RECONSTRUCTING 2D PLANAR COORDINATES USING LINEAR AND NON-LINEAR TECHNIQUES

Marcus Dunn[1], Jon Wheat[1], Stuart Miller[2], Steve Haake[1] and Simon Goodwill[1]

Centre for Sports Engineering Research, Sheffield Hallam University, UK[1]
International Tennis Federation, Roehampton, UK[2]

Planar analyses of sports movements are commonly performed to quantify player movements from video. Lens distortions, which are common when using wide filming angles at competitive sports events, can influence the accuracy of reconstructed coordinates and derived metrics. This study describes a planar reconstruction method that accounts for lens distortion and compares reconstruction accuracy to 2D-DLT. The planar reconstruction method yielded improved reconstruction accuracy in wide angle filming conditions experienced at an international tennis event.

KEYWORDS: planar, reconstruction, method, calibration, 2d-dlt.

INTRODUCTION: The identification of player movements, specifically gait strategy, during competitive tennis is an important step for understanding tennis player-surface interactions. At competitive sports events, camera field-of-view is typically non-perpendicular, resulting in perspective projected images. Direct Linear Transformation (DLT) (Abdel-Aziz & Karara, 1971) is used extensively in sport biomechanics to reconstruct position data from camera images. However, Dainis & Juberts (1985) reported that DLT reconstruction error at the outer 10% of an image was 100% greater than at the image centre. Modified versions of DLT, to account for symmetrical lens distortion, yielded only small accuracy improvements, i.e. 0.1 mm (Challis & Kerwin, 1992), and remain largely unadopted. However, small accuracy gains might reflect controlled filming conditions, as image distortion at the lens' optical centre is zero (Bradski & Kaehler, 2008). Two-dimensional (2D) analyses of sports movements, where motion is considered to occur on a single plane, are commonly performed to obtain basic performance metrics, e.g. step length, step frequency, etc. A planar modification of DLT, termed 2D-DLT, calculates eight DLT coefficients necessary to reconstruct the 2D position of a point on a plane (Walton, 1981 cited by Kwon, 1999). The accuracy of 2D-DLT reconstruction is dependent on a number of factors. A minimum of four calibration points are necessary to calculate the DLT coefficients. Increasing the number of calibration points has been shown to reduce 2D-DLT reconstruction error (McLean et al., 2004). Further, reconstructed points are assumed to exist within the area defined by calibration points. Brewin and Kerwin (2003) demonstrated trends of greater reconstruction error for points located outside of calibration points. Finally, although lens distortion can be calculated using 2D-DLT (Feng et al., 2004), current implementations of 2D-DLT do not account for image distortions due to the lens. In competitive sport scenarios, lens distortion, required by wide filming angles, affects point reconstruction accuracy. A non-linear camera calibration technique (Zhang, 1999) can be used to define a camera-plane model describing camera position, orientation, focal length and lens distortion. Planar image points can be reconstructed when the filming location requires lens distortion. The purpose of this study was to assess the accuracy of a novel, planar reconstruction method for reconstructing planar coordinates in a restricted sport environment, in comparison to 2D-DLT.

METHOD: Part 1 – Data collection: A single high-definition video camera with a regular lens (Everio GZ-HD40EK, JVC, Japan), operating at 25 Hz, was mounted on a tripod at the Roland Garros (RG) Qualifiers (Paris, 2011). The camera field-of-view was zoomed-out, i.e. wide filming angle, to capture the tennis court as well as baseline and sideline areas (Figure 1A). Camera shutter speed, aperture and focal length were set manually and then locked. To perform a planar calibration (Zhang, 1999), the camera was panned (180°) on the tripod to film (~120 s) a checkerboard held in different positions (<4 m) and orientations relative to the camera. The camera was then panned back to the desired field-of-view without making any alterations to intrinsic camera parameters, i.e. zoom, focal length, etc. Deinterlaced still images (50 Hz) of the checkerboard were extracted and processed using Matlab-based software (Bouguet, 2010). Checkerboard corners were extracted and intrinsic camera parameters calculated. Extrinsic camera parameters were defined by digitising the uv coordinates of four known points, i.e. court line intersections, during post-calibration filming. Calibration points were assumed to identify a plane and remain stationary in relation to the camera. Extrinsic camera parameters reveal that camera elevation was 8.6 m, resultant translation was 26.2 m and camera azimuth was 52.8º to the court's positive X axis (Figure 1A). Pinhole camera geometry illustrates the relationship of these parameters for the camera-plane model; a detailed explanation is provided by Zhang (1999).

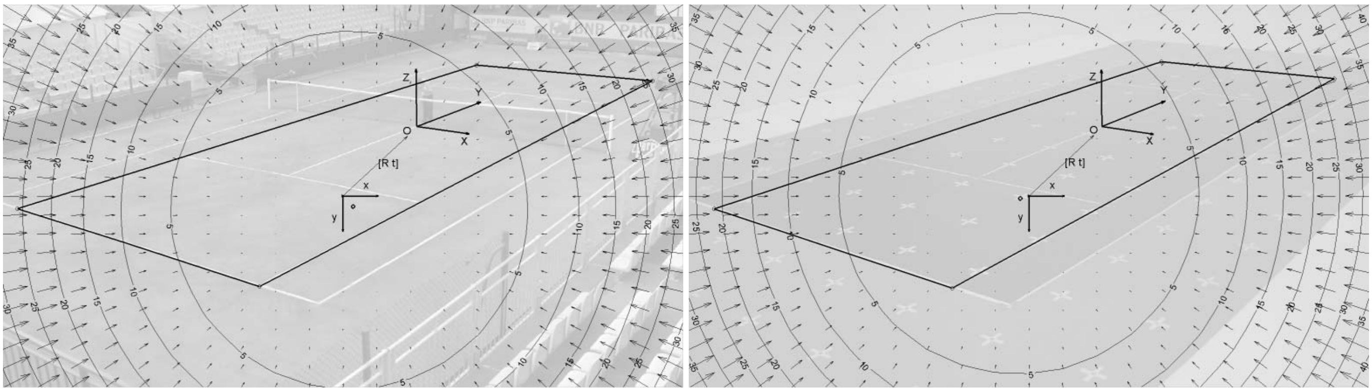

Figure 1: Camera perspectives of real (A) and model (B) tennis courts in filming conditions experienced at Roland Garros. Images illustrate perspective projection and lens distortion (arrows and rings, values in pixels) required to capture the whole court, i.e. wide filming angle. $[\boldsymbol{R}\ \boldsymbol{t}]$ illustrates the homography between court (XYZ) and camera (xy) coordinate systems, where $\boldsymbol{R}$ is a 3×3 matrix of direction cosines and $\boldsymbol{t}$ is a 3×1 translation vector, i.e. $[t_x t_y t_z]^T$.

Part 2 – Modelling: Tennis courts at international competitions are typically inaccessible. A 1:30 scale model of a tennis court was created using CorelDRAW (Graphics Suite 12, Corel, USA), printed on size A0 paper and affixed to a level, planar surface (Figure 1B). Reconstruction points (n=162) and calibration points (court line intersections; n=21) were printed on the paper (represented by crosshairs) to aid manual digitising. Extrinsic camera parameters from the real world camera calibration – described in part 1 – were used to position and orientate the same camera – used in part 1 – in relation to the scale model. Camera shutter speed, aperture and focal length were set manually and locked; the tennis court model was then filmed for 5 s. Camera calibration – described in part 1 – was then performed to determine intrinsic and extrinsic camera parameters.
Extrinsic camera parameters reveal that camera elevation was equivalent to 8.8 m, resultant translation was equivalent to 26.8 m and camera azimuth was 51.8º to the court's positive X axis (Figure 1B). Similar camera position, field-of-view and lens distortion, i.e. Figure 1, illustrate the efficacy of the approach. Reconstruction points in view were manually digitised at a sub-pixel resolution on five occasions. Standard error of the mean was less than 0.4 pixels for all uv image coordinates. Raw image coordinates of reconstruction points were reconstructed using existing 2D-DLT routines (Meershoek, 1997); the number of calibration points passed to the 2D-DLT routine was incremented from four to 15, i.e. observable court line intersections. The same uv coordinates were then passed to the planar reconstruction method using only the initial four calibration points. Raw image coordinates of reconstruction points were first normalised to the camera coordinate system (Bouguet, 2010). Assuming coplanarity, any normalised point in the camera image can subsequently be reconstructed:

$$z_{inv} = 1 \,/\, \big(n \cdot (P - C) \,/\, n \cdot ([x_n, y_n, 1]^T - C)\big) \quad (1)$$
$$[xy]^T = [x_n, y_n] \,/\, z_{inv} \quad (2)$$
$$[xy0]^T = [[xy]^T, 1 \,/\, z_{inv}] \quad (3)$$

$$\begin{bmatrix} X \\ Y \\ 0 \end{bmatrix} = [R]^T \begin{bmatrix} x & - & t_x \\ y & - & t_y \\ 0 & - & t_z \end{bmatrix} \quad (4)$$

where n is the court or plane normal vector, P is the court or plane origin, C is the camera origin, x_n and y_n are normalised horizontal and vertical image coordinates respectively.

Reconstructed coordinates were sorted to identify internally and externally located points in relation to calibration points, i.e. inside or outside of court markings. The root mean square error (RMSE) between reconstructed and world coordinates was calculated for the X, Y, i.e. net and centreline respectively (Figure 1), and resultant (R) directions with the following:

$$RMSE = \sqrt{\sum_{i=1}^{N}(X_{iR} - X_{ir})^2 \,/\, N} \quad (5)$$

where X_{iR} is the world coordinate, X_{ir} is the reconstructed coordinate and N is the number of points used.

RESULTS: For the Roland Garros model, RMSE in the R direction (rRMSE) for planar reconstruction were between 13.5 – 98.2 and 43.4 – 88.6 mm lower for internal and external reconstruction points respectively, when compared to 2D-DLT (Table 1). Further, for 2D-DLT, RMSE in the Y axis was the largest component of rRMSE, ranging between 38.2 – 116.4 and 64.5 – 113.6 mm greater than planar reconstruction, for internal and external reconstruction points respectively (Table 1).

Table 1: RMSE (mm) for X, Y and R directions using 2D-DLT (incremented calibration points) and planar reconstruction in a high lens distortion scenario

		2D-DLT						Planar
		4	6	8	12	14	15	
RG Internal (n=60)	X	61.0	45.7	39.7	40.7	40.2	39.8	55.9
	Y	154.5	81.0	80.4	80.7	76.3	71.0	38.1
	R	166.1	93.0	89.7	90.3	86.3	81.4	67.7
RG External (n=64)	X	75.3	70.0	67.9	69.5	69.5	71.6	69.7
	Y	148.5	116.5	114.4	99.4	104.0	105.8	34.9
	R	166.5	135.9	133.0	121.3	125.1	127.8	78.0

DISCUSSION: This study describes a method for reconstructing planar coordinates in restricted sport filming conditions. Planar reconstruction analyses have previously been concerned with small calibration objects viewed in optimal conditions (Brewin and Kerwin, 2003; McLean et al., 2004). However, when filming markers on a football pitch, Alcock et al. (2009) reported 2D-DLT reconstruction errors of 0.35 ±0.27 m using an elevated camera (7 m) positioned 3 m from the halfway-touchline intersection. Such close-range filming would require a wide filming angle, i.e. zoomed-out, and as such, induce image distortion due to the lens. Roland Garros represented a realistic sport filming scenario, where low camera elevation (8.6 m) and small resultant translation (26.2 m) required a wide filming angle and thus, image distortion due to the lens. Internal point rRMSE for 2D-DLT were between 13.5 – 98.2 mm greater than for planar reconstruction. Further, for external reconstruction points, rRMSE for 2D-DLT were between 43.4 – 88.6 mm greater than for planar reconstruction.
Greater 2D-DLT rRMSE for internal points demonstrates the importance of accounting for lens distortion in wide angle filming conditions, as such coordinates would normally be considered suitable for analysis. Further, greater 2D-DLT rRMSE for external points reflect previously observed trends (Brewin & Kerwin, 2003) and highlight reports that DLT accuracy can degrade by 100% for points outside the centre 90% field-of-view (Dainis & Juberts, 1985). The RMSE for 2D-DLT in the court's Y axis, for both internal and external reconstruction points, demonstrate the impact of neglecting lens distortion with 2D-DLT,

since greater lens distortion was present in this direction (Figure 1B). This highlights that, in noisy scenarios, reconstruction accuracy using DLT and 2D-DLT can be limited, as lens distortion cannot be considered in linear equation solving (Tsai, 1987). Current data demonstrate improved reconstruction accuracy for 2D analyses using the planar reconstruction method and demonstrate the applicability of planar reconstruction in tennis, as player motion frequently occurs outside of calibration points. The planar reconstruction method yielded encouraging results in realistic sport filming conditions and warrants development for future use in sport biomechanics.

CONCLUSION: This study describes a method for reconstructing planar coordinates from video footage in compromised but realistic sport filming scenarios. The planar reconstruction method yielded improved reconstruction accuracy (resultant direction) for both internal and external reconstruction points in wide angle filming conditions. The planar reconstruction method will be useful to sport biomechanists when filming is restricted by camera location, a limited number of calibration points exist and when performer motion occurs outside of calibration points.

REFERENCES:

Abdel-Aziz, Y.I. & Karara, H.M. (1971). Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. In: *ASP Symposium on Close-Range Photogrammetry*. Falls Church, VA, USA, pp 1 – 18.

Alcock, A., Hunter, A. & Brown, N. (2009). Determination of football pitch locations from video footage and official pitch markings. *Sports Biomechanics*, 8(2), 129-140.

Bouguet, J.Y. (2010). Camera Calibration Toolbox for Matlab [online]. Available from: http://www.vision.caltech.edu/bouguetj/calib_doc/ [Accessed 10 October, 2010].

Bradski, G. & Kaehler, A. (2008). *Learning OpenCV*. O'Reilly Media, CA, USA.

Brewin, M.A. & Kerwin, D.G. (2003). Accuracy of scaling and DLT reconstruction techniques for planar motion analyses. *Journal of Applied Biomechanics*, 19, 79-88.

Challis, J.H. & Kerwin, D.G. (1992). Accurate assessment and control point configuration when using the DLT for photogrammetry. *Journal of Biomechanics*, 25(9), 1053-1058.

Dainis, A. & Juberts, M. (1985). Accurate remote measurement of robot trajectory motion. In: *Proceedings of International Conference of Robotics and Automation*. St. Louis, MO, USA, pp. 92-99.

Feng, W., Li, J. & Yan, L. (2004). Creation of distortion model for digital camera (dmdc) based on 2d Dlt. In: 20th *International Society for Photogrammetry and Remote Sensing*. Istanbul, Turkey, pp. 979-984.

Kwon, Y-H. (1999). Object plane deformation due to refraction in two-dimensional underwater motion analysis. *Journal of Applied Biomechanics*, 15, 396-403.

McLean, S.P., Vint, P.F., Hinrichs, R.N., DeWitt, J.K. Morrison, B. & Mitchell, J. (2004). Factors affecting the reconstruction accuracy of 2D-DLT calibration. In: *28th Annual Meeting of the American Society of Biomechanics*. Portland, OR, USA, pp. 246 – 247.

Meershoek, L. (1997). Matlab routines for 2-D camera calibration and point reconstruction using the DLT for 2-D analysis with non-perpendicular camera angle [online]. Last accessed 20 January 2011 at: http://isbweb.org/software/movanal.html

Tsai, R. (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *Journal of Robotics and Automation*, 3(4), 323-344.

Zhang, Z. (1999). Flexible camera calibration by viewing a plane from unknown orientations. In: *International conference on computer vision*. Corfu, Greece, pp. 666-673.

Acknowledgement: This research was part-funded by the ITF, Roehampton, UK.