A wearable sensing system for timing analysis in tennis

Lars Büthe[1], Ulf Blanke[1], Haralds Capkevics[1] and Gerhard Tröster[1]

***Abstract*— Wearables find in sports one of their main applications. In recent years, many wearable devices have been commercially released such as the Babolat Play or Sony Smart Tennis Sensor that detect and classify different types of tennis shots and provide a performance analysis to the player. However, available devices focus on a single technical element of tennis only - the shot. As tennis performance is the result of a full body coordination and timing of the movement, the present work wants to take a broader view at the tennis player performance and include the simultaneous work of legs and arms with the goal to time elements of movement. We design a sensor system with three inertial measurement units, one attached to each foot as well as one at the racket. We develop a pipeline to detect and classify leg and arm movement and implement a gesture recognition for the shooting arm based on LCSS (longest common subsequence). The algorithm distinguishes between forehand and backhand (with topspin and slice, respectively) as well as a smash. Footwork is first segmented into potential steps and then classified by a support vector machine between shot and side steps. In the person-dependent case the algorithm achieved 87% recall and 89% precision. The step recognition algorithm has been able to detect 76% of the steps with a classification accuracy of 95%. Based on these results timing information within the shooting state can be robustly obtained which is crucial for a thorough analysis of the whole shot.**

I. INTRODUCTION

Understanding a player's state of motion performance in sports is crucial to identify the shortcomings and to improve the technique. Analyzing video recordings which are often used for playback in professional training to assess the player's performance are a tedious and time-consuming task. Not only professionals profit from a more efficient analysis during their day-to-day training, but it can also support amateur players to gain insight into their skill level and understand aspects for improvement. Static systems based on image and video analysis have the advantage of high accuracy, but have a high setup cost and are not flexible with regard to location. Quite often also an expert user of the system is required.

With advances in microelectronics, wearables recently gained significant attention in commercial applications in healthcare [1], entertainment [2], or sports [3]. Available as a commodity device, they became popular in the sports area. During the performance of the sports, detailed data about limb rotation velocity, acceleration, and orientation can be captured at high frequency rates. Making the wearer's activity visible, their application promises to make sports analysis more efficient. In order to not overwhelm the user with the data, relevant statistics are extracted using machine learning techniques, which have been widely explored for the last decade. Popular methods for gesture recognition that have been applied include Hidden Markov Models (HMM) and Dynamic Time Warping (DTW). Such analysis has already been performed for several sports, e.g. ice hockey [4] or rowing [5], where data of body-worn IMUs is being processed.

Specifically for tennis, the machine learning methods detect and classify a basic set of shot classes such as forehand, backhand, or serves [6], [7], [8], [9]. With the combination of inertial and visual sensing a high accuracy of greater than 90% [10] can be achieved. To gain additional insight into other parameters such as fatigue and endurance, others added physiological sensors to capture vital signs of the player such as heart and respiration rate [11]. The same authors also use inertial sensing - combining accelerometer, gyroscopes, and magnetometer - to detect the occurrence of shots [12]. The authors of [13] use three gyroscopes (chest, upper arm, wrist) to assess and classify the particpants' skills during a serve.

Beyond academic research, a few devices have been made commercially available. The most prominent ones are certainly Babolat Play[1] and Sony Smart Tennis Sensor[2]. Here, an IMU with a wireless transmitting device is attached or integrated on the racket and data is sent to a smartphone for further analysis.

Fig. 1: Capturing coordination of hitting arm along with the footwork of the player with three IMUs.

Surprisingly, none of the previous work in research nor in product development have been considering the footwork

[1]Lars Büthe, Ulf Blanke, Haralds Capkevics and Gerhard Tröster are with the Electronics Laboratory, ETH Zurich, 8092 Zurich, Switzerland lars.buethe@ife.ee.ethz.ch

[1]http://en.babolatplay.com
[2]http://www.sony.com/electronics/smart-devices/sse-tn1w

978-1-5090-3087-3/16/$31.00 ©2016 IEEE

as integral part of the gameplay, although influencing the overall performance [14].

Orchestration of the body posture, i.e. timing foot steps and the arm gesture, is paramount for a successful shot. Especially for beginners timing and coordinating different body parts is a challenge. Therefore, we take a broad view on the body coordination in this work. We extend to a simple sensor setup that is able to capture a shot and the timing of arm *and* foot coordination leading to the shot (Figure 1). We present a pipeline to (1) detect and classify the type of the shot, and (2) detect left and right foot steps. Finally the system visualizes the shot and step detections on a timeline with timing information to the user. This can help the player to understand how well the timing of his footwork is aligned with the racket swing.

In the next Section II we present the different types of shots and steps that shall be recognized. In Section III we detail the sensor system as well as the classification pipeline and evaluate our approach in Section IV. We also show a potential application view that can be displayed to the user based on the system described in this paper.

II. Tennis Movement Types

In tennis a variety of shot types exist, depending on the player's position on the court and his direction to the ball. Moreover, as the footwork influences the nuances of the shot, we selected relevant movement types around the execution of the shot. Most of the recent work has been done to identify high-level shot types, i.e. distinguish between a forehand and backhand shot [9], [12]. In this work, we want to focus also on classifying the shots in more detail by discriminating between a topspin and a slice ball. We are also not aware of any works where the steps around a shot and their timing with regard to the shot have been investigated at all. Here, we present a relevant sub selection of classes during the gameplay both for arm and footwork, and which we focus on in the recognition phase in Section III.

A. Shot classes

Five different classes of shot strokes shall be recorded and detected:

- *Forehand Topspin (FH-TS)* - Brushing over the ball with the racket and thus giving it forward spin from the player's dominant side.
- *Forehand Slice (FH-SLICE)* - Carving the ball with the racket and thus giving it backspin from the player's dominant side.
- *Backhand Topspin (BH-TS)* - Brushing over the ball with the racket and thus giving it forward spin from the player's non-dominant side.
- *Backhand Slice (BH-SLICE)* - Carving the ball with the racket and thus giving it backspin from the player's non-dominant side.
- *Smash* (SMASH) - Hitting the ball above the head with a serve-like motion from the player's dominant side.

Other shot types such as volleys, half-volleys, drop shots and flat shots which are less common were not included in this first study.

B. Step classes

For the footwork we defined two different classes of steps:

- *Shot steps* - Basic forward-like steps when approaching the ball and stepping through the ball during a shot, these can also be cross steps.
- *Side steps* - More rapid sideways foot movements, normally parallel to the net.

The step classes resemble the basic movements during a regular tennis match. This useful technical information allows to identify how well the player has approached the ball before each shot and if the player has performed the sidestepping back to his *home* position after each shot.

III. Sensing and Recognition System

To capture the movement of tennis players, we employed a wearable system capturing motion data of locomotion and shot movement. We feed the data into two separate classifiers to detect and classify shots and footwork into the categories described in previous section II.

A. Sensor setup

We designed a wearable sensor system which is easy to apply. As sensors, we selected commercially available EXLs3[3] IMU modules. They have a weight of 22 g and feature 3-axis accelerometer, 3-axis gyroscope and 3-axis magnetometer and allow a sampling rate of up to 200 Hz. During our data recording, the data is sampled from the accelerometers with a range of ± 16 g and from the gyroscopes with a range of ± 500 dps. The data is immediately transmitted wirelessly via Bluetooth to a data capturing device (in our case a laptop) where it is timestamped to ensure a synchronization between the data streams of the three sensors. The data is then post-processed through Matlab.

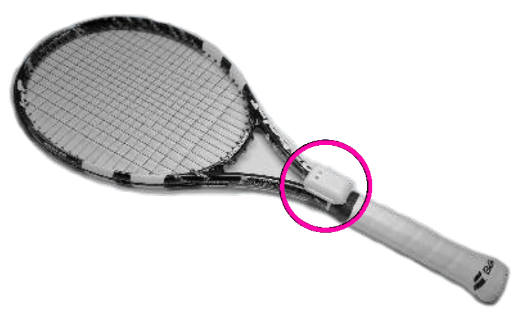

Fig. 2: Placement of EXLs3 IMU sensor on tennis racket.

We use one sensor to capture the shot type by taping it to a standard tennis racket (Babolat Pure Drive) in a way that it is unobtrusive to the player (Figure 2). The low weight of the sensor does not influence the feeling or handling of the tennis racket (weight: 317 g). It was observed that a

[3] http://www.exelmicroel.com/eng_electronic_medical-wearable-technology-exl-s3_module.html

firm connection between racket and sensor is vital to the quality of the obtained data, to accurately detect the moment of the shot and achieve a correct measurement of acceleration values. Additionally two more sensors are attached to either shoe of the player, to gather data on the movement of the feet (Figure 3). For a future commercial product a complete integration of the sensors into racket and shoes could be envisioned. This would mitigate any risk of misplacement of the sensors and resulting wrong orientation in case the sensors would be placed manually on the body of the player.

Fig. 3: Placement of two EXLs3 IMU sensors on shoes.

B. Shot detection and classification

We want to be able to extract the timing information of the shot, i.e., information about the entire movement sequence for performing a shot. For classifying the shot movement we employ a LCSS (Longest Common Subsequence) algorithm. LCSS [15], [16] is an algorithm for obtaining the similarity measure of two curves. It is similar in its principle to Dynamic Time Warping (DTW) in that it provides a way to find the best non-linear alignment of two sequences. The main advantage of LCSS over DTW is that it is more robust against outliers or noise in the data. In addition to this, this algorithm also addresses data recorded at different sampling rates or having different speeds, similar motions in different space regions, different sequence lengths and can be used online. We use 3D-gyroscope data from the tennis racket (Figure 2). Initial experiments showed that during play movement, the gyroscope signal is more suitable than accelerometer data, as it allows to separate shot movements more clearly. Before applying LCSS we first symbolize, i.e. we discretize the data, using a kMeans algorithm. As a result of the similarity measurement of the cluster assignment timeseries, we obtain a similarity score for segments voting for a specific class. Since detected potential segments can overlap the same actual segment, we perform a simple non-maxima-surpression by simply selecting that returned segment with the highest similarity score, which overlaps the ones with a lower score.

C. Step detection and classification

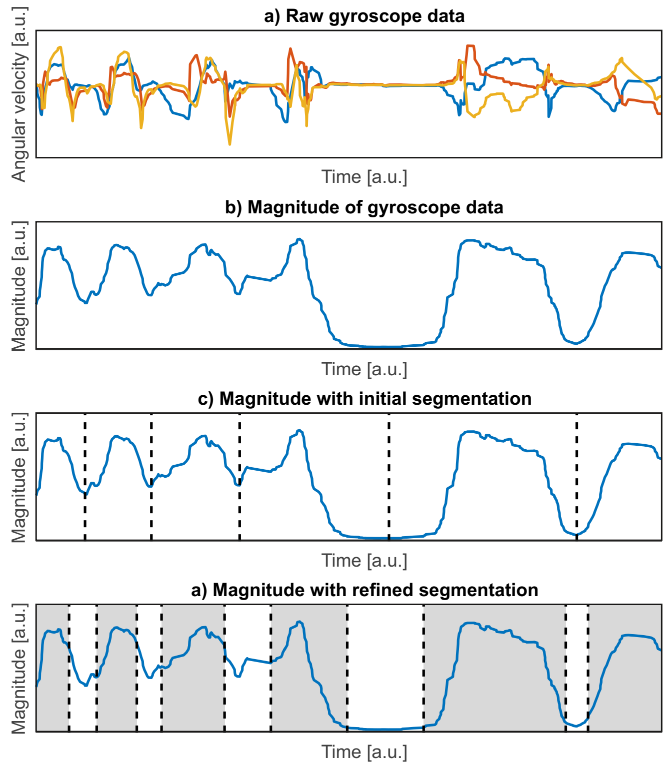

Fig. 4: Segmentation of segments of interest for classifying foot movement. a) raw gyroscope signal of side steps, b) magnitude of the gyroscope signal, c) initial segmentation of magnitude using minimal peaks (dashed line), d) the refined segments (highlighted) by growing stationary segments.

To classify the type of steps performed during playing, we capture gyroscope data from a foot mounted sensor (Figure 3). We first segment a continuous stream of gyroscope data by step detection. Obtained motion segments - between steps - are fed into a binary classifier to classify the steps types previously described, *shot steps* and *side steps*. To identify steps we make use of a technique in deadreckoning [17], [18]. Usually acceleration data can be used to robustly detect steps, i.e., zero-velocity moments, in which the foot rests on the ground. However, as in tennis the foot moves quickly, stepping can be harder to detect using accelerometers. Due to the spring behaviour in the signal and the foot not resting sufficiently long, we switched to the data of the gyroscope sensors (Figure 4a)), which experimentally yielded improved performance. We identify these moments of zero-velocity by first calculating the magnitude for the 3-dimensional gyroscope signal (Figure 4b)). Then we apply a peak detection algorithm[4] to extract minima peaks. We assume these peaks to represent the moments of the foot being stationary on ground, which segment the data into motion segments of interest (Figure 4c)). As the stream of data can have elongated minima, e.g., while the player is standing, we refine the motion segments. We grow the

[4] *peakdet.m* from Eli Billauer

stationary moments as in Figure 4d). As a result we obtain segments of interest during which the foot is moving in air. For these segments, we calculate common features such as mean, variance or the spectral energy for both acceleration and gyroscope data. We then feed these features into a support vector machine for classification.

IV. Evaluation

Based on the recognition pipeline described in the previous Section III, we separately evaluate the performance of classifying shots and footwork respectively. To this end we recorded a training session of several amateur and expert players. We use common evaluation metrics such as precision and recall.

A. Dataset

For the dataset four subjects recorded different shot types, step types as well as complete shots. Three subjects were right-handed male amateurs, whereas a fourth subject was a left-handed male expert. The data collection was divided into three main groups: racket swing recordings (without any foot movement), foot movement recordings (without any racket swing) and full shot recordings (using the full body technique of approaching the ball, striking it and returning to the initial position). Each of the five shot types was repeated 10 times while the sensor was attached to the opposite face of the racket where it hit the ball. For the foot movement recordings, a single dataset contained five steps for the shot steps and three steps for the side steps, respectively, each repeated 10 times as well. This leads to a total of 640 recorded steps (400 in the shot step class and 240 in the side step class). The full shot with racket and foot data acquisition was recorded three times per subject. Table I lists the shot and step types together with the respective amount of repetitions.

TABLE I: Data set composition with four different players performing the given shots and steps.

Category	Sub-element	Repetitions
Racket	Forehand Topspin	10
	Forehand Slice	10
	Backhand Topspin	10
	Backhand Slice	10
	Smash	10
Feet	Shot steps	10
	Side steps	10
Racket + Feet	Full shot	3

B. Results

The following paragraphs will describe and explain the achieved results, first focusing on the shot classification, subsequently the step classification before finally analyzing a full shot composed of racket swing and steps.

Shots. In a first step, the recorded shots are evaluated with a leave-one-out cross validation to identify the user-dependency of the classification. Figure 5 shows the respective confusion matrices for each subject where subject #1 is the left-handed player. It can be clearly seen that our pipeline is strongly user-dependent. The accuracy for subject #1 is $0\,\%$ which can be expected as this subject performed the shots left-handed. However, even if the cross-validation of this subject is left out, the recall values for the remaining subjects is around $60\,\%$, meaning that a large fraction of shots are not detected. The values for recall and precision are $0.49 \pm 0.22\,\%$ and $0.49 \pm 0.04\,\%$, respectively. This allows the conclusion that the implemented algorithm does not work exceptionally well when trained with data from users that are not part of the testing set and when the reference dataset does not contain many subjects. Already small nuances in the execution of a certain shot influence the classification behaviour.

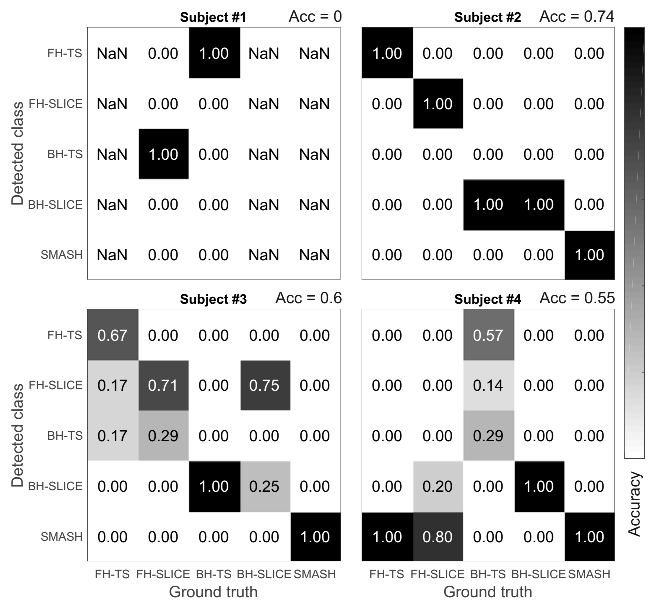

Fig. 5: Shot classification results and respective accuracy with leave-one-out cross validation for each subject (person-independent evaluation).

Nevertheless, an implementation scenario exists where our system is trained with the data of a specific user that will also be using it. We therefore evaluated the algorithm in a second step where it is trained with user-dependent data. A single shot of each class of each subject is selected, removed from the testing dataset and instead used for the training. The results in Figure 6 show a great increase in the accuracy, reaching a value as high as $94\,\%$. The values for recall and precision are $87\,\%$ and $89\,\%$, respectively. Only parts of the backhand topspin shots are misclassified as backhand slices which could also be attributed to an imprecise execution of the shot during the data recording session.

Steps. Out of the 640 recorded steps, the foot movement

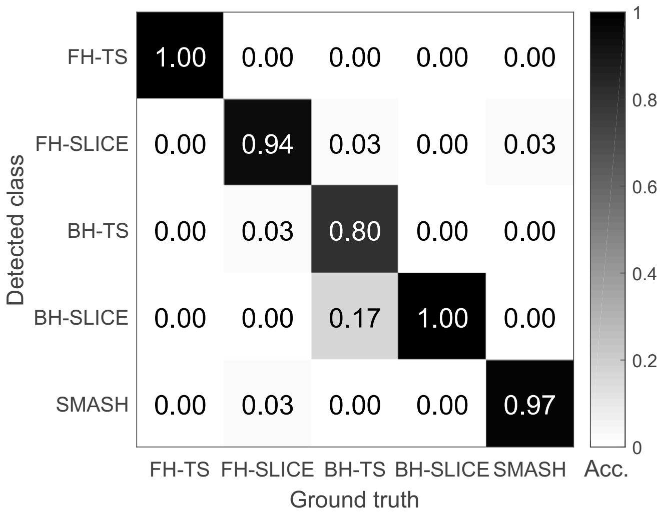

Fig. 6: Shot classification results with person-dependent training and testing.

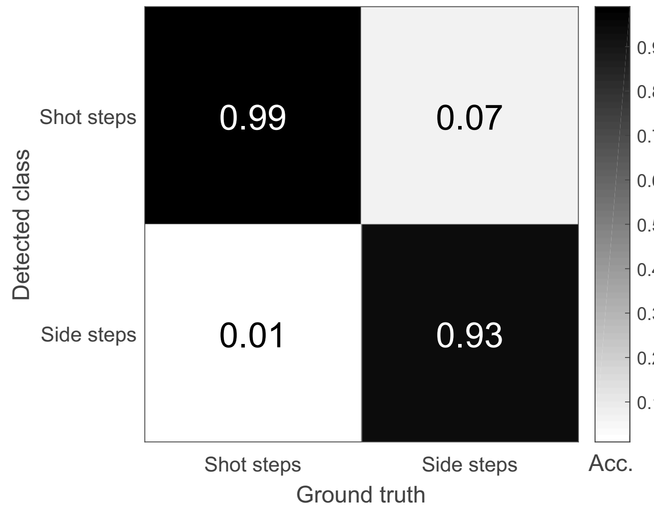

Fig. 7: Step classification results for shot and side steps.

detection algorithm was able to recognize 76 % as steps. The side steps showed a detection ratio of 96 % while the shot steps were only recognized as steps in 63 % of the cases. This poor detection ratio for the shot steps can be attributed to the subjects not making their steps distinct enough.
However, when feeding the detected steps in the classification algorithm, the resulting confusion matrix shows good results (Figure 7). The overall accuracy of correctly classified steps is around 95 %. These results are achieved by selecting 10 % of the data as training data and testing it on the remaining samples.
As can be seen from Table II even a small amount of samples for testing (1 %) already achieve a good accuracy of around 79 %. The best accuracy can be achieved when selecting 10 % of the data as training data. The good classification results of the steps show that the steps are quite user-independent and individual differences in how a step is performed do not have any significant influence on the outcome.

TABLE II: Accuracy of step classification for different number of steps used of training.

Number of training steps	Ratio	Accuracy
5	1 %	0.79
9	2 %	0.84
24	5 %	0.80
48	10 %	0.95

Complete shot. With the above classification of the shots as well as the steps, a combination of the data as well as a timing analysis is possible. This is depicted exemplarily for one of the recorded full shots (racket swing as well as foot work) of the left handed subject in Figure 8. The moment when the racket hits the ball with a forehand topspin is depicted with a dashed line at the time 0. Before the shot, the player makes three shot steps towards the ball. An additional shot step after hitting the ball is made when stepping through with the swing of the racket and transferring the weight forward. Subsequently, three side steps are made to return to the initial position.

The recording shows a relatively long break after the shot steps before starting with the side steps. This can be attributed to the fact that the recording was not performed in a game situation. In a real game, this timespan is expected to be much shorter which the side steps starting immediately after the completion of the shot.

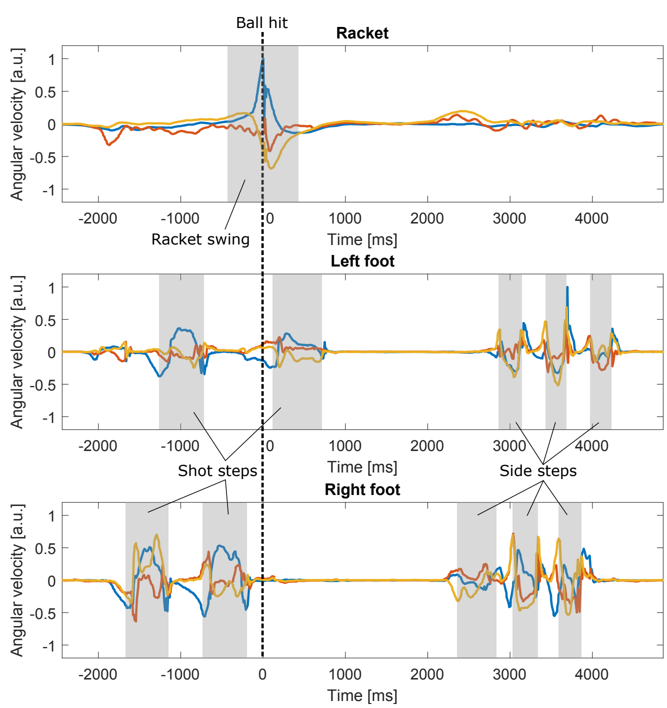

Fig. 8: Example use case of full body tennis assistant with annotated racket swing and steps, using the gyroscope data.

V. CONCLUSION

We presented a wearable sensor system consisting of only three IMUs that allow to capture the racket and foot movement of a tennis player. The recorded data can be used to independently classify the type of shot as well as the type of step of the player. We recorded and analyzed a data set which consisted of four different players. It could be seen that the data for the shots are highly user-dependent. With a user-dependent approach where only a single shot of each class was used for training, the accuracy could be raised to $94\,\%$. This leads to the conclusion that each of the four players has his own distinct technique for a specific shot which influences the classification behaviour. However, we believe that with a larger dataset of more subjects, we will be able to better cover the different variations in techniques allowing for a user-independent approach. The step detection on the other hand showed in a user-dependent analysis an accuracy of $95\,\%$ which allows the conclusion that the steps are similar across our subjects.

With these two classifications and a timing display of each event, a simple tennis assistant can be set up. This assistant allows to visualize when each single step as well as the shot happened. From this, an expert player or the trainer of an amateur player will be able to deduct and suggest improvements in the timing for each single shot type. With more training data of expert players an app could be envisioned that could automatically give recommendations on the timing to the player without the need of an outside trainer analyzing the shot.

ACKNOWLEDGMENT

This work was supported by the EU Marie Curie Network iCareNet under grant number 264738. The authors thank L. Nguyen-Dinh and M. Hardegger for providing the LCSS implementation in Matlab.

REFERENCES

[1] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," *Journal of neuroengineering and rehabilitation*, vol. 9, no. 1, p. 1, 2012.

[2] B. Mortazavi, S. Nyamathi, S. I. Lee, T. Wilkerson, H. Ghasemzadeh, and M. Sarrafzadeh, "Near-realistic mobile exergames with wireless wearable sensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 2, pp. 449–456, March 2014.

[3] S. Coyle, D. Morris, K. T. Lau, D. Diamond, and N. Moyna, "Textile-based wearable sensors for assisting sports performance," in *Wearable and Implantable Body Sensor Networks, 2009. BSN 2009. Sixth International Workshop on*, June 2009, pp. 307–311.

[4] M. Hardegger, B. Ledergerber, S. Mutter, C. Vogt, J. Seiter, A. Calatroni, and G. Tröster, "Sensor technology for ice hockey and skating," in *Wearable and Implantable Body Sensor Networks (BSN), 2015 IEEE 12th International Conference on*, June 2015, pp. 1–6.

[5] F. Gravenhorst, A. Muaremi, F. Kottmann, R. Sigrist, N. Gerig, C. Draper, and G. Tröster, "Strap and row: Rowing technique analysis based on inertial measurement units implemented in mobile phones," in *International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, 2014.

[6] M. Petkovic, W. Jonker, and Z. Zivkovic, "Recognizing strokes in tennis videos using hidden markov models." in *VIIP*, 2001, pp. 512–516.

[7] P. Kelly and N. O'Connor, "Visualisation of tennis swings for coaching," in *Image Analysis for Multimedia Interactive Services (WIAMIS), 2012 13th International Workshop on*, May 2012, pp. 1–4.

[8] D. Connaghan, C. O Conaire, P. Kelly, and N. E. O'Connor, "Recognition of tennis strokes using key postures," in *Signals and Systems Conference (ISSC 2010), IET Irish*, June 2010, pp. 245–248.

[9] H. Shah, P. Chokalingam, B. Paluri, N. Pradeep, and B. Raman, "Automated stroke classification in tennis," in *Proceedings of the 4th International Conference on Image Analysis and Recognition*, ser. ICIAR'07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 1128–1137. [Online]. Available: http://dl.acm.org/citation.cfm?id=2394970.2395083

[10] C. Ó Conaire, D. Connaghan, P. Kelly, N. E. O'Connor, M. Gaffney, and J. Buckley, "Combining inertial and visual sensing for human action recognition in tennis," in *Proceedings of the First ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams*, ser. ARTEMIS '10. New York, NY, USA: ACM, 2010, pp. 51–56. [Online]. Available: http://doi.acm.org/10.1145/1877868.1877882

[11] D. Connaghan, S. Hughes, G. May, P. Kelly, C. Conaire, N. O'Connor, D. O'Gorman, A. Smeaton, and N. Moyna, "A sensing platform for physiological and contextual feedback to tennis athletes," in *Wearable and Implantable Body Sensor Networks, 2009. BSN 2009. Sixth International Workshop on*, June 2009, pp. 224–229.

[12] D. Connaghan, C. O Conaire, P. Kelly, and N. E. O'Connor, "Recognition of tennis strokes using key postures," in *Signals and Systems Conference (ISSC 2010), IET Irish*, June 2010, pp. 245–248.

[13] A. Ahmadi, D. Rowlands, and D. A. James, "Towards a wearable device for skill assessment and skill acquisition of a tennis player during the first serve," *Sports Technology*, vol. 2, no. 3-4, pp. 129–136, 2009. [Online]. Available: http://dx.doi.org/10.1002/jst.112

[14] D. Knudson, D. Luedtke, and J. Faribault, "Performance: Excellence: How to analyze the tennis serve," *Strategies*, vol. 7, no. 8, pp. 19–22, 1994.

[15] A. Kuzmanic and V. Zanchi, "Hand shape classification using dtw and lcss as similarity measures for vision-based gesture recognition system," in *EUROCON, 2007. The International Conference on Computer as a Tool*, Sept 2007, pp. 264–269.

[16] M. Vlachos, G. Kollios, and D. Gunopulos, "Discovering similar multidimensional trajectories," in *Data Engineering, 2002. Proceedings. 18th International Conference on*, 2002, pp. 673–684.

[17] A. R. Jimenez, F. Seco, C. Prieto, and J. Guevara, "A comparison of pedestrian dead-reckoning algorithms using a low-cost mems imu," in *Intelligent Signal Processing, 2009. WISP 2009. IEEE International Symposium on*. IEEE, 2009, pp. 37–42.

[18] M. Hardegger, D. Roggen, S. Mazilu, and G. Tröster, "Actionslam: Using location-related actions as landmarks in pedestrian slam," in *Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on*, 2012, pp. 1–10.