

Recording of occurrences through image processing in Taekwondo training: first insights

Tiago Pinto¹, Emanuel Faria¹, Pedro Cunha¹, Filomena Soares¹, Vítor Carvalho², Hélder Carvalho³

¹Dep. Industrial Electronics, Algoritmi Research Centre, University of Minho, Guimarães, Portugal [a71889@alunos.uminho.pt; a72236@alunos.uminho.pt; id5514@alunos.uminho.pt; fsoares@dei.uminho.pt]

²School of Technology, IPCA, Barcelos & Algoritmi Research Centre, University of Minho, Guimarães, Portugal [vcarvalho@ipca.pt]

³Dep. Textil Engineering, 2C2T Research Centre, University of Minho, Guimarães, Portugal [helder@det.uminho.pt]

Abstract

Nowadays, the most used evaluation method of athlete's performance in the martial art Taekwondo is still performed manually, where the coach analyses the collected videos of the athlete training. This method besides being time consuming, it is prone to errors. Aiming the development of methods for improvement the training, this project intends to present a new method for recording occurrences and recognizing the movements of athletes in real time during the Taekwondo training. To achieve the purpose, it was used the Microsoft Kinect sensor fused with image processing technics. This project arises as a collaboration between the University of Minho, the School of Technology from the Polytechnic Institute of Cávado and Ave and the Sporting Club de Braga Taekwondo section, Portugal. It is the authors believe that the proposed system may improve the athlete's performance and the development of the Taekwondo training technics.

Keywords: Image processing, Taekwondo, martial arts, motion analysis.

Introduction

Taekwondo is a martial art sport that has gained popularity over the last two decades and it was introduced as an Olympic Sport in the year 2000.

The current methods of evaluating the athlete's performance are mostly made manually by the coach, by analysing videos from training and competition. This

process is time consuming and it is not efficient, because the coach must analyse each team trainee video at a time with the goal of providing athletes with relevant feedback to improve their performance.

The proposed system will provide a new method that allows analysing and quantifying the athlete's performance in real-time examining the athlete movements using image processing. The main aim is to establish a new level of training, helping the coach on giving real time feedback on the trainee's performance and improving the training time. This project arises from the collaboration between the University of Minho, the School of Technology from the Polytechnic Institute of Cávado and Ave, and the Sporting Club de Braga (SCB) Taekwondo team, Portugal.

This paper is organized in seven sections. In the second section, it is presented the state of the art of systems to aid athlete's performance in sports; in the third section the materials used for developing the proposed system are described; in the fourth section the system output requirements are presented; in the fifth section the solution developed is presented; in the sixth section the experimental results are shown; and finally, in the seventh section the discussion and conclusions are enunciated.

State of the Art

Accordingly, to the literature review there are some systems for analysing the performance of athletes in tennis, martial arts, among others sports [1-5].

In [1] the authors distinguished the different types of tennis strokes in a ball and classified them as serve, forehand or backhand by using vision and accelerometer measurements (in the athlete body). They tested those measurements independently and merged by fusing the collected data. The best results were obtained by the merged solution.

Also in [2] the merged data from visual and sensing information were used to recognise the movements in Taekwondo, essentially by dividing the movements in three parts (beginning, middle and end) and then comparing the results with pre-defined movements.

In [3] the analysis of gait patterns allowed the recognition on human activities such as walking, jogging, running and jumping with rates that reach 95.01% of accuracy, by considering the legs distance, the angle between them, velocity and axis movement.

A Kinect camera was explored in [4] to correct the errors between colour and deep images to create a virtual scenario in real time. The methods explored intend to correct the flickering phenomenon and unmatched edges. It is a simple approach presenting high speed velocity in image processing (around 10ms in an Intel Core i5-2500 with a GTX 670 and 8GB of RAM), better than other methods such as Telea and Criminisi with 100.4ms and 4596.7ms. The method presented also achieves better results than the fast-linear interpolation. With this method, it is possible to obtain real time processing of the frames received by the camera.

Regarding the image-processing field, in [5] it is presented a method that allows to primarily discard the irrelevant frames (the ones that do not have movements or do not fit the required parameters). Then, the trajectory of the particles is analysed so that the system can recognise the next position of a given moving part, such as an arm or leg.

With the works presented in [4] and [5], it is possible to obtain a faster acquiring and processing system that enables real-time and quality requirements.

Following this trend and to explore and overcome the fragilities detected in the literature, the proposal presented in this paper is to use computer vision or a combination of both computer vision and sensors information to detect and recognize movements in the Taekwondo training.

Materials

To run the proposed solution, it is needed at least one motion sensor and a computer capable of processing the software developed.

As for the motion sensor, it was selected the Kinect V1 for Windows. This version is the most suitable for development since it integrates a proper SDK (Software Development Kit) developed by Microsoft and it can get in depth images (the distance between the camera and the point being detected). These features along with the skeleton tracking allowed easy progress in the software development and better results regarding the processing of the acquired data.

Concerning the computer, connected to the Kinect sensor and where the developed software is processed, it has the following technical features:

- 4^o generation processor Intel® Core™ i7-4510U with 3.10 GHz clock;
- Windows 10 Home;
- 8GB of DDR3L RAM (1,600 MHz);
- AMD Radeon™ R7 M260 with 2GB VRAM dedicated and DDR3 video RAM.

The computer supports the installation of the required software and drivers, and collects the data from the motion sensor Kinect.

Requirements

The goal of the proposed system is to recognize the movements performed by the athlete in real time during training to give feedback on his/her performance (if the movement was correctly done or if it could have improvement). So, the three main requirements that the system must fulfil are:

- To recognize the Taekwondo movements;
- To register the movements and count the occurrences;
- To ensure the real-time processing.

Each frame acquired by the Kinect sensor must be processed to extract the information given, to posteriorly present the processed frames to the interface, showing the results to the trainer or athlete.

It is worth mention that the methods and hardware used should not disturb the athlete in any way, such as restrict his/her movements or bother him/her.

Solution

To recognize the movement made by the athlete, firstly it is acquired the athlete skeleton given by the Kinect sensor. It gives the information about each joint of the athlete, namely, the respective spatial coordinates (x, y, z) .

The Kinect sensor uses a Cartesian coordinate system centred at the Kinect sensor that for now on will be referred as CCS (Centred Coordinate System). The positive Y axis points up, the positive Z axis points where the Kinect is pointing, and finally the positive X axis is to the left where Kinect is pointing, which is presented in Fig. 1. The skeleton acquired identify a total of 20 joints and each coordinate of the respective joint in the human body. All joints available are represented in Figure 2.

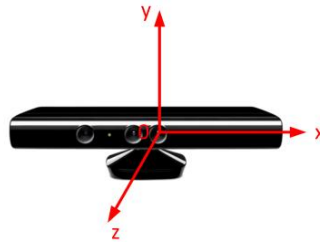


Fig. 1 Kinect Cartesian axis [6].

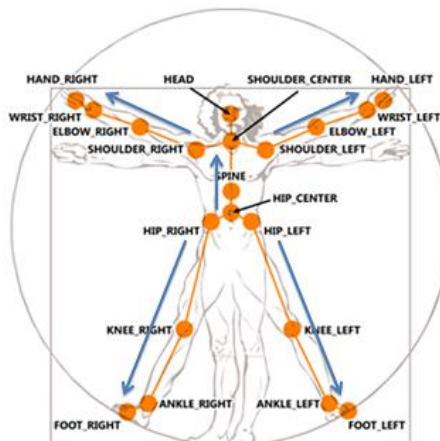


Fig. 2 Available skeleton joints [7].

Before testing the system in the real world, an evaluation was made on how precise and trustworthy the information given by the Kinect sensor is. In a first approach some data was verified, namely the coordinates of each joint, the angles between joints and the recognition of techniques.

In Figure 3 it is presented all the joints acquired by the Kinect sensor and the respective X, Y, Z coordinates in meters. All the coordinates, as explained above, are represented in the same CCS provided by Kinect sensor. Later, these coordinates were used to calculate the angles of each joint.

joint positioning	X	Y	Z
head	-0,97	1,02	2,57
chest	-0,73	0,80	2,80
left shoulder	-0,84	0,75	2,68
left elbow	-0,93	0,55	2,63
left wrist	-0,90	0,30	2,45
left hand	-0,77	0,46	2,52
right shoulder	-0,36	0,67	2,54
right elbow	-0,37	0,39	2,41
right elbow	-0,24	0,27	2,31
right hand	-0,34	0,05	2,26
spine	-0,68	0,49	2,74
hip	-0,68	0,45	2,67
left hip	-0,76	0,30	2,49
left knee	-0,88	-0,34	2,25
left ankle	-0,95	-0,75	2,02
left foot	-0,73	-0,77	1,94
right hip	-0,56	0,25	2,71
right knee	-0,57	-0,24	2,22
right ankle	-0,60	-0,70	1,97
right foot	-0,67	-0,76	1,94

Fig. 3 Cartesian coordinates of each joint: values x, y, z for each joint positioning acquired from the developed software.

The skeleton is drawn in the interface, to show what is captured by the Kinect sensor, and to see if the Kinect sensor is acquiring the correct athlete positioning. The angles between joints needed to qualify and quantify the movement in terms of success are then calculated. The angles are calculated using dot products of scalar product. The dot product is a value expressing the angular relationship between two vectors, that is, each joint is represented as a 3D vector. The dot product returns a value that is the result of the operation of two vectors with the same number of components, in this case, with the same number of coordinates. The dot product is thus the sum of each coordinate of the two vectors. For example, given two 3D vectors A and B, the respective dot is presented in equation 1. In Figure 4 it is represented how two vectors are related to each other in the dot product

$$A \cdot B = A_x \cdot B_x + A_y \cdot B_y + A_z \cdot B_z \quad (1)$$

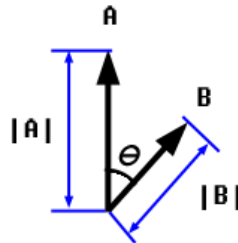


Fig. 4 - Geometric interpretation dot product [8].

The dot product can be obtained by the equation 2 where $|A|$ and $|B|$ represents the length of vector A and B, respectively, and θ is the angle between the two vectors,

$$A \cdot B = |A| \cdot |B| \cdot \cos(\theta) \quad (2)$$

By solving equation (2) in order of θ , the angle between the two joints is obtained.

As an example, the angle of the elbow depends on the orientation of the shoulder and the wrist; so firstly, it is calculated the dot product between elbow and shoulder, elbow and wrist, the respective dot products are subtracted and the respective angle is calculated.

After obtained the coordinates of each joint and the respective angles, the system can recognize the movement performed by the athlete.

It is worth noting that before implementing the movement recognition algorithm in the system, each movement was previously defined as how it should be undertaken. A movement is detected when a relative positioning of three joints occurs for a given number of frames. So, a typical fist happens when the elbow goes from a position that is closer to the body (that is, the elbow angle value is between 30° and 90° , and the Z coordinate of the shoulder with respect to the Z coordinate of the elbow is between 0 and 20 centimetres in module), to a position where the elbow angle is between the admissible range. This range is expressed by a maximum of 10% of the real angle value. For example, if the left arm is fully stretched out, the left real elbow angle should have a value of exactly 180° , but given this range the arm is considered as fully stretched if the elbow angle goes between 162° and 180° , approximately. Then the Z coordinate position of the wrist is lower than the elbow Z coordinate position and both lower than the shoulder. The movements are validated by comparing, in each frame, the current relative positioning and the angles of athlete joints with the reference values, the angles are validated with a given range, as expressed above, to identify the movement performed.

The system developed acquires the angles in the arms and legs, once in the practice of Taekwondo the fight movements are mainly performed by arms and legs. This option allows the system a faster processing speed of each frame.

Experiments and Results

To verify the accuracy of the system developed preliminary tests were performed in laboratory environment with non-athletes and later with Taekwondo athletes.

Figure 5 shows a laboratory test with a non-athlete; his skeleton and the arm angles were obtained by the dot product between joints, using the coordinate vectors. It was performed a fist movement and the respective values of the angles, at the end of the fist, were checked. The first column in Figure 5 corresponds to the shoulder angles, the second to the elbow angles, and the last to the wrist angles, for both left and right arms.

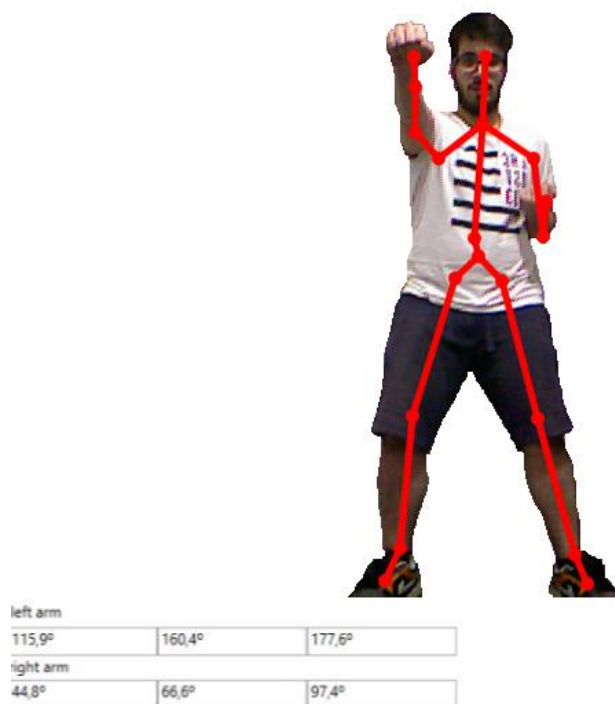


Fig. 5 - Angles calculated from the skeleton. The first column corresponds to the shoulder angles, the second to the elbow angles, and the last to the wrist angles, for both left and right arms.

The left arm is completely stretched out, that is, the respective elbow angle should be 180° , but the value given by the system is around 160° ; the other angles, namely the shoulder and wrist angles of both left and right arms, are in the admissible range. As shown in figure 5 the wrist is fully stretched so the corresponding angle should be 180° , but the value obtained by the system is 177.6° , then the value obtained is within the defined range.

A normal kick movement was performed to test the angle values. The test is presented in Figure 6. The first column corresponds to the hip angles, the second column to the knee angles, and the last column to the ankle angles.

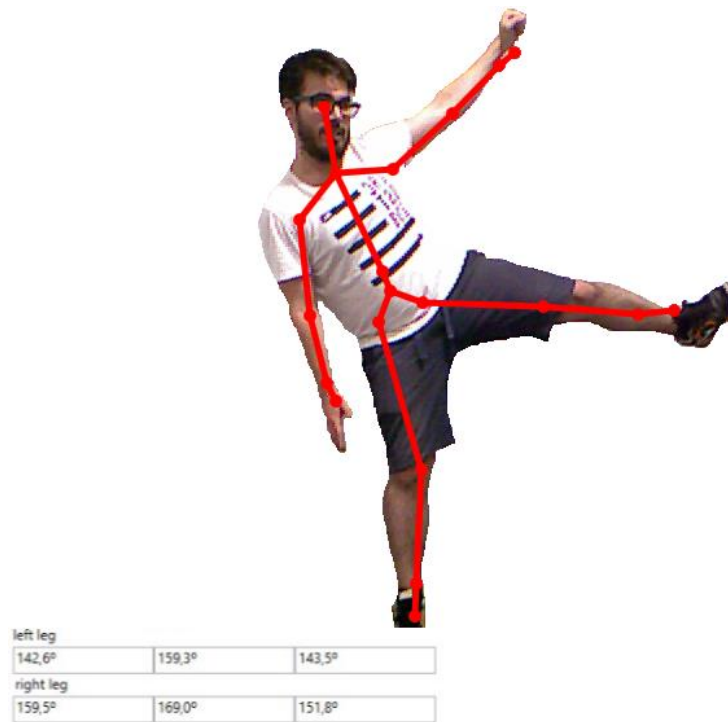


Fig. 6 - Kick movement. The first column corresponds to the hip angles, the second column to the knee angles, and the last column to the ankle angles.

The success rate for the fist movement is presented in Figure 7. From the twenty-eight fist movements detected by the system, 10 were performed correctly within the defined angle parameters for the fist movement (36% of correct movements).

normal fist	wrong fist	rate of success
10	18	36%

Fig. 7 - Success rate calculated by the developed system.

Ten sets of thirty punches each were performed by a Taekwondo athlete. The success rate increased from set to set, varying from 76% to 97%. An average value of 87% of success rate was achieved. The system helped the trainee to achieve a

better performance in the execution of the movement, with the real-time feedback, stimulating the athlete training improvement.

Several tests were performed moving both arms and legs. During these tests, it was verified that when the arm was fully stretched, that is, the angle of the corresponding elbow should be 180°, the maximum reading error of the system was 8% (corresponding to an angle of around 165°).

Discussion and Conclusions

This project aims to develop a system that allows the recognition and registry of the correct athlete movements, in real time, for the Taekwondo training technics.

In a sequence of ten tests of thirty punches each, the program identifies the punches with an average success rate of 87%. These results are like the withdrawn from [1] and [3], indicating that the proposed system is a suitable solution in recognizing the movements performed by the Taekwondo athlete.

A real-time response was achieved for the punch movements, which means the coach can observe the data from the program and see the evolution of the trainee by the percentage of success movements, correcting him/her when a low percentage is observed.

The first tests performed with the system were quite expressive in general, allowing determining the limitation of the system, and its possible improvements, in particular:

- The skeleton obtained by the camera did not always correspond to the real body, which sometimes makes the results unrealistic;
- The percentage of detected moments could be improved with better hardware and other methods of processing the frames received.

Considering the system hardware used, improvements can be made especially in the camera such as to replace the Kinect V1 for one with higher frames per second, better resolution and more range, for example the Kinect V2.

Before starting to analyse other movements such as leg movements, the improvements mentioned above need to be implemented, because leg movements are more complex and mix rotation of the torso with the angles of the joints.

Future versions of this work will have optimized identification of the movements and include a higher dataset of movements detection.

Acknowledgments

The authors would like also to express their acknowledgments to COMPETE: POCI-01-0145-FEDER-007043 and FCT – *Fundação para a Ciência e Tecnologia* within the Project Scope: UID/CEC/00319/2013. Pedro Cunha thanks FCT for PhD scholarship SFRH/BD/121994/2016.

Special thanks to Coach Joaquim Peixoto, as well as Sport Club de Braga Taekwondo section and to the national team, for allowing us to use their training

site, as well as seeing/participating in training to test the developed system, as in general for all the cooperation.

Additional thanks to Ricardo Guimarães from Vitória Sport Clube Taekwondo team for testing the first versions of the system and help with the calibration of the system.

References

- [1] Ó Conaire, C., Connaghan, D., Kelly, P., O'Connor, N. E., Gaffney, M., & Buckley, J. (2010, October). 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 (pp. 51-56). ACM.
- [2] Kwon, D. Y., & Gross, M. (2005, June). Combining body sensors and visual sensors for motion training. In Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology (pp. 94-101). ACM.
- [3] Gupta, J. P., Polytool, D., Singh, N., & Aemwal, V. B. (2014). Analysis of gait pattern to recognize the human activities. *IJIMAI*, 2(7), 7-16.
- [4] Lee, G. C., & Yoo, J. (2014). Real-time virtual-viewpoint image synthesis algorithm using Kinect camera. *Journal of Electrical Engineering and Technology*, 9(3), 1016-1022.
- [5] Narhare, A. D., Molke, G. V., & Kanphade, R. (2014). Motion Trajectory of Players using Optical Flow.
- [6] Kinect v1 SDK C++ - 3. Kinect Point Clouds. Retrieved from <http://homes.cs.washington.edu/~edzhang/tutorials/kinect/kinect3.html>
- [7] Tracking Users with Kinect Skeletal Tracking. Retrieved from <https://msdn.microsoft.com/en-us/library/jj131025.aspx>
- [8] Robert Dunlop (2005, June 26). Understanding the Dot Product. Retrieved from <http://www.mvps.org/DirectX/articles/math/dot/index.htm>