases with depth up to the value on the order of 55mm at the end of the sensor range. For reference, Fig. 3 also shows measurement errors obtained for all checkerboard corners in all views in Fig. 4. The measurement points are shown as green dots; the points that were used in calibration are highlighted in red.

It is important to underline that the error reduction that can be achieved by applying the described procedure is indeed a reduction of differences between reading of depth sensor and depth values obtained using the sensor's RGB camera. Therefore, well calibrated camera is an absolute prerequisite for a quality calibration.

IV. Sensor Accuracy

In this section, an analysis is made of achievable accuracy of both uncalibrated and calibrated sensor and the improvement introduced by calibration.

To this end, two possible scenarios are examined. In the first scenario, a use of only the depth sensor is assumed. In the second scenario, the concern is on accuracy of 3D coordinates inferred for a selected point in RGB image. Thus, in this case, a joint use of depth sensor and the RGB camera is explored.

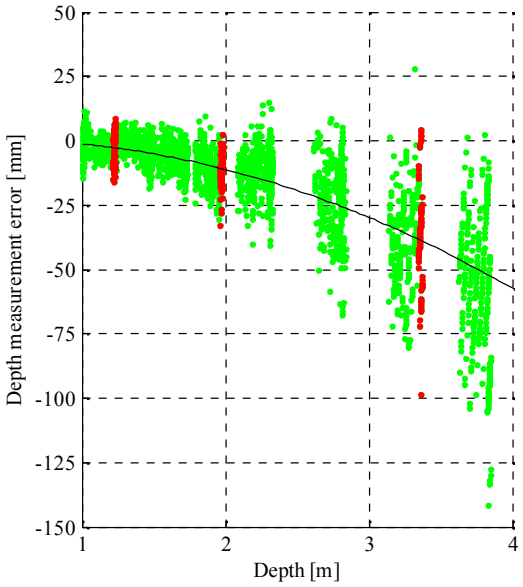


Figure 3. Depth measurement correction

Figure 4. Checkerboard views used in depth model calibration and accuracy analysis

In both examined scenarios, three levels of sensor modeling are investigated: the nominal sensor model, the sensor with calibrated RGB/IR cameras, and the case with the additional depth measurement model calibration.

The analysis is based on measurements conducted on the same checkerboard employed in sensor calibration, using the views of the checkerboard shown in Fig. 4 (with the exclusion of the three central views that were used in calibration of the depth model). As in the calibration case, measurement points correspond to inner 90 corners of the checkerboard.

For the purpose of analysis, the views were divided into nine groups, each containing two or three views of the checkerboard placed approximately at the same position with respect to the sensor but in different orientations. In this manner, a set of nine clusters of measurement points is obtained. The clusters were used to estimate RMS errors and standard deviations for different distances from the sensor.

A. Depth Sensor Accuracy

In this analysis, the accuracy of 3D coordinates $\mathbf{X}_{IR}(i,j,k)$ determined by the sensor for selected points $\mathbf{x}_{IR}^{(P)}(i,j,k)$ in its depth map is examined. The points $\mathbf{x}_{IR}^{(P)}$ are selected as projections of checkerboard corners and they are computed using the same procedure as in depth measurement calibration: first, the transform ${}^{RGB}\mathbf{T}_{C}(k)$ is determined for each view k, then (8) is applied to find the best guess $\hat{\mathbf{X}}_{IR}(i,j,k)$ of actual external coordinates from which $\mathbf{x}_{IR}^{(P)}(i,j,k)$ are calculated using (9). Sensor output is afterward generated by applying in order (10), (7), (3), and (4). (From the order of calculation it is seen that the factors affecting the accuracy are the IR camera and depth measurement algorithm.)

Resulting deviations of $\mathbf{X}_{IR}(i,j,k)$ from $\hat{\mathbf{X}}_{IR}(i,j,k)$ are clustered according to average depth and the obtained statistics (root mean square error ΔL and standard deviation σ) for different clusters is summarized in Table II and Fig. 5. It is seen that the application of nominal model produces large average errors, which are on the order of 35mm at the distance of 2m and on the order of 75mm at the distance of 3m. Calibration of the IR camera yields significant reduction of errors at shorter distances. (It could be noted here that almost the same results have been obtained after neglecting all deviations in parameters of the camera except for the focal length. This insensitivity is attributed to small camera distortion and the fact that the measurement points were always in the central region of the image.) However, the errors remain large at distances larger than 2.5m. In this range, calibration of depth model allows decreasing average errors more than two times compared to the case of calibrating only camera model and more than three times compared to the nominal sensor model.

B. Joint RGB Camera/Depth Sensor Accuracy

In the second analyzed scenario, it is assumed that the inputs are pixel coordinates $\mathbf{x}_{RGB}^{(P)}(i,j,k)$ of a point in RGB image for which it is requested to obtain the corresponding 3D coordinates $\mathbf{X}_{IR}(i,j,k)$. Compared to the first scenario, the only essential difference are unknown pixel coordinates $\mathbf{x}_{IR}^{(P)}(i,j,k)$. In this work, the task of finding coordinates $\mathbf{x}_{IR}^{(P)}$ corresponding to $\mathbf{x}_{RGB}^{(P)}$ was realized by using the considered (approximate) model of the RGB camera to transform $\mathbf{x}_{RGB}^{(P)}$ into undistorted normalized coordinates $\mathbf{x}_{RGB}$, generating a ray through the center of the RGB camera and the point $\mathbf{x}_{RGB}$, and finally finding an intersection of the ray through the undistorted depth map.

As in the first scenario, points $\mathbf{X}_{IR}(i,j,k)$ are further compared to $\hat{\mathbf{X}}_{IR}(i,j,k)$ and the statistics obtained for clusters of points is calculated. The results are illustrated in Fig. 6, where it is seen that the errors obtained with nominal parameters enlarged by approximately 50% compared to the first scenario. On the other hand, errors ob-

TABLE II.
DEPTH SENSOR ACCURACY INDICATORS

Depth [m]	Nominal reading		With calibrated camera		With calibrated camera and depth model	
	ΔL [mm]	σ [mm]	ΔL [mm]	σ [mm]	ΔL [mm]	σ [mm]
0.96	17.1	2.7	4.7	3.6	4.7	3.4
1.16	19.9	2.1	3.5	2.4	3.2	2.4
1.41	24.6	3.0	4.2	3.5	3.6	2.7
1.65	28.8	3.0	7.4	4.5	4.0	3.0
1.88	33.9	3.5	10.9	6.6	5.7	4.2
2.23	38.2	3.9	12.7	8.4	7.8	5.9
2.76	54.2	7.4	23.3	14.1	11.7	8.1
3.24	66.3	9.7	38.6	18.3	15.5	11.0
3.76	83.0	19.0	58.6	27.8	22.4	16.6

tained with calibrated RGB camera almost did not change, what is expected having in mind that the evaluation of $\hat{\mathbf{X}}_{IR}(i,j,k)$ and $\mathbf{X}_{IR}(i,j,k)$ was done using the same RGB camera model.

V. CONCLUSION

This study confirmed large RMS errors of considered sensors. The errors were on the order of 35mm (standard deviation on the order of 10mm) at the distance of 2m and on the order of 75mm (standard deviation on the order of 15mm) at the distance of 3m. When considering joint use of the depth sensor with its associated RGB camera, the effective RMS errors enlarged by 50%.

Errors could be reduced by calibration of sensor's camera and depth measurement model. The procedure proposed in [5] proved to be effective as it allowed reducing RMS errors more than three times compared to the case when nominal sensor model was employed.

Calibration of depth model was an important element of the overall calibration as it allowed reducing by 50% the RMS errors left after calibration of sensor's cameras.

REFERENCES

[1] Microsoft Corporation, *Kinect for Xbox 360*, online at http://-www.microsoft.com/education/ww/products/Pages/kinect.aspx

[2] Microsoft Corporation, *Kinect for Windows*, online at http://-www.microsoft.com/en-us/kinectforwindows/

[3] ASUSTeK, *WAVI Xtion, Intuitive living room experience*, online at http://event.asus.com/wavi/product/xtion.aspx

[4] PrimeSense, *Get your sensor*, online at http://www.primesense.-com/developers/get-your-sensor/

[5] B. Karan, "Calibration of depth measurement model for Kinect-type 3D vision sensors," in *Proc. 21st Int. Conf. Computer Graphics, Visualization and Computer Vision (WSCG 2013)*, Plzen, Czech Republic, June 2013.

[6] T. Dutta, "Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace," in *Applied Ergonomics*, vol. 43, no. 4, pp. 645–649, 2012.

[7] H. Gonzalez-Jorge, B. Riveiro, E. Vazquez-Fernandez, J. Martínez-Sánchez, and P. Arias, "Metrological evaluation of Microsoft Kinect and Asus Xtion sensors," in *Measurement: Journal of the International Measurement Confederation*, vol. 46, no. 6, pp. 1800-1806, 2013.

[8] N. Burrus, *Kinect calibration*, online at http://nicolas.burrus.-name/index.php/Research/KinectCalibration

[9] C. Zhang and Z. Zhang, "Calibration between depth and color sensors for commodity depth cameras," in *Proc. 2011 IEEE Int. Conf. Multimedia and Expo (ICME)*, Barcelona, Spain, July 2011, pp. 1-6.

[10] K. Khoshelham and S.O. Elberink, "Accuracy and resolution of Kinect depth data for indoor mapping applications," in *Sensors*, vol. 12, pp. 1437–1454, 2012.

[11] J. Smisek, M. Jancosek, and T. Pajdla, "3D with Kinect," in *Proc. 2011 IEEE Int. Conf. Computer Vision Workshops*, Barcelona, Spain, Nov. 2011, pp. 1154–1160.

[12] D. Herrera C., J. Kannala, and J. Heikkilä, "Joint depth and color camera calibration with distortion correction," in *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 34, no. 10, pp. 2058–2064, Oct. 2012.

[13] OpenNI Consortium, *OpenNI, the standard framework for 3D sensing*, online at http://www.openni.org/

[14] Microsoft Corporation, *Kinect for Windows SDK*, online at http://-msdn.microsoft.com/en-us/library/hh855347.aspx

[15] M. Draelos, N. Deshpande, and E. Grant, "The Kinect up close: Adaptations for short-range imaging," in *Proc. 2012 IEEE Conf. Multisensor Fusion and Integration for Intelligent Systems (MFI)*, Hamburg, Germany, Sept. 2012, pp. 251-256.

[16] A. Geiger, F. Moosmann, Ö. Car, and B. Schuster, "Automatic camera and range sensor calibration using a single shot," in *Proc. 2012 IEEE Int. Conf. Robotics and Automation (ICRA)*, Saint Paul, Minnesota, USA, May 2012, pp. 3936-3943.

[17] L. Shibo and Z. Qing, "A new approach to calibrate range image and color image from Kinect," in *Proc. 4. Int. Conf. Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, Nanchang, China, Aug. 2012, vol. 2, pp. 252–255.

[18] J.-Y. Bouguet, *Camera Calibration Toolbox for Matlab*, online at http://www.vision.caltech.edu/bouguetj/calib_doc

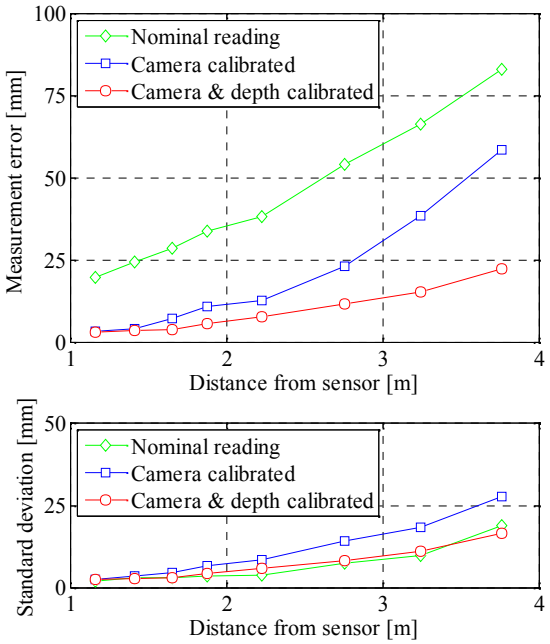


Figure 5. Depth sensor accuracy

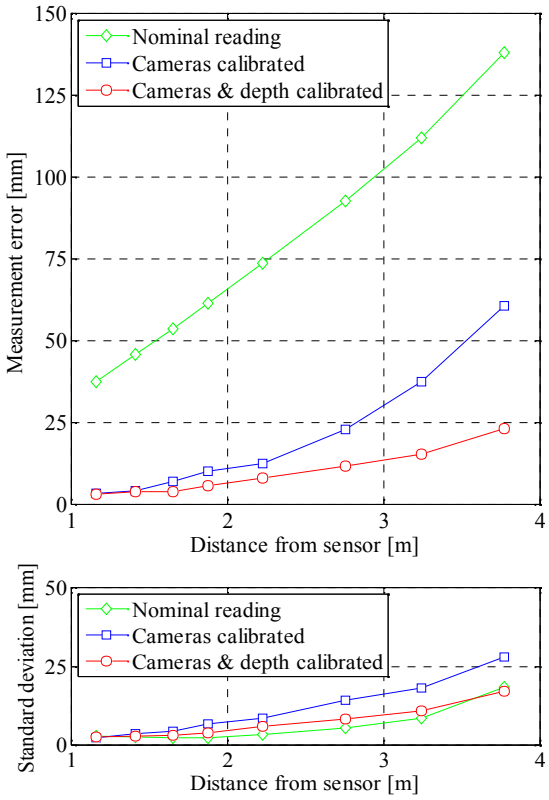


Figure 6. Joint RGB camera/sensor accuracy