Analyzing Tennis Game through Sensor Data with Machine Learning and Multi-Objective Optimization

Miha Mlakar
Jozef Stefan Institute
1000-SI Ljubljana, Slovenia
Miha.mlakar@ijs.si

Mitja Luštrek
Jozef Stefan Institute
1000-SI Ljubljana, Slovenia
mitja.lustrek@ijs.si

Abstract
Wearable devices are heavily used in many sports. However, the existing sports wearables are either not tennis-specific, or are limited to information on shots. We therefore added tennis-specific information to a leading commercial device. Firstly, we developed a method for classifying shot types into forehand, backhand and serve. Secondly, we used multi-objective optimization to distinguish active play from the time in-between points. By combining both parts with the general movement information already provided by the device, we get comprehensive metrics for professional players and coaches to objectively measure a player's performance and enable in-depth tactical analysis.

Author Keywords
Tennis; Wearable analytics; Shot detection; Optimization;

ACM Classification Keywords
I.2.6. Artificial Intelligence: Training

Introduction
The use of wearable sensors in sport is growing fast and can already be considered essential for success in some disciplines. In tennis the analytics started with

Peak_strength **calculation:**

- Calculate the absolute sum of angular speeds on the Roll and Yaw axes
- Raise it to the power 4, to emphasize higher values
- Apply Butterworth band-pass filter (1.5 and 25 Hz.)
- To get *Peak_strength* final value, set the lower peak to zero when two peaks are too close (min. distance = 1.3s set by a domain expert)

Calculation of other shot-specific features:

- Calculate the positions of intersections between the angular speeds on the Roll and Yaw axes
- Measure the times and calculate the surface areas between the current point and the previous and next intersection. These surfaces represent the accumulated speed before the impact with the ball (back swing) and after the impact (follow through).
- Use times, areas and ratios between them as features.

computer vision and sensors for measuring shots. However, both of these approaches have limitations for professional use. The sensors worn on the playing wrist or built into the tennis racquets deliver information about the shots [5] or enable the analysis and modeling of different shot techniques [7]. However, this information lacks context (under what circumstances and where on the court did a specific shot occur), so it is not sufficiently actionable, i.e. cannot be used for tactical preparations or to significantly improve players' game. Video analysis offers better information [1, 8], but cheap solutions provide low accuracy, while better solutions are extremely expensive because they require advanced cameras with complex software. Additionally, they are bound to specific courts, so the information is not available whenever needed by the player or coach.

Due to these limitations, devices worn on the torso, and equipped with accelerometers, gyroscopes and GPS receivers are emerging as the new approach. These devices are perfect for determining the effort, distance covered, sprints analysis and much more. A leading provider is Catapult Sports, whose S5 device is currently used by the best tennis player in the world Andy Murray. However, S5 offers no tennis-specific metrics. That is why in our research we add tennis-specific information to the metrics already available in the Catapult S5 system, to produce a comprehensive solution that enables professional players to make better tactical preparations and to improve their game.

Our approach consists of two steps. In the first step, we detect when a tennis shot occurs and which type of shot it is. In the second part, we focus on detecting when the players actually play points (active play) and when they are in-between points. This allows us to determine the actual net playing time and real distance covered, and adds context to shots which enable complex analysis like "Is the player playing weaker shots if the point is longer than 15 seconds?"

Data Acquisition

The Catapult S5 device was positioned high on the player's back attached to a tight shirt. The device contains a 3D accelerometer (frequency 100 Hz), 3D gyroscope (frequency 100 Hz), 3D magnetometer (frequency 100 Hz), and a GPS sensor returning latitude and longitude (frequency 10 Hz).

We recorded 5 different professional tennis players for 6 hours in total, including 1,373 shots. Each shot was labeled as a serve, forehand or backhand. As for detecting active play, we also manually labeled the beginning and end of each sequence of active play (rally). All the data were recorded during matches and none during predefined situations of practice sessions.

Methodology

For every data point obtained from our device, we extracted a number of features. These features were then used for shot and active play detection. Supervised machine learning was used for shot detection, and multi-objective optimization for detecting the beginning and end of rally.

We calculated several features for shot detection. In addition to the basic ones (average values, variances and standard deviations for each accelerometer and gyroscope axis for two different window sizes of 0.8 s and 1.2 s), we also calculated movement speed obtained from GPS coordinates, *Peak_strength* and other shot-specific features (explained in the sidebar).

Optimization rules

$(DV > p1)\ \&\ ((BV < p3)\ ||\ (FV < p4))$

Equation 1: The rule for detecting the beginning of a rally.

$(DV < p2)\ \&\ ((BV < p3)\ ||\ (FV < p4))$

Equation 2: The rule for detecting the end of a rally.

Results

	Precision	Recall
Forehand	91.5%	90.5%
Backhand	93.6%	90.6%
Serve	99.8%	98.2%
All	95.0%	93.1%

Table 1: Precision and recall for detecting types of tennis shots

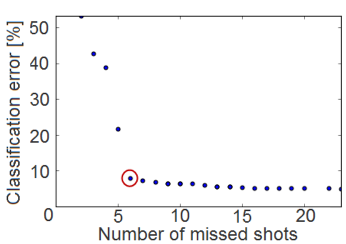

Figure 1: The final front showing the best solutions based on the classification error and the number of shots outside of detected active play.

The main idea when detecting the start (end end) point of a rally is that at this point the difference between the intensity of movement before and after would be the largest. This intensity was calculated as the sum of variance for all accelerometer axes for a window size that was subject to optimization. So for each data point, we calculated the back overall variance (BV), the forward overall variance (FV) and the difference between these two (DV = BV – FV).

We set two rules (Equation 1 and 2 in the side bar) for detecting the beginning (and end) of a rally. A data point is marked as the beginning (or end) of a rally if it satisfies the defined rules. The rule parameters $p1$, $p2$, $p3$ and $p4$ were determined through optimization in addition to the windows size.

Evaluation

The evaluation was divided into two parts. First evaluated how well we can detect if a shot occurred and which type of shot it was. For building the model we used the Random Forest algorithm. Each model consisted of 10 decision trees, the minimum number of samples required to split an internal node was 8, and the minimum number of samples required at a leaf node was 4. Since the results obtained with cross validation on sensor data can be misleading [2], we used the leave-one-player-out approach (LOPO), where we used one player's data for testing and the data from the other players for training. This approach enabled us to estimate the accuracy of the models for previously unseen players with different shot techniques. When evaluating the models, we classified each data entry as a type of shot or no-shot. With this approach almost all the data points were classified as no-shots, so calculating the classification accuracy would be useless. We therefore focused on the precision and recall.

For the optimization of rules defining the start and top of active play, we used the well-known evolutionary multi-objective optimization algorithm NSGA-II [1]. The population size was set to 25, the stopping criterion was set to 10,000 solution evaluations, and the tournament selection was used. We optimized two objectives, the classification error (for every data point, we detected if it was correctly predicted as play or no-play) and the number of shots that were missed (shots that were not inside active play sequence).

Results

The average precision and recall for detecting shot types presented in Table 1. The main sources of errors are fast unnatural body rotation movements and special events that occur during the play. An example from our data set is a player warming up doing very similar body movements as during shots.

For the optimization problem, the final front of non-dominated solutions showing a tradeoff between objectives can be seen in Figure 1. The solution on a knee of the front labeled with a circle has the classification accuracy of 93% and only 6 shots were left out of the predicted active play.

By combining the classified shot types, detected active playing phases and locations from the GPS, we can calculate several useful metrics that help remove subjectivity from the game and allow for objective evaluation of different tactical approaches and training routines. An example of such a view can be seen in Figure 2, where we present the heat map of a player's

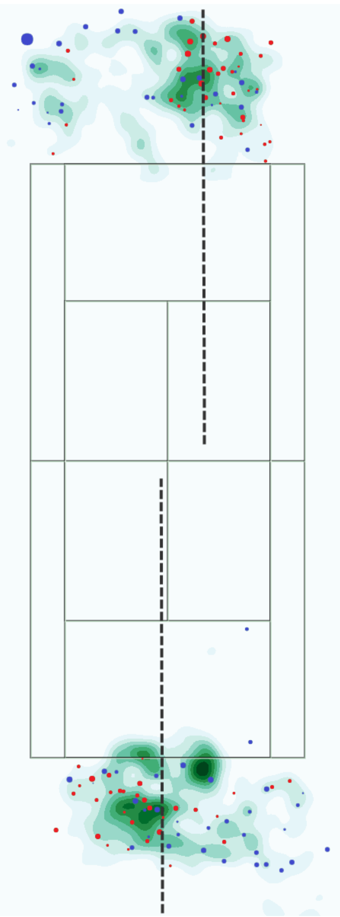

Figure 2: Heat map of a player's position during active play combined with shot locations (blue = forehand, red = backhand) and their *Peak_strength* (size of points).

position during active play and combine it with forehand and backhand shots as points of different color and size. We also included a dashed line that separates part of the court where more backhands are played from the one where more forehands were played. We can see that the player played more aggressively on the bottom side, thus his heat map is closer to the baseline. On the top side, we can see a less aggressive approach allowed a player to play more forehands and thus was able to control the game by playing more often with his better shot (forehand).

Conclusion

In this article we presented a two-step approach for analyzing wearable sensor data for professional tennis players. Firstly, we detected and classified different shot types, and secondly, we distinguished the active playing phases from the time in-between points. By combining the procedures, players can get a unique perspective on their game which enables objective analysis in the tactical and physical sense.

For the future, we plan to equip both players with the same type of sensor, and by measuring the time difference between their shots and by calculating the distance between them, we will be able to calculate the average speed of ball and thus additionally quantify the quality of each shot.

Acknowledgements

We would like to thank Catapult Sports for their sound advice and cooperation, and the national research agency (ARRS) for providing the research grant.

References

1. Deb, Kalyanmoy, et al. "A fast and elitist multi-objective genetic algorithm: NSGA-II." IEEE transactions on evolutionary computation 6.2 (2002): 182-197.
2. Hammerla, Nils Y., and Thomas Plötz. "Let's (not) stick together: pairwise similarity biases cross-validation in activity recognition." Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM, 2015.
3. Liaw, Andy, et al. "Classification and regression by RandomForest." R news 2.3 (2002): 18-22.
4. Liberty, Edo, Kevin Lang, and Konstantin Shmakov. "Stratified sampling meets machine learning." International Conference on Machine Learning. 2016.
5. Pei, Weiping, et al. "An embedded 6-axis sensor based recognition for tennis stroke." IEEE International Conference on Consumer Electronics (ICCE), 2017.
6. Renò, Vito, et al. "A technology platform for automatic high-level tennis game analysis." Computer Vision and Image Understanding (2017).
7. Yang, Disheng, et al. "TennisMaster: an IMU-based online serve performance evaluation system." Proceedings of the 8th Augmented Human International Conference. ACM, 2017
8. Yu, Xinguo, et al. "A trajectory-based ball detection and tracking algorithm in broadcast tennis video." Image Processing, 2004. ICIP'04. 2004 International Conference on. Vol. 2. IEEE, 2004.