Tennis Stroke Classification: Comparing Wrist and Racket as IMU Sensor Position

Christopher J. Ebner
christopher.ebner@students.fh-hagenberg.at
Department of Mobility & Energy, University of Applied Sciences Upper Austria
Hagenberg, Austria

Rainhard Dieter Findling
rainhard.findling@aalto.fi
Ambient Intelligence Group, Department of Communications and Networking, Aalto University
Espoo, Finland

ABSTRACT

Automatic tennis stroke recognition can help tennis players improve their training experience. Previous work has used sensors positions on both wrist and tennis racket, of which different physiological aspects bring different sensing capabilities. However, no comparison of the performance of both positions has been done yet. In this paper we comparatively assess wrist and racket sensor positions for tennis stroke detection and classification. We investigate detection and classification rates with 8 well-known stroke types and visualize their differences in 3D acceleration and angular velocity. Our stroke detection utilizes a peak detection with thresholding and windowing on the derivative of sensed acceleration, while for our stroke recognition we evaluate different feature sets and classification models. Despite the different physiological aspects of wrist and racket as sensor position, for a controlled environment results indicate similar performance in both stroke detection (98.5%-99.5%) and user-dependent and independent classification (89%-99%).

CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing**; *Visualization*; • **Computing methodologies** → *Machine learning*; • **Applied computing** → *Life and medical sciences.*

KEYWORDS

machine learning, tennis stroke detection, tennis stroke recognition, wearable sensors

ACM Reference Format:
Christopher J. Ebner and Rainhard Dieter Findling. 2019. Tennis Stroke Classification: Comparing Wrist and Racket as IMU Sensor Position. In *17th International Conference on Advances in Mobile Computing & Multimedia (MoMM '19), December 2–4, 2019, Munich, Germany.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3365921.3365929

1 INTRODUCTION

Wearable sensors can help tennis players analyze their tennis stroke technique, improve training effectiveness, and prevent injuries [2–4, 12, 14, 19, 27, 30]. Wearable technology is on the rise across sports in general, and in tennis its use is already allowed even during competitions [8], which in turn allows for reviewing critical information in a game during set breaks [33]. Quantifying the external workload, which is the amount of work performed by an athlete, is fundamental for the training process in sports. In tennis, the external hitting load (also called shot count, which means counting the shots and potentially also the type of shots) is an important part of the external workload. The current standard in tennis to measure the external hitting load is manual counting – which is time intensive, prone to human error and bias, and in some cases might not be possible at all [11, 29]. There are commercial products which provide counting and various other statistics for tennis. However, studies have indicated (cf. [16]) that the counting and classification performance of such systems might still be rather weak.

From a technical perspective, tennis strokes are mainly sensed with computer vision or wearable motion sensors [31]. The former can be expensive, complex, and their access can be limited [13]. Extracting relevant information about tennis from video sources is furthermore computationally expensive [15]. Wearable sensors however are already widely used in several sports [13] in general, still get integrated into more sport applications [33], have a great potential to detect sport movements [9], and have already been proven to be effective for biomechanic sports analysis [13]. Wearable sensor for tennis are thereby either placed in the racket handle, on the strings, on the grip, or on the player's wrist [18].

While previous work used sensors either on racket or wrist, those positions have not yet been directly compared to each other. This paper closes this gap: we compare wrist and racket as sensor positions for stroke detection and classification. To understand differences in data from wrist and racket we visualize and compare acceleration and angular velocity of 8+3 well-known types of tennis strokes. For stroke detection we employ a peak detection, thresholding, and windowing on the acceleration derivative [18]. For stroke recognition we evaluate 4 machine learning models with 3 feature sets, of which we use the results to assess the applicability of wrist and racket as tennis sensor position.

In our evaluation we utilize video data of a semi-professional tennis player to allow for error-free alignment and marking of the phases of the considered strokes (e.g. backswing, forward swing, the impact point of the ball, follow through, and the start and end of the execution [17, 31]) in the sensed acceleration and angular velocity data. Based on this synchronization, we use inertial measurement unit (IMU) sensor data of multiple players to quantify detection and segmentation performance, in which players vary the execution speed and the target location of their tennis strokes.

MoMM '19, December 2–4, 2019, Munich, Germany

ACM ISBN 978-1-4503-7178-0/19/12…$15.00
https://doi.org/10.1145/3365921.3365929

The corresponding dataset consists of 5 tennis players performing 8 different stroke types multiple times, resulting in a total of about 2000 stroke samples. For stroke classification, to compare wrist and racket as sensor position, each is assessed with multiple models and feature sets. Model training thereby relies on cross validation for parameter tuning and selection. Evaluation results are then presented in a comparative manner for both positions in a user-dependent and independent settings. Summarizing, the contribution of this paper are:

- We record and utilize data of semi-professional tennis players for our evaluation: of one player for combined video and IMU data, and of 5 players for extended IMU data.
- We visualize acceleration and angular velocity for 8+3 well-known tennis stroke types for both wrist and racket as basis for investigating their distinct stroke characteristics.
- We compare the performance of tennis stroke detection and segmentation for data sensed on wrist and racket.
- We extract 3 different feature sets and employ 4 different machine learning models to classify the extracted tennis strokes. We then use the best suited model to compare the stroke classification performance of wrist and racket.

2 RELATED WORK

In this section we review previous work on assisting tennis with wearable sensors, including tennis stroke detection and recognition with IMU sensors and machine learning, as well as other approaches utilizing concepts useful for our goal.

Video-based approaches were the first to be used in sport applications [13]. Video footage thereby serves feature extraction and classification of different stroke types [10, 24, 26, 32]. With IMU sensors becoming more popular, Ó Conaire et al. [22] used 6 on-body worn WIMUS sensors and various video cameras to detect and classify strokes, and to compare accelerometer and video as data for tennis. As video processing for tennis can be expensive, has a limited access, and as the size and costs of IMUs have reduced substantially, since then IMU sensors have become more widely used for sport applications [13].

Several previous work has utilized IMUs for detection and classification of tennis strokes. Also related domains have addressed similar problems with wearable sensors. For example, in table tennis, Blank et al. [5] detect and classify 8 different stroke types with a miPod sensor, which is mounted on the front-end of the handle of the table-tennis racket. In the tennis domain itself, there is different related work that classifies various tennis stroke types using IMU data and different sensor positions for the data collection. The majority of approaches use a sensor strapped on the wrist of the dominant arm [4, 6, 18, 20, 28, 29]. Connaghan et al. [10] use a similar position instead: on the middle of the forearm. Besides wrist, data sensed from the racket was used as well. Zhao et al. [31] place a sensor above the handle, and Pei et al. [23] embedded it into the handle. Mlakar and Luštrek [21] place a sensor on the upper back region of the handle. Three sensors are used by Büthe et al. [7] for stroke classification by placing a sensor on the racket and one on each foot, to additionally detect steps of the player. Ó Conaire et al. [22] use multiple sensors and place six sensors on both lower legs, both forearms, the chest, and the back.

As sensors, both self-built prototypes [18, 20, 22, 23] and commercial sensing systems [4, 6, 7, 21, 28, 29] have been used. Most frequently, the IMUs used process acceleration within $\pm 16g$, and the gyroscopes used a range of $\pm 2000 \frac{\text{deg}}{\text{s}}$ [18, 23, 29, 31]. Anand et al. [4] and Srivastava et al. [28] use a smartwatch sensor with an acceleration range of $\pm 8g$, Brzostowski and Szwach [6] one with $\pm 4g$. Büthe et al. [7] record the gyroscope data with a range of $\pm 500 \frac{\text{deg}}{\text{s}}$ instead of $\pm 2000 \frac{\text{deg}}{\text{s}}$. The utilized sampling rates range from 25Hz [28] over 100Hz [4, 6, 18, 21, 23, 31], 120Hz [22], 200Hz [7], to 500Hz [29].

For distinguishing strokes, different stroke types have been considered in literature. The most common are serve (also called smash in many publications), forehand, and backhand [6, 10, 18, 21, 22]. Further types include topspin and backspin, or topspin and slice [4, 7, 28]. Pei et al. [23] further distinguish serve and smash as two different strokes. Zhao et al. [31] divide their strokes into serve, groundstrokes, and volley. Whiteside et al. [29] distinguish 8 well-known strokes: serve, forehand, forehand slice, forehand volley, backhand, backhand slice, backhand volley, and smash. Additional to those, true negative samples (non-strokes) are recorded and trained as an additional stroke type.

There are different approaches to detect strokes in an IMU data stream. One is peak detection on acceleration data with a predefined threshold. Zhao et al. [31] use x-axis data, which is aligned with the handle of the racket, and consider spikes over 9g as a stroke. Connaghan et al. [10] and Ó Conaire et al. [22] set a threshold for the calculated acceleration magnitude of 3g and 8g. In addition to such thresholding, Whiteside et al. [29] add rules to eliminate false negative detections. Those include a minimum distance to adjacent maxima of 1.25 s, or alternatively gyroscope resultant values within 0.06 s on both sides of the acceleration maxima exceeding $\pm 400 \frac{\text{deg}}{\text{s}}$. Pei et al. [23] use a sliding window approach, where calculated feature values that defined the volatility of the selected data represented a stroke only if a defined threshold is reached. Where the previous approaches mainly use the lobe of the accelerometer data, Anand et al. [4] additionally consider the calculated energy of the gyroscope data for their threshold shot detection approach. The authors find that the shape of the acceleration on the x-axis and the g-energy data of the shots in tennis and squash have a particular form. This leads to Dynamic Time Warping (DTW) on a reference template being effective in detecting their strokes. Büthe et al. [7] use only the gyroscope data and discretize them with k-means. Kos and Kramberger [18] calculate the sum of the two-point derivative of every axes on acceleration data. After averaging, thresholding is used, in which each peak in the signal represented a stroke.

For segmenting strokes, most approaches use a simple time window around the detection point [4, 10, 22, 23, 29, 31]. To extract features from a detected stroke for subsequent classification, Zhao et al. [31] extract the mean, standard deviation (std), skewness, kurtosis, minimum, and maximum of each axis of the acceleration and gyroscope data, and additionally use the amplitude of both, which results in a total of 38 features. Whiteside et al. [29] extract 40 features including min, max, median, integral, and discrete value at the time of impact of every axis from the acceleration and gyroscope data, as well as their resultant. Bigger sets of features with subsequent feature selection are considered by Brzostowski and Szwach [6] and Anand et al. [4]. The former include Mel-Frequency

Cepstrum Coefficients (MFCC) of the acceleration data and a principal component analysis (PCA) for dimension reduction. The latter uses the data of the accelerometer and gyroscope to derive about 2000 features, including statistical features, pairwise correlation coefficients, and shape-based features for predefined stages in detected stroke. The authors apply a correlation-based feature selection which reduces their feature set to 300-500 distinct features.

Most stroke classifications use machine learning. However, other approaches exist as well. Kos and Kramberger [18] and Pei et al. [23] both use rule based methods. The former utilize minimum and maximum values of gyroscope data at the bi-point for their classification decision, the latter instead combine accelerometer, gyroscope, and the converted gravity data in Euler angles. Büthe et al. [7] and Büthe et al. [7] vote for a specific class with the longest common subsequence (LCSS) algorithm, while Srivastava et al. [28] at the first stage classify with DTW between forehand, backhand, and serve, and then use quaternion DTW (QDTW) to distinguish slice, flat, and topspin. Frequently used machine learning models for stroke classification include Naive Bayes (NB) [10, 31], support vector machine (SVM) [22, 29, 31], decision and classification tree (CT) [29, 31], random forest (RF) [21, 29, 31], k-nearest neighbor (KNN) [6, 22, 29], neural network (NN) [4, 29] and logistic regression (LG) [4, 6]. Further models include AdaBoost [31] and discriminant analysis [29].

In all previous work, no comparative visualization or interpretation of IMU-sensed data of different tennis strokes – with wrist and racket as sensor position – has been done yet. Visualizing those strokes would aid future research that explores further ways of recognizing them, for example by illustrating distinct characteristics in acceleration and angular velocity that define stroke types. Furthermore, no comparison between wrist and racket as sensor position for IMU sensors in tennis has been done yet. Since those two positions bring different sensing capabilities, it would be interesting to quantify their applicability for tennis stroke detection and classification.

3 APPROACH

Our method for detecting and classifying tennis strokes with IMU sensors comprises of data acquisition with acceleration and gyroscope sensing, stroke visualization, stroke detection and segmentation, and stroke classification. The corresponding details of each are stated in this section.

3.1 Data Acquisition

To record tennis strokes we use two XSens MTw-38A70G20 sensors. Their housing dimension are 34.5 x 57.8 x 14.5 mm (W x L x H), their weight is 27 g, and they are connected to one XSens Avinda Station for data transfer. We mount one sensor on the throat of the racket, and attach one to the wrist of the hand holding the racket (Fig. 1). For the wrist we use the sensor strap holder provided by XSens. For the racket, we use a self-built 3D printed sensor holder. The target of this holder was to bring the sensor as near as possible to the rackets balance point. To avoid distortion, the holder is constructed sturdily and strapped as tight as possible using four 4.6 mm laces. In contrast, the sensor on the wrist is strapped less tight to not

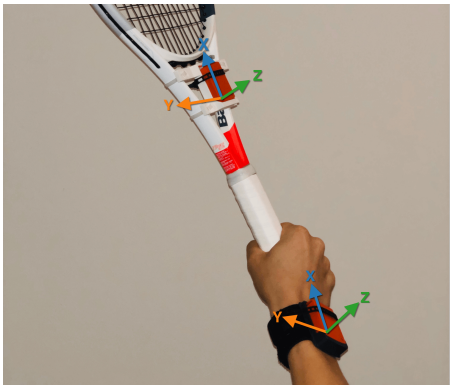


Figure 1: Sensor setup on wrist and racket.

constrict the arm of the player. For all our recordings we use a Babolat Pure Strike 16/19 racket.

For data validation and synchronization, we further record individual sequences of strokes with a 1080p 240 Hz video camera and a 720p 30 Hz video camera. The first serves as primary data source for precise synchronization of stroke parts with data sensed from the IMU sensors with one selected player. The second serves validating the sensor input data over the whole recording phase with all players to account for sensor hardware problems, players performing different than the planned strokes, and alike. To synchronize IMU sensor data with video data we use a synchronization gesture. In it, the player hits a tennis ball two times with the racket while facing the camera. This procedure is used at the start and the end of the recording.

Besides video, we record 13 data-streams, including the acceleration, angular velocity sensed by the gyroscope, free-acceleration, magnetic field, all in 3 axis, as well as the pressure. However, in preliminary experiments we found acceleration and angular velocity to best suited for our approach, which agrees with findings from literature, which is why we utilize only those two data sources in our evaluation. The gyroscope value range is $\pm 1200 \frac{\text{deg}}{\text{s}}$, the accelerometer value range is $\pm 160 \frac{\text{m}}{\text{s}^2}$, and their sampling rates are 100 Hz.

In our approach we collect four different types of data sets (Tab. 1). The *stroke visualization data set (DS1)* represents different stroke types, the *ball impact data set (DS2)* variations of ball speed and stroke speed of one specific stroke type, the *varied strokes data set (DS3)* the varying stroke execution and target ball direction of different stroke types, and the *stroke classification data set (DS4)* the data for the subsequent machine learning.

We consider two different sets of strokes in our approach. The *big stroke type set (ST1)* contains 11 different stroke types: forehand (fh), backhand one- (bh1h) and two-handed (bh2h), forehand volley (volleyFh) and slice (sliceFh), backhand volley (volleyBh) and slice (sliceBh), serve straight (serveStr), slice (serveSli), and kick (serveK), and smash (smash). In the *small stroke type set (ST2)*, the one- and two-handed backhand strokes are grouped into one class (backhand), as well as the three serve types are group to one class (serve). Therefore, ST2 contains only 8 instead of 11 different stroke classes.

Table 1: Data sets used in our evaluation.

	DS1	DS2	DS3	DS4
Nr. of stroke types	11	1 (fh)	8	8
Repetitions	5	5	20	50
Stroke movement	constant	slow, fast	variation	variation
Ball impact	constant	no, slow, fast	variation	constant
Video recording	240 Hz, 30 Hz	240 Hz, 30 Hz	240 Hz, 30 Hz	30 Hz
Different players	1	1	1	5

3.1.1 Stroke visualization data set (DS1). DS1 contains the acceleration and angular velocity data of five tennis strokes recordings, for each stroke type considered in ST1 performed 5 times, which results in a total of 55 samples. Data is recorded with a semi-professional tennis player. Each recording includes the synchronisation gesture and the corresponding stroke five times, and is also recorded in video (240 Hz). The ball is thrown to the player with constant speed by hand, and the player performs a representative stroke for the specific type to answer the ball.

3.1.2 Ball impact data set (DS2). DS2 differs from DS1 in that only one stroke type (forehand) is performed by the player. The different recordings vary in the movement of the stroke itself, which can either be light or strong, and the impact speed of the incoming ball, which can either be slow, fast, or no ball. This results in 6 different possible settings of ball impact.

3.1.3 Varied strokes data set (DS3). For DS3 the 8 stroke types of ST2 are recorded with a synchronization gesture by the same player as DS1 and DS2. For each recording the player performs 20 repetitions of the corresponding stroke while varying stroke movement and target direction of the ball.

3.1.4 Classification data set (DS4). DS4 contains data of 5 experienced players, which each perform each stroke type in ST2 around 50 times. 4 of the players are male and one is female. The mean age of the participants is 23 years (std 4 years) and their height is 179.2 cm (std 4.0 cm). All players play two-handed backhands and the tennis grips vary between participants and stroke type (from continental grip to western grip). Their ITN-Austria rating, which indicates the skill level of tournament player from 1 (high) to 10 (low), ranges from 2.1 to 8.7 (mean 5.6, std 2.4). DS4 contains around 400 stroke samples per player, and a total of about 2000 stroke samples. All players did vary their stroke movements, and for serve class of strokes, they used all three stroke subtypes. The balls were tossed to the player with a ball machine, which results in minimal variation in ball speed and direction. Due to the amount of samples in DS4, errors in data recording, such as players performing a different stroke type, performing a non-relevant movement with the arm, or twirling the racket in the hand, are present in DS4. For data cleaning we manually synchronize and compare video and IMU data and remove errors. Errors not removed include if a player rotates the racket in the hand during the recording. This leads to having IMU data of one sensor having a rotated orientation, resulting in a rotated coordinate system of sensed data. To revert this rotation we invert the y- and z- axis of the accelerometer and gyroscope streams for rotated samples.

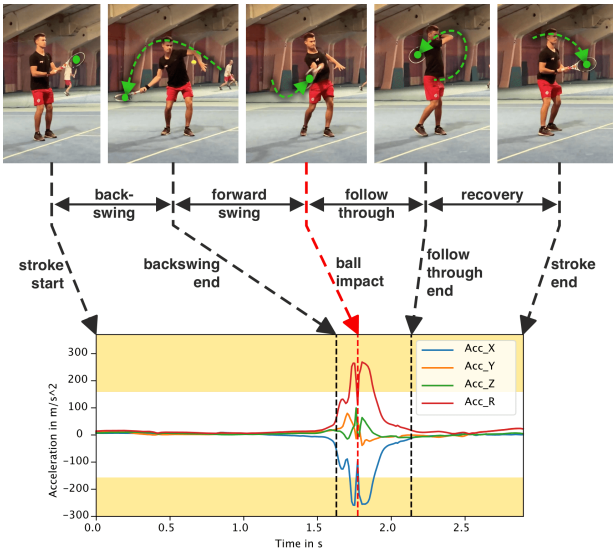


Figure 2: Images of the video footage at the detected stroke phase times.

3.2 Stroke Visualization

The goal of our data visualization of tennis strokes for wrist and racket is to better understand distinguishing key characteristics in their acceleration and angular velocity. We deem this to be important for achieving good subsequent preprocessing, in both our own work as well as future work in tennis stroke recognition. For our visualizations we use DS1 and choose one representative stroke of every stroke type. For each of those stroke samples we extract the exact start and end time of stroke phases and label them in the corresponding 240 Hz video footage of the stroke (Fig. 2). Those include stroke start, backswing end, ball impact, follow through end, and stroke end. Since the utilized sensors are only specified to work up to 16g, acceleration outside that area – which possibly delivers unreliable values – is marked with a yellow background.

3.3 Stroke Segmentation

Our stroke detection and segments automatically finds and extracts strokes from only the continuous stream of IMU sensor data, without requiring video data. In our recordings we found acceleration to reliably show a significant value change in a short period of time when the ball hits the racket. We therefore apply peak detection on the derivative averaged accelerometer data. At first we calculate the derivative of all three accelerometer axes individually, then average the axes. On the resulting one-dimensional stream we use a max value detection algorithm to extract all ball impact points. Peaks thereby need to be apart from each other by a predefined window to avoid the same peak being detected multiple times. Peaks further need to have a minimum height, which we solve with a threshold a peak needs to surpass. To obtain a robust threshold, we derive our threshold on empirical basis from DS3, in which stroke samples are performed with varying style, speed, and ball direction. The determined threshold value is $\pm 15 \frac{m}{s^3}$ (Fig 3, threshold depicted as horizontal red line), which in DS3 leads to 100% correctly detected

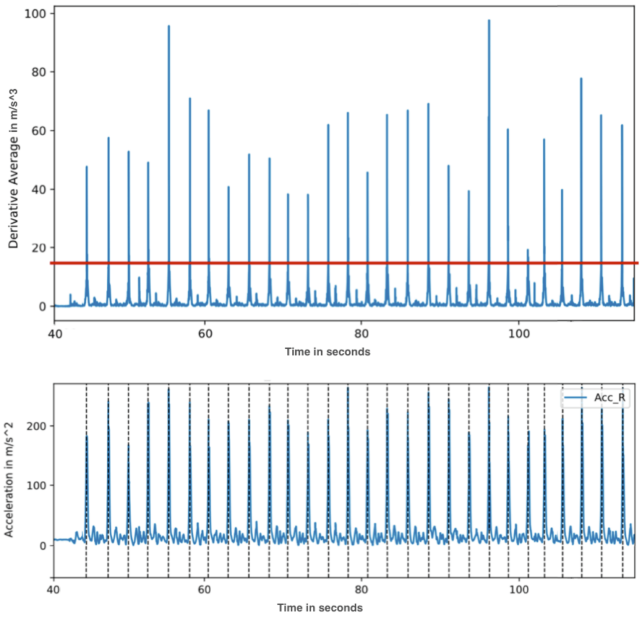


Figure 3: Stroke detection with a backhand slice sample sensed on the racket. Upper figure: derivative of averaged acceleration (blue) and peak threshold (red). Lower figure: acceleration (blue) and the detected peaks (dashed black) depicting the ball impact points.

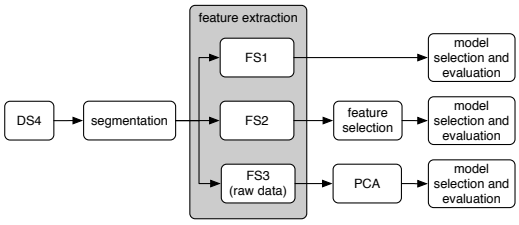


Figure 4: Stroke classification processing chain.

ball impact points for both sensor positions, which furthermore have been assessed to also agree 100% with the video recordings of DS3. While the upper part of Fig. 3 depicts the acceleration derivative, the lower parts depicts the actual acceleration of strokes detected with our peak detection and thresholding. After detecting stroke peaks we segment the IMU data around the peak to obtain stroke samples. For this segmentation we use a 1 s window which results in samples containing 0.5 s of data both before and after detected peaks.

3.4 Stroke Classification

Based on our stroke detection we employ a classification to recognize the stroke type. In our classification evaluation we evaluate a total of 4 different models with 3 different feature sets for both wrist and racket as sensor position (Fig. 4).

For this we at first detect strokes in DS4 and extract both accelerometer and gyroscope data for our samples. We then extract 3 feature sets (FS1-FS3) for each sample, which we subsequently individually evaluate and compare results of. For FS1 we extract the same features as Whiteside et al. [29], which include min, max, median, integral, and the discrete sensor value at ball impact time. FS2 is an extended version of FS1 that also includes standard deviation (std), skewness, kurtosis, innerquartile range, frequency bands, as well as the zero and mean crossing rate. In contrast to FS1 and FS2, FS3 consists of only the raw accelerometer and gyroscope sensor values. Applying this feature extraction on all 6 considered axes of tennis stroke samples together with the two corresponding magnitudes, this leads to samples with FS1 containing a total of 40 features, 480 features with FS2, and 800 features with FS3. We subsequently perform feature selection on FS2 and FS3. With FS2 we employ a mutual info classifier [25] which reduces the features to 100. With FS3 we employ principal component analysis (PCA) for dimension reduction. We thereby use a threshold of 99% of variance preserved in the data, which results in 75 principal components remaining as the features in FS3.

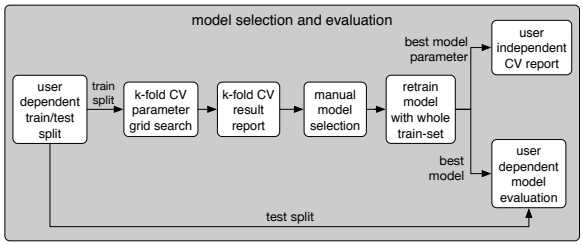


Figure 5: Model selection and evaluation processing chain.

Table 2: Hyparparameter ranges used during grid search.

Model	Hyperparameter range
SVM linear	C=[3^{-20}, 3^{2}]
SVM rbf	C=[3^{-7}, 3^{6}], γ=[3^{-10}, 3^{-1}]
KNN	neighbors=floor([1.5^{0}, 1.5^{9}])
CART	max_depth=[2, 15]

We use multi-stage data partitioning in our model training and evaluation (Fig. 5). We at first do a 80/20 train/test split for a user-dependent model evaluation due to the limited amount of participants in our data. Within the training partition we employ a 10-fold cross validation (10CV) with a hyperparameter grid search for each model type (Tab. 2) to determine the best suited hyperparameter set per model for our classification. This is done for both wrist and racket as sensor position.

The hyperparameter sets that yield the highest stroke classification accuracy in the 10CV are selected separately for wrist and racket. They then are re-verified with a double-CV using a second 10CV on the same data with different partitions to compensate for selection bias. Subsequently, the two models are trained with the whole training set for user-dependent evaluation, which is then done with the whole test set. For the user-independent evaluation we use a leave-one-subject-out CV on all data and average the results over the corresponding left-out participants.

4 EVALUATION, RESULTS, AND DISCUSSION

4.1 Stroke Visualization

We use the stroke data visualization example (see appendix, Fig. 9 and 10) with the data of one player as both basis to investigate if different strokes might be hard to detect or to classify in our work, as well as basis for future work to investigate further stroke characteristics. The player who originated the data uses the western grip [1] for forehand strokes and a standard one- and two-handed grip otherwise.

The accelerometer values of the x-axis in general are higher than the other two axes, independent of the stroke type. We assume this is caused by the centrifugal force of the arm rotation over the whole stroke phase. The strongest accelerometer and gyroscope changes furthermore seem to occur in the forward swing and follow through phase of a stroke. Those phases vary together from 0.5 s-1.08 s (mean 0.78 s, std 0.17 s), with only the time of the smash exceeding 1 s. As expected, major peaks in acceleration data are visible when the ball hits the racket (marked with dotted red lines in the figures).

Because of the major value change at the ball impact point all strokes seem easily detectable. A time window of 1 s second (0.5 s before and after the ball impact point) further seems to cover all major data changes corresponding to the stroke.

4.1.1 wrist vs. racket. Acceleration sensed on the racket seems to in general be stronger than on the wrist. At the visualizations of the overhead stroke types, which include the three different serves and the smash, the acceleration peaks of the wrist are higher than the one of the racket. Nevertheless, because the maximal recording value of the sensors is 16g and the curve of the forward swing and follow through is broader, we assume that the data of the racket also exceeds the wrist values.

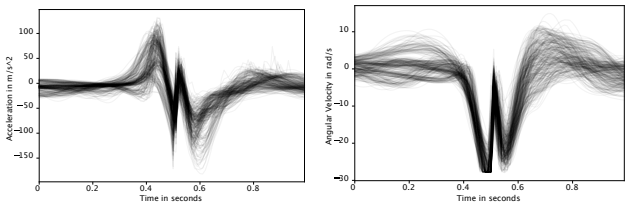
4.1.2 fh vs. bh1h. We recognize, that the racket acceleration data of the z-axis from the forehand is characteristically similar to the corresponding backhand data, but horizontally mirrored. This might be, due to the ball hitting the racket on the other side of the hitting surface to which the z-axis of the sensor is normal to. The gyroscope racket data show similar characteristics between the two strokes for the z-axis and the same mirrored occurrence as mentioned before for the y-axis. The gyroscope data from the wrist shows the same characteristics as the racket data.

Whereas the characteristics of specific axes are mirrored as well in the acceleration as in the gyroscope data, the forehand seems to be well distinguishable from the one-handed backhand.

4.1.3 bh1h vs. bh2h. These two stroke types have very similar characteristics in all data streams. The acceleration data of the z- and y-axis from the two-handed backhand have in this case higher values. This can be spotted in both sensor data, racket and wrist.

Because of the similarities of these two stroke types, we assume to be able to group them to one stroke type backhand (bh).

4.1.4 serveStr, serveSl, serveK, and smash. Also, these four stroke types look very similar, based on their characteristics. The types also have high acceleration values on the x-axis which exceeds the 16g sensor range.



(a) z-axis accelerometer (b) y-axis gyroscope

Figure 6: Segments backhand strokes from all participants.

In order to create the two stroke types serve and smash for the classification process, the different serves in our assessment can be grouped to one stroke type, but the distinction between serve and smash could be a problem.

4.1.5 volley vs. normal. In general, in the volley stroke types we see lower sensor values than in the normal forehand and backhand graphs. The acceleration z-axis values of the volley strokes have the same characteristics as the values of the normal strokes. It seems that the volley strokes have more similar max values over all three axes, whereas for the normal stroke types the x-axis of the acceleration data shows more dominance. For the gyroscope data we observed, that for the forehand the x- and z-axis have other characteristics than the normal strokes whereas for the backhand strokes only the data of the racket z-axis show differences.

For these two types we expect to be able to distinguish between them, regarding the max acceleration values in relation to all three axes and because of the different characteristics of the gyroscope data.

4.1.6 slice vs. volley. Compared to the volley strokes, the stoke type slice shows very similar characteristics over all data streams.

Because of the similarities we assume problems to differentiate between these two stroke types.

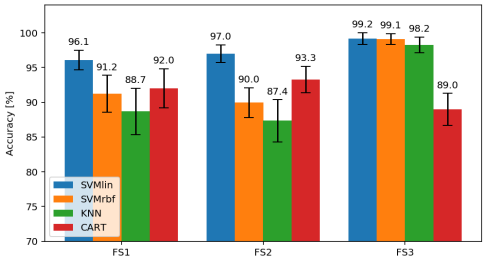
4.2 Stroke Segmentation

For the evaluation of the stroke detection and segmentation we use DS4. We visualize the detected ball-impact points together with the acceleration data of the recordings. From those we derive true positives (TP) as the correct detections of strokes, false positives (FP) as erroneous detections of non-strokes, as well as false negatives (FN) as erroneous non-detection of strokes. The resulting accuracies are then calculated on the corresponding TP, FP and FN rates for both wrist and racket as sensor position.

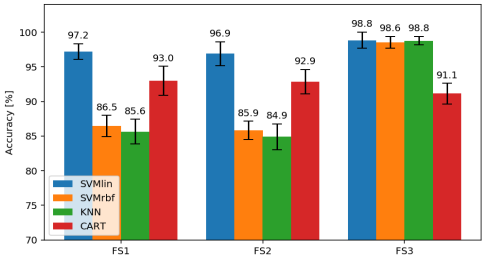
With our approach we therefore are able to obtain 98.47% detection accuracy for wrist and 99.44% racket as sensor position over all stroke types (Tab. 3). When distinguishing stroke types two groups become visible: the first includes forehand volley, backhand volley, and the slice backhand with wrist data. Those show a higher FN-rate (FNR) caused by the threshold being too high to detect certain strokes. The second contains serve and smash with racket data, which show higher FP-rate (FPR). This is caused by the racket hitting the floor or the body of the player after the follow through. The corresponding impact is wrongly detected as a ball impact.

Stroke type	wrist	racket	wrist & racket
fh	100%	99.7%	99.9%
bh	100%	100%	100%
serve	100%	97.8%	98.9%
volleyFh	95.8%	100%	97.9%
volleyBh	94.9%	99.7%	97.3%
sliceFh	99.3%	100%	99.6%
sliceBh	98.3%	99.7%	99.0%
smash	99.5%	98.7%	99.1%
average	98.28%	99.44%	98.86%

Table 3: Stroke detection accuracies for different stroke types (rows) and sensor positions (columns).



(a) Models for wrist data



(b) Models for racket data

Figure 7: Cross-validation results for model selection.

After the stroke detection, a fixed sized window around the ball impact point is segmented from the used data streams. The resulting strokes over all players and samples are visualized in an example for backhand strokes in Fig. 6(a), which shows the z-axis of the accelerometer, and Fig. 6(b), which shows the y-axis of the gyroscope. All samples show clear alignment and key characteristics when centered at the ball impact point (the clearly visible peak). The same is applied to all other stroke types, which allows us to generalize from Fig. 9 and 10 to other players.

4.3 Stroke Classification

For our classification evaluation we use strokes extracted from DS4.

Fig. 7(a) and 7(b) depict the best accuracy per model type from the 10CV evaluation from which the best model is chosen for wrist and racket as sensor position. A linear SVM model performs best for both positions with $C = 3^{-8}$, using FS3. The average accuracies of FS1 and FS2 are nearly equal, with 91.3% and 91.0%. However, using FS3 outperforms both with with 96.6%, effectively reducing the error to less than half. The mean accuracy over all wrist models is 93.4%, while racket achieves a comparable value with 92.5%.

We then evaluate the linear SVM models in a user-dependent manner on the test set (Fig. 8(a) and 8(b)), and in a user-independent manner with leave-subject-out-cross-validation over all data (Fig. 8(c) and 8(d)). The user-dependent test set yields an accuracy of 99.55% for the wrist. 2% of the volley forehands and the slice forehands are wrongly confused with each other. In comparison, racket test data yields an accuracy of 98.23%, with some confusion between backhand volleys and slices as well as between serves and smashes.

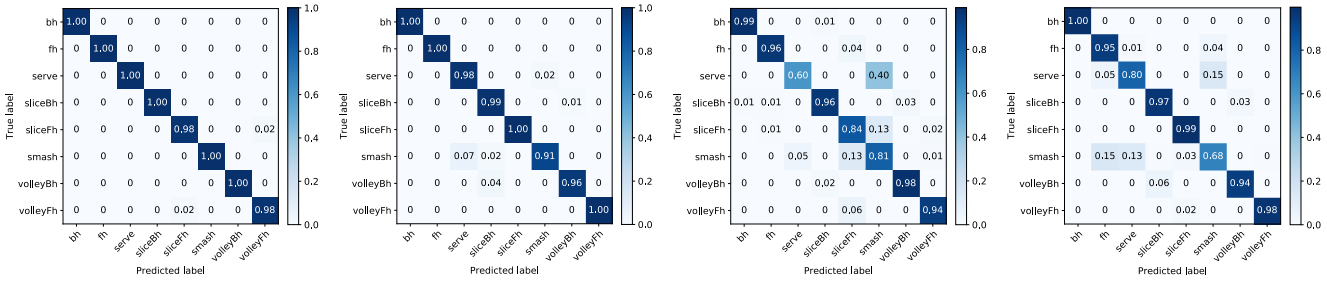
In contrast, the user-independent set yields worse results. This is expected due to the classification being applied to users unseen during training, which indicates the ability to generalize to new players. Wrist results indicate an accuracy of 89.30% and confusion is mostly between serve and smash: with wrist data, 40% of serves are wrongly classified as smashes. Further confusion is between forehand slice and smash, as well as forehand slice and volley. In comparison to wrist, racket as sensor position seems to be slightly more robust: especially less serves are confused (15% as smashes, 5% as forehands). However, more smashes are confused as serves (13%) and forehands (15%).

4.4 Summary

Our visualization of well-known tennis strokes indicates a clear rapid acceleration change at the time of the ball impact to the racket. This applies to all stroke types and for both wrist and racket as sensor position. A 1 s window around the ball impact (0.5 s before to 0.5 s after) covers the characteristic value changes of acceleration and angular velocity data corresponding to the stroke. The agreement of multiple samples of the same stroke type, over different players, is largest at the ball impact point. When comparing wrist and racket, acceleration and angular velocity values are in general lower for wrist than racket. When comparing strokes, forehand and one-handed backhand are well distinguishable. However, one and two-handed backhands are more difficult to distinguish. This also applies to distinguishing serves from smashes and slices from volleys.

With our peak detection based stroke detection and segmentation, all 8 considered stroke types can be detected well for both the wrist and racket sensor position, with a mean accuracy of 98.9% (std 0.9%). False positives are mostly caused after serves and smashes, when the racket hits the body or the floor after the follow through. False negatives are caused by a too high threshold, which mostly affects volleys or light slices.

For stroke type classification we found a SVM with a linear kernel to be most effective, together with raw acceleration and angular velocity features and PCA for dimension reduction to 75 features per stroke sample. This setup outperformed other, hand crafted feature sets by around 6% accuracy. A subsequent validation with user-dependent test data yields around 98% accuracy for both sensor positions. This indicates that both wrist and racket are comparable when models are trained for the individual user. A further user-independent evaluation yields accuracies around 90%, where the racket seems to perform slightly better than the wrist. This indicates that if models are not trained for individual users but are meant to generalize to new users, then a higher error of around 10% should



(a) User-dependent wrist (b) User-dependent racket (c) User-independent wrist (d) User-independent racket

Figure 8: Confusion matrices for wrist and racket, assessed in a user-dependent and user-independent way.

be expected for both wrist and racket. The most frequent confusion thereby occurs between serve and smash, smash and volley, and slice and volley.

5 CONCLUSION

In this paper we visualized acceleration and angular velocity data of 8+3 well-known tennis stroke types, and compared wrist and racket as sensor position for stroke classification. We recorded 11 different stroke types for both sensor positions. For stroke detection, we use a peak detection on the averaged derivatives of the data of the three acceleration axes. Those peaks occur at the ball impact time and are clearly visible and detectable. We then segment a 1 s window around the peak, which in our data yields around 2000 stroke samples over all types. Our stroke classification compares 4 machine learning models types (SVM linear and radial, KNN, and CART) and 3 feature sets (2 hand-crafted and one with raw sensor values and subsequent PCA).

Our results indicate that tennis strokes in general can be detected well for both wrist and racket (nearly 99% detection accuracy). In our setup we used a single threshold for both positions, which lead to a small amount of false positives. These could likely be avoided in future work using two independent thresholds. In stroke classification, our user-dependent results indicate that strokes can be distinguished well (around 98% for both wrist and racket). This indicates a) that if models are trained for individual users recognition can be expected to be good, and b) that neither wrist nor racket did bring significant advantages over each other for stroke classification with our evaluation setup. Our user-independent test indicates a lower accuracy of around 90% for both wrist an racket, where racket slightly outperforms wrist. This indicates that the generalization of models beyond the users they are trained for (e.g. new tennis players using the system without training) should be expected to cause significantly higher error rates, likely due to differences in execution of tennis strokes. Future work could extend this investigation with a dataset of more players and different skill levels, as well as different play styles. We expect such data to contribute to the generalization ability of the resulting models. It could also investigate if there are stroke types beyond the ones considered in our work for which either racket or wrist do bring significant advantages in stroke type recognition, which seem to not be the case for the ones considered in our work.

ACKNOWLEDGMENTS

We appreciate funding in the frame of the ChistEra RadioSense Big Data and process modeling for Smart Industry (BDSI) project.

REFERENCES

[1] Geoffrey D. Abrams, Per A. Renstrom, and Marc R. Safran. 2012. Epidemiology of musculoskeletal injury in the tennis player. *British Journal of Sports Medicine* 46, 7 (2012), 492–498. https://doi.org/10.1136/bjsports-2012-091164 arXiv:https://bjsm.bmj.com/content/46/7/492.full.pdf

[2] Amin Ahmadi, David Rowlands, and Daniel Arthur James. 2009. Towards a wearable device for skill assessment and skill acquisition of a tennis player during the first serve. *Sports Technology* 2, 3-4 (2009), 129–136. https://doi.org/10.1080/19346182.2009.9648510 arXiv:https://doi.org/10.1080/19346182.2009.9648510

[3] Amin Ahmadi, David D. Rowlands, and Daniel A. James. 2010. Development of inertial and novel marker-based techniques and analysis for upper arm rotational velocity measurements in tennis. *Sports Engineering* 12, 4 (01 Aug 2010), 179–188. https://doi.org/10.1007/s12283-010-0044-1

[4] Akash Anand, Manish Sharma, Rupika Srivastava, Lakshmi Kaligounder, and Divya Prakash. 2017. Wearable Motion Sensor Based Analysis of Swing Sports. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, IEEE, 261–267. https://doi.org/10.1109/ICMLA.2017.0-149

[5] Peter Blank, Julian Hüssbach, Dominik Schuldhaus, and Bjoern M. Eskofier. 2015. Sensor-based Stroke Detection and Stroke Type Classification in Table Tennis. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers (ISWC '15)*. ACM, New York, NY, USA, 93–100. https://doi.org/10.1145/2802083.2802087

[6] Krzysztof Brzostowski and Piotr Szwach. 2018. Data Fusion in Ubiquitous Sports Training: Methodology and Application. *Wireless Communications and Mobile Computing* 2018 (2018), 1–14. https://doi.org/10.1155/2018/8180296

[7] Lars Büthe, Ulf Blanke, Haralds Capkevics, and Gerhard Tröster. 2016. A wearable sensing system for timing analysis in tennis. In *2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 43–48. https://doi.org/10.1109/BSN.2016.7516230

[8] Valentina Camomilla, Elena Bergamini, Silvia Fantozzi, and Giuseppe Vannozzi. 2018. Trends Supporting the In-Field Use of Wearable Inertial Sensors for Sport Performance Evaluation: A Systematic Review. *Sensors* 18, 3 (Mar 2018), 873. https://doi.org/10.3390/s18030873

[9] Ryan Chambers, Tim J. Gabbett, Michael H. Cole, and Adam Beard. 2015. The Use of Wearable Microsensors to Quantify Sport-Specific Movements. *Sports Medicine* 45, 7 (01 Jul 2015), 1065–1081. https://doi.org/10.1007/s40279-015-0332-9

[10] Damien Connaghan, Phillip Kelly, Noel E. O'Connor, Mark Gaffney, Michael Walsh, and Cian O'Mathuna. 2011. Multi-Sensor Classification of Tennis Strokes. In *SENSORS, 2011 IEEE*. IEEE, 1437–1440. https://doi.org/10.1109/ICSENS.2011.6127084

[11] Emily E. Cust, Alice J. Sweeting, Kevin Ball, and Sam Robertson. 2019. Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance. *Journal of Sports Sciences* 37, 5 (2019), 568–600. https://doi.org/10.1080/02640414.2018.1521769 arXiv:https://doi.org/10.1080/02640414.2018.1521769 PMID: 30307362.

[12] R Dhinesh, Preejith Sp, and Mohanasankar Sivaprakasam. 2018. Tennis Serve Correction using a Performance Improvement Platform. In *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)*. IEEE, 1–7. https://doi.org/10.1109/SeGAH.2018.8401370

[13] Hugo G. Espinosa, Jim Lee, and Daniel A. James. 2015. The inertial sensor: A base platform for wider adoption in sports science applications. *Journal of Fitness*

Research 4, 1 (2015), 34. https://doi.org/10.3390/sports6020034

[14] Takahiro Ishikawa and Toshiyuki Murakami. 2015. An Approach to 3D Gyro Sensor Based Motion Analysis in Tennis Forehand Stroke. In *IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 002354–002359. https://doi.org/10.1109/IECON.2015.7392454

[15] Holger Junker, Oliver Amft, Paul Lukowicz, and Gerhard TrÃűster. 2008. Gesture spotting with body-worn inertial sensors to detect user activities. *Pattern Recognition* 41, 6 (2008), 2010 – 2024. https://doi.org/10.1016/j.patcog.2007.11.016

[16] Eliza M. Keaney and Machar Reid. 2018. Quantifying hitting activity in tennis with racket sensors: new dawn or false dawn? *Sports Biomechanics* 0, 0 (2018), 1–9. https://doi.org/10.1080/14763141.2018.1535619 arXiv:https://doi.org/10.1080/14763141.2018.1535619 PMID: 30540215.

[17] Duane Knudson and Bruce Elliott. 2004. *Biomechanics of Tennis Strokes*. Springer US, Boston, MA, 153–181. https://doi.org/10.1007/978-1-4419-8887-4_7

[18] Marko Kos and Iztok Kramberger. 2017. A Wearable Device and System for Movement and Biometric Data Acquisition for Sports Applications. *IEEE Access* 5 (2017), 6411–6420. https://doi.org/10.1109/ACCESS.2017.2675538

[19] Marko Kos and Iztok Kramberger. 2018. Tennis Stroke Consistency Analysis Using Miniature Wearable IMU. In *2018 25th International Conference on Systems, Signals and Image Processing (IWSSIP)*. IEEE, 1–4. https://doi.org/10.1109/IWSSIP.2018.8439382

[20] Marko Kos, Jernej Ženko, Damjan Vlaj, and Iztok Kramberger. 2016. Tennis Stroke Detection and Classification Using Miniature Wearable IMU Device. In *2016 International Conference on Systems, Signals and Image Processing (IWSSIP)*. IEEE, 1–4. https://doi.org/10.1109/IWSSIP.2016.7502764

[21] Miha Mlakar and Mitja Luštrek. 2017. Analyzing Tennis Game Through Sensor Data with Machine Learning and Multi-objective Optimization. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers (UbiComp '17)*. ACM, New York, NY, USA, 153–156. https://doi.org/10.1145/3123024.3123163

[22] Ciarán Ó Conaire, Damien Connaghan, Philip Kelly, Noel E. O'Connor, Mark Gaffney, and John Buckley. 2010. 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 (ARTEMIS '10)*. ACM, New York, NY, USA, 51–56. https://doi.org/10.1145/1877868.1877882

[23] Weiping Pei, Jun Wang, Xubin Xu, Zhengwei Wu, and Xiaorong Du. 2017. An Embedded 6-axis Sensor based Recognition for Tennis Stroke. In *2017 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 55–58. https://doi.org/10.1109/ICCE.2017.7889228

[24] Milan Petkovic, Willem Jonker, and Zoran Zivkovic. 2001. Recognizing Strokes in Tennis Videos Using Hidden Markov Models. In *IASTED International Conference on Visualization, Imaging and Image Processing, VIIP 2001*. Citeseer, 512–516.

[25] Brian C. Ross. 2014. Mutual Information between Discrete and Continuous Data Sets. *PLOS ONE* 9, 2 (02 2014), 1–5. https://doi.org/10.1371/journal.pone.0087357

[26] Hitesh Shah, Prakash Chokalingam, Balamanohar Paluri, Nalin Pradeep, and Balasubramanian Raman. 2007. Automated Stroke Classification in Tennis. In *Image Analysis and Recognition*, Mohamed Kamel and Aurélio Campilho (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 1128–1137.

[27] Manish Sharma, Rupika Srivastava, Akash Anand, Divya Prakash, and Lakshmi Kaligounder. 2017. Wearable Motion Sensor Based Phasic Analysis of Tennis Serve for Performance Feedback. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 5945–5949. https://doi.org/10.1109/ICASSP.2017.7953297

[28] Rupika Srivastava, Ayush Patwari, Sunil Kumar, Gaurav Mishra, Laksmi Kaligounder, and Purnendu Sinha. 2015. Efficient Characterization of Tennis Shots and Game Analysis using Wearable Sensors Data. In *2015 IEEE SENSORS*. IEEE, 1–4. https://doi.org/10.1109/ICSENS.2015.7370311

[29] David Whiteside, Olivia Cant, Molly Connolly, and Machar Reid. 2017. Monitoring Hitting Load in Tennis Using Inertial Sensors and Machine Learning. *International journal of sports physiology and performance* 12, 9 (2017), 1212–1217.

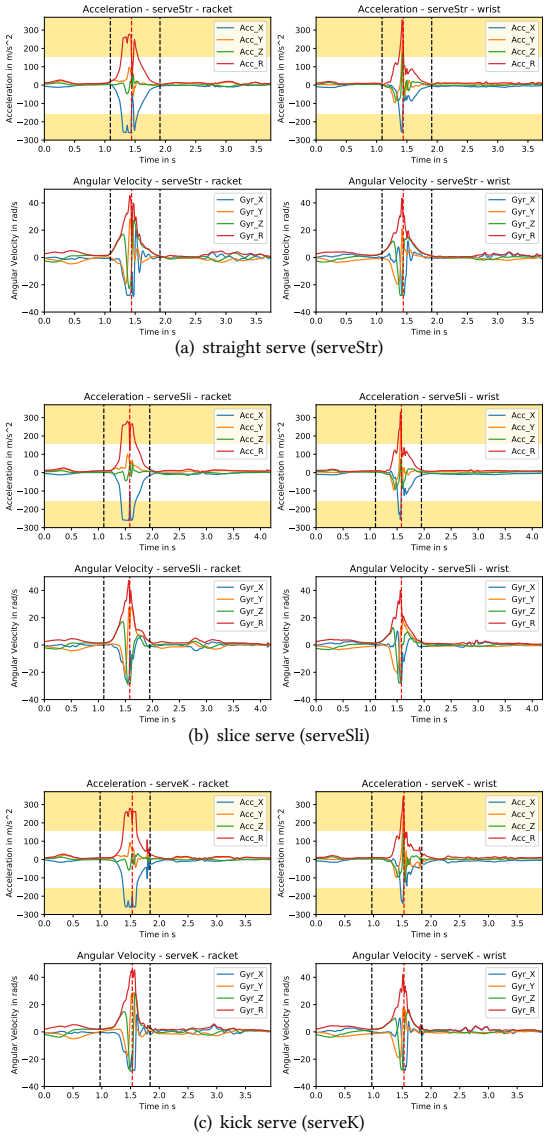
[30] Disheng Yang, Jian Tang, Yang Huang, Chao Xu, Jinyang Li, Liang Hu, Guobin Shen, Chieh-Jan Mike Liang, and Hengchang Liu. 2017. TennisMaster: An IMU-based Online Serve Performance Evaluation System. In *Proceedings of the 8th Augmented Human International Conference (AH '17)*. ACM, New York, NY, USA, Article 17, 8 pages. https://doi.org/10.1145/3041164.3041186

[31] Hongyang Zhao, Shuangquan Wang, Gang Zhou, and Woosub Jung. 2019. TennisEye: Tennis Ball Speed Estimation Using a Racket-mounted Motion Sensor. In *Proceedings of the 18th International Conference on Information Processing in Sensor Networks (IPSN '19)*. ACM, New York, NY, USA, 241–252. https://doi.org/10.1145/3302506.3310404

[32] Guangyu Zhu, Changsheng Xu, Qingming Huang, and Wen Gao. 2006. Action Recognition in Broadcast Tennis Video. In *18th International Conference on Pattern Recognition (ICPR'06)*, Vol. 1. IEEE, 251–254. https://doi.org/10.1109/ICPR.2006.205

[33] Mounir Zok. 2014. Inertial sensors are changing the Games. In *2014 International Symposium on Inertial Sensors and Systems (ISISS)*. IEEE, 1–3. https://doi.org/10.1109/ISISS.2014.6782518

APPENDIX

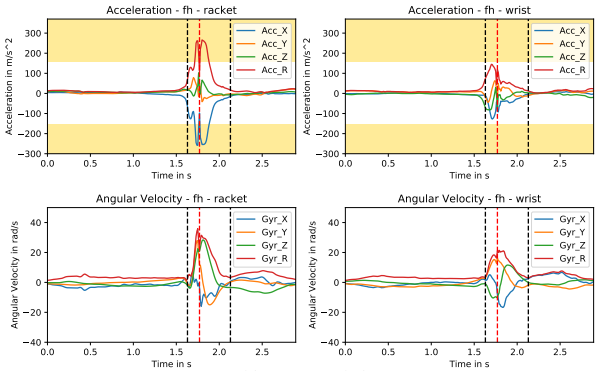


(a) straight serve (serveStr)

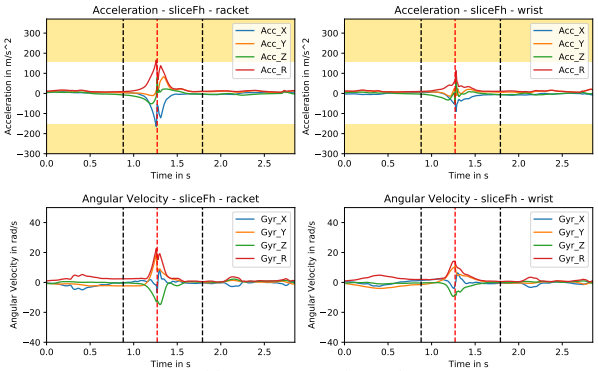
(b) slice serve (serveSli)

(c) kick serve (serveK)

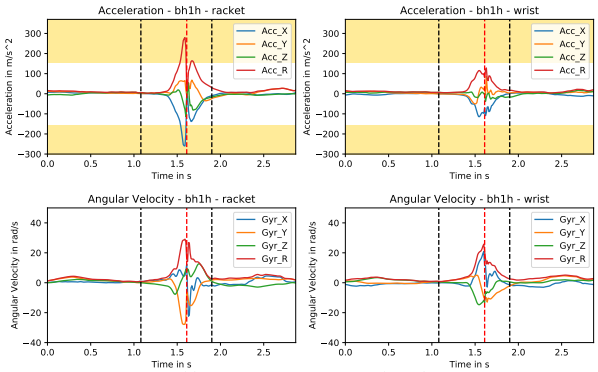
Figure 9: Visualization of the three serve types.



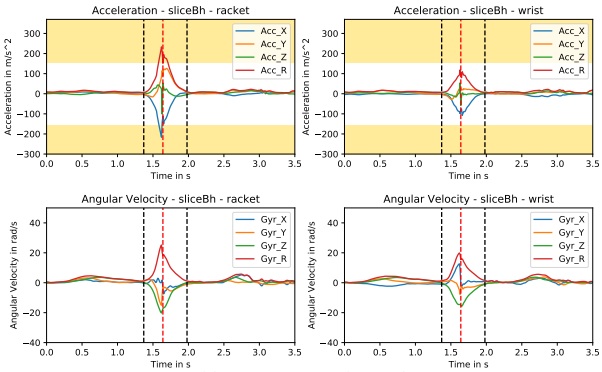
(a) forehand (fh)



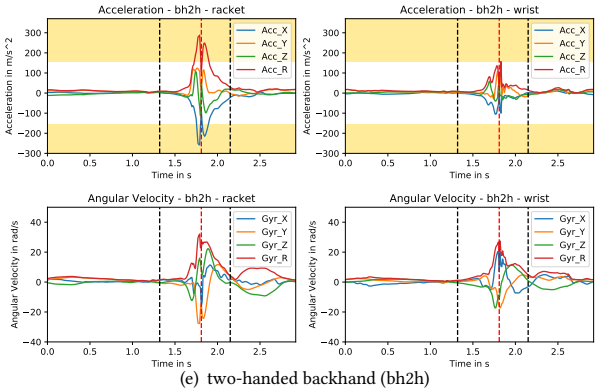
(b) forehand slice (sliceFh)



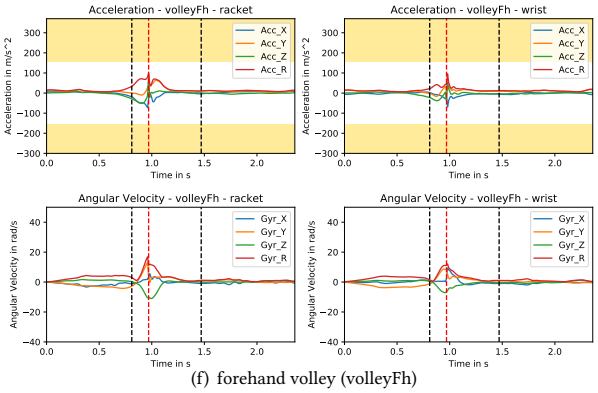
(c) one-handed backhand (bh1h)



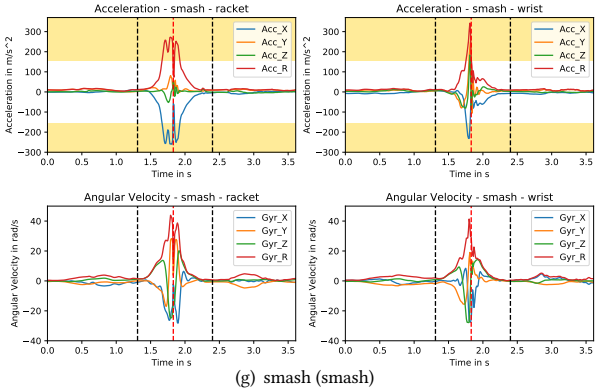
(d) backhand slice (sliceBh)



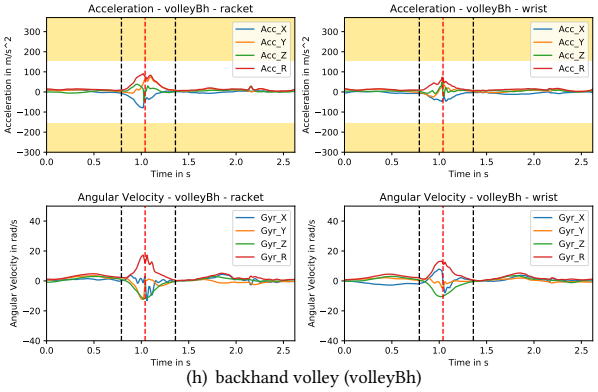
(e) two-handed backhand (bh2h)



(f) forehand volley (volleyFh)



(g) smash (smash)



(h) backhand volley (volleyBh)

Figure 10: Visualization of all 11 stroke types excluding serve.