VISUALISATION OF TENNIS SWINGS FOR COACHING

Philip Kelly and Noel E. O'Connor

CLARITY: Centre for Sensor Web Technologies, Dublin City University, Ireland

ABSTRACT

As a proficient tennis swing is a key element of success in tennis, many amateur tennis players spend a considerable amount of time and effort perfecting their tennis stroke mechanics, hoping to create more accuracy, consistency and power in their swing. In order to achieve these aims effectively a number of independent aspects of technique need to be addressed, including forming a correct racket grip, shot timing, body orientation and precise follow–through. Outside of a one-to-one coaching scenario, where constant professional feedback on technique can be provided, keeping all aspects of technique in mind can overwhelm amateur players. In this work, we have developed a set of visualisation tools to augment the development of amateur tennis players between dedicated one-to-one coaching sessions in the area of technique, timing and body posture. Our approach temporally aligns an amateur player's swing dynamics with that of an elite athlete, allowing direct technique comparison using augmented reality techniques.

1. INTRODUCTION

Many amateur tennis players spend a considerable amount of time and effort attempting to perfect their tennis stroke mechanics. To date, the best way to achieve these aims is with practice and regular consultations with expert coaching specialists, who can provide explicit guidance on the specific features in an athlete's technique that require attention and alteration. However, keeping true to all aspects of technique learned between one-to-one coaching sessions can be difficult, especially for newcomers to the sport. In this work, we have developed visualisation tools that can augment the development of amateur tennis player technique, timing and body posture *between* dedicated one-to-one coaching sessions. Our approach, analyses the implicit movements of a test subject's tennis stroke using information gleamed from a single accelerometer, worn on a player's dominant forearm, and compares it to a database of temporally aligned similar shot motions from elite athletes. The motions from both the elite and amateur players are rendered using augmented reality visualisation techniques, allowing the amateur player's motion to be directly compared to the elite player's performance. Depending upon the alignment technique used, differences in either technique, ball contact point, timing or follow through can be visualised and compared.

The paper is organised as follows: Section 2 gives an overview of the system described in this work. A description of the data collection undertaken in this work is described in Section 3. Section 4 details three techniques for the temporal alignment of both amateur and elite player stroke motion dynamics. In section 5 qualitative and quantitative evaluation of the stroke alignment stage of this work is presented. Finally, conclusions and future work are outlined in section 6.

2. SYSTEM OVERVIEW

In collaboration with Tennis Ireland [1], we are striving to provide technological solutions to real problems encountered by tennis coaches. As part of this project, we have instrumented an indoor tennis court with a data-gathering infrastructure that includes nine networked cameras positioned around the court. In previous works we have used this infrastructure for research into automatic ball and player tracking [2], synchronisation [3], automatic camera selection [4], stroke detection [2] and stroke recognition [5].

In this work, we combine and extend a number of modules from these previous works into a single system aimed at augmenting the development of amateur tennis players with respect to technique, timing and body posture outside of dedicated one-to-one coaching sessions. The approach aims to achieve these goals via the use of augmented reality visualisation techniques, where the implicit movements of the amateur player are directly compared and visualised against those of an elite athlete performing a similar shot action. The examination of different aspects of technique can be achieved by varying the technique employed to temporally align the motion of amateur and elite players, as described in Section 4. An overview of our approach is now given.

Our system records an amateur player's game play using nine time-synchronised [3] IP cameras, plus a data stream from a single accelerometer worn on the player's dominant forearm. Tennis shots from a game are automatically detected and classified as either serves, forehands and backhands from the accelerometer readings using Connaghan's [5] machine learning approach. For each amateur shot of interest to be analysed, a search through a pre-captured database of similar shots performed by elite athletes is made. The database shot

This work is supported by the European Commission under contract FP7-247688 3DLife.

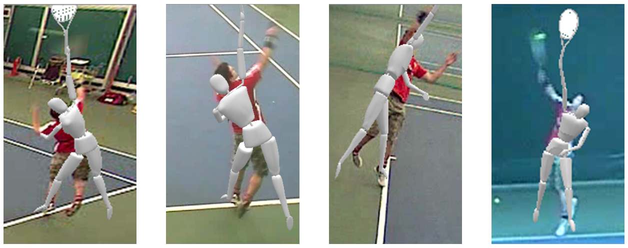


Fig. 1. Examination of posture differences during a tennis serve from four angles using augmented reality techniques.

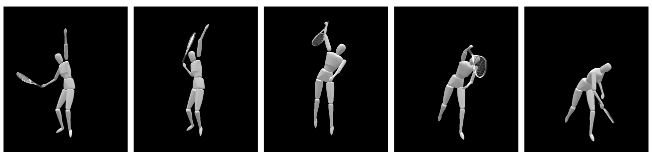


Fig. 2. Sample motion–capture frames of a tennis serve.

with the most similar inertial movements to the amateur's motion dynamics is selected, time aligned and rendered with the amateur player's motion from any of the camera viewpoints surrounding the court using augmented reality techniques.

Example system output is presented in Figure 1, where a player's posture during a serve stroke movement is examined. The amateur player's pose is shown from four of the camera court viewpoints, while the time aligned elite athlete pose is overlaid using a virtual 3D avatar. The avatar is correctly positioned on the court using the location information obtained from the automatic tracking of the amateur player via the approach of Ó Conaire [2]. The orientation heading of the avatar is set according to the ball trajectory [2] as it leaves the racket. Manual fine tuning of both the avatar's position and orientation can also be made. In Figure 1, differences between tennis serve posture can be seen, for example the amateur player extends and hits the ball when it is over his right shoulder, while the professional player hits the ball directly overhead. Using different temporal alignment methodologies, comparison of player timing or follow through can also be examined using similar visualisation techniques.

3. DATA CAPTURE

In order to compare the stroke movement dynamics of an arbitrary amateur player to that of high-skilled player, a database of elite player tennis shots needs to be captured. One approach to acquiring this data would be to capture each shot type in multiple court positions and orientations using the pre-existing camera infrastructure. However, this approach is hugely time consuming, as the elite player will need to perform each shot in *every* court position that will be plausibly required within the augmented reality sequences. In addition, if extra cameras are added, or the camera locations are altered, then the whole elite player data corpus would have to be re-captured. In this work, we take an alternative approach, capturing the 3D motion of a small number of elite athlete shots (20 serves, 43 forehands and 40 backhands) in a lab environment using a 12 camera 250 Hz Vicon infra-red motion capture system [6]. Using this approach, a 3D model of the elite player motion (such as the one presented in Figure 2) can be rendered virtually at any position and orientation in the court. In addition, the avatar can be easily scaled appropriately so that the system can be used with amateur players of different heights.

4. STROKE ALIGNMENT

In order to visually benchmark a given player's shot dynamics to an elite athletes, both strokes must be temporally aligned so that any timing variations incurred by contrasting swing tempos are eliminated. In this work, we apply one of three approaches for temporal alignment, depending upon which area of technique we wish to highlight in the augmented reality output sequences. For each of the three alignment methodologies, the amateur and elite player movements are temporally aligned using the data acquired from the worn tri-axial accelerometer data values streamed from the amateur athlete during match play. The first stage in each of the three alignment methodologies combines the tri-axial inertial readings into a single data stream via

$$R(t) = \sum_{i=1}^{3} ||A(t,i)||^2 \qquad (1)$$

where $R(t)$ is the combined value from each of the three accelerometer axes, $A(t,i)$, at time t. A similar raw inertial data stream, $V(u)$, at time u is obtained from the elite player motion during shots and both streams are used to align the two athlete motions. Although, $V(u)$ could have been captured from elite athlete shots during the Vicon capture, in this work we obtain $V(u)$ directly from the motion capture data using *virtual accelerometers* [7], with the virtual device matching the position and orientation of the real world sensor placed on the amateur athlete. Using virtual accelerometers is advantageous as they allow the system to be easily adapted for use with different range, number or position of amateur–worn accelerometers. Example $R(t)$ and $V(u)$ data streams from a forehand shot are provided in Figures 3(a) and (b) respectively. The two main acceleration peaks in either stream represents the two major movements inherent in a forehand shot; (1) drawing back the racket in anticipation of the shot, and (2) moving the racket forward to hit the ball.

The first, and most basic, approach to temporal alignment ensures that the final major acceleration peaks from $R(t)$ and $V(u)$ occur at the same temporal location, as shown in Figure 3(c). Let r^p and v^p be these two locations. Empirical tests have shown that this approach consistently aligns the apex of both shots temporally (i.e. aligns the time of ball hits). This alignment approach was employed for the results of Figure 1.

The second approach selects the time offset, o, between -0.5 and $+0.5$ seconds that minimises the similarity error between 1.5 seconds of real and *normalised* virtual acceleration data, $\hat{V}(u)$, centred around r^p and v^p, i.e.

$$o = arg \min_{o \in [-0.5, +0.5]} S(o) \tag{2}$$

where the similarity error is defined as

$$S(o) = \sum_{t=-0.75}^{+0.75} ||R(r^p + t) - \hat{V}(v^p + t + o)|| \tag{3}$$

and $V(u)$ is normalised so that the total virtual accelerations incurred over the 1.5 second period is equal to the total sum of real acceleration data samples. This single timing offset, o, maximises the energy between the amateur and elite athlete shots, and can be used to illustrate differences in shot timing with regards to the major movements in a shot. In the illustrative example of Figure 3(d), an offset of -0.1 seconds is selected. Figure 3(e) shows that the length of time the amateur player draws back the racket in anticipation of the shot is far too long, and should be shortened to that of an elite player.

Although each of the two previous alignment techniques have their benefits, neither approach can be used to directly compare posture during an entire shot sequence, as the timing variations between the two swings are not fully eliminated. Our third approach achieves this aim by applying a Dynamic Time Warping (DTW) [8] approach based on Dynamic Programming (DP) to obtain the *optimal* non-linear mapping between two shots acceleration readings. To apply the DTW approach, a similarity matrix is created where each amateur acceleration sample, $R(r^p + t)$, is compared to each virtual sample, $\hat{V}(v^p + u)$, using

$$M(t, u) = ||R(r^p + t) - \hat{V}(v^p + u)|| \tag{4}$$

for $t, u \in [-0.75, +0.75]$. The DTW algorithm finds the optimal sequence of (t, u) pairs so that the total cost, T, of the path through the similarity matrix is minimised, where

$$T = \sum_{t=-0.75}^{+0.75} M(t, u) \tag{5}$$

This processes essentially matches each amateur accelerometer sample, $R(r^p + t)$, to a single virtual sample, $\hat{V}(v^p + u)$, in a non-linear, but optimal, manner using the path, i.e. (t, u) pair matches, such as the path shown in red in Figure 3(f). The alignment of accelerometer data using DTW can be seen in Figure 3(g). Once temporally aligned using this approach, a direct alignment between each of the amateur player video sequences frames and the elite player avatar poses can be made.

Finally, returning to the system description of Section 2, the choice of the single elite athlete shot to be rendered from the database, alongside the amateur shot motion, is the selected as the one that minimises, T, from Equation 5.

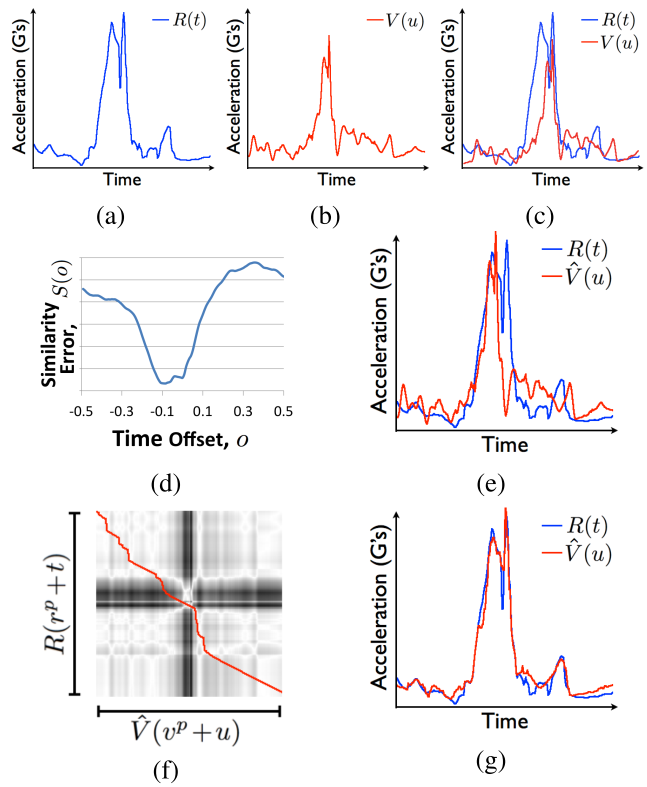
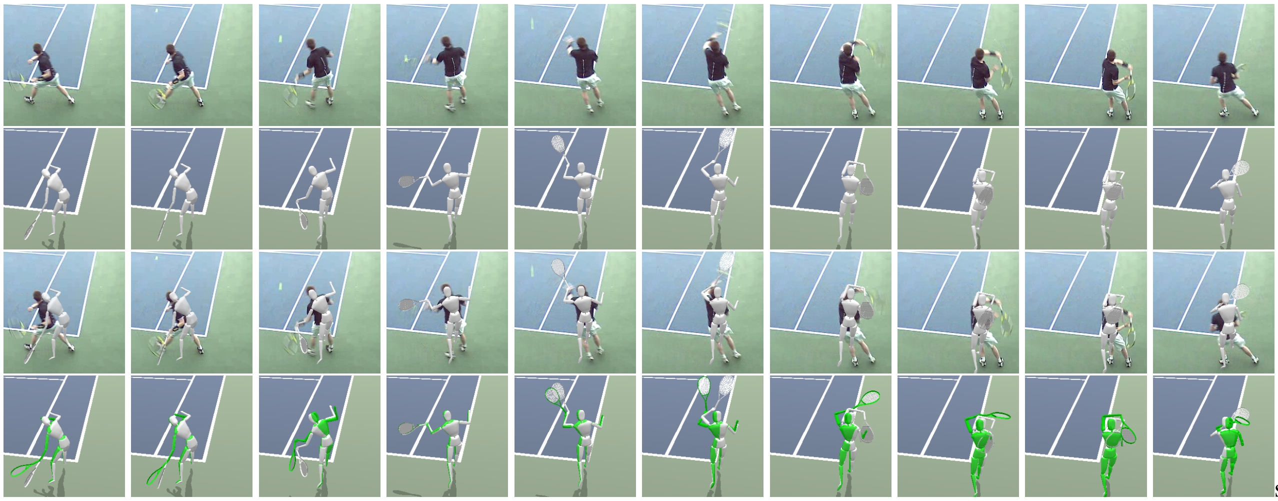


Fig. 3. Forehand stroke alignment; (a) Amateur player accelerometer readings; (b) *Virtual* accelerometer stream; (c) Peak alignment; (d) Peak similarly of *real* and *virtual* inertial data through entire stroke; (e) Best time offset alignment; (f) DTW path trough similarity matrix; (g) DTW alignment.

5. EXPERIMENTAL RESULTS

Once temporally aligned (using any of the three approaches), comparisons between amateur and elite player technique can be visualised. Rows 1–3 of Figure 4 provide qualitative output of the system for a forehand shot, aligned using the DTW approach. As can be seen, the amateur and elite player swings are well time aligned throughout the entire movement, allowing an examination of posture differences to be made. Row 4 of this figure adds a second avatar (in green) that is rendered using the first approach to temporal alignment approach (i.e. time aligning major acceleration peaks). This second alignment technique illustrates the smoother, more fluid motion dynamics, of the elite athlete who starts their forehand motion earlier, and finishes later, than the amateur player. The inexperienced player tends to *snatch* at the ball and rush the follow through, which can result in decreased shot accuracy.

Finally, we quantitatively evaluate the accuracy of the DTW approach in terms of matching amateur and elite athlete acceleration values. We captured a number of test game sequences using 4 different players and obtained 485 example shots, see Table 1 for details on the shot types collected including their



s

Fig. 4. Example output of the system for a tennis forehand.

Table 1. Test shot details. Accl = acceleration in m/s^2.

Shot Type	Number	Max Accl.	Average Accl.
Serves	148	110.36	14.81
Forehands	204	79.26	12.07
Backhands	133	79.17	13.64

Table 2. Error between real and virtual IMU accelerations.

Shot Type	Max. Error	Average Error (St.D.)
TO Serves	182.37	11.28 (18.54)
TO Forehands	51.01	8.14 (8.24)
TO Backhands	46.01	8.34 (7.85)
DTW Serves	34.83	5.20 (6.18)
DTW Forehands	26.09	3.63 (4.32)
DTW Backhands	26.49	3.63 (4.12)

average and maximum acceleration values. For each shot, we applied two alignment approaches; (1) the second approach of Section 4, a time offset, o, minimises the similarity error over the two entire shots (labelled TO in Table 2), and (2) the DTW based approach (labelled DTW in Table 2). From these results it can be seen that on average for the 148 service shots, the maximum difference between the elite and amateur athlete's accelerometer readings drop significantly from 182.37 to 34.83 m/s^2 using the DTW approach. In addition, a 46% and 300% reduction in the average error and standard deviation in this error was also achieved. Similar results were also obtained for both forehands and backhands.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we presented visualisation approach for the identification and elimination of faults in a tennis player's stroke mechanics. In order to achieve these goals the implicit movements of a test subject's stroke are aligned and compared to a players of a higher skill level. Future work will involve the incorporation of the DTW process into the machine learning algorithms for shot classification and evaluation studies of the software by domain expert coaches and athletes.

7. REFERENCES

[1] "Tennis ireland," http://www.tennisireland.ie/.

[2] C. Ó Conaire, P. Kelly, D. Connaghan, and N. E. O'Connor, "Tennissense: A platform for extracting semantic information from multi-camera tennis data," in *International Conf. on Digital Signal Processing*, 2009.

[3] Anil Aksay, Vlado Kitanovski, Karthikeyan Vaiapury, Efstathios Onasoglou, Juan Diego Prez Moneo Agapito, Petros Daras, and Ebroul Izquierdo., "Robust 3d tracking in tennis videos.," in *Engage Summer School*, 2010.

[4] P. Kelly, C. Ó Conaire, C. Kim, and N. E. O'Connor, "Automatic camera selection for activity monitoring in a multi-camera system for tennis.," in *ACM/IEEE International Conference on Distributed Smart Cameras*, 2009.

[5] D. Connaghan, P. Kelly, and N. E. O'Connor, "Multi-sensor classification of tennis strokes," in *IEEE Sensors*, 2011, pp. 1437–1440.

[6] "Vicon," http://www.vicon.com, 2012.

[7] Ronit Slyper and Jessica Hodgins, "Action capture with accelerometers," in *ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, 2008, pp. 193–199.

[8] Eamonn Keogh and Chotirat Ann Ratanamahatana, "Exact indexing of dynamic time warping," *Knowledge and Information Systems*, vol. 7, no. 3, pp. 358–386, 2005.