The Study of Improving the Accuracy In the 3d Data Acquisition of Motion Capture System

Changho Han, Soonchul Kim, and Choonsuk Oh,
Department of Information and Communications, SunMoon University
liberman@paran.com, scvictor@naver.com, csoh@sunmoon.ac.kr

Abstract

We introduced the motion capture system using two CCD cameras recently, but could not show any better accuracy than a system using PSD camera. In this paper, we propose two kinds of method to improve the accuracy of the 3D acquisition of the motion capture system using CCD cameras. The applied methods are a distortion removal and z-axis adjustment. We show the result how much the accuracy on the 3D acquisition system improved through comparing with the previous system

Keywords: *3d Data Acquisition, PSD camera, Motion Capture System*

1. Introduction

There are many motion capture systems to obtain 3D data [1] from performances or something moving and the technology of them is growing higher yearly. Specially, the motion capture system of VICON is famous of the precision motion tracking systems, serving customers and CG animation applications in film, visual effects, computer games, and broadcast television. It assisted Image works with installation, provided some custom mo-cap processing tools and provided support on set, which featured over 200 VICON MX40+ cameras in a volume capable of capturing a performer's full facial, body and finger movement, along with various marker props, which included everything from silverware to swords and set pieces. The F40 boasts a resolution of 4 megapixels, captures 10-bit grayscale using 2352 x 1728 pixels, can capture speeds of up to 2,000 frames per second, and can get great data at speeds up to 136,000 markers per second. The typical accuracy of a VICON MX is in the 0.1mm range. However, this equipment's price is very high, so many cheap products is developing now.

In this paper, we show the motion capture system using by active markers and CCD cameras at section 2, and describe the proposed methods to improve accuracy at section 3. The experiments and the result of the efficient on the applied methods will be uncovered at section 4. In the last section, we will discuss about the performance of the proposed methods.

2. Overview of the motion capture system

The existing motion capture system consists of two CCD cameras and the LED markers which has no any cable or radiation control board. The first step in this system is preprocessing to obtain 2D data from the real-time captured stereo images. In the next step, it will be reconstructed 3D data[10] by using disparity algorithm. When 3D data has extracted successfully from the original images, they will be used in the motion recognition system.

Figure 1. Operating procedure of the motion capture system

2.1 Preprocessing

To distinguish markers point in two obtained stereo images by CCD cameras, we used segmentation algorithm simply. At this time, it should be careful of illumination when the system captures performance's motion. Because the scan range of CCD camera is available in visible radiation area. Unfortunately, under the natural light, it cannot recognize any positions of motions. As you see the figure 2, the original input images are dark. The four points of each image is extracted by segmentation algorithm. Four points will be used the motion recognition system that all work in lab. The system is for boxing so it requires at least four of the data points.

Figure 2. Two original stereo images

If you want to get data points for full body animation, you will need more than four points and have to find correspondence points of markers in the other side to reconstruct 3d data points. In the real situation, it may lose the correspondence points because of some kind of noise. Then, we fail to get 3d points from the two images. When we experiment the acquiring 3d data, we disturbed by the around reflective stuff, natural light, etc.

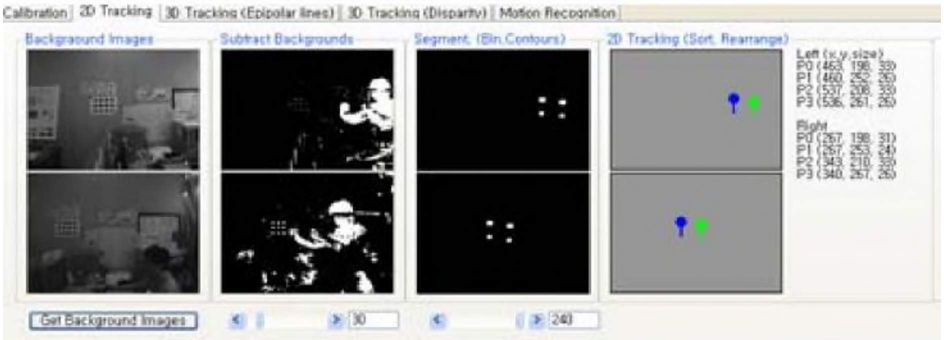


Figure 3. Preprocessing and acquiring 2d data points

As shown in figure 4, the marker consists of LEDs, batteries, and diffuser. It does not have any control hardware system. The diffuser has used to avoid veiling reflections. If a 3d point is not appeared in two images of CCD camera at the same time, it will be ignored, because all points have to have correspondent points in the opposite images. The visible range of marker is within about 125 degree at horizontally.

Figure 4. Marker's images

2.2 Calibration

Two cameras are mounted on a bar. There are some essentials conditions. Two cameras point parallel, perpendicular to the bar. Each camera can slide sideways and be clamped. It needs for the facility to make fine pointing adjustments of one camera to correct for minor misalignments.

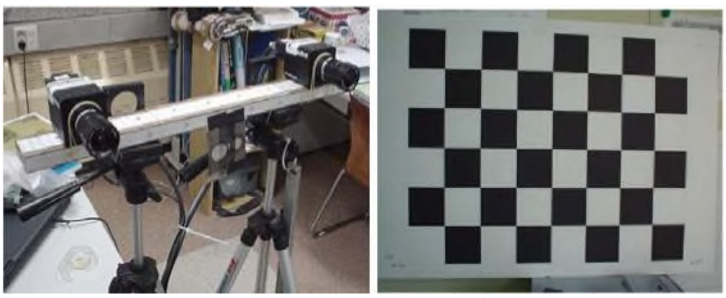


Figure 5. Stereo camera and calibration pattern

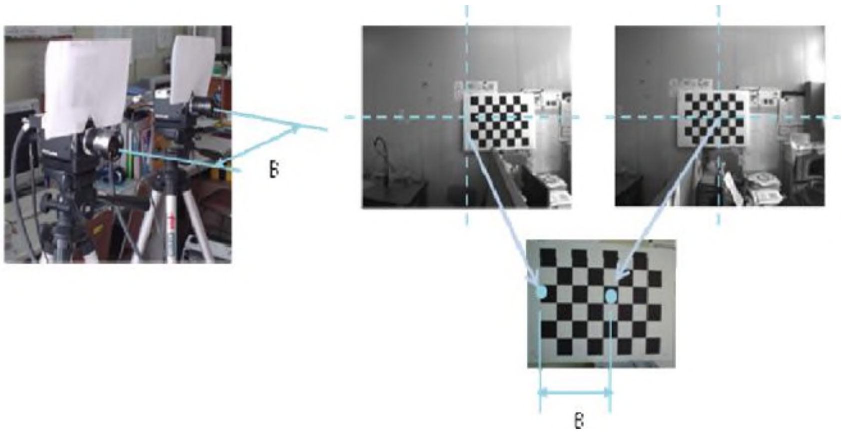


Figure 6. Calibration for parallel setting

To prove two cameras are parallel, we adjust the camera position according to the two points which are the each center point of captured stereo images from chess board. If the distance of two points we found are as same as it between cameras possibly, it will be satisfied the parallel position condition.

2.3 Acquiring 3d data points

After getting 2d data points through the preprocess-ing, it can be calculated 3d data points by using disparity algorithm[2,3,4,5,6,7]. For test, we get 2d data points from the chess board. The number of data points is 35, 7x5 totally. Since a chessboard is used to calibrate a camera[8], this OpenCV function is useful to find where the chessboard is. Finding a chessboard is often a difficult procedure as the contrast between the black and white squares has to be clear and even a little obstruction on the edges of the chess board is enough for OpenCV to fail to find the chess board.

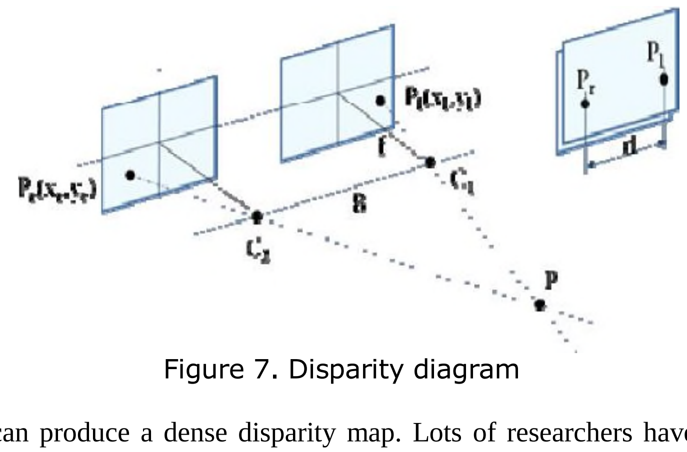


Figure 7. Disparity diagram

Stereo vision can produce a dense disparity map. Lots of researchers have proposed about disparity maps, ex. Area-based stereo methods, window based stereo methods, Bayesian model of stereo matching, cooperative stereo algorithm, etc. However, we utilized a disparity method to 3d disparity space with basic theory as shown in figure 8.

At P(x,y,z) in the real world system, x and y are calculated by average of $_{Pl(x1,yl)}$ and $P_r(x_r,y_r)$. The z value can be calculated by disparity algorithm. Of course is not exactly getting the accuracy depth from that equation. It added some experiment parameters to the equation to get more closed depth data.

3. Proposed algorithms for improving accuracy

There are two kind methods to reduce error of the existing motion capture system. One is to remove distortion of lens, the other is to adjust depth at z-axis.

3.1. Removing distortion

Since lens distortion[11] affects where world points are imaged, the direction of the ray along which a pixel is projected, estimation of lens distortion in a device can significantly improve calibration results, especially for wide-angle lenses. The most important type of distortion is radial, which increases with distance from the center of distortion in the image. The center of distortion is usually located at or near the principal point. In general, the amount of radial distortion is inversely proportional to the focal length of the lens. Most usually assume distortion models that contain radial and tangential distortions. The latter effect is also called decentering distortion [14]. If (x_d, y_d) is a point of a distorted image and (x_u, y_u) is the corresponding point of the undistorted image,

$$x_u = x_d\left(1 + \sum_{i=1} K_i r^{2i}\right) + \left(2P_1 xy + P_2(r^2 + 2x^2)\right)\left(1 + \sum_{i=1} P_{i+2} r^{2i}\right)$$
$$y_u = y_d\left(1 + \sum_{i=1} K_i r^{2i}\right) + \left(P_1(r^2 + 2y^2) + 2P_2 xy\right)\left(1 + \sum_{i=1} P_{i+2} r^{2i}\right) \quad (1)$$

$$x_u = (1 + K_1 r^2 + k_2 r^4) x_d + (2P_1 xy + P_2(r^2 + 2x^2))$$
$$y_u = (1 + K_1 r^2 + k_2 r^4) y_d + (P_1(r^2 + 2y^2) + 2P_2 xy) \quad (2)$$

where $\bar{x} = x_d - C_x$, $\bar{y} = y_d - C_y$, $r^2 = \bar{y}^2 + \bar{x}^2$. C_x and C_y are the optical center. K_i and P_i are the parameters of the radial and tangential distortions, respectively. Actually, the higher order terms on the right-hand of (1) are ignored. The constants that obtained by OpenCV calibration function in this system are k_1=-0.2457, k_2=0.1113, p_1=0.000987, p_2=-0.0000735.

3.2. Adjust depth

As second method, we adjusted the distance with a scale and ellipse equation that looks like fisheye lens. As you see the table 1, it needs change the distance of measured data at z-axis. The equation 3 shows relation between z' and z in 2d, 3d geometry[9].

$$z' = \sqrt{(z^2 - x^2)}$$
$$z' = \sqrt{(z^2 - x^2 - y^2)} \quad (3)$$

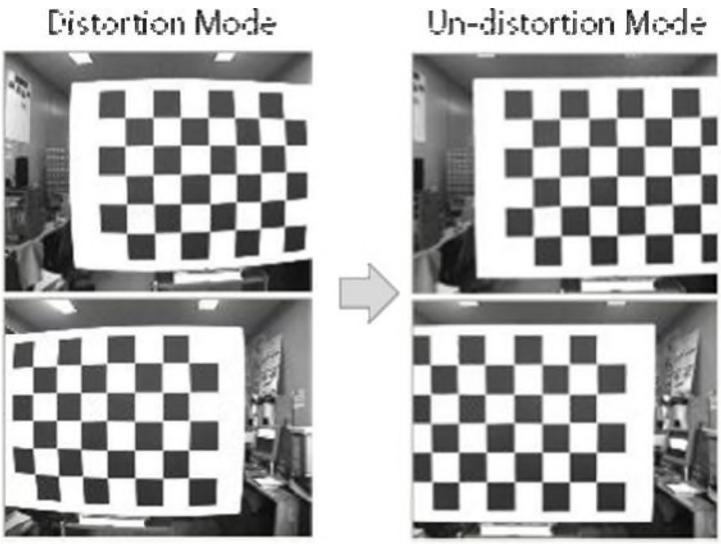


Figure 8. The result of un-distortion image

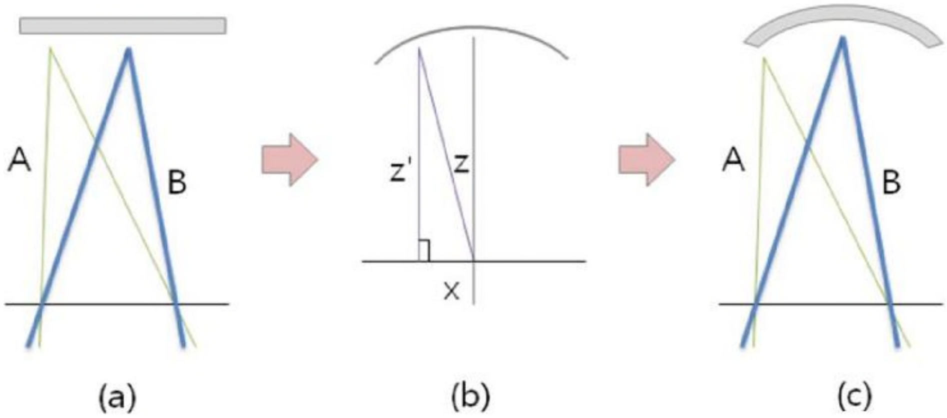


Figure 9. Adjust depth at z-axis

Table 1. The measured and world distance

World Z	Distortion Z	Distortion Scale	Un-Distortion Z	Un-Distortion Scale
750	772.6	0.971	732.0	1.025
1000	1003.6	0.996	972.2	1.028
1250	1234.8	1.012	1209.3	1.034
1500	1467.7	1.022	1444.0	1.039
1750	1702.7	1.028	1684.4	1.038
2000	1952.8	1.024	1934.1	1.034
2250	2170.1	1.037	2155.1	1.044

4. Experiments & Result

We tested our method on several different situations, three of which are showed below.

Table 2. The errors of the accuracy of motion capture.

	Removing Distortion			Adjusting Distance			Removing and Adjusting		
Dist(mm)	x	y	z	x	y	z	x	y	z
750	9.4	5.4	26.0	18.2	14.4	30.7	5.9	3.9	87.0
1000	4.1	5.6	40.4	8.2	10.7	26.2	4.5	4.6	64.4
1250	5.7	5.0	57.5	8.1	7.5	28.2	5.5	5.2	66.1
1500	5.2	4.7	64.8	8.7	6.3	38.6	4.2	8.1	59.8
1750	5.4	5.9	78.0	6.2	6.8	45.3	6.6	8.0	50.2
2000	9.7	4.2	92.0	10.8	5.7	49.8	5.9	7.1	72.3
2250	9.7	7.0	116.0	11.0	7.2	72.3	5.5	7.4	66.0
max	9.7	7.0	116.0	18.2	14.4	72.3	6.6	8.1	87.0

According to the our experiments, we can show that the accuracy is exceed the previous system three times when comparing the result of adjusting distance to it of previous system in the Table 3. It does not take into account the calibration error when we hold the chess board by hand. However, If you suppose to use this system for game or animation[12][13] which is not need a large scale area, about 750'-1250mm, you can get the accuracy is 20'-30mm.

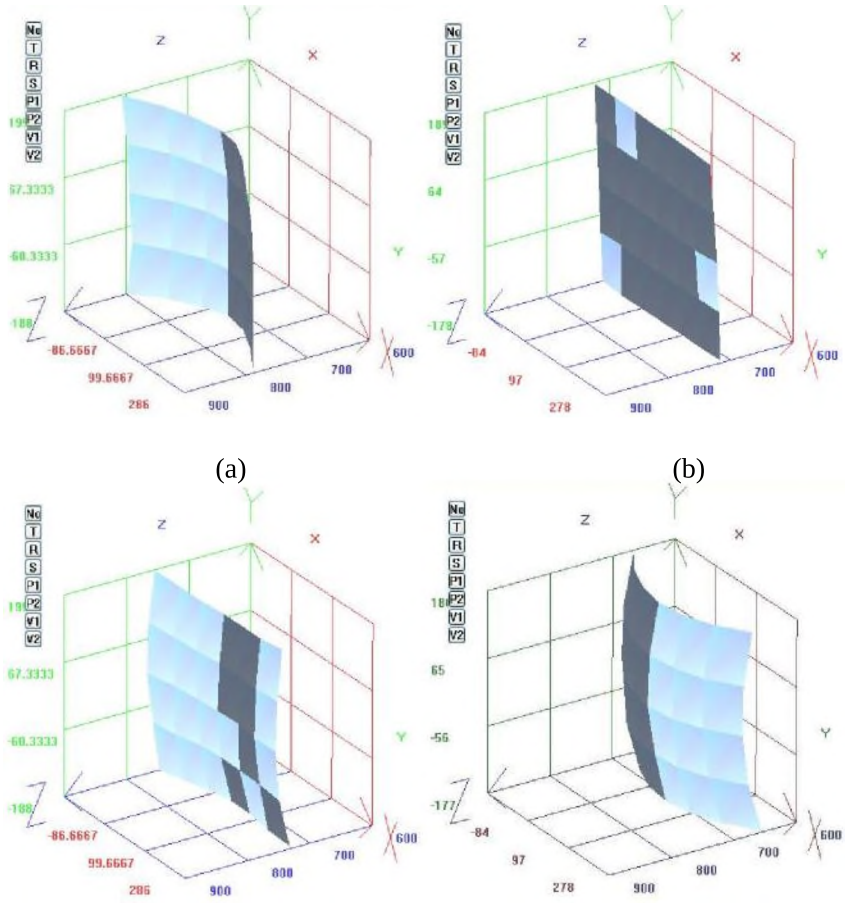
Table 3. The maximum errors of the accuracy of motion capture system.

Measurement Distance(mm)	Previous system	Removing Distortion	Adjusting Distance	Removing +Adjusting
750	96	26	30	87
1000	79	40	26	64
1250	96	58	28	66
1500	123	65	38	60
1750	147	78	45	50
2000	164	92	49	72
2250	179	116	72	66

5. Conclusion and Future work

In this work, the two main methods by which the accuracy for the previous motion capture system are proposed and tested. The method of removing distortion consumes lots of time, because of computing of whole images. It may reduce the estimating time later. The other method of adjusting distance is more effective than it of removing distortion.

One drawback of this system is that it does not robust when capturing motion data, because of natural light. In the next research, we will change the working range of illumination such as infrared ray so that the capturing data is robust.



(a) (b)

(c) (d)

Figure 10. The result of 3d data from chess board; (a) the previous motion capture system, (b) after removing distortion, (c) after adjusting distnace, (d) after removing and adjusting.

Acknowledgement

This work was supported by a Grant-in-Aid for Strategy Technology Development Programs from the Korea Ministry of Knowledge Economy (Robot Component Technical Program; No.10030817).

References

[1] Trucco E., Verri A. 1998. Introductory Techniques for 3-D Computer Vision, Prentice-Hall (eds).
[2] Lucas B., Kanade T. "An Iterative Image Registration Technique with an Application to Stereo Vision", in *Proc. of 7th International Joint Conference on Artificial Intelligence (IJCAI)*, 1981, pp. 674-679.
[3] Gilbert S´ebastien, Lagani`ere Robert, "Registration of a Moving Rigid Object Using a Stereoscopic Vision Setup", in *Proc. of the 3rd IEEE International Workshop on Haptic, Audio and Visual Environments and their Applications - HAVE 2004*, 2004, pp. 171-175.
[4] K.Kim, W.Woo, '3D Camera Tracking from Disparity Images,' *VCIP*, pp. 1381-1388, 2005.
[5] Y.K.Yu, K.H.Wong, S.H.Or and M.M.Y.Chang, "Recursive recovery of position and orientation from stereo image sequences without three-dimensional structures", *in Proc. IEEE CVPR,* New York, Jun. 2006.

[6] Kanade, T. 1994. "Development of a video-rate stereo machine", In Proc. of ARPA Image Understanding Workshop. Monterey, CA, 549-558.
[7] C. L. Zitnick and T. Kanade, 'A volumetric iterative approach to stereo matchingand occlusion detection,' CMU Technical Report CMU-RI-98-30, 1998.
[8] Chang Shu, Alan Brunton, and Mark Fiala. Automatic grid finding in calibration patterns using Delaunay triangulation. Technical Report NRC-46497/ERB-1104, National Research Council of Canada, Institute for Information Technology, 2003.
[9] Hartley, R. and Zisserman, A. 2000. Multiple view geometry in computer vision. Cambridge University Press, Cambridge, UK.
[10] DiFranco, D., Cham, T., and Rehg, J. 2001. Reconstruction of 3-D figure motion from 2-D correspondences, In Proc. of Computer Vision and Pattern Recognition, Kauai, HI, vol. 1: 307-314.
[11] D. Brown, "Decentering distortion of lenses," Photogrammetric Eng., vol. 32, no. 3, pp. 444–462, 1966.
[12] W. Frey, M. Zyda, R. McGhee and W. Cockayne, "Off-The-Shelf, Real-Time, Human Body Motion Capture for Synthetic Environments," Technical Report NPSCS-96-003, Computer Science Dept., Naval Postgraduate School, Monterey, California 93943-5118, USA.
[13] Chai, J., Hodgins, and J. K., 'Performance Animation from Low-dimensional Control Signals, ' *ACM Transactions on Graphics* (SIGGRAPH 2005).

Authors

Choonsuk Oh
He was born in Seoul, Korea on October 11, 1954. He received the B.S. degree in Electronics Engineering from Sogang University, Seoul, Korea in 1980, the M.S. degree in Electrical Engineering and Computer Engineering from Marquette University, Milwaukee, in 1986, and the Ph. D. degree in Electrical and Computer Engineering from the University of Arizona, Tucson, in 1992. He joined the Department of Electronics, Information and Communication Engineering at Sun Moon University, Korea, in 1993, where he is presently a Full Professor. His current research interests include image and signal processing, automated visual inspection and integrated optical devices.

Chang Ho, Han
He received the B.S., M.S. degrees in electronic engineering from the CheongJu University, Korea, in 1993 and 1995, respectively. He is currently pursuing the Ph.D degree in department of electronics, Sunmoon University. His research interests are BLU inspection, Biometrics, and motion capture.

Soon Chul, Kim
He received the B.S. degree in Information and Communication Dpt. From Sunmoon University in 2008. He is now M.S. candidate student in Sunmoon University. He research interests include image processing, motion capture, inspection automation.