

Joint Kinect and Multiple External Cameras Simultaneous Calibration

Yajie Liao¹, Gongfa Li^{1,3}, Zhaojie Ju², Honghai Liu², Du Jiang³

1. Key Laboratory of Metallurgical Equipment and Control Technology (Wuhan University of Science and Technology), Ministry of Education, China 430081

E-mail: liaoyajie123@126.com (Y.L.); ligongfa@wust.edu.cn (G.L.)

2. School of Computing, University of Portsmouth, Portsmouth PO1 3HE, UK

E-mail: zhaojie.ju@port.ac.uk (Z.J.); honghai.liu@port.ac.uk (H.L.)

3. Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan, China 430081

E-mail: jiangdu1231@163.com (D.J.)

Abstract: Camera calibration is a crucial problem in many applications, such as 3D reconstruction, structure from motion, object tracking and face alignment. Numerous methods have been proposed to solve the above problem with good performance in the last few decades. However, few methods are targeted at joint calibration of multi-sensors, which normally is a practical issue in a real-time system. In this paper, we propose a novel method that can simultaneously calibrate relative poses of a Kinect and three external cameras. By optimizing the final cost function, the joint calibration of multiple devices is constructed. Furthermore, the method is tested in a practical platform, and experiment results show that the accuracy of proposed method is higher than the manufacturer's calibration.

Key Words: Joint Calibration; Camera Calibration; Kinect; Computer Vision

1 INTRODUCTION

Camera calibration is a process of estimating intrinsic parameters and extrinsic parameters of camera [1]. It makes the measurement of distances in a real world from their projections on the image plane possible [2]. Thus, With the continuous development of computer/machine vision, camera calibration is believed to be widely applied in 3D reconstruction [3,4], structure from motion [5], object tracking [6-8] and gesture recognition [9,10], etc.

With the development of computer vision, more and more cameras that can acquire 3D information have been proposed, such as stereo cameras and Time of Flight (TOF) cameras. On 4 November 2010, With the launch of low-cost Microsoft Kinect sensors, 3D depth cameras are increasingly attracting researchers due to their versatile applications in computer vision [11]. Kinect was originally developed to improve the game player's experience, enhance human-computer interaction, it actually an RGB-D sensor which provides Synchronized RGB color and depth images. The image capture device of the Kinect include a color camera and a depth sensor which consists of the infrared (IR) projector combined with the IR camera. The experimental results show that Kinect was more accurate than the TOF depth sensor, and close to a medium-resolution stereo camera. However, it is well known that these parameters vary from device to device, and that the factory presets are not accurate enough for many applications [12]. To deal with the above issue, N. Burrus [13] obtained basic Kinect calibration algorithms by using

OpenCV camera calibration, but it calibrated only the intrinsic parameters of the infrared camera. On the other hand, Hirotake et al. [14] tried to independently calibrate the intrinsic of the depth sensor and color camera, and then register both in a common reference frame. Herrera et al. [15] proposed a high-precision color camera to assist Kinect calibration, which can achieve a high accuracy. In addition, Zhang et al. [16] augmented Herrera's work with correspondences between the color and depth images. However, their methods were generally limited to a single external camera, and could not be effectively employed in multiple devices.

For this reason, this paper presents a novel method, which can simultaneously calibrate a Kinect, three external cameras, and their relative positions. By optimizing the final cost function, the joint calibration of the depth sensor in Kinect and multiple external high-pixel color cameras is realized. Experiment results show that the proposed joint calibration methods can achieve a good performance. It does not only reduce the dependence on the number of input pictures, but also improves the accuracy of calibration.

2 CALIBRATION MODEL

2.1 Color camera projection model

In this paper, the intrinsic model of the color camera is similar to that in [17], which includes the pinhole model with radial and tangential distortion coefficients. It is assumed that the color camera coordinates is $X_C = [x_c, y_c, z_c]^T$, and it can be normalized by $\mathbf{x}_n = [x_n, y_n]^T = [x_c / z_c, y_c / z_c]^T$, which can be solved by the following formula [18]:

This work is supported by National Nature Science Foundation under Grant 51575407, 51575338, 61273106, 51575412 and the EU Seventh Framework Programme under Grant No. 611391.

$$\begin{aligned} x_{ct} &= x_n(1 + k_1r^2 + k_2r^4 + k_3r^6) \\ y_{ct} &= y_n(1 + k_1r^2 + k_2r^4 + k_3r^6) \end{aligned} \quad (1)$$

Similarly, tangential distortion happens when the camera lens is not perfectly parallel to the image plane [19]. It can be solved by the following formula:

$$\begin{aligned} x_{ct} &= x_n + [2p_1x_ny_n + p_2(r^2 + 2x_n^2)] \\ y_{ct} &= y_n + [p_1(r^2 + 2y_n^2) + 2p_2x_ny_n] \end{aligned} \quad (2)$$

wherein $r^2 = x_n^2 + y_n^2$, (x_{ct}, y_{ct}) represents the corrected coordinate point. k_1, k_2, k_3 and p_1, p_2 are the radial and tangential distortion coefficients, respectively. Therefore, $\mathbf{k} = [k_1, k_2, p_1, p_2, k_3]$ is used to represent the distortion coefficients. In addition, $\mathbf{k}_c = [k_{c1}, k_{c2}, p_{c1}, p_{c2}, k_{c3}]$ and $\mathbf{k}_d = [k_{d1}, k_{d2}, p_{d1}, p_{d2}, k_{d3}]$ represent the distortion coefficients of the color and depth cameras, respectively.

Then, the image coordinates can be obtained by this formula:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f_x & 0 \\ 0 & f_y \end{bmatrix} \begin{bmatrix} x_{ct} \\ y_{ct} \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \quad (3)$$

where $f = (f_x, f_y)$ is the focal length and $P_0 = (u_0, v_0)$ is the principal point of the image coordinate $P = (u, v)$. The same model can be applied to the color and external cameras. In this paper, the subscript c and d are used to distinguish the same parameter in the color camera and the depth camera, respectively. For example $f_c = (f_{cx}, f_{cy})$ represents the focal length of the color camera.

2.2 Depth camera intrinsic

The transformation relation between the depth camera coordinates and the depth image coordinates is similar to the model that used for the color camera [20]. According to the imaging principle of the depth sensor, the relation between the obtained disparity value d_k and the depth value z_k can be expressed as:

$$z_k = \frac{1}{c_1d_k + c_0} \quad (4)$$

wherein c_1, c_0 are part of the intrinsic parameters of the depth camera to be calibrated. If the measured value of disparity d is directly substituted into equation (4) for calibration (i.e. the disparity distortion correction is not performed). The depth information in the observation process produces a fixed error which could be corrected by adding a spatially varying offset Z_δ to effectively reduce the re-projection error [21], where the depth value z_{kk} can be re-expressed as:

$$z_{kk} = z_k + Z_\delta(u, v) \quad (5)$$

In order to improve the calibration accuracy, the method that in [20] is used to correct the original disparity d directly, [20] took the errors of all pixels from planes at several distances and normalized them. It can be found that the normalization error satisfies exponential

decay [22]. Therefore, a distortion model that uses an attenuated spatial offset to counteract the increasing disparity error can be constructed, and expressed as:

$$d_k = d + D_\delta(u, v)\exp(\alpha_0 - \alpha_1d) \quad (6)$$

where d is the uncorrected disparity value obtained from Kinect, D_δ is used to eliminate the influence of the distortion, and it represents the spatial distortion related to each pixel. α_0, α_1 models the decay of the distortion effect, d_k is the corrected disparity value.

According to the inverse of equation (4), it can be known that:

$$d_k = \frac{1}{c_1z_k} - \frac{c_0}{c_1} \quad (7)$$

and we use Guo's method that simplified the formula (6) by Taylor's formula[23].

$$\begin{aligned} d_k &= d + D_\delta(u, v)\exp(\alpha_0 - \alpha_1d) \\ &\approx d + D_\delta(u, v)(1 + \alpha_0 - \alpha_1d) \end{aligned} \quad (8)$$

Hence,

$$d = \frac{d_k - D_\delta - D_\delta\alpha_0}{1 - D_\delta\alpha_1} \quad (9)$$

The model for the depth camera is described by $L_d = \{f_d, P_{d0}, \mathbf{k}_d, c_0, c_1, D_\delta, \alpha_0, \alpha_1\}$, where the first three represent internal parameters of depth camera, and the last five are used to transform disparity-to-depth values.

3 JOINT CALIBRATION FOR MULTIPLE SENSORES

In this section, we present a novel method that can simultaneously calibrate a Kinect, three external cameras, and the relative poses among them. The block diagram of proposed calibration method is presented in Figure 1. The proposed calibration method that consists of three main consecutive steps:

- Selecting all the checkerboard corners by Zhang's method, and the four corners of the calibration plane are extracted in depth map to initially estimate the intrinsic parameters of depth camera
- Using Herrera's method to estimate the relative positions.
- Initializing the disparity distortion parameters, after that, substituting all the parameters into the new proposed cost function to calculate the nonlinear minimization.

In the workflow framework, step 1 and step 2 complete the initialization of the parameters. In step 3, providing the corresponding initial value, the nonlinear minimization of the parameters can be achieved by iteratively calculating the new cost function. When all the parameters meet the pre-set range, the joint calibration results are output, otherwise it will continue to iterate until the maximum number of iterations.

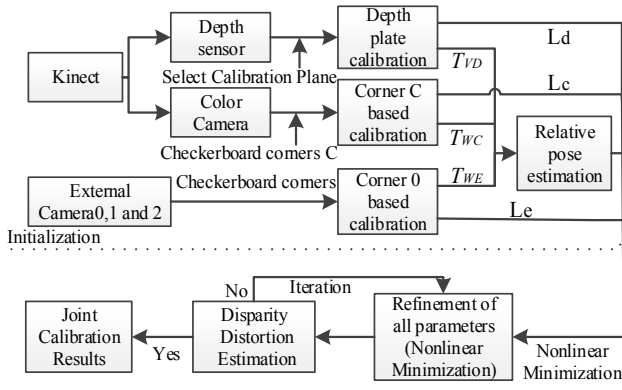


Fig 1. The illustration of work flow of the proposed method

3.1 Preparing corresponding image data and Preprocessing

This article relies on the experimental project. The existing experimental devices that require joint calibration are shown in Figure 2. Kinect was located in front of children who had Autism Spectrum Disorders (ASD), and the same place had an external color camera, which called Middle Camera (External Camera 0). Similarly, in the lower left corner and lower right corner were also fixed the others external color cameras, which called Left Camera (External Camera 1) and Right Camera (External Camera 2), respectively. They were fixed on the same rigid platform and did not change the relative position during the course of the experiment, and the color camera of Kinect was set to the origin of the experimental frame coordinate system.

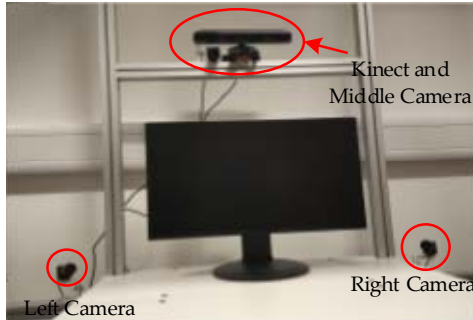


Fig 2. Relative Positions between Kinect and Three Cameras on the Framework

In the process of selecting the checkerboard corners, Zhang's method [24] was used to initialize parameters of the color camera in Kinect and three external cameras. Using a standard checkerboard grid with a width of 0.025 m, and there are 9 and 6 corner points in the x-axis and y-axis directions, respectively. When the number of input images is larger than 3, the unique solution of equation (3) can be found by Zhang's method [24]. In this paper, three sets of data are recorded at the distance of 0.8m, 1.6m and 2.4m away from the camera frame plane, respectively. Each set of data is divided into five pictures. Generally speaking, the corners of the checkerboard cannot be displayed in the depth image, we can only select the four corners of the calibration plate in the depth image. the plane formed by the four selected corners can only be used to initially guess the depth data of the calibration plate plane.

3.2 Relative Pose Estimation

In the relative position estimation, the color camera of Kinect was assumed the origin of the experimental frame coordinate system. All the equipment was fixed on the same rigid frame during the whole experiment. All the reference frames and transformations are illustrated in Figure 3. $\{D\}$, $\{C\}$, $\{W\}$, $\{V\}$ and $\{E0\}$, $\{E1\}$, $\{E2\}$ are the coordinate system of depth, color, checkerboard (world), calibration plate and external cameras, respectively. A point on a coordinate system can be transformed to another coordinate system by $T = \{R, t\}$, where R is the rotation variable, t is the translation variable. For example, T_{WC} represents the transformation from the checkerboard to the color camera coordinate system, and a point X_W in $\{W\}$ can be transformed into $\{C\}$ by the equation $X_C = R_{WC} X_W + t_{WC}$.

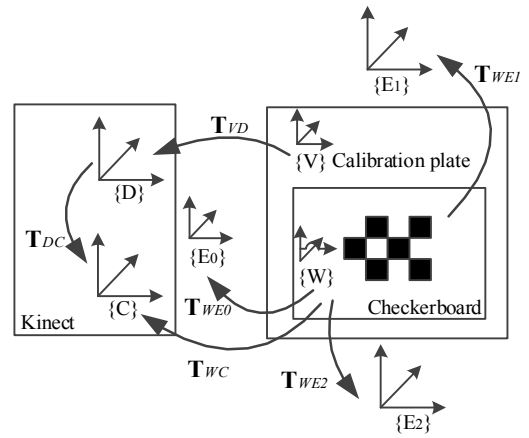


Fig 3. Reference frames and transformations. $\{D\}$, $\{C\}$ and $\{E0\}$, $\{E1\}$, $\{E2\}$ are the coordinate system of depth, color, and external cameras, respectively.

3.3 Nonlinear Minimization

Least square method is a basic, practical, widely used mathematical model [25], by minimizing the sum of squares of the error between samples and its reconstruct samples, and find the best cost function. During the camera calibration, the core of the calibration method aims to minimize the weighted sum of squares of the measurement re-projection errors over all parameters. The re-projection error for the color camera and external camera are the Euclidean distance between the measured corner position and its re-projected position, we assume that the re-projection positions of the color camera and the external camera are \hat{p}_c , \hat{p}_e respectively, and their actual measurement positions are p_c , p_e , respectively. For depth camera, the re-projection error is the difference between the original disparity measurement value d and the re-projection value \hat{d} (i.e. the estimated value of the original disparity) of the disparity. In the formula (4), we can get the original disparity estimated value \hat{d} . the method of [26] can be used to obtain the parameter Z_{kk} in equation (5), then, the original disparity measurement value d can also be obtained. At this point, we have a preliminary cost function:

Table1. Color Camera Intrinsic Parameters

	f_{cx}	f_{cy}	u_{c0}	v_{c0}	k_{c1}	k_{c2}	p_{c1}	p_{c2}	k_{c3}
C.C.	513.10 ±0.06	514.71 ±0.06	323.89 ±0.06	247.65 ±0.06	0.0436 ±0.0003	-0.1521 ±0.0016	0.0036 ±0.0000	0.0000 ±0.0000	-0.0175 ±0.0030
C.E.0	1611.02 ±12.99	1613.66 ±11.58	864.99 ±22.24	507.01 ±14.97	0.4011 ±0.1055	-1.9637 ±0.7412	0.0282 ±0.0034	0.0341 ±0.0084	2.0976 ±1.4582
C.E.1	1727.82 ±5.31	1698.45 ±4.73	432.49 ±18.42	244.32 ±13.85	-0.1442 ±0.0420	-0.0262 ±0.2397	-0.0139 ±0.0014	-0.0262 ±0.0030	0.2329 ±0.4790
C.E.2	1611.02 ±12.99	1613.66 ±11.58	864.99 ±22.24	507.01 ±14.97	0.4011 ±0.1055	-1.9637 ±0.7412	0.0282 ±0.0034	0.00341 ±0.0084	2.0976 ±1.4582

This table shows the focal length (f_{cx}, f_{cy}), the principal point (u_{c0}, v_{c0}) and the distortion coefficient $\mathbf{k}_c = [k_{c1} k_{c2} p_{c1} p_{c2} k_{c3}]$, respectively, wherein C.C. and E.C. represents Color and External Camera, respectively.

$$c = \frac{\sum \|\hat{p}_c - p_c\|^2}{\sigma_c^2} + \frac{\sum (\hat{d} - d)^2}{\sigma_d^2} + \frac{\sum \|\hat{p}_e - p_e\|^2}{\sigma_e^2} \quad (10)$$

where σ_c^2, σ_d^2 and σ_e^2 are the variances of the measurement error of color camera, depth camera and external camera respectively. Obviously, the formula (10) does not comply fully with our requirements. For example, some external camera parameters are completely not used. Thence, equation (10) needs to be modified.

First of all, taking into account the disparity distortion correction of the depth camera, the estimated value \hat{d} of the original disparity is replaced by \hat{d}_k which corrected by the equation (7), the measurement value d of original disparity is replaced by d_k which corrected by the equation (8). In the equation (8), the parameters D_δ and $\alpha = \{\alpha_0, \alpha_1\}$ are independent of all the other parameters, it only depends on the observed values of the pixel (u, v). Therefore, it can be optimized through least squares method individually, and the cost function of disparity distortion can be described as formula (11). Provided the initial values of D_δ and α , then, achieved the optimal solution by iteration.

$$c_d = \sum_{u,v} (\hat{d}_k - d_k) = \sum_{u,v} \left[\left(\frac{1}{c_1 z_k} - \frac{c_0}{c_1} \right) - (d + D_\delta(u, v) \exp(\alpha_0 - \alpha_1 d)) \right] \quad (11)$$

Secondly, the cost function (10) cannot achieve the simultaneous calibration of all external cameras. On this basis, we extend the intermediate term in Equation (10) that is associated with external cameras. Therefore, the modified new cost function is:

$$c = \frac{\sum \|\hat{p}_c - p_c\|^2}{\sigma_c^2} + \frac{\sum (\hat{d}_k - d_k)^2}{\sigma_d^2} + \frac{\sum \|\hat{p}_{ei} - p_{ei}\|^2}{\sigma_{ei}^2}, (i=0,1,2..) \quad (12)$$

It is not difficult to find, the formula (11) is the same as the corresponding intermediate term of the new cost function (12) and does not interact with the others parameters. Therefore, we can directly replace the corresponding initial value in equation (12), the nonlinear minimization of the parameters can be achieved by iteratively calculating the new cost function.

4 EXPERIMENTS

In order to verify the new method in the actual project, all the input images in this experiment comes from the same

database, which was collected and produced by our existing experimental equipment. All pictures are collected in the way described in section 3.1 and saved in JPG format. For comparison with Herrera's method, all depth images in this experiment were saved in the same PGM format as Herrera. In addition, since Herrera's method had a strong dependency on the number of input pictures, the results were random when the number of pictures was less than 20 [22], and the joint calibration method proposed in this paper only needs 15 pictures. The devices intrinsic parameters that calculated by proposed method as shown in Table 1 and Table 2, wherein C.C. represents Color Camera and E.C. represents External Camera.

Table2. Depth sensor intrinsic parameters

f_{dx}	f_{dy}	u_{d0}	v_{d0}	k_{d1}
592.54 ±0.00	588.83 ±0.00	321.05 ±0.00	236.02 ±0.00	0.0701 ±0.0000
k_{d2}	k_{d2}	p_{d2}	k_{d3}	c_0
-0.1596 ±0.0000	0.0034 ±0.0000	-0.0108 ±0.0000	0.0000 ±0.0000	3.28 ±0.000967
c_1	α_0	α_1		
-0.003016 ±0.00	2.4471 0.0718	0.0042 ±0.0001		

This table shows the focal length (f_{dx}, f_{dy}), the principal point (u_{d0}, v_{d0}), the distortion coefficient $\mathbf{k}_d = [k_{d1} k_{d2} p_{d1} p_{d2} k_{d3}]$, the depth parameters (c_0, c_1) and the depth distortion (α_0, α_1), respectively

In our results, each device corresponds to a unique set of values, which is what we expect. Because Herrera's calibration method was limited to a single external camera, and could not be effectively employed in multiple devices. That is, we can only calibrate each external camera one by one. Thence, in the actual calibration process, each of the different external cameras will correspond to a new set of Kinect data. How to choose from multiple sets of Kinect parameters is also a problem.

In this paper, Herrera's method is still used to calibrate the external camera 0, 1, 2, and there are three different sets of Kinect parameter values. We put each set of Kinect parameter values into the 3D reconstruction module,

respectively. By observing the effect of 3D reconstruction, and choosing the best group of values for Herrera's method. By analyzing the calibration results of the two methods for the same set of data, it can be obtained standard deviation of the re-projection error as shown in table 3. Here, the standard deviation of each re-projection error can be regarded as the actual value of the corresponding intermediate term after the nonlinear minimization by formula (12). Therefore, the actual value of c in equation (12) directly reflects the accuracy of the calibration. The smaller the value, the higher the calibration accuracy of the corresponding device. Which can be intuitively found that the results of these two methods are very close, both methods achieve accurate calibration, and the data shows that proposed joint calibration method is more accurate. Therefore, our method not only realize the joint calibration of the depth sensor and multiple external camera, but also improves the accuracy of calibration and reduces the dependence on the number of input images.

Table3. Standard deviation of re-projection error

	Herrera's Method			Proposed Method
C.C.	0.136091	0.141069	0.142754	0.136812
E.C.0	3.424092			3.416770
E.C.1		4.049291		4.093718
E.C.2			1.667060	1.617560
D.C.	0.799590	0.654227	0.654599	0.650924
c	9.935739			9.915784

Wherein C.C. represents Color Camera, E.C. represents External Camera and D.C. represents Depth Camera. c is the parameter in equation (12). To compare the data sets, the variances were kept constant ($\sigma_c = 0.01$ px, $\sigma_d = 0.90$ kud, $\sigma_{ei} = 0.30$ px).

5 CONCLUSION

Considering the problem that current research only focuses on the calibration of a single external camera instead of the calibration of multiple external cameras, we present a novel method that can simultaneously calibrate relative poses of a Kinect and three external cameras. By optimizing the final cost function, the joint calibration of multiple devices is efficiently constructed. At the same time, the validity and accuracy of the method are verified with comparative experiments. Experimental results show that the proposed method not only reduce the dependence on the number of input pictures, but also improves the accuracy of calibration.

REFERENCES

- [1] Z. Zhang, In Flexible camera calibration by viewing a plane from unknown orientations, Computer Vision, 1999 The Proceedings of the Seventh IEEE International Conference on, 666-673, 1999. [CrossRef]
- [2] J. Salvi, X. Armangué, J. Batlle, A comparative review of camera calibrating methods with accuracy evaluation, Pattern recognition, Vol.35, No.7, 1617-1635, 2002. [CrossRef]
- [3] X. Gong, Y. Lin, J. Liu, 3D LIDAR-camera extrinsic calibration using an arbitrary trihedron, Sensors, Vol.13, No.2, 1902-1918, 2013. [CrossRef]
- [4] A. Canessa, M. Chessa, A. Gibaldi, S. P. Sabatini, F. Solari, Calibrated depth and color cameras for accurate 3D interaction in a stereoscopic augmented reality environment, J Vis Commun Image R, Vol.25, No.1, 227-237, 2014. [CrossRef]
- [5] D. G. Choi, Y. Bok, J. S. Kim, I. Shim, I. S. Kweon, Structure-From-Motion in 3D Space Using 2D Lidars, Sensors, Vol.17, No.2, 242, 2017. [CrossRef]
- [6] N. Li, X. Zhao, Y. Liu, D. Li, S. Q. Wu, F. Zhao, Object tracking based on bit-planes, J Electron Imaging, Vol.25, No.1, 013032-013032, 2016. [CrossRef]
- [7] F. Lv, T. Zhao, R. Nevatia, Camera calibration from video of a walking human, IEEE T Pattern Anal, Vol.28, No.9, 1513-1518, 2006. [CrossRef]
- [8] X. Q. Xiang, Z. G. Pan, J. Tong, Depth camera in computer vision and computer graphics: an overview, J Frontiers Comput Sci Techn, Vol.5, No.6, 481-492, 2011. [CrossRef]
- [9] D. Chen, G. Li, Y. Sun, J. Kong, G. Jiang, H. Tang, Z. Ju, H. Yu, H. Liu, An Interactive Image Segmentation Method in Hand Gesture Recognition, Sensors, Vol.17, No.2, 2017. [CrossRef]
- [10] W. Miao, G. F. Li, G. Z. Jiang, Y. Fang, Z. J. Ju, H. H. Liu, Optimal grasp planning of multi-fingered robotic hands: A review, Appl Comput Math, Vol.14, 238-247, 2015.
- [11] Z. Zhang, Microsoft kinect sensor and its effect, IEEE multimedia, Vol.19, No.2, 4-10, 2012. [CrossRef]
- [12] P. Henry, M. Krainin, E. Herbst, X. Ren, D. Fox, In RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments, the 12th International Symposium on Experimental Robotics (ISER), 22-25, 2010. [CrossRef]
- [13] Burrus, N. Kinect Calibration. <http://nicolas.burrus.name/index.php/Research/KinectCalibration>, 2011.
- [14] H. Yamazoe, H. Habe, I. Mitsugami, Y. Yagi, In Easy depth sensor calibration, Pattern Recognition (ICPR), 2012 21st International Conference on, 465-468, 2012. [CrossRef]
- [15] D. Herrera, J. Kannala, J. Heikkilä, In Accurate and practical calibration of a depth and color camera pair, International Conference on Computer analysis of images and patterns, 437-445, 2011. [CrossRef]
- [16] C. Zhang, Z. Zhang, Calibration between depth and color sensors for commodity depth cameras, In Computer Vision and Machine Learning with RGB-D Sensors, 47-64, 2014. [CrossRef]
- [17] J. Heikkilä, Geometric camera calibration using circular control points, IEEE T Pattern Anal, Vol.22, No.10, 1066-1077, 2000. [CrossRef]
- [18] J. Weng, P. Cohen, M. Herniou, Camera calibration with distortion models and accuracy evaluation, IEEE T Pattern Anal, Vol.14, No.10, 965-980, 1992. [CrossRef]
- [19] J. Wang, F. Shi, J. Zhang, Y. Liu, A new calibration model of camera lens distortion, Pattern Recognition, Vol.41, No.2, 607-615, 2008. [CrossRef]
- [20] D. Herrera, J. Kannala, J. Heikkilä, Joint depth and color camera calibration with distortion correction, IEEE T Pattern Anal, Vol.34, No.10, 2058-2064, 2012. [CrossRef]
- [21] J. Smisek, M. Jancosek, T. Pajdla, 3D with Kinect, In Consumer depth cameras for computer vision, 3-25, 2013. [CrossRef]

- [22] C. Raposo, J. P. Barreto, U. Nunes, In Fast and accurate calibration of a kinect sensor, 3DTV-Conference, 2013 International Conference on, 342-349, 2013. [CrossRef]
- [23] L. P. Guo, X. N. Chen, B. Liu, Calibration of Kinect sensor with depth and color camera, J Image Graphics, Vol.19, No.11, 1584-1590, 2014. [CrossRef]
- [24] Z. Zhang, A flexible new technique for camera calibration, IEEE T Pattern Anal, Vol.22, No.11, 1330-1334, 2000. [CrossRef]
- [25] H. Lee, H. Rhee, J. H. Oh, J. H. Park, Measurement of 3-D Vibrational Motion by Dynamic Photogrammetry Using Least-Square Image Matching for Sub-Pixel Targeting to Improve Accuracy, Sensors, Vol.16, No.3, 359, 2016. [CrossRef]
- [26] K. L. Cheng, X. Ju, R. F. Tong, M. Tang, J. Chang, J. J. Zhang, A Linear Approach for Depth and Colour Camera Calibration Using Hybrid Parameters, J Comput Sci Techn, Vol.31, No.3, 479-488, 2016. [CrossRef]