Sheffield Hallam University

Single view silhouette fitting techniques for estimating tennis racket position

ELLIOTT, Nathan, CHOPPIN, Simon <http://orcid.org/0000-0003-2111-7710>, GOODWILL, Simon <http://orcid.org/0000-0003-0638-911X>, SENIOR, Terry <http://orcid.org/0000-0002-3049-5724>, HART, John <http://orcid.org/0000-0002-6142-4515> and ALLEN, Tom

Available from Sheffield Hallam University Research Archive (SHURA) at:

http://shura.shu.ac.uk/16383/

Published version

Repository use policy

CrossMark

TECHNICAL NOTE

Single view silhouette fitting techniques for estimating tennis racket position

Nathan Elliott[1,2] · Simon Choppin[1] · Simon Goodwill[1] · Terry Senior[1] · John Hart[1] · Tom Allen[3]

Abstract Stereo camera systems have been used to track markers attached to a racket, allowing its position to be obtained in three-dimensional (3D) space. Typically, markers are manually selected on the image plane, but this can be time-consuming. A markerless system based on one stationary camera estimating 3D racket position data is desirable for research and play. The markerless method presented in this paper relies on a set of racket silhouette views in a common reference frame captured with a calibrated camera and a silhouette of a racket captured with a camera whose relative pose is outside the common reference frame. The aim of this paper is to provide validation of these single view fitting techniques to estimate the pose of a tennis racket. This includes the development of a calibration method to provide the relative pose of a stationary camera with respect to a racket. Mean static racket position was reconstructed to within ±2 mm. Computer generated camera poses and silhouette views of a full size racket model were used to demonstrate the potential of the method to estimate 3D racket position during a simplified serve scenario. From a camera distance of 14 m, 3D racket position was estimated providing a spatial accuracy of 1.9 ± 0.14 mm, similar to recent 3D video marker tracking studies of tennis.

✉ Nathan Elliott
  n.elliott@head.com

1   Centre for Sports Engineering Research, Sheffield Hallam University, Sheffield S10 2NA, UK

2   HEAD Sport GmbH, Wuhrkopfweg 1, 6921 Kennelbach, Austria

3   The Manchester Metropolitan University, John Dalton Building, Chester Street, M1 5GD Manchester, UK

1 Introduction

Developments in science and technology can influence the game of tennis [1]. The International Tennis Federation (ITF) regularly reviews the rules, balancing equipment technological developments with preserving the nature of the sport. Ever-increasing advances in computing power, motion analysis software, sensors, and cameras combined with decreasing technology costs have also fostered the development of systems for measuring performance variables such as tennis ball velocity and spin. In 2014, the rules of tennis were updated to allow the use of player analysis technology (PAT) in competitive play [2]. PAT has the potential to provide players and coaches with key information for competition and training, for research as well as for sport broadcasting.

Obtaining racket motion during strokes is of interest as its speed and orientation at impact influence the rebound of the ball [3]. Sensors can be attached to a racket to measure its motion [4], although these add mass and elite players can distinguish differences in moment of inertia as small as 2.5% [5]. Recent work evaluating wireless inertial measurement units for measuring baseball bat swings indicates that current sensors may not be accurate for the full range of speeds experienced in play [6]. Stereo-calibrated cameras have been used to obtain racket motion in 3D, but most previous research, especially in field conditions, has relied on time-consuming manual digitisation of markers attached to the frame [7–9].

Markerless motion tracking systems have initially been developed to track ball and player movement in tennis, although these are often limited to two-dimensional (2D) analysis. Pingali, Jean and Carlbom [10] and Yan, Christmas and Kittler [11] applied temporal differencing techniques to automatically identify ball and player positions

isea
INTERNATIONAL SPORTS ENGINEERING ASSOCIATION

from broadcast television footage. Pingali, Jean, and Carlbom [10] used positional data to provide a statistical analysis of matches and Yan, Christmas and Kittler [11] developed a system designed for robustness rather than accuracy. Kelley et al. [12] developed a system to automatically measure ball spin rates and velocity from high-speed video footage. Dunn et al. [13] developed a semi-automatic method for identifying foot-surface contacts during match-play using a range of image processing techniques. The techniques utilised in these studies [10–13] cannot be applied directly to accurately track racket motion in 3D.

Marker-free techniques for analysis of movement in 3D often apply a Visual Hull [14], which uses silhouettes—typically extracted digitally—from a number of camera views to reconstruct a volume of interest [15–20]. Corazza et al. [17] tracked walking trials and gymnastic flips, and concluded that Visual Hull-based approaches for tracking human movements require at least eight cameras. Sheets et al. [18] used a Visual Hull generated from the silhouette views of eight high-speed cameras to measure racket and player motion during a serve. Strategic placement and calibration of eight cameras in relatively close proximity to the court and player is not well suited for analysis of racket movement during competitive play. A non-invasive markerless motion capture system utilising one stationary camera is desirable for measuring 3D racket movement, to limit cost, improve portability, reduce setup times, and facilitate use in competitive play.

Price and Morrison [21] applied silhouette fitting techniques to estimate the 3D position of particles using the view from one stationary camera. Initially, they calibrated six stationary cameras to obtain their poses within a common reference frame using a calibration object of known dimensions. The six calibrated cameras were then used to obtain six digital silhouette views of a stationary particle. The relative camera pose associated with each particle silhouette view was known from the calibration, and together, the group of silhouettes was defined as a calibrated set. Adopting methods from Forbes et al. [22], using the same six calibrated cameras, they merged calibrated sets—each with the particle oriented in a different stationary position—increasing the number of silhouette views in the calibrated set up to sixty.

Price and Morrison [21] then estimated the relative 3D position of the particle with respect to an additional stationary camera, which was not part of the calibrated set. They used the Levenberg–Marquardt routine to optimise the consistency between the calibrated set and a silhouette of the particle captured from the camera pose outside the calibrated set. Silhouette consistency was measured using the epipolar tangency constraint [21–24]. By adjusting the candidate relative poses, to minimise the epipolar tangency

error (ETE), an accurate estimate (within 5° of criterion orientation) of the pose of the particle with respect to the camera was made [21]. The accuracy of a pose estimate increases with the number of silhouettes in a set and the irregularity of the particle shape. More regular shaped particles can appear similar from different viewpoints, and therefore, the optimisation routine is more likely to converge to a local minimum away from the solution [21]. Elliott et al. [25] adapted methods from other authors [26, 27] to produce a 44-view calibrated set of a 1:5 scale model tennis racket using one camera. Single view fitting techniques adopted from Price and Morrison [21] were applied to estimate racket position to within ±2 mm. While these single view fitting techniques require prior construction of a calibrated set, they illustrate the potential for measuring racket motion on-court with one camera.

The aim of this paper was to validate single view fitting techniques to estimate the pose of a tennis racket. Applying these techniques to a tennis racket is particularly challenging, as it is more regular in shape than the particles studied by Price and Morrison [21]. An improved method is presented for constructing a more consistent calibrated set of tennis racket views in a laboratory. A view was removed from the calibrated set, and single view fitting techniques were applied to obtain the position of the racket using the silhouette associated with this view. The technique was then applied using computer generated silhouettes of a full size racket model in a simulated on-court scenario to estimate racket position during a simplified serve movement.

2 Methods

2.1 Construction of a calibrated set of silhouette views

A calibrated set of racket silhouette views was constructed in a common reference frame. A racket (Prince Warrior 100L ESP) was mounted upright at the centre of a Perspex board (400 × 300 × 4 mm); two cameras (Phantom Miro M110, Vision Research) were positioned to view the rig (Fig. 1a) and a two-dimensional planar calibration was performed for each camera with a checkerboard based on Zhang's algorithm [28], as per similar work [7, 18, 29–31]. Twenty-one stereo calibrations were performed with the master camera fixed and the slave camera in a different position each time. A common reference frame between all slave camera positions and the racket was obtained by digitising control points on the Perspex board and racket (orthogonal calibration object) in the image plane of the master camera (Fig. 1b). A silhouette view of the racket was extracted digitally using MATLAB's [32] image
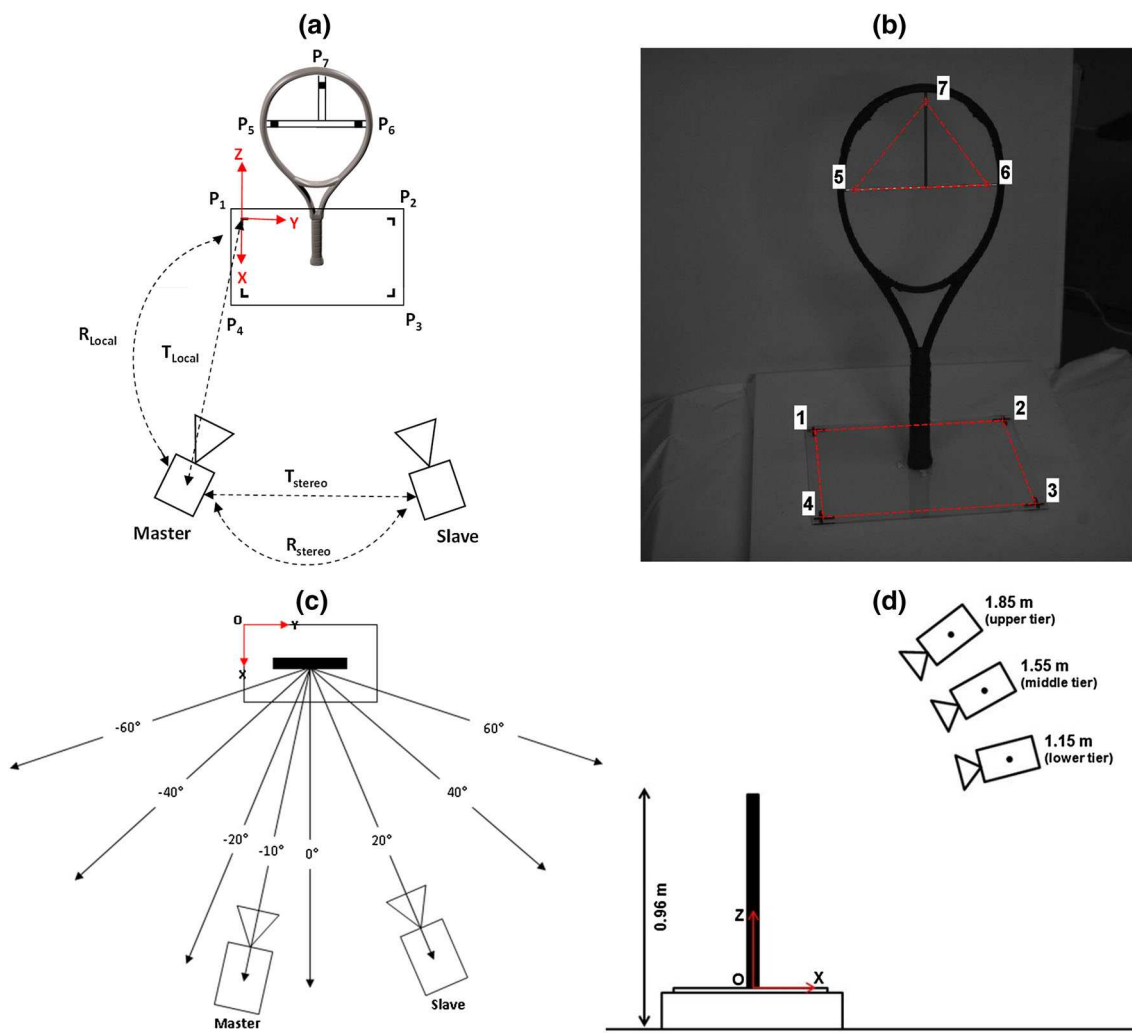
Fig. 1 **a** Schematic of the setup showing the master and slave camera, racket mounted on the perspex board and 7 control points, **b** seven manually digitised 2D image coordinates in the view of the master camera, **c** *top–down view* showing camera angles, and **d** *side-on view* showing racket and camera tier heights

processing toolbox from an image obtained from the slave camera in each position to form the calibrated set.

The racket was painted matt black and white sheets formed a contrasting backdrop to aid digital silhouette extraction (Fig. 1b). The Perspex board formed the *lower calibration plane* (global X–Y plane) and control points 1–4 consisted of machined grooves filled with black paint to form the corners of a rectangle measuring $\sim 360 \times \sim 290$ mm. The origin of the global coordinate system was set at control point 1, creating a left-hand reference frame (Fig. 1a). The racket formed the *upper calibration plane* and control points 5–7 were white marks painted on two black rods (diameter of 4 mm) fixed perpendicular to one another in the frame. A laser scanner (Metris ModelMaker D100) accurate to 0.050 mm was used to obtain the relative position of all control points, while also confirming that the calibration planes were

orthogonal and the face of the racket was parallel to the global Y–Z plane, within practical limits ($<0.5°$).

The master camera was positioned ~ 2 m from the racket, with the optical axis forming an angle of $\sim 10°$ with the global X–Z plane (Fig. 1c). The slave camera was positioned at angles from $\sim -60°$ to $\sim 60°$ in $\sim 20°$ increments at heights of ~ 1.15, ~ 1.55, and ~ 1.85 m to form three tiers (Fig. 1d). A suitable configuration for the calibrated set was found using simulations with computer generated silhouette views in Blender (v2.70) [33], with modifications to incorporate the practicalities of positioning cameras and undertaking a calibration, as detailed in Elliott [34]. The cameras were set at their maximum resolution of 1280×800 pixels, with an F-stop (aperture) of F22 and a shutter speed of 1/8th s. The checkerboard (8×8 squares each measuring 50×50 mm) was set in 50

orientations for each stereo calibration to maximise image coverage [28, 35].

Checkerboard images were passed to Bouguet's calibration Toolbox [36] in MATLAB [32] to obtain the intrinsic [focal length (fx, fy), principal point (cx, cy), and lens distortion] and extrinsic (rotation and translation in a common reference frame) camera parameters. Based on the findings of Elliott [34], the intrinsic parameters were estimated using a 4th order radial distortion model without the tangential component, and they were not recomputed when estimating the extrinsic parameters. The control points on the orthogonal calibration object were manually digitised ten times each (with short breaks to limit any learning effect), with the mean values passed to the Toolbox along with the corresponding world coordinates from the laser scan. The Toolbox used the control point coordinates from the manual digitisation and the laser scan, along with the intrinsic parameters, to compute the relative pose of the master camera with respect to the origin (control point P1), in a common reference frame with the racket. Averaged across all control points, the mean standard deviation from manual digitisation was 0.1 pixels, which equates to a relative calibrated camera pose error of less than 1 mm in translation and below 0.1° in rotation [34].

Using the extrinsic parameters from the stereo calibrations, rigid body transformations [21] were applied to obtain the slave cameras in a common reference frame with the master and racket. The calibration with the orthogonal object was performed once as it remained stationary along with the master camera, reducing uncertainty from manual digitisation in comparison to digitising the control points in an image from each slave camera position. Using an orthogonal object improved camera pose accuracy, in comparison to simply using the Perspex board [25] or racket as a planar calibration object, as detailed by Elliott [34]. MATLAB's [32] image processing toolbox was used to perform thresholding to digitally extract racket silhouettes from the slave camera's images and to segment polygonal silhouette boundaries [37–41]. The extracted boundary was plotted on the original image and its quality assessed visually. The threshold value was manually adjusted, to ensure the extracted boundary provided an accurate representation of the racket in the original image.

2.2 Estimating racket position from candidate relative camera poses

Each silhouette was removed from the calibrated set and its camera pose was estimated using an initial candidate relative pose. Estimates were then compared with a criterion, which was the pose of the camera (obtained from calibration) removed from the calibrated set. Tests with computer generated camera poses confirmed that reducing the number of silhouette views in the calibrated set by one did not influence the results [34].

Since an unloaded (not at impact) tennis racket frame has a fairly regular shape, the probability of the Levenberg–Marquardt optimisation routine converging to a local minimum is increased, resulting in a camera pose estimate on the wrong side of the object (antipodal view) [21]. The method works best if the initial candidate pose provided to the algorithm falls on the correct side of the racket, close to the true camera position. Each candidate pose was, therefore, created using a spherical coordinate system centred at the midpoint of the racket, with the radius corresponding to the known distance of the camera (obtained from calibration) taken from the calibrated set. The curved surface corresponding to the search region for candidate poses extended up to 30° either side (azimuthal angle) and 30° above and below (polar angle) the known camera pose, decreasing the likelihood of the antipodal view being found, as illustrated in Fig. 2.

A maximum of 100 candidate relative poses were used [21] and searches were terminated when the root-mean-squared (RMS) value of the ETE vector reached a threshold of 0.5 pixels. A threshold was required as the RMS ETE would not converge to zero due to inherent inconsistency in the set, as a result of small errors associated with calibration and silhouette extraction. If the threshold was not reached for any of the candidate relative poses, then the solution corresponding to the lowest ETE would be used. A threshold of 1 pixel did not always allow the optimisation to fully converge and reducing the value below 0.5 did not affect the solution.

Each camera pose estimate obtained with the view fitting techniques was used to reconstruct 106 3D coordinates on the racket face plane, using a camera-plane model [42, 43]. The reconstructed coordinates were compared with corresponding points on the racket frame surface obtained from the laser scan (criterion). As the coordinates extracted from the laser scan were not on the racket face plane, stereo triangulation was used for the reconstruction. Pixel projections of the coordinates were obtained using the calibration parameters, allowing for triangulation using the master camera (criterion) and each slave camera pose estimate from the view fitting techniques. The ability of the view fitting method to accurately reconstruct these coordinates was taken as the measure of how well racket position could be estimated.

2.3 Proof of concept for application to tennis

The methods described thus far were designed to develop a calibrated set configuration and validate a single view fitting method to estimate 3D racket position in a laboratory. Application of the method to play conditions requires
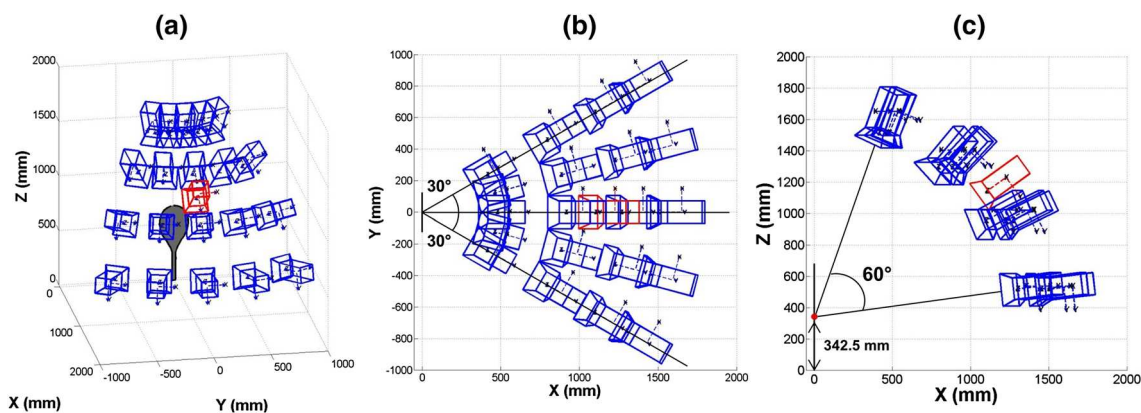
Fig. 2 Example of candidate relative poses generated using spherical coordinates showing **a** 3D view, **b** *top–down view*, and **c** *side view*. The central camera pose corresponds to the true location of the camera

development beyond the scope of this paper. For proof of concept without on-court testing, the calibrated set was simulated using computer generated camera poses and silhouette views [21, 41] of a full size racket model created in Blender (v2.70) [33]. Based on findings of Elliott [34], the simulated calibrated set was modified with the camera poses orientated randomly (not upright) about the optical axis. The random camera pose orientations were generated between −90° and 90° (camera poses were upright at 0°) using an inbuilt MATLAB [32] function. The calibrated set was used to estimate the 3D position of the racket model during a simplified simulated serve movement, using the camera pose in Fig. 3. The pose was similar to those used

by Choppin et al. [7], the camera was outside the court (which was full size) and should not be intrusive during play. The racket was located at the centre mark, with its face aligned with the global Y–Z plane. For simplicity, the racket model butt was set at the global origin as obtaining the relative position between the camera and racket was of interest. The court in Fig. 3 is for illustrative purposes.

To simulate motion during a simplified serve, the racket model was rotated about an axis 10.16 cm (4 inches) from the butt aligned with the global Y axis. This is the location of the axis of rotation used to define the swing weight of a racket [44]. The racket was rotated about the Y axis between −40° and 30° in 2° increments, with 0°

Fig. 3 Schematic to show the camera pose used to generate racket model silhouette images to estimate 3D racket position during a simulated serve
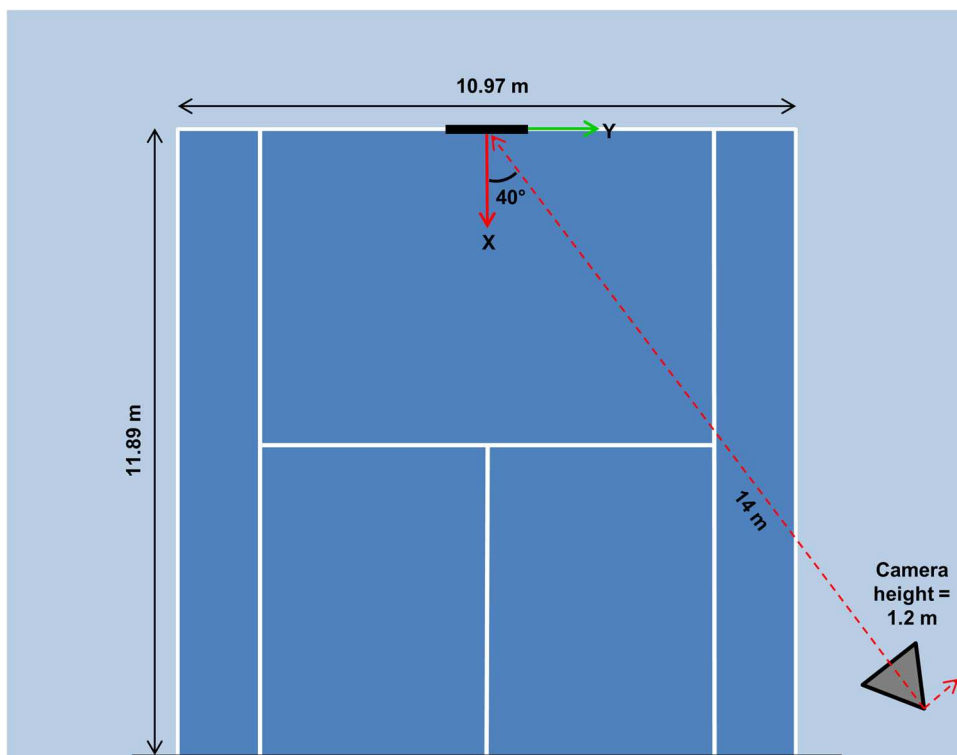
corresponding to upright. Silhouette images of the racket model were rendered every 2°, which for typical racket head speeds during a serve [18, 45–48]; a high-speed camera would need to operate at 200 frames per second (fps), so that sufficient silhouette images could be obtained. The algorithm was instructed to perform two optimisations; the first worked backwards from when the racket was oriented at 0° to −40°, the second worked forwards from 0° to 30°. An orientation of 0° was used as a starting point, because it was found that this position provided a more accurate pose initialisation [34]. Thus, for the first optimisation, with the racket orientated at 0°, the candidate relative pose was obtained using the method described in Sect. 2.2. This scenario requires the operator to provide the algorithm with an initial approximate distance between the camera and the racket, i.e., 14 m should be sufficient for baseline shots (Fig. 3). The following optimisations were then initialised using the camera pose estimate from the previous solution. The 3D racket positions were obtained using the camera pose estimates to reconstruct the 130 coordinates on its face plane in the Y, Z, and resultant dimensions for each angle [13]. Reconstruction results were validated against known 3D coordinates obtained from the racket model mesh.

3 Results

Table 1 shows the mean and standard deviation for the intrinsic parameters (fx, fy, cx, and cy) for the master and slave camera, averaged over the 21 stereo calibrations. Standard deviation values were less than 2.5 pixels, which is equivalent to 0.20 and 0.31% of the field of view in the horizontal and vertical directions, respectively. Mean and standard deviation for the resultant distance of the 21 slave camera centre to the butt of the racket were 1.5 ± 0.76 m. The RMS ETE for the calibrated set was 0.41 pixels (<0.1% of the racket length in the image), which was the same as that reported by Forbes [41].

Figure 4a, b shows views of the racket as seen by two cameras from the calibrated set. Pixel projections of the 3D

coordinates have been plotted on the images. The dots are the criterion (laser scan), while the crosshairs are projections of the reconstructed coordinates of the edges of the frame obtained using the camera pose estimate from the view fitting method. The visible crosshairs in Fig. 4a indicate lower reconstruction accuracy compared with Fig. 4b. In Fig. 4a, RMS error for reconstruction of the coordinates in the X, Y, and Z direction was 2.96, 1.28, and 4.05 mm, respectively, with a resultant of 5.18 mm. The camera pose associated with this view was located at $\sim 60°$ to the racket face plane normal (Fig. 4a). In Fig. 4b, RMSE for reconstruction in the X, Y, and Z directions was 1.13, 0.20, and 0.46 mm, respectively, with a resultant of 1.24 mm.

Figure 5 represents the error associated with the reconstruction of the coordinates on the racket using the camera pose estimate from the view fitting method, for all cameras in the calibrated set. Some camera poses corresponding to frontal and more side ($\sim 60°$) on views of the racket had lower positional accuracy. Camera pose estimate 5 ($\sim 40°$, middle tier as shown in Fig. 1d) produced the lowest reconstruction errors for the coordinates on the racket of 0.1, 0.1, and 0.3 mm in the X, Y, and Z directions, respectively, with a resultant of 0.33 mm. Mean and standard deviation for reconstruction error were 1.46 ± 1.13, 0.29 ± 0.32, and 1.03 ± 1.06 mm in the X, Y, and Z directions, respectively, and the resultant was 1.81 ± 1.58 mm, across all the camera pose estimates.

Figure 6 illustrates racket reconstruction errors for the simulated serve scenario. The racket face plane is perpendicular to the local X axis (Fig. 6a) and aligned with the local Y and Z axes (Fig. 6b). Averaged across all racket positions, mean and standard deviation for reconstruction errors in the Y and Z directions for coordinates on the racket face plane were 0.26 ± 0.17 and 1.93 ± 0.13, respectively, with a resultant of 1.96 ± 0.14 mm. Reconstruction error in the Z direction contributed a larger component of error during the simplified simulated serve. This is because out-of-plane 3D position estimation is difficult when one camera is used in a view fitting [21].

4 Discussion

Single view silhouette fitting techniques were able to accurately reconstruct points on the surface of a racket frame with a mean reconstruction error of <1.5 mm in all three principal directions, equating to a mean resultant reconstruction error of ±2 mm. Unlike previous markerless motion capture methods applied to tennis [18, 20]; the work presented here lays the foundations for a method which would require an accurate calibrated set and not a visual hull. An RMS ETE value of 0.41 pixels was

Table 1 Mean ± standard deviation for the intrinsic parameters (fx, fy, cx, and cy) (pixels) for the master and slave cameras over the 21 stereo calibrations

Intrinsics	Master camera	Slave camera
fx (pixels)	1804.30 ± 1.72	1781.60 ± 1.30
fy (pixels)	1802.20 ± 2.31	1779.50 ± 1.26
cx (pixels)	642.64 ± 0.75	651.53 ± 0.72
cy (pixels)	440.76 ± 0.84	399.30 ± 0.93

Fig. 4 Racket views (zoomed in) corresponding to (**a**) camera pose 1 (∼60°, *upper tier*) and **b** camera pose 13 (∼−20°, *upper tier*) from the calibrated set. The *dots* are exact projections of points obtained from the laser scan of the racket and the crosshairs are projections of coordinates reconstructed using pose estimates

(a)   **(b)**
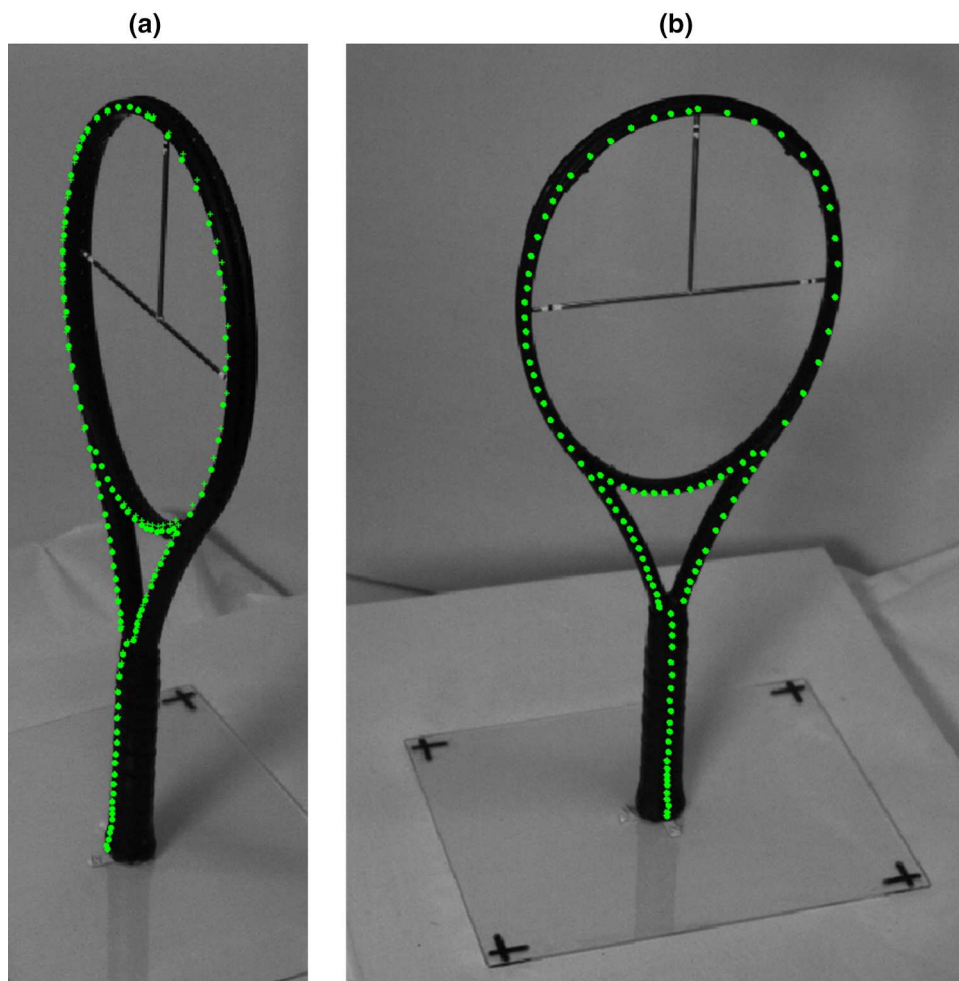
measured for the calibrated set, consistent with values reported by other authors for different objects and calibration methodologies [41, 49]. The consistency of a calibrated set is dependent on the accuracy of camera calibration and the quality of silhouette extraction [41]. Future work will, therefore, look to improve the calibration process and silhouette extraction techniques to increase the accuracy of racket pose estimates.

The largest resultant reconstruction errors were found for views from the front (4.63, 0.66, and 2.46 mm in the *X*, *Y*, and *Z* directions, respectively, for camera pose estimate 11, see Fig. 5c) and side (2.96, 1.28, and 4.05 mm in the *X*, *Y* and *Z* directions, respectively, for camera pose estimate 1, see Fig. 5c) of the racket, which may be due to calibration accuracy and the range of poses in the calibrated set. Using stereo calibration methods, the range of views in the calibrated set was constrained by the position of the master camera, as it was not possible to achieve adequate checkerboard coverage when the convergence angle between the two cameras was large (>∼60°). The convergence angle was low (∼10°) when obtaining frontal

silhouette views, which can result in inaccurate estimations of depth in stereo calibrations [35, 50]. Estimating the pose of a racket from a frontal silhouette view is particularly challenging due to its reflective symmetry, which can be accounted for by having a wide range of views in the calibrated set [21]. While future work could focus on how best to position the master camera to achieve a consistent calibrated set with a wider range of views, the best option may be to merge a number of sets [21, 22, 41, 49] produced with the master in different locations. Simulations utilising computer generated views could assist in determining the most suitable camera poses in the calibrated set, as per Elliott [34], taking into consideration the practicalities of physically reproducing the setup.

Choppin, Goodwill, and Haake [7] reported an accuracy of ±2.5 mm when reconstructing marker positions on a tennis racket using two stereo-calibrated video cameras. In relation to a tennis stroke, the reconstruction errors corresponded to a mean angular error of ±1° and velocity error of ±0.5 m s^{-1} for elite tennis players measured during practice. While the initial results presented here for the

Fig. 5 RMSE (mm) for reconstruction in the **a** *X*, **b** *Y*, and **c** *Z* directions of 3D points on the racket frame surface using camera pose estimates
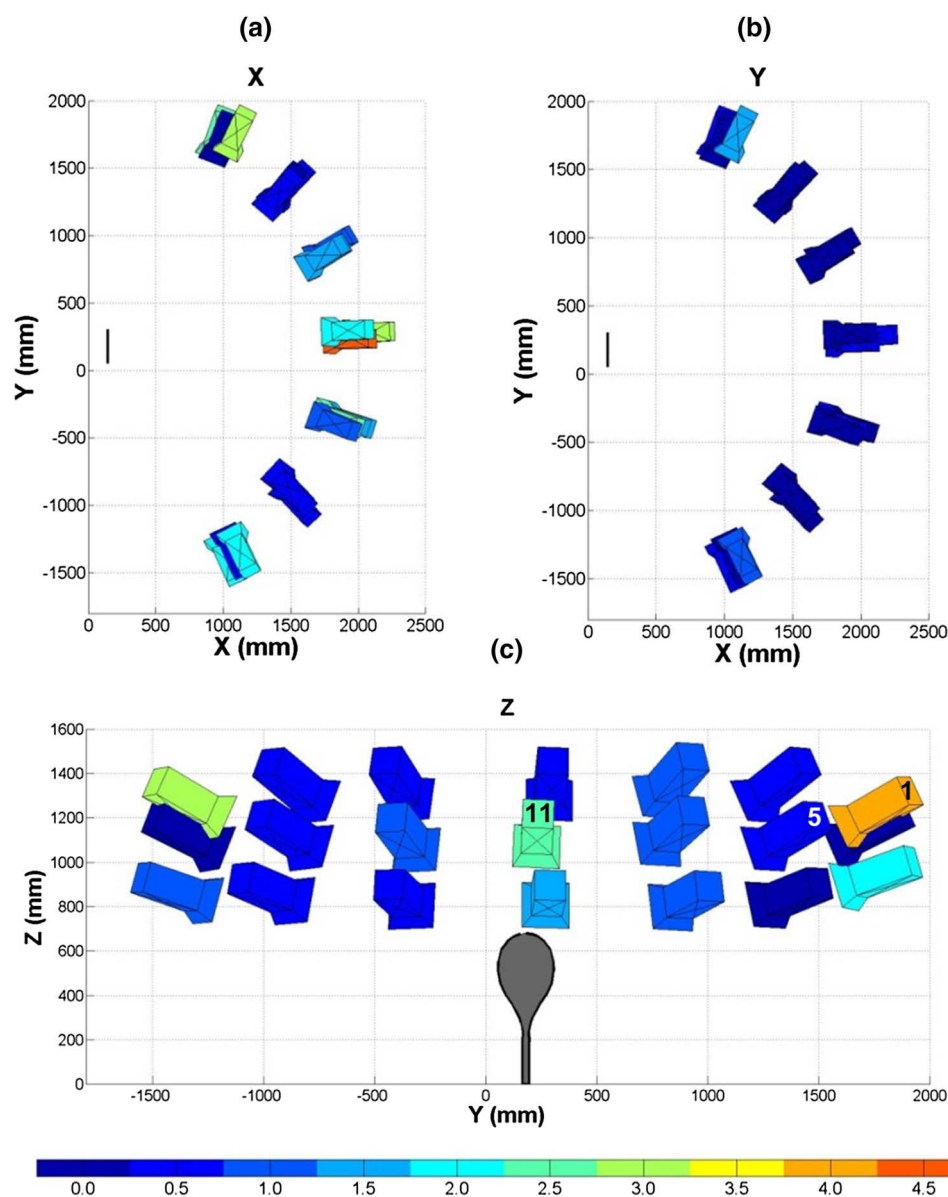
silhouette fitting techniques are promising, it is not possible to compare these directly with the values reported by Choppin and colleagues [7]. They reconstructed markers on the racket in a much less constrained scenario during a practice session at the Wimbledon qualifying tournament, with the racket up to 14 m from the camera. In the current study, reconstruction was performed in the laboratory from ~ 1.5 m from the racket and position estimation uncertainty may increase as the camera moves further away.

To trial the method before application to real-play conditions, the 3D position of a full size model racket was estimated from computer generated silhouette views in Blender (v2.70) [33] captured from a camera distance of 14 m. On average, 3D racket position was reconstructed to within 1.96 ± 0.14 mm during a simplified

simulated serve, comparable with measurements obtained in the laboratory, although the computer generated set contained no calibration error. All reconstruction errors were lower than the accuracy criteria of 15 ± 10 mm achieved by the markerless method developed by Corazza et al. [17], which was used by Sheets et al. [18] and Abrams et al. [20] to measure tennis serve kinematics in practice conditions.

The previous studies have reported racket head velocities of 34.8 m s^{-1} [45], 38.6 m s^{-1} [46], 43.2 m s^{-1} [47], 46 m s^{-1} [48], and 26.1 m s^{-1} [18], during a flat serve. The lower velocity reported by [18] is because the centre of volume rather than the tip of the racket was tracked. A spacing of 0.005 s between frames coupled with special data smoothing procedures [48, 51] is often used to track
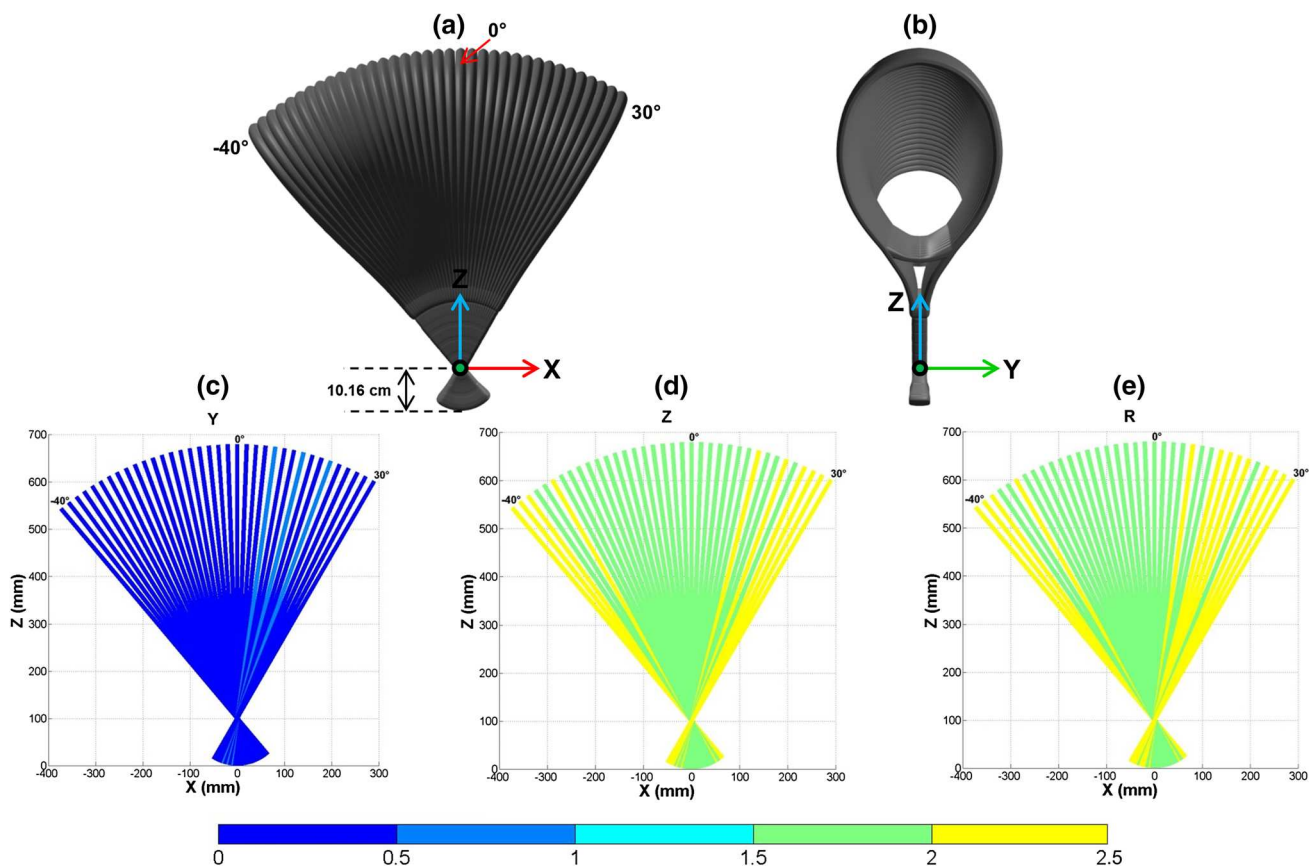
Fig. 6 a Side and **b** frontal views of the racket model when rotating about an axis 10.16 cm (4 inches) from the butt and reconstruction error in the **c** Y, **d** Z, and **e** resultant directions obtained using a calibrated set orientated randomly, during a simplified simulated serve

racket velocity around impact [18, 45, 46]. In the current study, the tip of the racket model was displaced by a resultant distance of 4 mm between each 2° rotation about the axis 10.16 cm from the butt (Fig. 6). This level of accuracy can be expected to translate to a real serve with a racket tip velocity ranging between 35 and 46 m s^{-1} captured with a camera operating with at least 200 fps. The results of the current study regarding application of a view fitting method to estimate 3D racket position are, however, limited to a simplified simulated serve movement, computer generated data, and silhouettes which were not extracted from noisy backgrounds in the field.

As robust silhouette extraction is important for obtaining accurate pose estimates with single view fitting techniques, images should be obtained at the highest possible resolution when generating a calibrated set [34, 41]. In the current study, racket silhouette images for the calibrated set were obtained using a white background to simplify boundary extraction. A robust method for digitally extracting racket silhouettes during tennis strokes is now required for view fitting against a calibrated set generated in a controlled environment. Extracting racket silhouettes from a tennis stroke poses a particular challenge due to

occlusion of the handle by the hand and the noisy background.

5 Conclusions

A markerless method capable of accurately estimating 3D racket position with one camera under controlled conditions has been presented. Development of a calibration method provided the relative pose of a camera with respect to a racket which was used to create a laboratory-based calibrated silhouette set. The set consisted of 21 camera poses in a semispherical configuration and its inconsistency was less than 0.1% of the mean length of the racket in a silhouette image. Using this set, mean racket position was reconstructed to within ±2 mm. Tests with computer generated camera poses and silhouette views allowed 3D racket position to be estimated during a simplified serve scenario. From a camera distance of 14 m, 3D racket position was estimated providing a spatial accuracy of 1.9 ± 0.14 mm. Further work will focus on developing the techniques presented here to measure 3D racket movements during play conditions. Thereafter, combining them

with single camera, ball tracking software could make for a useful tool to track racket motion for performance evaluations in research and for application with coaches and players.

References

1. Haake SJ, Allen TB, Choppin S, Goodwill SR (2007) The evolution of the tennis racket and its effect on serve speed. In: Miller S, Capel-Davies J (eds) Tennis science and technology 3. International Tennis Federation, London, pp 257–271
2. ITF (2016) 2016 ITF rules of tennis. http://www.itftennis.com/officiating/rulebooks/rules-of-tennis.aspx. Accessed 22 June 2016
3. Allen T, Choppin S, Knudson D (2016) A review of tennis racket performance parameters. Sports Eng 19(1):1–11
4. Mitchell SR, Jones R, King M (2000) Head speed vs. racket inertia in the tennis serve. Sports Eng 3(2):99–110
5. Brody H (2000) Player sensitivity to the moments of inertia of a tennis racket. Sports Eng 3(2):145–148
6. Lyu B, Smith L (2017) Evaluation of wireless bat swing speed sensors. Sports Eng. doi:10.1007/s12283-017-0224-3:1-6
7. Choppin S, Goodwill SR, Haake SJ (2011) Impact characteristics of the ball and racket during play at the Wimbledon qualifying tournament. Sports Eng 13(4):163–170
8. Knudson D, Bahamonde R (1999) Trunk and racket kinematics at impact in the open and square stance tennis forehand. Biol Sport 16(1):3–10
9. Bahamonde R, Knudson D (2003) Kinetics of the upper extremity in the open and square stance tennis forehand. J Sci Med Sport 6(1):88–101
10. Pingali GS, Jean Y, Carlbom I (1998) Real time tracking for enhanced tennis broadcasts. In: Computer vision and pattern recognition, pp 260–265
11. Yan F, Christmas W, Kittler J (2005) A tennis ball tracking algorithm for automatic annotation of a tennis match. British machine vision conference, pp 619–628
12. Kelley J, Choppin SB, Goodwill SR, Haake SJ (2010) Validation of a live, automatic ball velocity and spin rate finder in tennis. In: Sabo A, Kafka P, Litzenberger S, Sabo, C (eds) The Engineering of Sport 8. Proced Eng 2(2), pp 2967–2972
13. Dunn M, Haake S, Wheat J, Goodwill SR (2014) Validation of a single camera, spatio-temporal gait analysis system. Proced Eng 72:243–248
14. Laurentini A (1994) The visual hull concept for silhouette-based image understanding. IEEE Trans Pattern Anal Mach Intell 16(2):150–162
15. Mündermann L, Corazza S, Andriacchi TP (2006) The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. J NeuroEng Rehabil 3(1):1
16. Corazza S, Mündermann L, Chaudhari AM, Demattio T, Cobelli C, Andriacchi TP (2006) A markerless motion capture system to study musculoskeletal biomechanics: visual hull and simulated annealing approach. Ann Biomed Eng 34(6):1019–1029
17. Corazza S, Mündermann L, Gambaretto E, Ferrigno G, Andriacchi TP (2010) Markerless motion capture through visual hull, articulated icp and subject specific model generation. Int J Comput Vision 87(1–2):156–169
18. Sheets AL, Abrams GD, Corazza S, Safran MR, Andriacchi TP (2011) Kinematics differences between the flat, kick, and slice serves measured using a markerless motion capture method. Ann Biomed Eng 39(12):3011–3020
19. Ceseracciu E, Sawacha Z, Fantozzi S, Cortesi M, Gatta G, Corazza S, Cobelli C (2011) Markerless analysis of front crawl swimming. J Biomech 44(12):2236–2242
20. Abrams GD, Harris AH, Andriacchi TP, Safran MR (2014) Biomechanical analysis of three tennis serve types using a markerless system. Br J Sports Med 48(4):339–342
21. Price M, Morrison G (2007) Estimating 3D particle motion from high-speed video for simulation validation. Eng Comput 26(6):658–672
22. Forbes K, Voigt A, Bodika N (2003) Using silhouette consistency constraints to build 3D models. In: Proceedings of the fourteenth annual symposium of the pattern recognition association of South Africa (PRASA)
23. Wong K-YK (2001) Structure and motion from silhouettes. PhD, University of Cambridge
24. Bradski G, Kaehler A (2008) Learning OpenCV: computer vision with the OpenCV Library. O'Reilly Media, San Mateo
25. Elliott N, Choppin S, Goodwill S, Allen T (2014) Markerless tracking of tennis racket motion using a camera. In: James D, Choppin S, Allen T, Wheat J, Fleming P (eds) The Engineering of Sport 10. Proced Eng 10(72):344–349
26. Azevedo TC, Tavares JMR, Vaz MA (2009) 3D object reconstruction from uncalibrated images using a single off-the-shelf camera. Adv Comput Vis Med Image Process 13:117–136
27. Tavares JMR, Azevedo TC, Vaz MA (2008) 3D reconstruction of external anatomical structures using computer vision. In: Proceedings of the 8th international symposium on computer methods in biomechanics and biomedical engineering – CMBBE 2008. Porto, Portugal
28. Zhang Z (1999) Flexible camera calibration by viewing a plane from unknown orientations. In: proceedings of the 17th international Conference on Computer Vision. Corfu, Greece, pp 666–673
29. Allen T, Haake S, Goodwill SR (2009) Comparison of a finite element model of a tennis racket to experimental data. Sports Eng 12(2):87–98
30. Choppin S, Goodwill S, Haake S (2010) Investigations into the effect of grip tightness on off-centre forehand strikes in tennis. Proc Inst Mech Eng Part P: J Sports Eng Technol 224(4):249–257
31. Driscoll H, Kelley J, Kirk B, Koerger H, Haake S (2014) Measurement of shoe-surface interaction metrics during in situ performance analysis. Sports Eng 18(2):105–113
32. MATHWORKS (2014). MATLAB (2014b). http://uk.mathworks.com/help/optim/ug/lsqnonlin.html. Accessed 29 June 2016
33. Blender v2.70 (2014). https://www.blender.org. Accessed on 30 June 2016
34. Elliott N (2015) Camera calibration and configuration for estimation of tennis racket position in 3D. PhD, Sheffield Hallam University
35. Kelley J (2014) A camera calibration method for a hammer throw analysis tool. In: James D, Choppin S, Allen T, Wheat J, Fleming, P (eds) The engineering of sport 10. Proced Eng 10(72):74–79
36. Bouguet JY (2010) Camera calibration toolbox for MATLAB. http://www.vision.caltech.edu/bouguetj/calib_doc/. Accessed 10 Nov 2011
37. Matusik W, Buehler C, McMillan L (2001) Polyhedral visual hulls for real-time rendering. In: Proceedings of 12th Eurographics Workshop on Rendering, June 2001, pp 115–125

38. Lazebnik S (2002) Projective visual hulls. Master's Thesis, University of Illinois at Urbana-Champaign
39. Lazebnik S, Boyer E, Ponce J (2001) On computing exact visual hulls of solids bounded by smooth surfaces. Proc IEEE Comput Soc Confer Comput Vis Pattern Recogn 1:1–156
40. Lazebnik S, Furukawa Y, Ponce J (2007) Projective visual hulls. Int J Comput Vision 74(2):137–165
41. Forbes K (2007) Calibration, recognition, and shape from silhouettes of stones. PhD, University of Cape Town
42. Dunn M, Wheat J, Miller S, Haake S (2012) Reconstructing 2D planar coordinates using linear and non-linear techniques. In: Proceedings of the 30th international conference on biomechanics in sports, pp 381–383
43. Dunn M (2014) Video-based step measurement in sport and daily living. PhD, Sheffield Hallam University
44. Cross R, Bower R (2006) Effects of swing-weight on swing speed and racket power. J Sports Sci 24(1):23–30
45. Elliott B, Marsh T, Blanksby B (1986) A three-dimensional cinematographic analysis of the tennis serve. Int J Sport Biomech 2(4):260–271
46. Chow J, Carlton L, Lim YT, Chae WS, Shim JH, Kuenster A, Kokubun K (2003) Comparing the pre-and post-impact ball and racquet kinematics of elite tennis players' first and second serves: a preliminary study. J Sports Sci 21(7):529–537
47. Reid M, Elliott B, Alderson J (2007) Shoulder joint loading in the high performance flat and kick tennis serves. Br J Sports Med 41(12):884–889
48. Reid MM, Campbell AC, Elliott BC (2012) Comparison of endpoint data treatment methods for the estimation of kinematics and kinetics near impact during the tennis serve. J Appl Biomech 28(1):93–98
49. Wong K-YK, Cipolla R (2004) Reconstruction of sculpture from its profiles with unknown camera positions. IEEE Trans Image Process 13(3):381–389
50. Choppin S (2008) Modelling of tennis racket impacts in 3D using elite players. PhD, University of Sheffield
51. Knudson D, Bahamonde R (2001) Effect of endpoint conditions on position and velocity near impact in tennis. J Sports Sci 19(11):839–844