The accuracy of badminton player tracking using a depth camera

Jonathan Wheat[a,*], Ben Lane[a], Simon Choppin[a]

[a] *Centre for Sports Engineering Research, Collegiate Hall, Collegiate Campus, Sheffield Hallam University, Sheffield, S10 2BP*

Abstract

In notational analysis player position is often used to infer performance metrics. As an alternative to manual methods of player tracking – which are costly and time consuming – automatic camera-based systems have been used. However, distinguishing the player from their surroundings is difficult to achieve using only colour/intensity information. Depth cameras improve player segmentation by creating a clear separation between player and environment. Our aim was to investigate the suitability of a consumer depth camera – the Microsoft Kinect – to track player position in a Badminton context. In one experiment, tracking accuracy was assessed in a representative, on-court environment. Player position was determined using the depth camera with the participant stood at known locations, and moving through known trajectories (marked on the floor of the court). In the second experiment, player position estimated using the depth camera was compared to that determined using a three-dimensional motion capture system, during simulated badminton play in a laboratory environment. In both experiments, the depth camera system tracked the location of the player with an accuracy that compares favourably with other automatic/semi-automatic tracking systems – RMSE of 25 cm on position and 0.21 ms^{-1} on velocity. Depth data makes automatic algorithms simple and robust compared to traditional video camera based tracking algorithms. Further work is required to develop algorithms for multi-person tracking – in badminton doubles or squash play, for example.

Keywords: person tracking, Kinect, validity, badminton

1. Introduction

Determining the location of people in a three-dimensional (3D) space is important in sport and in monitoring activities of daily living. For example, 3D position has been used to detect falls in older people

*Corresponding author, telephone: +44 (0) 114 225 4330
Email addresses: j.wheat@shu.ac.uk (Jonathan Wheat), s.choppin@shu.ac.uk (Simon Choppin)

(Rougier *et al.*, 2011). In sports applications location data must be captured quickly and accurately to be useful. This information is used to monitor distances, speeds and positions at specific points during play. As such, quantitative analysis of player activity is now considered an important aspect of coaching in sport (Barris and Button, 2008). Tracking players in racket sports such as squash, tennis and badminton has received attention (Han and de With, 2007); fewer players and a restricted playing area make player tracking systems particularly suited to these sports.

Existing player tracking systems are primarily based on one or more video cameras (Pingali *et al.*, 1998; Mauthner *et al.*, 2007; Xing *et al.*, 2011), although GPS (or similar) based methods have also been used (Portas *et al.*, 2007; Pino *et al.*, 2007). An example of a video-based system is the SAGIT/squash computer tracking software. A single video camera – with a wide angle lens – is mounted on a gantry above the squash court, with the entire court in the field of view. After background subtraction, algorithms extract each player from the image of the court and an approximation of their position in the floor plane is tracked in each video frame. A review of vision based motion analysis in sport is provided by Barris and Button (2008). GPS systems rely on sensors – which must be attached to skin or clothing – to track position. GPS sensors are reliant on satellite signals and have low spatial resolution, rendering them unsuitable for use indoors or over small areas. As such, GPS tracking studies have focused on sports such as football (Portas *et al.*, 2007; Pino *et al.*, 2007) and skiing (Seifriz *et al.*, 2002). Body-worn sensors, however, might be prohibited during competition.

Papers that detail person tracking techniques rarely present details of tracking accuracy with respect to a ground truth measure – although video and GPS systems have been directly compared (Edgecomb and Norton, 2006). Often, a technique is deemed successful if the algorithms do not fail and are robust in the tracking of individuals. However, Vučković *et al.* (2010) quantified the accuracy of the SAGIT/Squash system. The authors demonstrated what they deemed to be acceptable accuracy by having participants stand at known locations on the court and move through predetermined movement paths. Error in distance covered by the player in one minute ranged between 1.33 m and 21 m (around 10%), depending on the position and nature of the movements being performed (E.g. Whether or not the racket was swung around the player).

Video-based methods employ computer vision algorithms in order to detect people within a scene. Although many methods succeed in detection, video image analysis is complex; it is dependent on, for example, ambient lighting and camera position. Furthermore, estimates of position are often given in only two dimensions, with points normally projected onto the floor plane.

The Microsoft Kinect is a consumer depth camera (originally designed for gesture recognition in computer games) which has found favour with researchers from many different fields (for example, physiotherapy, robotics and virtual reality (Clark *et al.*, 2012; Stowers *et al.*, 2011; Rydén *et al.*, 2011)). The depth data it provides – through the use of pseudo structured light techniques (Scharstein and Szeliski, 2003) – can be used in combination with its colour video camera to identify the location of a person's external boundary in 3D (Mirante *et al.*, 2011). The ability to capture position in three dimensions is an advantage in sports where vertical position of the player changes significantly, for example, in badminton the player may jump to smash shots. The strong edge given by the depth camera serves for person detection while colour information from the video camera could assist in person identification (using clothing colour, for example). However, the limited range of the depth camera (0.8 - 10 m) means its player tracking capabilities (as a single camera) are limited to sports such as squash, badminton or, to an extent, tennis.

The aim of this study was to investigate the accuracy with which a depth camera (Microsoft Kinect) could be used to automatically track a player during simulated badminton singles play. Two experiments were conducted. The first explored the accuracy with which centre-of-mass (COM) measurements could be taken in a representative, on-court setting – replicating many of the procedures implemented by Vučković *et al.* (2010). The second – performed in a laboratory – compared the 3D location of the participant's COM measured using a Kinect and a 3D motion capture system (taken as a gold-standard).

2. Methods

2.1. Participant

One male, recreational badminton player volunteered to participate in the study and provided written informed consent before data collection began. The participant was free from any neuromuscular/skeletal injury that might affect completion of the protocol. The Faculty of Heath and Wellbeing Research Ethics committee approved all procedures.

2.2. On-court testing

2.2.1. Experimental Setup

A standard badminton court in the University sports hall was used for data collection. A Kinect was placed on a tripod at a height of approximately 1.5m, 2.3 m behind the back court line (see Figure 1). A distance of 2.3 m behind the back line was chosen as a compromise between the Kinect's field of view and the resolution of its depth data – which decreases with distance. Furthermore, the position complies with the Badminton World Federation's recommendation that there is at least 2 m clear space surrounding the outer lines of the court (BWF, 2011). Additional markings were added to the court using tape: a 5×5 grid of points (equally distributed over the full width and length of the court) and a 2 m diameter circle at the geometric centre of the court.

2.2.2. Procedure

The participant completed two tests. During the first, they stood on each of the grid points for one minute. Participants were instructed to stand as still as possible, facing away from the Kinect, with feet approximately shoulder width apart and with the grid point at the centre of stance. The second test was a dynamic trial in which participants moved around a circular trajectory marked on the floor of the court. Starting from a position closest to the Kinect, participants completed laps of the circle at three self-selected speeds: walking, running and sprinting. Three laps of the circle were recorded at each speed.

2.2.3. Global axes orientation

A court-fixed coordinate system was defined with the $x - axis$ aligned to the length of the court, $y - axis$ aligned to the width, and $z - axis$ aligned vertically, the origin of the system was at the centre of the baseline. To correctly orientate the axes, a manually selected region of the floor was used to define a plane. A principal component analysis (PCA, see Daffertshofer *et al.* (2004) for a description) was used to define three orthogonal axes. The $x-$ and $y - axes$ lay in the floor plane and the $z - axis$ was perpendicular. Correct alignment of the co-ordinate system was achieved by visually aligning the $x-$ and $y - axes$ with court markings (through rotation about the $z - axis$ and translation in the floor plane). Custom-written code enabled us to perform this manually using recorded Kinect video frames.

2.2.4. Data Analysis

For each frame of all trials, the Kinect returned a three-dimensional point cloud representing the scene in the field-of-view. We transformed the point cloud into the court-fixed coordinate system and segmented the participant by assuming that any point 200 mm above the floor (the vertical z component) belonged to the participant. A threshold of 200 mm was used to avoid the influence of noisy points from the floor, sometimes present at large distances (∼8 m) from the kinect. The position of the participant was determined by calculating the 'centre of mass' (COM) of the resulting point cloud; each three-dimensional point was assigned an equal, nominal, mass.

Although the Kinect determined the three-dimensional position of the participant, our reference locations were in the plane of the floor. As such, the position of the participant's COM was projected onto the floor plane ($x - y$ plane of the court coordinate system)by dismissing the value of z. When analysing the position of the participant at each grid location (see Figure 1), we calculated the root mean square error (RMSE) between the participant's location and the reference location using equation 1:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2} \tag{1}$$

where N is the number of data points, $\hat{x}$ and $\hat{y}$ are the estimated positions, and x and y are the reference positions.

For the trials in which the participant moved along a circular trajectory, we assessed agreement between the participant's movement and the reference trajectory using equation 2.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{r}_i - r)^2} \tag{2}$$

where N is the number of data points, $\hat{r}$ is the estimated radial distance and r is the radius of the circle marked on the floor of the court (2 m).

2.3. *Laboratory based testing*

2.3.1. *Experimental Setup*

During all trials, data were collected concurrently from a Kinect and a 12 digital-camera motion capture system (MAC: Motion Analysis Corporation, Santa Rosa, CA, USA). The MAC cameras were positioned around a 4 m x 3 m x 2.5 m measurement volume and calibrated in accordance with the manufacturer's specifications. RMS error of the MAC system is reported at < 2 mm for moving participants (Richards, 1999).

To measure COM position with the MAC, markers were attached to the participant at relevant anatomical landmarks. In a static trial, 43 retro-reflective markers were used; after this step 10 markers were removed for ensuing movement trials.

The sampling frequency of the MAC was 300 Hz; during all trials, the sampling frequency of the Kinect was approximately 30 Hz. The Kinect and MAC were event synchronised using two simultaneous events: The illumination of a red LED – visible in the Kinect colour camera – and the generation of a +5 V signal – an analogue input to the MAC. A button press triggered both events which were used to time-align the systems.

All testing was performed within an area equivalent to one eighth of a standard badminton court – chosen due the size of the available testing space. The testing region was marked out on the laboratory floor within the motion capture system's calibrated measurement volume.

The testing area served as each of the four quarters of *one half* of a badminton court: front forehand, back forehand, front backhand and back backhand. Each was tested in turn and the Kinect was positioned to account for the quarter's relative position on the court. For a back quarter, the Kinect was positioned 2.3 m away from the back of the testing area, for a front quarter, the Kinect was positioned 5.65 m away from the back of the testing area. For a forehand quarter the Kinect was aligned with the left-hand extremity of the testing area, for a backhand quarter the Kinect was aligned with the right-hand extremity of the testing area. This arrangement ensured that the Kinect was always located the equivalent of 2.3 m behind the middle of the court's back line.

2.3.2. Global axes orientations

A court-fixed coordinate system, common to the MAC and Kinect was defined with the $x-axis$ aligned to the length of the court, $y-axis$ aligned to the width, and $z-axis$ aligned vertically, the origin of the system was at the centre of the baseline. With the MAC system, three retro-reflective markers were placed on the floor in an 'L' arrangement; the position of the origin was then translated appropriately using Badminton court dimensions. For the Kinect, the previously described method (for on court testing) was used.

2.3.3. Procedure

At the start of procedures we gave the participant time to read the information sheet, ask any questions, and sign the informed consent form. Participants changed into tight fitting Lycra shorts and a t-shirt and the retro-reflective markers were attached to the body using double sided tape. Once prepared the participant stood in the anatomical position for a static calibration trial lasting approximately two seconds. Subsequently, four movement trials were collected.

In each trial the participant completed a badminton specific movement drill in the testing area, with movements adapted from those presented by Ooi *et al.* (2009). The participant started in a corner of the testing area – equivalent to the court centre – and moved to a point near the perimeter before returning to their start point. These intermediate points were: Three corners of the testing area and three mid-points on the testing area perimeter.

The four court quarters were tested in the following order: Front forehand, Front backhand, Back backhand and Back forehand. Movements in each quarter were separated with a rest period of approximately one minute.

2.3.4. Data analysis

Raw marker coordinate data from the motion capture system were filtered using a fourth order, zero-lag, low-pass Butterworth filter at 10 Hz. We used the default geometric model in Visual3D (C-Motion, Germantown, MD, USA) to estimate whole-body COM location – the method is similar to that presented by Hanavan (1964). Truncated cones were used to model every segment but the head, which was modelled as an ellipsoid. After assuming uniform density within each body segment (Dempster, 1955), segment mass and COM position were estimated for each segment and therefore the whole-body. The position of the whole-body COM in the court-fixed co-ordinate system was taken as the ground truth location.

An estimate of the location of the participant's COM in the court-fixed coordinate system was also made using the Kinect. Three-dimensional points that lay within a bounding box (coincident with the measurement volume of the motion capture system) were assumed to represent the participant – the minimum height of the bounding box was set to 400 mm above the floor to avoid noisy points from the floor. This threshold was twice that used for on-court testing. It was increased due to the higher level of observed noise from the floor – it wasn't clear why noise was greater in this instance. Each three-dimensional point was assigned an equal, nominal, mass and the centre of mass of the points was assumed to represent the centre of mass of the participant.

The 300 Hz COM data from the MAC system were re-sampled to the same time points as the 30 Hz Kinect data using cubic spline interpolation (the sampling frequency of the Kinect was sometimes erratic). We calculated the root mean square error (RMSE) between the Kinect and ground truth for speed and position (x, y and z) of the participant's COM. This was done for each court region. In addition, we compared estimates of the total distance moved by the participant during each movement drill – the position data from the Kinect were pre-filtered using a simple moving average filter (with a 10 point radius).

3. Results

3.1. On-court testing

For static estimations of position (according to grid positions) the error in Kinect estimation increased with distance from the back court line (Figure 1), with a minimum error of 19 mm (Euclidean) and maximum error of 395 mm. As the player stood at each position for a minute, the standard deviation of measured position was calculated in each case. This was lowest (30 mm) at the back court line and increased to a maximum (577 mm) towards the front right corner of the court (Figure 1). Four grid points could not be resolved due to the field of view of the Kinect, these were the two points, closest to the Kinect, on the left and right court boundaries.

[Figure 1 about here.]

In the moving trial (circular trajectory), RMSE deviation from the circle was measured as 196 mm, 252 mm and 276 mm in the walking, running and sprinting trials, respectively. The recorded traces in each case are shown in Figure 2.

[Figure 2 about here.]

3.2. Laboratory testing

Table 1 contains the RMSEs for participant COM position and speed compared to the MAC system in addition to the total distance moved by the participant during each trial, as given by the MAC and Kinect. A comparison of participant COM location in the floor plane, derived from the Kinect and MAC systems is presented in Figure 3.

[Table 1 about here.]

[Figure 3 about here.]

Figure 3 shows regions in the 'back' portions of the court where Kinect tracking has failed – due to the player being outside of the view range. These data were excluded when calculating given RMS values.

4. Discussion

We investigated how accurately the Microsoft Kinect can determine the position of a person in three dimensional space. As a relevant example situation, we chose to track the on-court movements of badminton players. However, there are many more equally important contexts, such as monitoring the location of an older person in a residential home. Badminton is a very dynamic activity and presents challenging conditions in which to track the location of a player. The experiments reported in this document investigated the accuracy of the Kinect-based system in two experimental settings: in the lab and in the 'field' (on a badminton court).

It was difficult to define a 'true' player position in the – more representative – badminton court setting. However, our approach is similar to that employed by Vučković *et al.* (2010) who investigated the accuracy of the SAGIT system in squash play. The lab-based testing enabled the use of an accurate 'gold standard' player position; determined using a complex, expensive three-dimensional motion capture system. We could find no studies in the literature which have used this approach to investigating the accuracy of person tracking algorithms. Agreement between the measurement systems in estimating participant position in the x and y directions (in the floor plane) was generally good (Table 1). Results of the lab based testing also indicated good agreement between the motion capture and Kinect-based systems in estimating the speed of movement of, and total distance travelled by, the participant Table 1.

Estimates of participant location using the Kinect were best when the participant was less than 6 m away. The results at these distances compare favourably with the results presented by Vučković *et al.* (2010); they presented a lower bound on error of 110 mm (in the centre of the field of view)and an upper bound on error of 420 mm (at the extremities of the field of view). In the present study RMSE errors ranged between 64 mm and 190 mm at distances below 6m . Errors in participant position were larger – RMSE 85-280 mm – for locations further away from the Kinect – distances greater than 6 m . The Kinect's limited resolution at these depths is a likely cause of this increased error.

Depth values for the Kinect are returned in 'bins' on an 11-bit scale – 2048 possible depth values. The real world distance between bins increases at larger depths (further away from the Kinect). Close to the Kinect ($\sim$1 m) the distance between bins is approximately 1 mm, towards the extremities of the depth range ($>$ 9 m) the distance between bins is $>$ 0.1 m.

Establishing the required level of accuracy from a tracking system of this nature depends on the desired application from the coach or athlete. Error on instantaneous position affects the level of measurement precision when investigating – for example – where a player should be on court at a particular time. If locating the player within a particular region of the court (a quarter etc.) is sufficient, then the current system demonstrates appropriate accuracy with RMSE of position within 280 mm in the worst cases. RMSE error in instantaneous velocity was reported as 0.21 ms^{-1} in the worst case, around 7% of maximum. Error in total distance covered was 1.4 m over 19 m in the worst case, 7.6%. These errors compare favourably with the data presented by Vučković *et al.* (2010).

When a squash player moved around a circular reference trajectory (radius 2.7 m) Vuckovic et al. reported errors of 170 mm, 320 mm and 500 mm for walking, running and sprinting respectively. In the current study, RMSE errors were slightly greater for walking (196 mm) but lower for the two running speeds (252 mm for running and 276 mm for sprinting). Other studies which report the magnitude of tracking errors include an investigation of beach vollyball players (Mauthner *et al.*, 2007) and a 2 camera tracking system in handball (Per *et al.*, 2002). Marginally larger errors were reported by (Mauthner *et al.*, 2007) - approximately 225-350 mm - than those presented in the present study, again indicating the quality of the Kinect-based tracking system. Per *et al.* (2002) reported errors in velocity of 0.07 to 0.35 ms^{-1} depending on smoothing. Errors in position of 0.36 m were reported. Again, the values reported in this study compare favourably.

When moving in the circular path, the radius of the participant's trajectory was systematically smaller than the reference trajectory (Figure 2) – similar to the results presented by Vučković *et al.* (2010). This was caused by the participant leaning their upper body towards the centre of the circular path. This issue becomes more pronounced at faster speeds (Figure 2), as would be expected.

Unlike previous tracking systems reported in the literature (Vučković *et al.*, 2010; Mauthner *et al.*, 2007; Per *et al.*, 2002) which report only two-dimensional position of a participant, the Kinect-based tracking system estimates three-dimensional position. This is important in analysing activities such as badminton in which a great deal of movement takes place in the vertical direction, with players jumping to hit high, and lunging to hit low. Furthermore, movements in the vertical direction can affect the accuracy of some two-dimensional tracking systems; for example in situations where the camera used for tracking is not perpendicular to the floor/ground plane (Mauthner *et al.*, 2007). The RMSEs presented in Table 1 suggest good agreement between the Kinect-based tracking system and the motion capture system in the vertical direction (z). However, accounting for the range of movement of the participant in the vertical direction, tracking was quite poor. Tracking in the vertical direction was more accurate in the front court quadrants – when the participant was a relatively large distance away from the Kinect – the opposite of what might be expected. A simple tracking algorithm was used in this study that did not differentiate between the player and the racket. When the participant was relatively close to the Kinect, the racket appeared in the point cloud data, affecting the centre of mass calculation. The racket was tracked less often when testing the back of the court and may explain why vertical position was estimated more accurately in these quadrants.

The view range of the Kinect resulted in two 'black spots' of tracking – located in the back corners of the court. Extending the range of the current set-up by moving the Kinect further backwards would further degrade the data at the front of the court and may corrupt the data to a greater extent than currently. A suitable solution to this problem may result from further developments in hardware, from a system with a wider view range and equally low cost.

The algorithms used for estimating the location of the badminton player were deliberately simple. Even with the simple tracking algorithm, the participant was *never* misidentified when the participant was in the Kinect's field-of-view (all tracking was automatic without the requirement for any manual intervention). Improvements could be made by, for example, developing algorithms to ignore the racket and employing more sophisticated segmentation methodologies (Mirante *et al.*, 2011) to track two players in badminton doubles play. The added depth information provided by the Kinect makes tracking multiple participants simpler than traditional two-dimensional image based analysis.

It is important to note that, in this study the environment was free from objects that might obscure the field-of-view of the Kinect. For other person tracking applications, e.g. tracking people in the home, such objects might be present. More work is required to explore the difficulties associated with tracking people in more cluttered environments.

In conclusion, the data presented here suggest that the depth camera based system (Microsoft Kinect) can be used to track the location of a badminton player. The accuracy of the system compares favourably with previous player tracking systems. Also, the depth data make automatic algorithms simple and robust compared with traditional video camera based tracking algorithms. Further work is required to develop algorithms for multi-person tracking in badminton doubles or squash play, for example. Work should also explore refining the tracking algorithms to enable more accurate tracking of the player in the vertical direction.

References

Barris, S. and Button, C., 2008. A review of vision-based motion analysis in sport., *Sports medicine (Auckland, N.Z.)*, 38 (12), 1025–43.

BWF, 2011. *BADMINTON WORLD FEDERATION BWF HANDBOOK II (Laws of Badminton & Regulations)*, Badminton World Federation.

Clark, R., Pua, Y., Fortin, K., Ritchie, C., Webster, K., Denehy, L., and Bryant, A., 2012. Validity of the microsoft kinect for assessment of postural control, *Gait & Posture*.

Daffertshofer, A., Lamoth, C.J.C., Meijer, O.G., and Beek, P.J., 2004. PCA in studying coordination and variability: a tutorial., *Clinical biomechanics (Bristol, Avon)*, 19 (4), 415–28.

Dempster, W., 1955. Space requirements of the seated operator, *in: WADC Tech. Report*, Dayton, OH.: Wright-Patterson Air Force Base, chap. Space requ, 55–159.

Edgecomb, S. and Norton, K., 2006. Comparison of global positioning and computer-based tracking systems for measuring player movement distance during australian football, *Journal of Science and Medicine in Sport*, 9 (1), 25–32.

Han, J. and de With, P., 2007. A unified and efficient framework for court-net sports video analysis using 3-d camera modeling, *in: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 6506, 15.

Hanavan, E., 1964. A mathematical model of the human body, Ohio: Aerospace Medical Division, 64–102.

Mauthner, T., Koch, C., Tilp, M., and Bischof, H., 2007. Visual Tracking of Athletes in Beach Volleyball Using a Single Camera, *Journal of Computer Science*, 6, 21–34.

Mirante, E., Georgiev, M., and Gotchev, A., 2011. A fast image segmentation algorithm using color and depth map, *in: 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2011*, IEEE, 1–4.

Ooi, C.H., Tan, A., Ahmad, A., Kwong, K.W., Sompong, R., Ghazali, K.A.M., Liew, S.L., Chai, W.J., and Thompson, M.W., 2009. Physiological characteristics of elite and sub-elite badminton players., *Journal of sports sciences*, 27 (14), 1591–9.

Per, J., Bon, M., Kovai, S., ibila, M., and Deman, B., 2002. Observation and analysis of large-scale human motion, *Human Movement Science*, 21 (2), 295 – 311.

Pingali, G., Jean, Y., and Carlbom, I., 1998. Real time tracking for enhanced tennis broadcasts, *in: Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on*, IEEE, 260–265.

Pino, J., Martinez-Santos, R., Moreno, M., and Padilla, C., 2007. Automatic analysis of football games using gps on real time, *Journal of Sports Science and Medicine*, 10.

Portas, M.D.M., Rush, C.J.C., Barnes, C.A.C., and Batterham, A.M.A., 2007. Method comparison of linear distance and velocity measurements with global positioning satellite (gps) and the timing gate techniques, *Journal of Sports Science and Medicine*, 6 (10).

Richards, J.G., 1999. The measurement of human motion: A comparison of commercially available systems, *Human Movement Science*, 18 (5), 589–602.

Rougier, C., Auvinet, E., Rousseau, J., Mignotte, M., and Meunier, J., 2011. Fall detection from depth map video sequences, *Toward Useful Services for Elderly and People with Disabilities*, 121–128.

Rydén, F., Chezeck, H., Kosari, S., King, H., and Hannaford, B., 2011. Using kinect and a haptic interface for implementation of real-time virtual fixtures, *in: Proceedings of the 2nd Workshop on RGB-D: Advanced Reasoning with Depth Cameras (in conjunction with RSS 2011)*.

Scharstein, D. and Szeliski, R., 2003. High-accuracy stereo depth maps using structured light, *in: Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, IEEE, vol. 1, 195–202.

Seifriz, F., Mester, J., Krmer, A., and Roth, R., 2002. The use of gps for continuous measurement of kinematic data and for the validation of a model in alpine skiing., *Int. J. Comp. Sci. Sport*, 1 (1).

Stowers, J., Hayes, M., and Bainbridge-Smith, A., 2011. Altitude control of a quadrotor helicopter using depth map from microsoft kinect sensor, *in: Mechatronics (ICM), 2011 IEEE International Conference on*, IEEE, 358–362.

Vučković, G., Perš, J., James, N., and Hughes, M., 2010. Measurement error associated with the SAGIT/Squash computer tracking software, *European Journal of Sport Science*, 10 (2), 129–140.

Xing, J., Ai, H., Liu, L., and Lao, S., 2011. Multiple player tracking in sports video: A dual-mode two-way bayesian inference approach with progressive observation modeling, *Image Processing, IEEE Transactions on*, 20 (6), 1652 –1667.

List of Figures

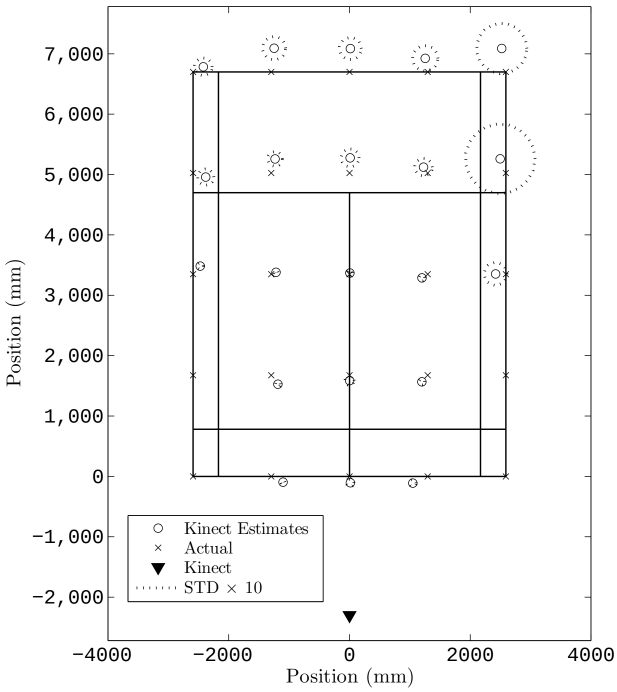

Figure 1: Positions of the participant estimated by the Kinect-based system at each of the predetermined reference grid positions on the badminton court. The standard deviation of prediction (× 10) is shown to illustrate the variation in prediction according to court position

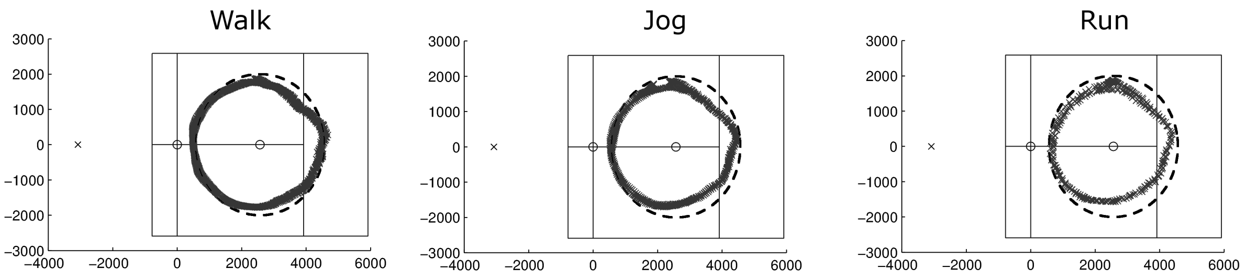

Figure 2: Circular trajectories measured using the Kinect with the participant moving at different speeds. The cross hairs indicate the trajectory of the player over three laps. The dotted line indicates the reference trajectory. The cross indicates the location of the Kinect

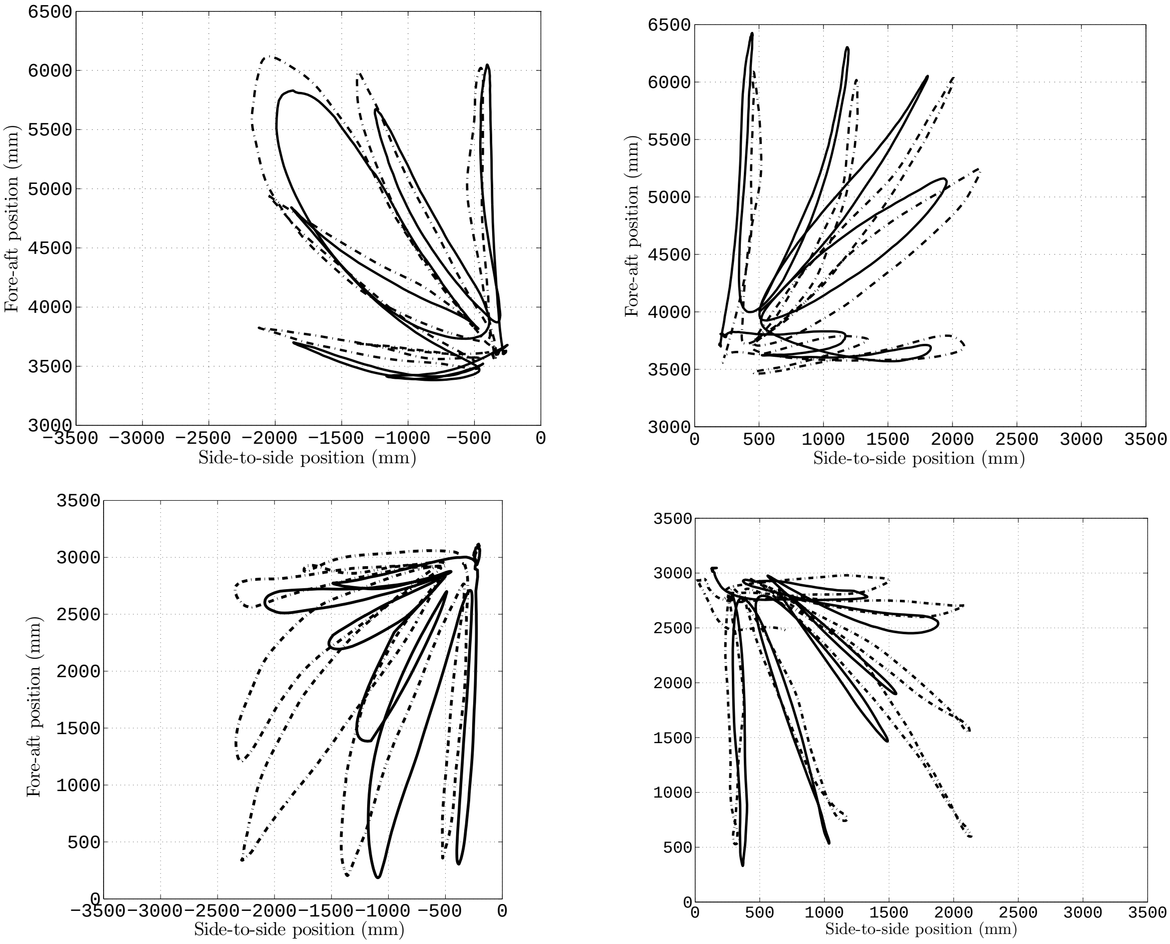

Figure 3: Centre of mass position the court floor plane. Top left: backhand front court, top right: forehand front court, bottom left: backhand back court, bottom right: forehand front court. Dashed line - MAC, Solid line - Kinect

List of Tables

Table 1: The distance travelled by the participant – estimated using the Kinect-based system and the MAC motion capture system– and the RMSE of COM position (cm) and speed (m/s)

Court Region	**Distance Travelled** (m)			**RMSE** (cm, ms^{-1})			
	Kinect	**MAC**	**Difference**	**X**	**Y**	**Z**	**Speed**
Front Forehand	21.4	21.2	0.2	25	9	6	0.21
Front Backhand	21.2	20.9	0.3	16	11	5	0.21
Back Forehand	17.6	19.0	-1.4	13	9	8	0.18
Back Backhand	18.5	19.5	-1.0	11	16	6	0.20