A Sensing Platform for Physiological and Contextual Feedback to Tennis Athletes

Damien Connaghan, Sarah Hughes, Gregory May, Philip Kelly, Ciarán Ó Conaire,
Noel E. O'Connor, Donal O'Gorman, Alan F. Smeaton and Niall Moyna
CLARITY: Centre for Sensor Web Technologies,
Dublin City University, Dublin 9, Ireland
connagha@eeng.dcu.ie

Abstract—In this paper we describe our work on creating a multi-modal sensing platform for providing feedback to tennis coaches and players. The platform includes a fixed installation around a tennis court consisting of a video camera network and a localisation system as well as wearable sensing technology deployed to individual athletes. We describe the various components of this platform and explain how we can capture synchronised multi-modal sensor data streams for games or training sessions. We then describe the content-based retrieval system we are building to facilitate the development of novel coaching tools. We provide some examples of the queries that the system can support, where these queries are chosen to be suitably expressive so as to reflect a coach's complex information needs regarding tennis-related performance factors.

I. INTRODUCTION

It is well known that if one is to excel in a given sport, then peak physical conditioning plays a crucial role. A keen understanding of tactics and strategy is also extremely important to put this conditioning to the best use. Tennis is no exception, and indeed requires significant proficiency in both areas. For a tennis player to perform at their best, they must have just the right mix of aerobic and anaerobic endurance, explosive strength and power, speed off the mark and agility. The amount of strength, speed, agility and flexibility conditioning a player is prepared to undertake has been linked to the standard they play at [4]. Designing and implementing an effective training regime for tennis requires a solid understanding of the many variables critical to optimal performance. In order to achieve this, the coach must (among other capabilities); (a) possess the ability to accurately monitor the performance of an athlete during training in order to maximise his/her effectiveness; (b) be able to reveal tactical, strategic or physiological performance issues through efficient post-training or post-game analysis; (c) complete long-term statistical performance analysis to gauge his/her improvement; and (d) relay this information effectively to the athlete.

In this paper we describe our ongoing work to develop a sensing platform, consisting of both wearable and fixed infrastructure components, upon which we can develop coaching tools that provide both *contextual* and *physiological* feedback to tennis coaches and athletes. By contextual feedback, we mean information on the player's location at a given time instant, his/her movement over a given time period as well as information on specific actions performed relevant in a tennis scenario e.g. a specific shot played (serve, return, etc) or the nature of a segment of a game (e.g. rallies, games, sets). In the current version of the platform, physiological feedback refers to heartrate and respiration, although we plan to extend this in the near future with more sophisticated sensing technology. Specifically, we plan to integrate a recently developed wearable smart fabric for on-body sweat analysis [1].

II. SENSOR TECHNOLOGIES

In collaboration with Tennis Ireland , the national governing body for the sport of tennis in Ireland, we have developed the TennisSense system at their coaching headquarters. The facilities provided by Tennis Ireland include four indoor hard-courts, three outdoor hard-courts and three indoor clay-courts, and these facilities are in frequent use by both amateur and elite players. As part of the TennisSense project, we have instrumented one of their indoor hard-courts with a data-gathering infrastructure for use as a test-bed for sports and health research. The infrastructure is designed to act as a counter-point to on-body sensing that we have deployed to individual tennis players. In this section, we will outline the sensor technologies incorporated into this combined platform. The goal is to provide synchronised sensing of both the *person*, via wearable sensing technology, and the *place*, via a custom-built sensing infrastructure.

A. Visual Sensing

The infrastructure includes nine IP cameras positioned around the court, with pan, tilt and zoom (PTZ) capability (see figure 1 for the location of the cameras around the court and a sample of the different camera views available). The two cameras at the centre of the baseline at either end of the court are AXIS 215 PTZ cameras, which are positioned 2.8 metres above the court and have very high zoom functionality, as well as physical pan and tilt. The high zoom is useful for obtaining a front view of the opposing side, or for focussing on the feet of the player from behind the court baseline. The other seven cameras are AXIS 212 PTZ cameras which have wide angle lenses (140°) and include fast digital PTZ functionality by subsampling from a high-resolution sensor.
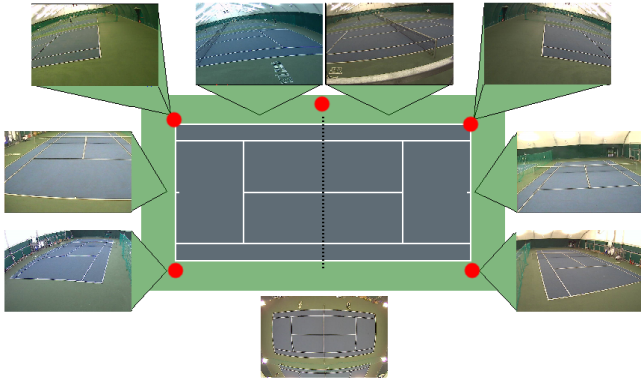
Fig. 1.   Camera and Ubisense base-station locations around the court.

Fig. 2.   Foster Miller Vest.

Six of these cameras are located at 3.5 metres above the court, while one overhead camera is positioned at 13.8 metres from the ground. All the cameras have a resolution of 640×480 and a frame rate of 30Hz. In addition to capturing raw video, the content of a video sequence can be automatically structured by employing image processing techniques (as described in [3]) to temporally track the tennis ball and to automatically segment the training session (or match) into individual *shots* or *rallies*, where a rally is a sequence of shots within a point.

B. Localisation

The infrastructure also includes a Ubisense spatial localisation system that facilitates the real-time 3D localisation and tracking of small wireless *tags* to within a $\pm 15cm$ degree of accuracy. This system requires five sensor base-stations to be fixed and calibrated around the court (the positions of these base-stations are indicated by red dots in figure 1). Using these sensors, each player's 3D position can be obtained via the triangulation of ultrawideband (UWB) radio pulses from small ($38mm \times 39mm \times 16.5mm$), light ($25g$) Ubisense tags typically carried by the players in their pocket or shoe. The Ubisense framework provides player 3D positions, in addition to facilitating the capture of further player statistics such as average velocity and distance traveled, that augments video sequences with contextual data.

C. Body Sensors

Foster Miller (an independent company but part of the QinetiQ, Inc. group) has recently developed a t-shirt based Ambulatory Physiological Monitoring System which monitors vital signs of a person during activity and transmits the data wirelessly to a remote station.

Three passive sensors are embedded into a form-fitting t-shirt. Circumferential wires embedded in the shirt make up a respiration sensor which senses the expansion and contraction of the chest caused by breathing. A passive temperature sensor called a thermistor is placed in proximity to the skin just under the right arm. Each of these sensors transmits data through an e-textile band in the shirt.

The data is received by a electronic processing hub, which is approximately the size of small cell phone. This processing hub contains a printed circuit board (PCB) and onto the printed circuit board is mounted an additional sensor, an accelerometer, which reports data about the movement of the subject. The PCB also includes a small radio transmitter that broadcasts the desired data to a nearby receiver for display and storage.

The Foster Miller vest has a number of potential applications in tennis. Physiological information can be used to enhance performance during training and competition. A simple example would be providing the player/coach with the average percentage of maximal heart rate (Breathing rate consistently accounts for the greatest amount of variation in an individuals perception of effort during exercise. This is not surprising considering that many individuals unconsciously use breathing rate to regulate the pace of many daily activities. For example, when we climb a flight of stairs we are more aware of changes in our breathing rate that heart rate. In addition, breathing rate is closely linked to lactic acid levels in the blood during exercise. A novel aspect of the FM vest is the ability to measure ambulatory breathing rate. This will for the first time allow sport scientists to develop training programs and design match strategies based on breathing rate.

III. DATA CAPTURE

For our initial experiments on combining multiple sensor modalities, we decided to capture a relatively small amount of sensor readings, using two experienced tennis players. Two 21 minute competitive training sessions were captured using all sensor modalities outlined in section II. Each training session was broken down into seven distinct three minute segments as follows; (1) the players remaining stationary in the centre of the court at the start of each training session, in order to provide a benchmark of each player's physiological status when at rest; (2) the players participating in a number of light rallies, whereby each player was constrained to stand within two opposing service boxes; (3) a period whereby these rallies were extended in the third segment to more intensive full court rallies; (4/5) one player would remain stationary and hit the balls from the left to the right side of the court in an attempt to raise their opponents heart rate and respiration to greater
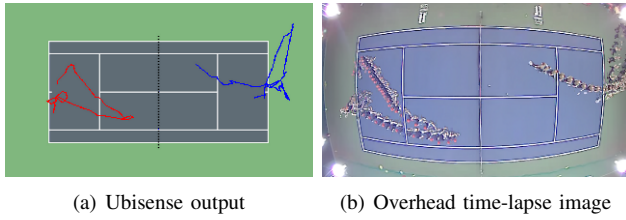
(a) Ubisense output   (b) Overhead time-lapse image

Fig. 3. Camera and Ubisense alignment
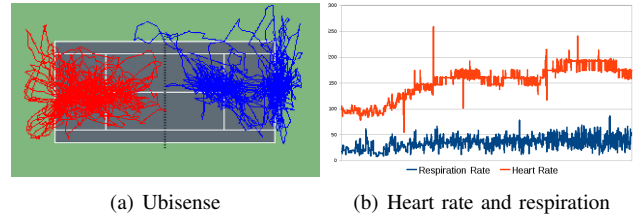

(a) Ubisense   (b) Heart rate and respiration

Fig. 4. Sensor traces for Player A

levels; (6) a second light rally session; and finally (7) a period of rest.

A. Setup and Synchronisation

The first stage in the data capture process is to pan, tilt and zoom the nine cameras appropriately for the training session at hand. This may involve positioning the cameras to provide as much coverage of the court as possible, or focusing some or all of the cameras on specific regions of the court to provide as many viewpoints of player technique as possible. The camera parameters can be altered appropriately via a graphical user interface (GUI) developed for the use of the tennis coaches. In addition, the coaches can store a variety of preset camera positions and zoom-levels into the system to semi-automate this process for future training sessions. Using the same system, all nine video streams (or a sub-set of them) can be controlled, streamed (as either MPEG-4 or MJPG) and stored on to a single computer on the same subnet as the camera network. Due to this system of starting and terminating all recorded streams from a single computer, the alignment of video sequences is trivial as long as no drift between the nine recorded sequences occurs. In the TennisSense infrastructure, visual examination of a variety of test sequences (some up to two hours in length) have revealed no visual indication of drift.

Ubisense data is also controlled, streamed and stored on the same computer via the GUI interface. However, unlike the video streams, the Ubisense tags have a non-uniform sampling rate (this rate increases in proportion to the amount of movement detected by the tags). In order to align the video and Ubisense data, the internal clock of the recording computer is used to index the offset between time at which a Ubisense reading is obtained and the start time of the video recordings. An example of the alignment between video and Ubisense can be seen in figure 3. In figure 3(a) the Ubisense readings from a 20 second sequence taken from the *full-court rallies* segment of the first training session can be seen. In this figure, the red and blue tracks indicate the locations of the two individual players on the tennis court. Figure 3(b) illustrates the activity in the video sequences via a time-lapse image of the corresponding two players activities during the corresponding 20 second sequence. The close alignment between the two sensor modalities can be clearly seen.

Physiological data was also captured for the two training sessions via a second computer using the Foster Miller vest. Data was streamed and stored using proprietary software

provided with the FM vest and therefore had to be manually aligned with the other data sources. This alignment was made by starting the data capture from the camera / Ubisense software before that of the FM vest, which was subsequently initialised in conjunction with a well defined visual cue (i.e. via a clapper-board movement). At the end of the training session, a second visual cue was provided to indicate the end of the FM capture, after which the video and Ubisense data capture was terminated. Finally, the temporal length of the captured FM data was examined and compared to the corresponding length of time between the two visual cues to determine if there was any possibility of drift between the readings of the different sensor modalities. It was found that the two time lengths corresponded well, and as such it was inferred that drift did not occur. In future work it is envisioned that the FM data streams will be incorporated into the same recording software as is used for the video and Ubisense data capture. In such an application, all physiological data will be time-stamped and aligned in a similar manner to that of the Ubisense and video data.

B. Test Corpus

In total, 42 minutes of aligned video, contextual and physiological data was obtained. Figure 4 displays the total sensor traces for both the Ubisense, heart rate and respiration readings. It can be seen from these readings that both a large area of the court was covered during the session and that the FM readings for both the heart and respiration rates increased significantly during the training period. In addition to these sensor readings, further contextual data was obtained by employing the image processing techniques, described in [3], to automatically segment the training session into individual shots and rallies. The results of this technique led to 375 shots and 72 rallies in the first training session, and 369 shots and 85 rallies in the second session.

After wearing the sensors, the players commented that the gear was easy to wear and didnt feel like it hampered the play in any way. They were interested to see changes in heart rate and breathing as the game progressed and added that it could help to understand why certain points are lost.

IV. RETRIEVAL SYSTEM

The contextual and physiological sensor information can be used to facilitate both efficient browsing of the associated video, as well as advanced querying of the other data sources. We believe that both are essential components of next

generation coaching tools. We now describe the content-based retrieval system we have built to facilitate queries against the various data sources, and provide some examples of how the system interface is used to build queries.

A. System Interface

The query retrieval interface (see figure 5(a)) provides the coach with a visual tool to build queries from multiple data sources, such as the FM Vest, Ubisense and video analysis in an intuitive manner. The interface has three main sections: playback panel, multi-modal query builder panel and query/results panel.

1) Playback Panel: This panel provides the coach with a variety of playback options for viewing the video and sensor data results returned from input queries.

a) Video Playback: The couch can watch the query results from nine different camera angles. To view a result from a particular angle, the coach can click on the result and select a camera angle by clicking on a red camera icon on the tennis court image, to the right of the video in figure 5(a). He/she can also pause, fast-forward or rewind, or can choose to watch the video in slow motion.

b) Ubisense Court Traces: The localisation of each player is tracked with the Ubisense tags, which each player carries. The playback of Ubisense data is synchronised with the selected video playback sequence and may be viewed on the tennis court picture in figure 5(a).

c) Physiological Data Graph: When the coach is viewing the video results, the physiological graph (see the bottom of figure 5(a)) shows each player's heart rate and respiration. The data in this graph is synchronised with the selected video playback sequence. The graph displays the current physiological data (with respect to the current video position) at the black line in the centre of the graph. There are two tabs, one for respiration and the other for heart rate. Data from one or more players can be displayed in the graph simultaneously by clicking on the player radio button below the tabs.

2) Multi-modal Query Builder Panel: The multi-modal query panel is the main area whereby coaches can build queries from the multiple data sources. It has three tabs: *video analysis*, *physiological* and *localisation*. Each tab has a range of parameters that can be altered to build queries from one or more of these features.

a) Video Analysis Tab: The analysis video panel, see figure 5(b) has two tabs, *shot hits* and *rallies*, which can be used to build an individual or combined query. The shot hits tab allows the coach to find all shots that occurred inside or outside a given area, which is defined by drawing a rectangle onto the tennis court as shown in figure 5(e). The rally tab allows the coach to search the video for all rallies of duration x seconds or greater. When the coach clicks *Add to Query*, the relevant search criteria is appended to the Query/Results panel (see section IV-A3).

b) Physiological Data Tab: The FM Vests record data from each of the two players, so queries based on physiological sensors can be drawn from both players. A physiological query

can be sought from any time interval using the time range pointer on figure 5(c). A player's heart rate or respiration rate can be queried above or below a given value, by entering this value into the respective text box.

c) Localisation Data Tab: Each player wears a Ubisense tag that can be used to calculate various geometric-based performance factors such as speed, distance traveled and the area covered for a given time period. This opens up the possibility to browse all the video sequences where a given player matched such criteria. Each of the relevant factors can be queried from any time interval by entering a value into their respective text box in figure 5(d).

3) Query/Results Panel: The Query/Results panel has two tabs, *Query* and *Results*, as shown in figure 5(f) and 5(g) respectively. The Query tab gives the coach a textual description of the overall query as it is built from multiple data sources. This helps the coach to conceptualize the final query from multiple data sources. In the query panel the overall query is grouped into three sub-sections (video / physiological / localisation) and each sub-section can be selectively turned on or off before the final query is run. To exclude a sub-query from the overall query, the checkbox beside the sub-query is simply unchecked. The Results tab lists the retrieval results that best match the submitted query. Each result gives a video starting time, rally length or shot hits (if applicable).

B. Sample Queries

In this section, we illustrate the use of the retrieval system via descriptions of how a number of queries can be built and submitted to the system. Currently we have not performed any field trials or quantitative/qualitative evaluations of the system. We plan to do so in the near future with a large group of tennis players at different levels of proficiency. The first three examples are of querying using specific modalities from *either* the heart rate from Foster Miller vest, distance from Ubisense, or match content structuring from visual content analysis. The fourth example combines all three of these sensor inputs into a multi-modal query.

1) Physiological Query: Find all the occurrences where the heart rate of Player A went above 150BPM in a rally of ten shots or more. Using the multi-modal query panel, the coach inserts the value ten into the Video rally tab and the respective heart rate of +150 for Player A into the Physiological Sensor tab. Video analysis that has tracked the ball can also detect when a shot is a serve or return and from this the rally length can be determined. The query is then submitted and the results shown on the Query/Results panel on figure 5(g).

2) Location-based Query: Return all performance factors after Player B has covered distance x since a give time instant This query will return the average statistics after a player has covered a given distance. Player distance can be calculated using the data from the Ubisense tags. A typical result from this query would include the average heart rate, respiration rate, how many serves or returns the player made, and what percentage of time the ball was in play.
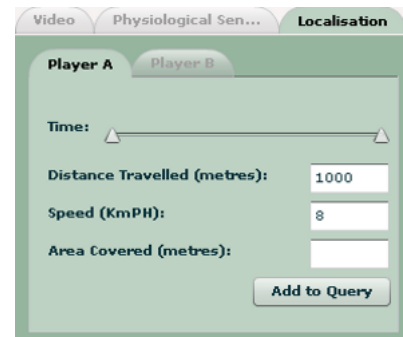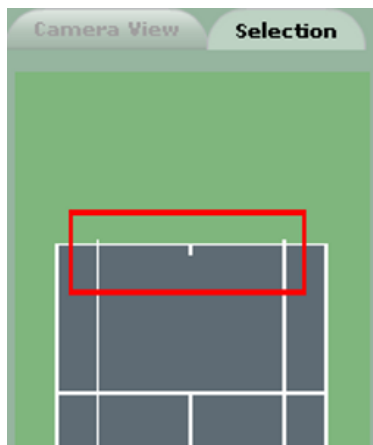
(a) Overall Interface
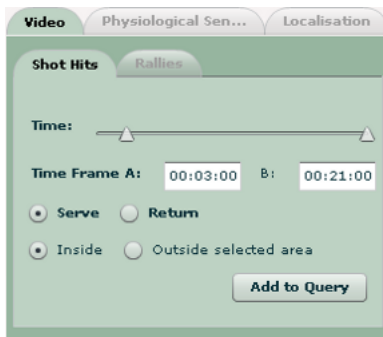
(b) Video Analysis Query Tab

(c) Physiological Data Query Tab

(d) Localisation Data Query Tab

(e) Shot Hits Selection Area

(f) Multi-modal Query

(g) Results

Fig. 5.   Content-based retrieval system

3) Visual Query: For each rally, what is the total area covered? What is the average speed of the players? And what was the distance covered? The results here are calculated from a combination of the Ubisense data and the video analysis. First, all the rallies are detected from the video analysis, then the Ubisense data is analysed whenever a rally was in progress.

4) Comparing Two Sets of Performance Factors: Generate a comparison of the performance factors for Player A from two different intervals in the game. A good example in this case is when the coach wishes to view the performance factors for a given player for the first kilometre traveled (or first X mins or N rallies or the last two times the player's heart rate was above 150) to the last kilometre (or last Y mins or M rallies) to chart the player's performance profile over the course of a game.

V. CONCLUSIONS AND FUTURE WORK

In this paper we have described the sensing platform that we are developing for providing feedback on a range of different performance factors in tennis. The platform consists of a wearable component for physiological sensing and a fixed installation of visual and RF-based localisation sensing. We described the content-based search and analysis engine we are building to both generate the required feedback and present it to coaches in an accessible manner. Finally, we have provided some illustrative examples of the types of queries that can be generated on this multi-modal sensor data.

Whilst physical conditioning is a key performance consideration, the literature also reveals that in younger players technical stroke production appears to influence rankings more than physical ability [2]. As such it has been suggested that their training should concentrate on effective and efficient stroke mechanics, improving technique and ball placement, with less emphasis on physical conditioning. Clearly this points to a potential need for facilitating *biomechanical* feedback as well as contextual and physiological feedback. This is currently not integrated into our sensing platform and forms a key aspect of our plans for future work. Recent advances in robust markerless motion capture based on wearable sensing technology, such as the multi-accelerometer vest reported in [5], present exciting opportunities in this regard.

REFERENCES

[1] S. Coyle, Y. Wu, K.T. Lau, D. Kim, S. Brady, G. Wallace, and D. Diamond. Design of a wearable sensing platform for sweat analysis. In *Personalised healthcare (pHealth)*, 2007.

[2] M.S. Kovacs. Applied physiology of tennis performance. *British Journal of Sports Medicine*, 40:381–386, 2006.

[3] 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 Conference on Digital Signal Processing(DSP) (under review)*, 2009.

[4] P.E. Roetert, S.W. Brown, P.A. Piorkowskil, and R.B. Woods. Fitness comparisons among three different levels of elite tennis players. *Journal of Strength and Conditioning Research*, 10(3):139–143, 1996.

[5] R. Slyper and J. Hodgins. Action capture with accelerometers. In *ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, 2008.