International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-4, Issue-11)

Research Article | **November 2015**

Setup and Calibration of a Camera-Based Stereoscopic Vision System

Ioannis Pachoulakis, Konstantinos Kourmoulis
Department of Informatics Engineering, Technological Educational Institute of Crete,
Greece

Abstract—

The present work reports on the accuracy and error assessment of a camera-based, stereoscopic vision system to be used in tandem with other methods to capture and analyse the gait of people with specific kinetic problems. The system employs (a) a pair of Panasonic NV-GS500 DV video cameras in a topology that enables stereoscopic vision and (b) Simi Motion 2D/3D. The videos captured by the cameras are fed to Simi Motion V7.5 and key markers on a calibration structure are used to map points visible in the left and right videos from image screen coordinates to 3D coordinates in the adopted space. Proper calibration and consistent tracking can lead to very small random errors of the order of a few mm. However, systematic errors such as those introduced by camera lens deformation can be substantial. Still, these can be properly assessed and accounted for.

Keywords— Stereoscopic vision, motion capture, skeleton tracking, Parkinson's disease, 3D visualization

I. INTRODUCTION

Computer stereoscopic vision is a process that allows the 3D reconstruction of a scene that is recorded by multiple cameras from different vantage points simultaneously. The process transforms the coordinates of any single point of interest (e.g., a marker) that is visible by more than one cameras from 2D camera coordinates to an adopted 3D world coordinate system. It requires proper calibration of the imaging system using a calibration structure of known geometry placed at a known location and at a known orientation within the adopted 3D coordinate system.

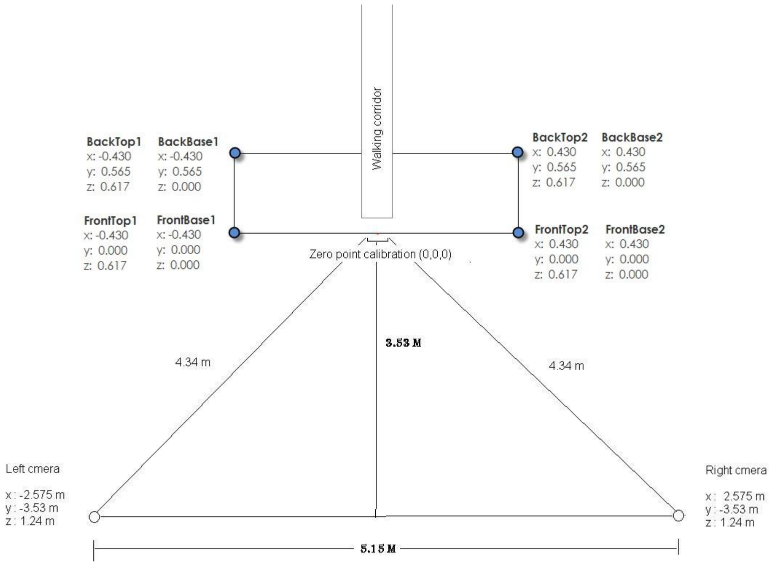
Our stereoscopic vision setup utilizes two Panasonic NV-GS500 Digital Video (DV) camcorders (f3.3-39.6) with resolution 720x576 which record at 25 frames per second using the YUV colour space and the 4:2:2 Chroma sub-sampling scheme. In addition, Simi Motion 2D/3D v7.5 from Reality Motion Systems, Germany offers a complete suite of image processing-based techniques for professional 2D/3D motion analysis, offering a complete range of recording, processing, editing and analysis of motion data.

II. CALIBRATION PROCEDURE

The floor-projected geometry for our stereoscopic recordings appears in Fig. 1 and is the most expansive geometry that can be accommodated in the designated space for the recordings. The two Panasonic cameras, shown at the bottom of Fig. 1, were placed on tripods at a height of 1.24 meters from the floor, symmetrically with respect to a “walking corridor” (along which a subject will be instructed to walk) at a distance of 5.15m between them. The walking corridor bisects the baseline joining the cameras. The adopted right-handed Cartesian coordinate system employs a positive x-axis pointing to the right, a positive y-axis pointing along a walking corridor and in a direction away from the cameras and a positive z-axis aiming up from the floor. The origin of the coordinate system (x,y,z)=(0,0,0) is clearly designated in Fig. 1. The face of the calibration structure closest to the cameras is centred at the origin. This same coordinate system will be used consequently by our motion tracking software to accurately track any marker in 3D space.

The angle between the optical axes of the recording cameras should be between 60 and 120 degrees with optimal results expected at an angle of approximately 90^{o}, according to e.g., [1] and [2]. Based on the available designated space for the shootings, we opted to (a) use the smallest focal distance (f3.3) on the Panasonic cameras to allow patients walking along the corridor to fit in the camera field-of-view and (b) use an opening angle of just 72.2^{o}, which is well within the suggested limits.

Fig. 1 Floor-projected geometry. The (left and right) camera positions are shown at the bottom. The cameras are placed on tripods at a height of 124cm from the ground and aim at the “walking corridor” along which a subject is instructed to walk. The adopted right-handed coordinate system employs a positive x-axis pointing to the right, a positive y-axis pointing along the corridor away from the cameras and a positive z-axis aiming up (from the floor). The face of the calibration structure closest to the cameras is centred at (x,y,z)=(0,0,0) as shown in the figure.



Calibrating the system requires placing an object of known dimensions (typically called a "calibration cube", but in practice it commonly is an orthogonal parallelepiped) at the point of origin and information regarding the distance of each camera relative to the object and inserting the coordinates of its vertices into the system. Our calibration structure, shown in Fig. 2, is an orthogonal parallelepiped with length 86cm, width 56.5cm and height 71,7cm. Proper calibration requires at least eight calibration points that are visible from both cameras, however increasing their number and distribution to encompass the entire 3D space where the subject is supposed to move, i.e., the useful length of the walking corridor, results in positional measurements of higher accuracy. For this first approach we were eager to search the level of accuracy obtained using the eight corners of the given calibration structure (Fig. 2).

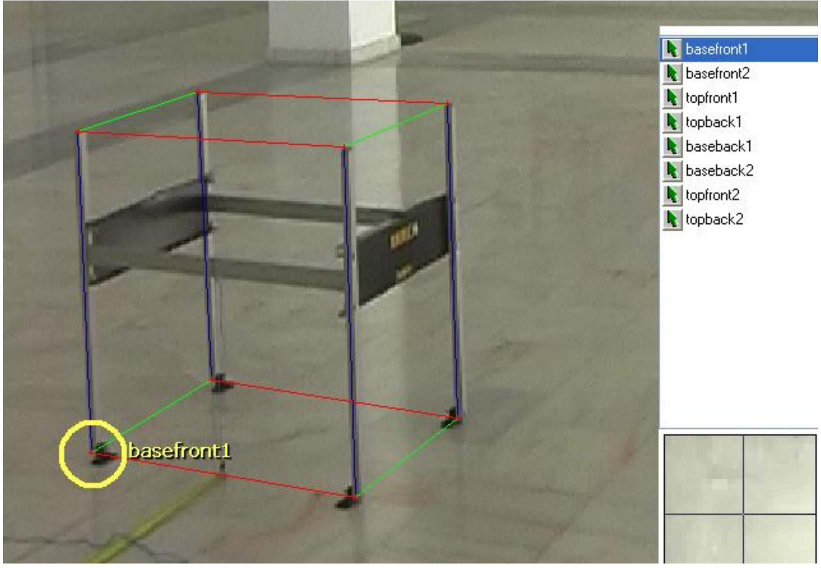
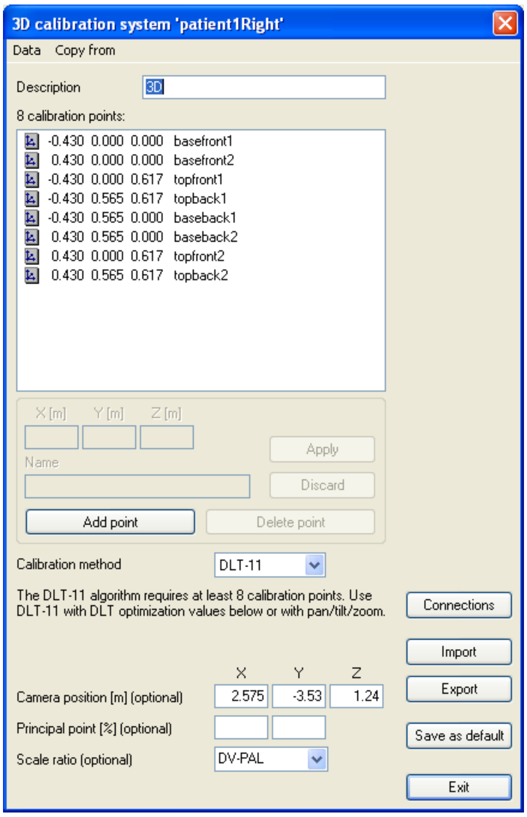


Fig. 2 Calibration structure with 8 calibration points defining an orthogonal parallelepiped.

Next, the actual 3D coordinates in the adopted Cartesian coordinate system for all 8 calibration points (corners of the calibration structure) were fed (in meters) into Simi Motion and were associated with their corresponding locations for each camera. This association process entailed selecting a calibration point and then clicking on the location of that calibration point on that camera's image, then moving to the next calibration point and so on. For example, the top part of the Simi Motion window shown in Fig. 3 lists the 8 calibration points for the right camera. The system is now able to use this information as a reference to accurately calculate the 3D coordinates of any given point on the video region. It should be noted that calibration of extracted videos is not necessary on a per-frame basis, as the cameras are not panned, tilted or zoomed during a recording session. Therefore, calibration of a single frame is sufficient for the entire frame sequence in the video.

Further calibration settings in Simi Motion for the right camera appear in Fig 3 and employ the Direct Linear Transformation technique (e.g., [3], [4]) in its DLT-11 version detailed in [5] which determines pertinent internal and external parameters for the cameras which are necessary for the analysis using the 8 adopted calibration points that are shown in Fig. 3. DLT-11 employs the standard DLT parameters, without higher order optical distortion terms or de-centring distortion terms. Finally, each of the X, Y and Z coordinates for camera position are optional in the sense that, if they are unknown, they will be calculated during the calibration process. However, they are known for our geometry and they have been explicitly entered (for the right camera).



Description: 3D

8 calibration points:

X	Y	Z	Name
-0.430	0.000	0.000	basefront1
0.430	0.000	0.000	basefront2
-0.430	0.000	0.617	topfront1
-0.430	0.565	0.617	topback1
-0.430	0.565	0.000	baseback1
0.430	0.565	0.000	baseback2
0.430	0.000	0.617	topfront2
0.430	0.565	0.617	topback2

X [m] Y [m] Z [m]

Name

Calibration method: DLT-11

The DLT-11 algorithm requires at least 8 calibration points. Use DLT-11 with DLT optimization values below or with pan/tilt/zoom.

	X	Y	Z
Camera position [m] (optional)	2.575	-3.53	1.24
Principal point [%] (optional)			

Scale ratio (optional): DV-PAL

Fig. 3 Calibration cube settings showing the calibration method (DLT-11) being fed 8 calibration points corresponding to the cube edges. The camera position, in this case the right camera, appears near the bottom of the window.

III. ACCURACY DETERMINATION AND CONCLUSIONS

The accuracy of the calibration process was assessed by placing the calibration structure along the y-axis, i.e., along the “walking corridor” at the following distances from the point of origin of the adopted Cartesian coordinate system: y=0, y=3m, y=6m, y=9m, y=12m and y=15m. At each distance, the coordinates of two the bottom edges of the structure closest to the cameras (tagged “FrontBase1” and “FrontBase2”, the first of which is shown in Fig. 2), were measured and the resulting measurements recorded in Table I. For each base point (“FrontBase1” or “FrontBase2”), Table I lists:

- In the column group tagged “Actual Position”, X(m), Y(m) and Z(m) are the actual coordinates in meters of the relevant base point that are known by the exact placement of the structure at the designated distance Y from the point of origin of the adopted Cartesian coordinate system.
- In the column group tagged “Simi-calculated Positions”, x(m), y(m) and z(m) are the coordinates of the relevant base point that are calculated in using Simi Motion.

TABLE I ACTUAL AND MEASURED COORDINATES FOR THE TWO FRONT BOTTOM CORNERS OF THE CALIBRATION STRUCTURE AT INCREASING DISTANCE Y FROM THE ORIGIN

	Actual Position			Simi-calculated Positions			y-Y (m)
	X(m)	**Y(m)**	**Z(m)**	**x(m)**	**y(m)**	**z(m)**	
Front-Base1 (L) (diamonds in	-0.43	0	0	-0.417	-0.011	0.002	-0.011
	-0.43	3	0	-0.438	3.166	-0.01	0.166
	-0.43	6	0	-0.499	6.381	-0.005	0.381
	-0.43	9	0	-0.563	9.554	0.003	0.554
	-0.43	12	0	-0.631	12.719	0.008	0.719
	-0.43	15	0	-0.702	15.802	0.022	0.802
Front-Base2 (R) (squares in graphs)	-0.43	0	0	0.411	-0.005	0.001	-0.005
	-0.43	3	0	0.405	3.114	-0.006	0.114
	-0.43	6	0	0.334	6.246	0.006	0.246
	-0.43	9	0	0.257	9.389	0.01	0.389
	-0.43	12	0	0.181	12.468	0.02	0.468
	-0.43	15	0	0.096	15.499	0.034	0.499

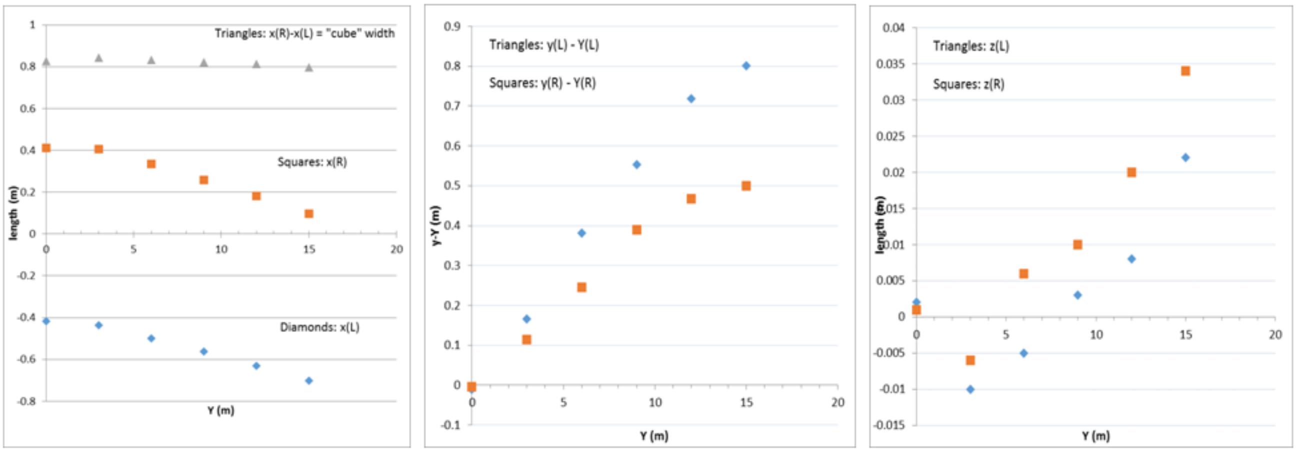


Fig. 4 Systematic positional errors for the bottom two corners of the calibration cube facing the cameras

For each base point, the differences [x-X], [y-Y] and [z-Z] are plotted in Fig. 4 with respect to the known distance Y of the calibration structure from the origin and express error estimates obtained simply by using the front bottom edges of the cube as markers. The left figure insert plots the x-coordinates of the left bottom calibration point (negative values) as diamonds and those of the right bottom corner calibration point (positive values) as squares, both of which drift to more negative values as Y increases in value, although their difference (the actual width of the cube) remains constant. Again for the bottom two corners of the calibration structure, the middle figure insert shows an increasing difference between the measured y coordinate and the actual Y value with increasing Y values. This is a systematic error which can be attributed to the geometric distortion introduced by the wide-angle lens (at f3.3) of the cameras that has not been appropriately taken care of in the calibration procedure due to the small size of the calibration structure used. An equally increasing trend appears in the measured z-coordinates.

One final quantity of interest in our setup is the angular resolution of our cameras at f3.3. This quantity was estimated by placing a brightly coloured orange ball of diameter 4.0cm (a typical width for the patches we used as markers) at increasing distances from one of our cameras set at f3.3 (3.8m, 4.74m, 5.15m and 5.88m) and measuring its size on the camera image in pixels. The respective numbers were 12, 10, 9 and 8 pixels. We thus estimate that, at approximately 9 to 10m from the camera, the entire ball will fit inside a single pixel, meaning that a "walking corridor" starting at the origin and extending for another 4-5m in the direction away from the cameras would not be affected.

REFERENCES

[1] M. Moessner, P. Kaps, and W. Nachbauer, "A method for obtaining 3-d data in alpine skiing using pan and tilt cameras with zoom lenses", *Ski Trauma and Skiing Safety*, vol.10, pp. 155-164, 1996.

[2] R. Bartlett, *Introduction to Sports Biomechanics – Analyzing Human Movement Patterns*, The Netherlands : Routledge, Taylor and Francis Group, 2007.

[3] Y. I. Abdel-Aziz, and H. M. Karara, "Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry," in *Proc. Symposium on Close-Range Photogrammetry*, Falls Church, VA, USA, pp. 1–18, 1971.

[4] R. Shapiro, "Direct linear transformation method for three-dimensional cinematography," *Res. Quart.*, vol.49, pp. 197–205, 1978.

[5] H. Hatze, "High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified dlt-approach," *Journal of Biomechanics*, vol. 21, pp. 533–538, 1988.