



Multi-Sensor Fusion for Online Detection and Classification of Table Tennis Strokes

Habiba Hegazy^{1*} Mohamed Abdelsalam¹ Moustafa Hussien¹
 Seif Elmosalamy¹ Yomna Hassan¹ Ayman Nabil¹ Ayman Atia^{2,3}

¹Faculty of Computer Science, Misr International University, Egypt

²HCI-LAB, Faculty of Computers and Artificial Intelligence, Helwan University, Egypt

³October University for Modern Sciences and Arts (MSA), Egypt

* Corresponding author's Email: habiba1611146@miuegypt.edu.eg

Abstract: Sports training generally focuses on speed of response and variety of strategies aimed at encouraging sustainable physical activity and improving learning skills for players, thus, enhance their performance and skills through matches. In this paper, we are presenting a methodology for multi-sensor fusion method of IR depth camera and smart band sensors using curve fitting to recognize and group the characteristics of strokes performed by players and adapt the training accordingly. The main aim of the methodology is to classify various techniques played in table tennis and enhance the strokes based on various body joints. Moreover, the main contribution of this paper is to experiment different sensors to get the most optimal tools in classifying players strokes. Also, to test various classification algorithms to get the optimal and best result possible. Overall, based on the experiments we have concluded that sensor fusion between internal sensors and IR depth camera has increased the classification results and robustness of the solution. The system's results indicate an average accuracy of 95% - 100%.

Keywords: Stroke classification, Stroke identification, Table tennis, IR depth camera, Hand gestures, Smart band, Accelerometer, Gyroscope, Inner room server, Multi-sensor, Sensor fusion.

1. Introduction

Competing sports are the major focusing of diverse audiences around the world. Table Tennis has over 300 million active participants globally and is a popular game. Anyone can get involved at any time and keep playing all the time. It is particularly good for developing awareness and movement. One of the main reasons behind this popularity is that people of all ages can play it. Table tennis is now a global sensation. The highest popularity of the game is in Asia. China is considered one of the best-performing countries in table tennis. Also, this sport not restricted for men only as according to sports and fitness industry association shows that 40% of total table tennis participants are female. Table tennis movements are categorized into basic and advanced stages. The basic stroke movements are push, drive, and topspin. The forehand stroke is generally played

from the forehand position, but players are often allowed to use their forehand stroke against the balls arriving at the centre. On the other side, the backhand drive is regularly played from the backside. Players are not normally allowed to play backhand strokes from their forehand, which may contribute to weak technique. Adjusting the wrist to different incoming balls can create optimal arcs. The arcs are formulated by the instant strength and speed of the player. Mostly there are lots of movements either forehand or backhand that are based on the correct movement of the wrist. One of the main wrist error movements is ball risen, which means that the racket is moving from the bottom to the top, the wrist is the joint that controls the racket's movement. In another way, shifting the wrist from the bottom to the top. Fig. 1 (a) indicates the incorrect movement.

In our attempt to increase the accuracy we have

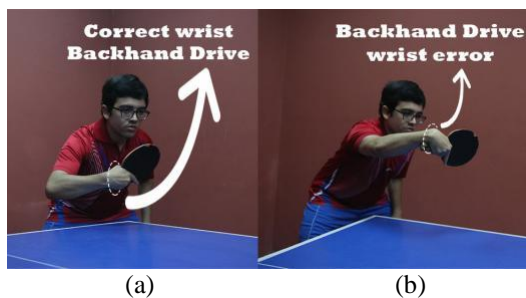


Figure. 1 Incorrect waist movement: (a) backhand drive and (b) forehand drive

started with our previous work [1], we have experimented with using an IR depth camera but it wasn't able to achieve accurate results on its own and face some limitations to detect the wrist movement in an appropriate way. In order to enhance the results, we are proposing a new methodology of using curve fitting for multi-sensor fusion to obtain better accuracies for detecting and classifying the stroke motion of the player. The details of the methodology will be explained in the following partitions. The paper is partitioned into four parts. Firstly, the related work where we present other researches in different domains related to the techniques used in the proposed system. Secondly, the methodology where we mainly explain the proposed approach and the techniques used. Thirdly, experiment where we show the proposed system contribution and results. Fourthly, a discussion where we investigate the basis of the introduced platform. Finally, summary and future studies where the results and outcomes of the research are presented.

2. Related work

This section is divided into several parts. Every part is mandatory to mention as it is used in our system.

2.1 Hand gestures

Table tennis is a relatively fast-moving sport that can involve the fastest responses in any other sport. The motion of the player's hand can extremely affect the strokes. It is enough to identify and recognize tennis strokes by capturing hand development. Therefore, recognition and identification of a table tennis strokes were treated as a problem for hand motion as table tennis relies on hand movement [2]. Hand motion identification and recognition is a well-known field for human computer applications. Research [3] introduced a very effective system of recognition and identification of hand movements with online feedback.

Hand motions can be distinguished using an IR depth camera too, Y. Chen, B. Luo, Y. Chen, G. Liang, and X. Wu [4] provide a reliable real time replacement to massive HCI devices using Kinect and SVM algorithms, which resulted in 95.42% precision. One of many sensors were used in the table tennis field to classify and recognize hand gestures was IMU. According to M. Kos, J. Zenko, D. Vljaj, and I. Kramberger [5] which indicates that using some simple calculations, you can reach a very good accuracy in general table tennis strokes. According to K. Xia, H. Wang, M. Xu, Z. Li, S. He, and Y. Tang [6] they proposed a framework of a wristband (Accelerometer and gyroscope sensors) and a hybrid multi-layer cluster model for activity detection by using different clustering algorithms, feature extraction and segmentation processes. Their proposed model achieved the highest overall accuracy with 86%. Moreover, according to Z. Bankosz and S. Winiarski [7] they proposed a system to be examined in the topspin forehand by using an IMU device which is a wearable inertial sensor to estimate and measure the movements of the player's hand and analysing different positions. There observation indicates that the highest variability is in the wrist joint extension. Overall, IR depth camera as a sensor was used lots in sports field and other fields too. It was important to get data for different joints but also sensors like accelerometer and gyroscope indicates to be really useful for small joints.

2.2 IR depth camera sensor background

As the fundamental concentration from the innovation and technology viewpoint in popular sports is the classification and detection of the player's developments. Identification of motions mostly completed by using a camera particularly an IR depth camera or a wearable sensor device. According to T. Le, M. Nguyen, and T. Nguyen [8] the usage of depth camera with the SVM algorithm to target the human posture recognition did achieve high precision in several interesting places. Their findings reveal that Kinect is a low-cost system that can track and identify multiple positions with high precision. Overall, the usage of IR depth camera got lots of limitations such as artificial vibration, which might affect the results of classifications especially of sensitive joints. Therefore, looking to attached sensors to body might overcome this limitation.

2.3 Accelerometer and gyroscope sensors background

The main two sensors placed in mobile phones, smartwatches, smart bands are an accelerometer, and

gyroscope sensors. There are many applications of accelerometers and gyroscopes to evaluate physical performance [9]. Assigning a hardware component to the racket is one of the methods using accelerometer and gyroscope for recording player's movements and that was the study and research method [10], they did experiment that using these sensors for detecting the movements can be easily and gives sufficiently precise accuracy. Using the same concept and methodology [11] reached to an accuracy of 95.7%. Research [12] attached the mobile phone into the player's wrist to use mobile accelerometer and gyroscope sensors. They achieved an average accuracy of 69.63% and 77.21%. They also created an online android program to examine the played strokes of the player. Several problems were faced while using an accelerometer and gyroscope without an IR depth camera. The main problems were the precise time and synchronization of different positions [10]. On the other hand, while using only accelerometer and gyroscope. P. Blank, J. Hossbach, D. Schuldhau, and B. Eskofier [11] faced a challenge of detecting events that can't be allocated to each of the types of stroke listed, this is because of the usage of accelerometer and gyroscope on a small joint which is the wrist. In other words, in table tennis to achieve a very correct and accurate movement, it is based on different joints and can't be classified using only a single joint. The data gathered from different sensors does contain noises, in order to remove them we need to filter the data.

2.4 Filtration process

The usage of IR depth cameras and wearable sensors leads to a high rate of distortion that affects the records of data and the system recognition [13]. Thus, filtering is the main pre-processing necessary for sensors. For Kinect, due to the light-based systems, the data would have noise. According to G. Li, H. Tang, Y. Sun, J. Kong, G. Jiang, D. Jiang, B. Tao, S. Xu, and H. Liu [14] a joint bilateral filter was utilized to label the collected images, therefore, increasing the quality of the images. On the other hand, for the wearable sensors (accelerometer and gyroscope) noise might be due to temperature change, shock, or vibration. The noise needed to be reduced well as it affects the classification [15]. Moreover, according to [16] Kalman filter did achieve a good rate of denoise different sensors (IR depth camera, accelerometer, and gyroscope). Kalman filter was utilized to monitor the joints of the body with the elimination of the noise from the unwanted pulses also decreases the difference in the middle place of the joint [17]. Right before the classification process,

the implementation of sensor fusion is needed to avoid problems of single sensor usage.

2.5 Sensor fusion

To avoid problems of single usage of IR depth camera or accelerometer and gyroscope as mentioned in [10, 11], sensor fusion was necessary, especially in a real-time application. According to [18] fusion between Kinect and internal sensors (accelerometer and gyroscope) did enhance the system's accuracy. This was achieved in different system states the position, velocity, and acceleration. Merging between Kinect and Internal sensors did help in increasing the classification accuracy to be approximately 98% [19]. In the field of tennis, [20] proposed an efficient real-time analysis system focused on the combination of motion and sensor vision. A detailed survey of various sensor fusion techniques utilized in detecting various human activities have been presented in a detailed survey [21]. K. Irie discussed the idea of sports swing measurements, he used an IMU and a camera to measure motions. Overall, he indicates that the method of sensor fusion decreased the errors significantly in inertial motion calculations and reasonable additional costs [22]. At the end of the process the classifier is one of the main methods in system which lots of key feature can depend on the type of classifier used.

2.6 Hand gesture classification algorithms

After signals intake and connection of the sensors, the remaining part is to classify the movements and strokes. Many and different algorithms were used in different hand gestures application. Using data mining and machine learning algorithm was a real focus approach in some researches. The usage of k-NN was one approach for the identification of hand movements and to increase the recognition process, where F. Mufarroha and F. Utamingrum [23] reached an accuracy of 80.77% into investigating hand gesture recognition. Moreover, classifying table tennis strokes Using SVM with an attached inertial sensor in a racket would reach a rate of 96.7% [11]. Another research implemented a methodology that differentiates different fixed hand gestures in a complex context by using the Naive Bayes classifier and Gabor filter reaching an accuracy of 90% [24]. Another important approach in hand gesture was the use of neural network algorithms. According to S. Shin and W. Sung [25] they used to classify hand gesture movements on accelerometer data using recurrent neural network (RNN) algorithm. They indicate that RNN-based solutions provide positive

performance and can be further enhanced by using more training details.

In addition, the dynamic time wrapping algorithm is one of the most used algorithms in the field of hand gestures, it acts as a key classifier for real-time systems. It is a strategy that calculates an ideal fit for two groupings of information with diverse imperatives. With the recognition of human behaviour, dynamic time warping and IR depth cameras have been used for human development due to their effect on contrast with speed or style while conducting activities [26]. [27] mentioned by the usage of DTW with Kinect they could reach an accuracy of 97.8% for recognizing hand gestures. According to [28] DTW results in high accuracy in classification of a system relies on a system combination between Kinect, accelerometer, and gyroscope.

Table 1 shows that generally the previous work that has been done in the field of table tennis was only focusing on the main stroke types and on classification of the motion using different wearable sensors or sensors attached to the racket, but neither of those work focused on the mistakes made while player's training based on different joints the only joint of focus was the Wrist and the feedback was given offline after the training session. Although, focusing on the player's different joints rather than the racket might potentially lead to more accurate results. In our previous work [1] we proposed a technique to recognize and identify the utility of movements executed by the players using different algorithms and focused on an IR depth camera. To

enhance the algorithms classification results in some table tennis strokes and based on observations of data IR depth camera got some limitations with small joints as wrist which is a very important joint in table tennis. Therefore, in this work we present a methodology that relies on the multi-sensor fusion of IR depth camera and smart band. Recognizing and group the characteristics of strokes performed by players and adapt the training accordingly depending on breaking down the data obtained from IR depth cameras and smart band sensors.

3. Methodology

The system is divided into different parts. Firstly, is the data acquisition where data is collected for different joints through different sensors. Secondly, the pre-processing part which consists of segmenting strokes, data filtration, and sensor fusion between IR depth camera, accelerometer, and gyroscope. Thirdly, the feature extraction part to identify key features in the data. Finally, the processing part where the classification is made to classify the strokes.

3.1 Data acquisition

Data acquisition is carried out using the IR depth camera SDK to distinguish various joints (Elbow, Shoulder, and Waist) as shown in Fig. 2 from the player's body along with the sensors readings from the smart band (accelerometer and gyroscope) worn by the player around his/her wrist. The data collected from each sensor are three-dimensional points (X, Y, Z). IR depth camera gets three-dimensional points for each joint. The data collected is transmitted to the inner room server connected to both sensors.

3.2 Pre-processing

Pre-processing step measures for automated learning help remove unnecessary and redundant data, increase the accuracy of learning, and improve the understanding ability of performance [29]. The data are collected from the sensor during this step and filtered before the feature extraction phase.

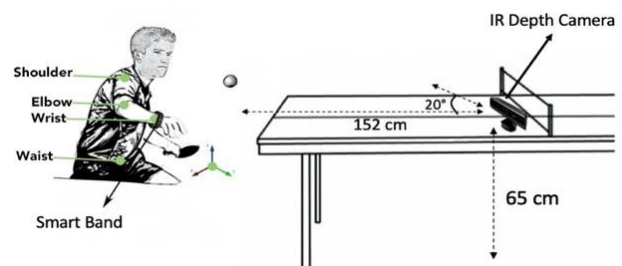


Figure. 2 The system tracks and records the player's joints and its movements related to table tennis strokes.

Table 1. System comparison

Ref. No.	Sensor used	Feed-back	Joint of focus	Stroke type
[7]	IMU device	Off-line	Wrist	Serve, forehand and backhand.
[11]	miPod sensor	Off-line	Wrist	Forehand and backhand push, block and topspin
[12]	Mobile Sensors	On-line	Wrist	General forehand and backhand
Prop. Sys.	IR depth camera and smart band sensors	On-line	Waist, Wrist, Shoulder and elbow	Forehand and backhand push, drive and topspin.

3.2.1. Segmenting strokes

The stroke should be segmented from the collected data stream after automated data collection. Over table tennis stroke stages, the program calculates the Euclidean distance among the primary starting point at the stage of passes to. The Euclidean distance is used to distinguish and detect each stroke from the sequence gained from the sensors. Euclidean distance tests the length of a section linking the two points.

3.2.2. Filtering data

Kalman filter is utilized for IR depth camera noise removal. Kalman filter utilized filter for noise elimination and motion tracking. It is used to improve the accuracy of joints position estimation of IR depth camera and displayed an acceptable pre-processing and accuracy enhancement as mentioned in [25], and fusing it with other sensors increases the accuracy as mentioned in [14]. Two types of noise affect the IR depth camera time-series data, which are because of high-frequency fluctuations and extracting joint position from the skeleton. Our collected data passed through the Kalman filter for improving joint position estimation of IR depth camera and removing the noise that affects the data like temperature change, gravity, vibration, etc. Moreover, the Kalman filter assists to remove the noise of the accelerometer and gyroscope that might be caused due to vibration or temperature. It displays an acceptable pre-processing and accuracy enhancement as mentioned in [13]. Kalman filter contains a set of mathematical formulas that predicts of the inner condition from a variety of distorted calculations for a linear dynamic system and provides an estimation of unknown variables, which appear to be more accurate than the ones dependent on a single calculation by calculating a combined distribution of probability around the variables in each time-frame. A linear stochastic differential equation is used to model a discernible time cycle.

3.2.3. Sensor fusion

Like most sensors they are not immune to measurement error. Besides the presence of drift, other events (such as rapid movements, occlusions, disturbance in the Bluetooth communication, optical or magnetic noise) can lead to false or missing measurement data from one of both sensors. Since the wrist is the most important joint in detecting and classifying the stroke, by using wearable sensors along with the IR depth camera, more measurement data of the wrist joint were added to the system to compensate for the missing or false data using an

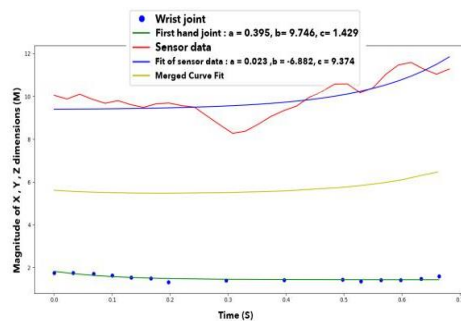


Figure. 3 Curve fitting and fusion between accelerometer, gyroscope and depth camera sensors.

algorithm that fuses information from wearable sensors and IR depth camera. After collecting and filtering the data from different sensors that contain data related to a single stroke. Due to the difference between the data frequency acquired from the IR depth camera and the wearable sensors. The process of the sensor fusion starts with fitting for every dimensional stream of data into a curve that is ideally matched to a variety of data points, this process is normally considered only as a part of the conditioning process. Curve correction is a form of regression analysis that is a collection of statistical processes to estimate the relationship between dependent variables, also referred to like the outcome variables. Then, we used multiple equations and models of linear and non-linear regression and compare them to determine which model provides the best fit for our data.

Fig. 3, illustrates a model of the process on the wrist joint data from the correct forehand drive stroke. The blue point is the position of wrist joint acquired from IR depth camera and the red curve is the data acquired from the sensors (Accelerometer/Gyroscope). The green curve shows the fitted curve for Kinect data, the same goes for the blue curve which shows the fitted curve for sensors data (Accelerometer/Gyroscope). The process of curve fitting gives us the ability to get data from both IR depth camera and sensors at the same exact time, then we merge them into one curve and that is the yellow curve. The algorithm utilizes the least-squares approach for optimization to approximate the solution of equation sets by minimizing the size of residual squares in each equation results. The default algorithm is Levenberg–Marquardt for unconstrained problems and Trust Region reflective algorithm if bounds are provided. In the end, it returns a list of these optimal values for the parameters, and a two-dimensional list contains the covariance calculated for these ideal values. The diagonals show the parameter estimation variance. The system uses these two lists in forming a set of a combined stream of data

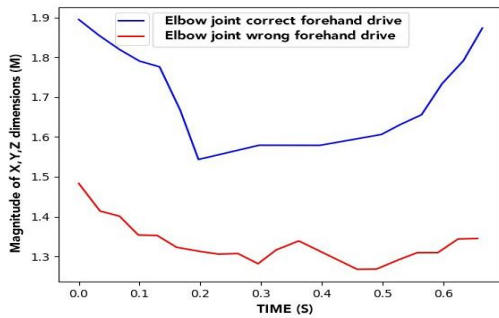


Figure. 4 Correct/wrong forehand drive graph (elbow joint).

for each dimension and estimates and predicts all the data series needed from all sensors in the same timestamp as shown in Fig. 4 graph.

3.2.4. Curve fitting methodologies

As curve fitting is our main methodology for sensor fusion, it was really important to obtain it with the best curve fit possible. The following Table 2 shows all equations tested in our system and the accuracy obtained along with the graph created by each equation. The usage of the curve fitting equations depends on the type of data [30]. According to the proposed system, the data is considered to be real-time dimensional data. At Table 2 the constraints is given by default as we have utilized Scipy library to implement the proposed system. Moreover, Fig. 5 shows that Eq. (1) at Table 2 was the most optimal equation used in the proposed system. This was determined by creating a method to calculating the distance between the curve points for each number between the range of constrains. The shortest overall distance between the points was considered to be the optimal.

3.3 Feature extraction

The selection of features increases the understanding of the mechanism and help to obtain high performance of learning algorithms [29]. Through this phase it was found that the observation of the collected data and the initial raw data set was reduced for the pro-cessing phase to more

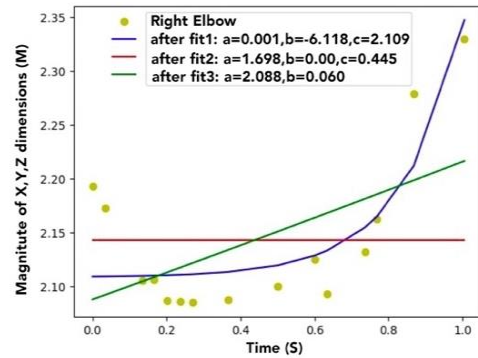


Figure. 5 Curve fitting equations graph.

manageable categories. Feature extraction is used to classify essential data characteristics and to decrease the number of resources available for evaluation without lacking specific or significant information. Multiple extractions of feature techniques were used such as magnitude, standard deviation, peek-to-peek amplitude, mean, median, minimum, and maximum.

3.4 Processing

The data extracted is subsequently recognized with the data set through FastDTW after the pre-processing process. The dynamic time warping (DTW) algorithm allows for the optimal synchronization of the double pair. It is also used to define time series parallels, classifications, and determine the correct regions between two periods. FastDTW is a DTW approximation that has a linear structure in time and space. FastDTW utilizes a multilevel approach which helps to resolve a solution based on a coarse resolution and refines the proposed solution. The FastDTW starts to make the cost matrix for every player's stroke and every dataset stroke. FastDTW proceeds the process of determining the minimum value between both points and their neighbours in the matrix between the test stroke focuses and each stroke. Furthermore, The FastDTW utilizes cost matrix analysis and greedy search to literally get the distance. FastDTW ordinarily begins by including a point from the upper left of the cost grid cell to recognize the two strokes. The FastDTW guarantees that the correct balance between each stroke in the dataset and the stroke tried is accomplished. At last, FastDTW starts to scan for the base parity separation. Consequently, as classified the stroke mark potentially acquired from the dataset. Although, in order to test the efficiency of FastDTW algorithm, we compared the algorithm with k-NN, SVM, logistic naïve bayes and RNN which were implemented in their default structures and used in previous work [4, 23-25]. The comparison of the

Table 2. Curve fitting equations

Eq.	CF Equations	Constraints	Graph Fig. 6 Curve
1	$a \times (e^{-b \times x})$	No constraints	Blue Curve
2	$+ c$	$0 < a \leq 3.0;$ $0 < b \leq 1.0;$ $0 < c \leq 0.5;$	Red Curve
3	$a \times (e^{b \times x})$	Initial guesses $a = 2.088; b = 0.060$	Green Curve

usage of algorithms will be shown in experiment 2 Table 4.

4. Experiments

This section presents equipment setup for the system and the data collected. Moreover, it also presents the implementation and assessment of the use of different sensors through our system and FastDTW, k-NN, SVM, Naive Bayes and RNN algorithms for classification of table tennis strokes.

4.1 Equipment setup

The first-generation IR depth sensor, accelerometer, gyroscope, and the software were used to capture motion in our trainer system. The IR depth camera, accelerometer, and gyroscope, along with their time stamps, were used for tracking the (X, Y, Z) coordinates of the four joints of the body. The IR depth camera was located on the tennis table, 65 cm height from the floor and the player is 152 cm standing away, as shown in the Fig. 2.

4.2 Data collection

Overall, the data obtained was collected over 800 trials from 8 different professional players (6 male and 2 female) on two keystrokes that are forehand drive and backhand drive. Each player repeated every stroke 50 times with a total of 100. The trials are divided into correct and incorrect strokes equally. The wrong strokes distinguish between various wrist errors, such as wrong approaches of movement and wrong gestures of the hand. Data obtained by our own built software and joint coordinates delivered for classification to the inner room server via socket programming to minimize time consumption, avoid leak of data and preserve the integrity of data. subsequently the other eight players executed the movement more than once to evaluate the platform and get the feedback in real-time.

4.3 Experiment 1 - Usage of different sensors

As we concerned in our latest published paper [1] to achieve valuable results and obtain precise accuracy, we used an IR depth camera to detect the skeletal joints (Wrist, Elbow, Shoulder, Waist) in direction X, Y, Z. We noticed that many different strokes depend on the movement of the wrist, considering the wrist to be a very sensitive joint. Also, as mentioned previously, the IR depth camera had several limitations with small and sensitive joints in detection as shown in the IR depth camera part on the graph Fig. 6. Therefore, the ability to detect wrist

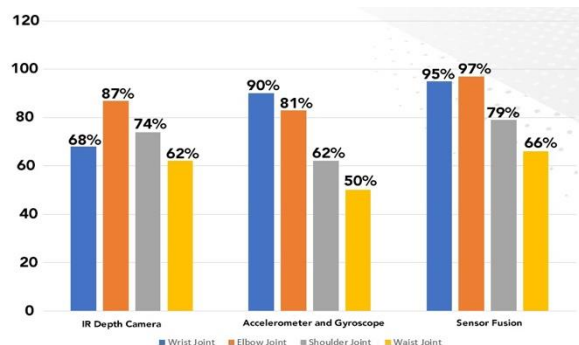


Figure. 6 Joints accuracy with different sensors

joint more accurately will increase our classification results.

To overcome this problem and avoid these limitations that can affect the player's experience, we decide to use a smart band including accelerometer and gyroscope sensors which are suitable for relatively small movements of the wrist, which must be analysed in real-time. These sensors detected the wrist's motion easily and give adequately precise accuracy. On the other hand, using accelerometer and gyroscope sensors can't be satisfying for detecting and classifying other different large joints like shoulder and waist as shown in Accelerometer and Gyroscope part in the graph Fig. 6. Therefore, affecting the achievement of good accuracy and classification of the data. In this experiment we asked 4 professional players (2 male / 2 female) to perform multiple wrong strokes in each joint. Each player performed a total 25 wrong stroke for every joint for three times, once using only an IR depth camera and once using Accelerometer and Gyroscope only and last using both as sensor fusion. Sensor fusion between both sensors did overcome the drawbacks of both sensors and raised the accuracy of detecting and classification for each joint as shown in the Sensor fusion part in the graph Fig. 6. It raised a lot especially in the wrist and elbow joints comparing to the IR depth camera or Accelerometer and Gyroscope alone. Duo to the limitations of the sensors to the distance between the further joints (Shoulder/Waist), the accuracy was not raised a lot but still better than any single sensor. Which helped to develop the identification and classification of swing motion of the player. Using sensor fusion did increase the classification accuracy as shown in Experiment 2.

4.4 Experiment 2 - Sensors fusion algorithm comparison

In this Table 4, indicate the outcomes which were performed to approximate and measure the efficiency of distinguishing main points algorithms utilizing 4 major algorithms illustrated in the former work [9,

14]. With these algorithms, the most critical point in an algorithm we are searching for is that it combines between high accuracy low classification time. The examination was achieved and tested on the dataset we gathered and made. Each algorithm calculated on the dataset and checked and attempted many times to obtain the average percentage. Table 3 is a confusion matrix based on the data tested 240 samples (30% of the total data gathered) that illustrates how well the system could identify strokes and assign it to be correct or incorrect related to our data-set.

5. Discussion

Table 3 illustrates how the system was successful in the detection and classification of strokes. The system did have some mistakes in detecting some movements and that's because some mistakes while performing the stroke was in the shoulder or waist movement. As both movements contain less error

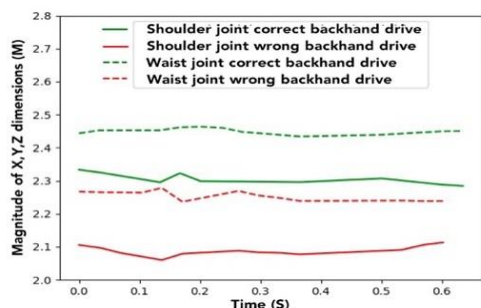


Figure. 7 Correct/wrong forehand drive graph (shoulder/waist joints)

Table 3. Confusion matrix of the testing data (fastDTW)

	CBD	CFD	WBD	WFD
CBD	58	0	3	0
CFD	0	60	0	1
WBD	2	0	56	0
WFD	0	0	1	59

Table 4. Different classification algorithm comparison

	Fast-DTW	k-NN	SVM	Naïve Bayes	RNN
CBD	96.7 %	96.7%	90 %	85 %	98.3 %
CFD	100 %	98.3%	91.7 %	76.6 %	100 %
WBD	93.3 %	95 %	81.6 %	80 %	96.7 %
WFD	98.3 %	93.3%	83.3 %	71.6 %	100 %
Average Accuracy	97.2 %	95.6%	86.7 %	78.3 %	98.8 %
Average time	1.38 sec	4.14 sec	4.89 sec	2.76 sec	6.41 sec

ratio as shown in graph Fig. 7, so the system might confuse between the movements. Despite these limitations, FastDTW and RNN achieved the best classification accuracy 97.2% and 98.8% as shown in Table 4, and since our system is time-sensitive to provide real-time feedback we depended on FastDTW. The FastDTW algorithm obtained an acceptable accuracy average with the lowest time taken of 1.38 second per 1 stroke because it analyses various time series to measure the association with a related behaviour between two-time intervals, where time and speed vary significantly.

6. Conclusion

To sum up, in this study we have proposed a system to train table tennis players and increase their performance. The system was built on the basics of strokes in table tennis: backhand drive, forehand drive. Based on our studies, the system explored that by the usage of FastDTW algorithm and utilization of an IR depth camera and wearable sensors, the system results in an accuracy of 97.2%. The system was able to achieve a high classification accuracy in an acceptable time for a real-time system. Our future work aims at adding more table tennis strokes to the system data set, which can be detected and classified by the proposed method. We also planning to extend the systems functionalities, using new algorithms in the analysis of the table tennis strokes.

Conflicts of Interest

The authors of the paper titled by “Multi-Sensor Fusion for Online Detection and Classification of Table Tennis Strokes” certify and identify that they have NO conflict of interest with any personal circumstances or interest that may be perceived as inappropriately influencing the representation or interpretation of paper and research results. Also, the authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this paper.

Author Contributions

Conceptualization, Ayman Atia and Habiba Hegazy; Methodology, Moustafa Hussien and Habiba Hegazy; Validation, Ayman Atia, Habiba Hegazy and Moustafa Hussien; Data collection, Mohammed Abdelsalam and Moustafa Hussien; Writing original draft preparation, Seif Elmosalamy, Mohammed Abdelsalam, Habiba Hegazy and Moustafa Hussien; Writing—review and editing, Habiba Hegazy, Ayman Atia and Youmna Ibrahim;

Supervision, Ayman Nabil and Ayman Atia. Testing application with players, Seif Elmosalamy and Mohammed Abdelsalam.

References

- [1] H. Hegazy, M. Abdelsalam, M. Hussien, S. Elmosalamy, Y. Hassan, A. Nabil, and A. Atia. "Online detection and classification of in-corrected played strokes in table tennis using IR depth camera", In: *Proc. of Computer Science Conf. on Emerging Data and Industry*, pp. 555–562, 2020.
- [2] R. Liu, Z. Wang, X. Shi, H. Zhao, S. Qiu, J. Li, and N. Yang, "Table Tennis Stroke Recognition Based on Body Sensor Network", In: *Proc. of International Conf. on Internet and Distributed Computing Systems*, pp. 1–10. 2019.
- [3] P. Molchanov, X. Yang, S. Gupta, K. Kim, S. Tyree, and J. Kautz. "Online detection and classification of dynamic hand gestures with recurrent 3D convolutional neural networks", In: *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, Las Vegas, NV, pp. 4207–4215, 2016.
- [4] Y. Chen, B. Luo, Y. Chen, G. Liang, and X. Wu. "A real-time dynamic hand gesture recognition system using Kinect sensor", In: *Proc. of IEEE International Conf. on Robotics and Biomimetics (ROBIO)*, pp. 2026–2030, 2015.
- [5] M. Kos, J. Zenko, D. Vlaj, and I. Kramberger, "Tennis stroke detection and classification using miniature wearable IMU device", In: *Proc. of International Conf. on Systems, Signals and Image Processing (IWSSIP)*, Bratislava, pp. 1–4, 2016.
- [6] K. Xia, H. Wang, M. Xu, Z. Li, S. He, and Y. Tang, "Racquet sports recognition using a hybrid clustering model learned from integrated wearable sensor", *Sensors Journal*, Vol. 20, No. 6, pp. 1638, 2020.
- [7] Z. Bankosz and S. Winiarski, "Using wearable inertial sensors to estimate kinematic parameters and variability in the table tennis topspin forehand stroke", In: *Proc. of International Conf. on Applied Bionics and Biomechanics*, pp. 1–10, 2020.
- [8] T. Le, M. Nguyen, and T. Nguyen. "Human posture recognition using human skeleton provided by Kinect", In: *Proc. of the International Conf. on Computing, Management and Telecommunications*, 2013.
- [9] H. Wang, L. Li, H. Chen, Y. Li, S. Qiu, and R. Gravina. "Motion Recognition for Smart Sports Based on Wearable Inertial Sensors", *EAI International Conf. on Body Area Networks*, pp. 114–124, 2019.
- [10] E. Boyer and F. Bevilacqua and François Phal and S. Hanneton and U. Staps. "Low-cost motion sensing of table tennis players for real time feedback", *International Journal of Table Tennis Sciences*, Vol. 8, No. 1, pp. 1–4, 2013.
- [11] P. Blank, J. Hossbach, D. Schuldhaus, and B. Eskofier, "Sensor-based stroke detection and stroke type classification in table tennis", In: *Proc. of the 2015 ACM International Symposium Conf. on Wearable Computers* pp. 93–100, 2015.
- [12] W. Viyanon, V. Kosasaeng, S. Chatchawal, and A. Komonpetch, "Swingpong: analysis and suggestion based on motion data from mobile sensors for table tennis strokes using decision tree", In: *Proc. of the 2016 International Conf. on Intelligent Information Processing*, pp. 1–6, 2016.
- [13] A. Chatterjee and V. Govindu, "Noise in structured-light stereo depth cameras: Modeling and its applications", *ArXiv Journal*, Vol. 1, No. 1, pp. 1–31, 2015.
- [14] G. Li, H. Tang, Y. Sun, J. Kong, G. Jiang, D. Jiang, B. Tao, S. Xu, and H. Liu, "Hand gesture recognition based on convolutional neural network", *The Journal of Networks, Software Tools and Applications*, Vol. 22, No. S2, pp. 2719–2729, 2017.
- [15] J. Du, C. Gerdtnan, and M. Linden, "Signal quality improvement algorithms for mems gyroscope-based human motion analysis systems: A systematic review", *Sensors Journal*, Vol. 18, No. 4, pp. 1123, 2018.
- [16] S. Moon, Y. Park, Dong Wook Ko, and Il Hong Suh, "Multiple Kinect sensor fusion for human skeleton tracking using Kalman filtering", *International Journal of Advanced Robotic Systems*, Vol. 13, No. 2, pp. 1–10, 2016.
- [17] P. Das, K. Chakravarty, A. Chowdhury, D. Chatterjee, A. Sinha, and A. Pal, "Improving joint

- position estimation of Kinect using anthropometric constraint based adaptive Kalman filter for rehabilitation”, *Journal of Biomedical Physics and Engineering Express*, Vol. 4, No. 3, pp. 1–29, 2018.
- [18] S. Feng and R. Murray-Smith, “Fusing Kinect sensor and inertial sensors with multi-rate Kalman filter”, In: *Proc. of IET Conf. on Data Fusion & Target Tracking: Algorithms and Applications*, Liverpool, UK, Vol. 14, pp. 1–8, 2014.
- [19] Z. Ahmad and N. Khan, “Human action recognition using deep multilevel multimodal fusion of depth and inertial sensors”, *IEEE Sensors Journal*, Vol. 20, No. 3, pp. 1445–1455, 2019.
- [20] C. Zhang, F. Yang, G. Li, Q. Zhai, Y. Jiang, and D. Xuan, “Mv-sports: A motion and vision sensor integration-based sports analysis system”, In: *Proc. of IEEE Conf. on Computer Communications*, Honolulu, HI, pp. 1070–1078, 2018.
- [21] A. Aguilera, R. Brena, O. Mayora, E. Molino-Minero-Re, and L. Trejo, “Multi-sensor fusion for activity recognition—a survey”, *Sensors Journal*, Vol. 19, No. 17, pp. 3808, 2019.
- [22] K. Irie, “A loop-closure-based inertial motion capture, with application to sports swing measurements”, In: *Proc. of 2020 IEEE/SICE International Symposium on System Integration*, Honolulu, HI, USA, pp. 693–699, 2020.
- [23] F. Mufarroha and F. Utaminigrum, “Hand gesture recognition using adaptive network based fuzzy inference system and k-nearest neighbour”, *International Journal of Technology*, Vol. 8, No. 3, pp. 559–567, 2017.
- [24] T. Ashfaq and K. Khurshid, “Classification of hand gestures using Gabor filter with Bayesian and naive bayes classifier”, *International Journal of Advanced Computer Science and Applications*, Vol. 7, No. 3, pp. 276–279, 2016.
- [25] S. Shin and W. Sung, “Dynamic hand gesture recognition for wearable devices with low complexity recurrent neural networks”, In: *Proc. of IEEE International Symposium on Circuits and Systems (ISCAS)*, Montreal, QC, pp. 2274–2277, 2016.
- [26] I. Pernek, K. Hummel, and P. Kokol, “Exercise repetition detection for resistance training based on smartphones”, *Personal and Ubiquitous Computing Journal*, Vol. 17, No. 4, pp. 771–782, 2012.
- [27] A. D. Calin, “Gesture Recognition on Kinect Time Series Data Using Dynamic Time Warping and Hidden Markov Models”, In: *Proc. of International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, Timisoara, pp. 264–271, 2016.
- [28] K. Srinivas and M. Rajagopal, “Study of hand gesture recognition and classification”, *Asian Journal of Pharmaceutical and Clinical Research*, Vol. 10, No. 1, pp. 25–30, 2017.
- [29] S. Khalid, T. Khalil, and S. Nasreen, “A survey of feature selection and feature extraction techniques in machine learning”, In: *Proc. of Science and Information Conf.*, pp. 372–378, 2014.
- [30] S. Chekanov, *Scientific Data Analysis using Jython Scripting and Java*. Vol. 1, Springer-Verlag, London 2010.