Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 170 (2020) 466–473

Procedia Computer Science

www.elsevier.com/locate/procedia

The 11th International Conference on Ambient Systems, Networks and Technologies (ANT)
April 6 - 9, 2020, Warsaw, Poland

iKarate: Improving Karate Kata

Bassel Emad[a,*], Omar Atef[a], Yehya Shams[a], Ahmed El-Kerdany[a], Nada Shorim[a], Ayman Nabil[a], Ayman Atia[b,c]

[a] *Faculty Of Computer Science, Misr International University (MIU), Cairo, Egypt*
[b] *Faculty Of Computer Science And Artificial Intelligence, Helwan University (HU), Cairo, Egypt*
[c] *Faculty Of Computer Science, October University for Modern Sciences and Arts (MSA), Cairo, Egypt*

Abstract

Karate is a martial art that can be practiced using hands and feet to deliver and block strikes. Karate moves must be done in a certain way, many moves are practiced incorrectly during training. In this paper, we present a fully functional system which Karate players, coaches, judges and clubs could use. The system helps in capturing Karate moves using Kinect v2 sensor and analyzing these moves using F-DTW. Real-time feedback and a report are displayed to the users using a simple GUI (Graphical User Interface) that the users can easily use, to learn how a mistake was made, and how to fix it/learn from it. F-DTW was used for proving the concept, and an average accuracy of 90% was achieved. This Paper is mainly concerned about the moves named: Age-Uke, Mae-Geri, Gedan-Barai and Soto-Uke.

Keywords: Classification; Kinect; Karate Coaching; Recognition; Analysis; Real-time Feedback; Human Movement Analysis

1. Introduction

Karate moves are combination of physical moves. Its methods are qualitative not quantitative which makes it hard for students to perform the perfect version of the motion and harder to judge it. Kids nowadays may find difficulties learning those moves at a young age and since the training may consist of a large number of students so the trainer himself may not be able to focus on every detail of every student's move, which could result in taking more time to learn and master the move or it may lead to learning incorrectly from the beginning.

Despite the importance of the sports field and the fact that skill sports are well-learned by video observation [17]. There are no suited viewpoint videos for learning Karate movements. In Addition, many people do not focus on

* Corresponding Author. Tel.: +20-0100-5550-477.
E-mail address: Bassel1600604@miuegypt.edu.eg

10.1016/j.procs.2020.03.090

working in the Karate field. So, it was a good opportunity to continue the researches, by developing a system that would make a difference in learning skill based sports in not only karate, but in general. Furthermore, kids playing Karate at the beginning do not get the attention needed from the coach who should focus on the little mistakes made by the kids.

The main goal of this paper is to capture the moves of the performers in real time, analyse those moves and give them a feedback report to enhance their technique or alert them if they are performing a move or a stance incorrectly. One of the challenges we faced while comparing and analysing the captured motion, is that we should take into consideration that the activities might be performed with different speed, body proportions such as (Limbs length) and initial position of the students [12]. Another challenge is the real-time feedback, giving the users a feedback and a report on their moves whether it was right or not in real-time is essentially important after the move is performed. The report includes tips on how to execute the move correctly the next time. As shown below in figure 1, the main focus in this paper is going to be on the Age-Uke, Mae-Geri, Gedan-Barai and Soto-Uke.

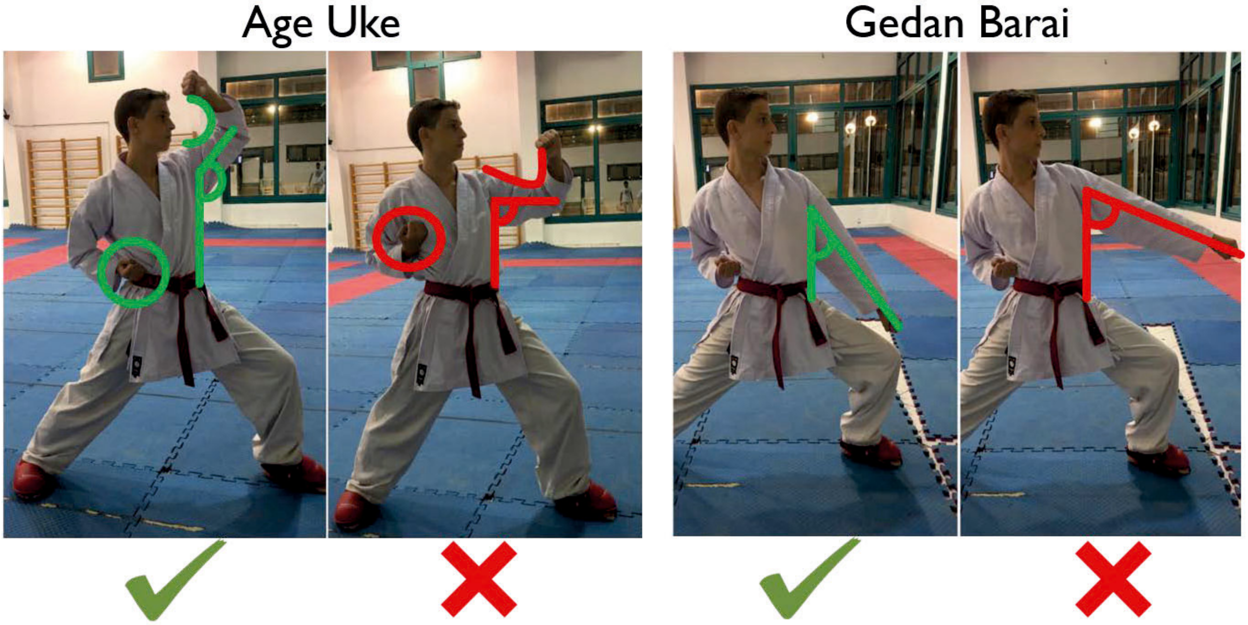

Fig. 1. Age-Uke And Gedan-Barai Right And Wrong

The rest of the paper is organized as follows: Section 2 explains the related work and other approaches. Section 3 explains our proposed approach and the methods that were used. Section 4 shows the experiments and results reached in this paper. Section 5 is the discussion of the system. Finally section 6 is the conclusion and the future work.

2. Related Work

T. Hachaj et al. [6] and [9] talked about using Kinect sensors to classify Karate move and how to obtain the highest results. Firstly, [6] T. Hachaj et al. proposed statistical differences between Kinect v1 and Kinect v2 recognition in some of the Karate moves. Their research [6] was done to evaluate the effectiveness of both Kinects in Karate motion recognition and to find the best suited sensor. They made it clear that Kinect v2 would be more accurate and improve recognition, rather than Kinect v1. Secondly, [9] proposed a calibration procedure of three Kinect sensors that integrates the data into one skeleton. The problem was to find a way to increase Karate motion recognition accuracy and effectiveness of fusion of body joints gathered from different sensors. They reached a positioning of the three sensors to improve the non-classified techniques by 48%. Also A. D. Calin, [2] compared the efficiency of different classifiers tested on six data-sets obtained from both Kinect v1 and v2. To evaluate the accuracy of the two Kinect Sensors and analyse the gesture recognition accuracy of several classifiers. He reached high accuracy but with a time period of 65.93 seconds using Multilayer Perceptron algorithm and Kinect v2.

P. Alborno et al. [1] and N. T. Thanh et al. [18] developed a system for evaluating and giving a score based on the quality and performance of each move. In [1] P. Alborno et al. proposed a method to measure the quality of karate moves. The analysis and evaluation of full human body movements' qualities were an important problem facing people creating sports systems lately. They reached a solution by studying how much the limbs are synchronized during relevant motion phases. On the other hand, N. T. Thanh et al. [18] proposed a performance scoring system that applies the data of the images from the Kinect. The aim was to make a standard model to be used in any system around the world for Vietnamese Traditional martial arts.

T. Hachaj et al. in [12] proposed evaluation and visualization technique for advanced motion analysis. Their paper evaluated the comparison, analyses and visualization method of the similarities and differences between three dimensional trajectories of human body joints in Karate movements. Their interest was in investigating the differences in human moves, due to the imperfect imitation problem. E. Escobedo-Cardenas and G. Camara-Chave in [4] Developed a method for hand gesture recognition using the Kinect. They made this approach to overcome the problems of hand gesture recognition using Sensor and video based tools. T. Hachaj et al. in [7] proposed actions descriptions with maximally three key-frames. Their aim was to make motion recognition in low-dimensional feature space and the selection of proper features for modelling multiple human actions. Their selection of proper features set to model human actions in low-dimensional space, will benefit in prioritizing the feature selection for better recognition. Y. Choubik and A. Mahmoudi in [3] developed a real-time human body poses classification technique. Each karate move is a sequence of poses, which make Poses Recognition important in this system. Their problem was applying machine learning algorithms to classify human body poses in real time. T. Hachaj et al. in [8] developed a video annotation method that enables both numerical and categorical features calculation. Their aim was to create an efficient system for learning Karate.

From [6] we could conclude that Kinect v2 is always more reliable than Kinect v1 in terms of accuracy, due to its better calculations of legs joints positions. Also, [4] and [7] they worked on selecting important features to improve the recognition of hand and body gestures, and they reached an accuracy of 88.38% and 88% respectively. In [12] and [3] they focused on solving the imperfect imitation problem and different positions that will cause errors during classifications, using different algorithms and approaches. While in [2] they made a well detailed comparison on different classifiers on six data-sets taken from Kinect v1 and v2. Those comparisons will help in selecting the most suitable hardware and algorithm to achieve the highest accuracy in the system. Moreover, the quality of karate movements was mentioned in [18] and [1], they worked on making a measuring technique for the quality of movements, one of these techniques was evaluation of limbs positioning. Lastly, in [9] the paper proposed a way for using three Kinects for improving the accuracy but this could only work in case of using GDL (Gesture Description Language).

3. Methodology

The system is composed of one or multiple Kinects. The Kinect(s) would be facing the user while he/she performs a sequence of moves. Then the frames and the skeleton are extracted from the Kinect, After that the pre-processing, enhancement, saving data on the cloud proceeds simultaneously with computing the key frames, feature extraction and finally the classification. After every move has been performed, the practitioner is presented with the move name and whether it was done correctly or not as shown in figure 2. The performers is given a score to know how good their performance of the move after it has been analyzed. The score evaluation is based on the practitioners' motion while performing the move and their speed. Dynamic analysis of the movement gives real-time feedback and a report to the practitioner or the coach, making the application more interactive. The report contains the player name, age, weight, height, belt color, move name and duration, how well the player preformed the move, how to improve the user's performance and if any mistake were made it will be shown in the report.

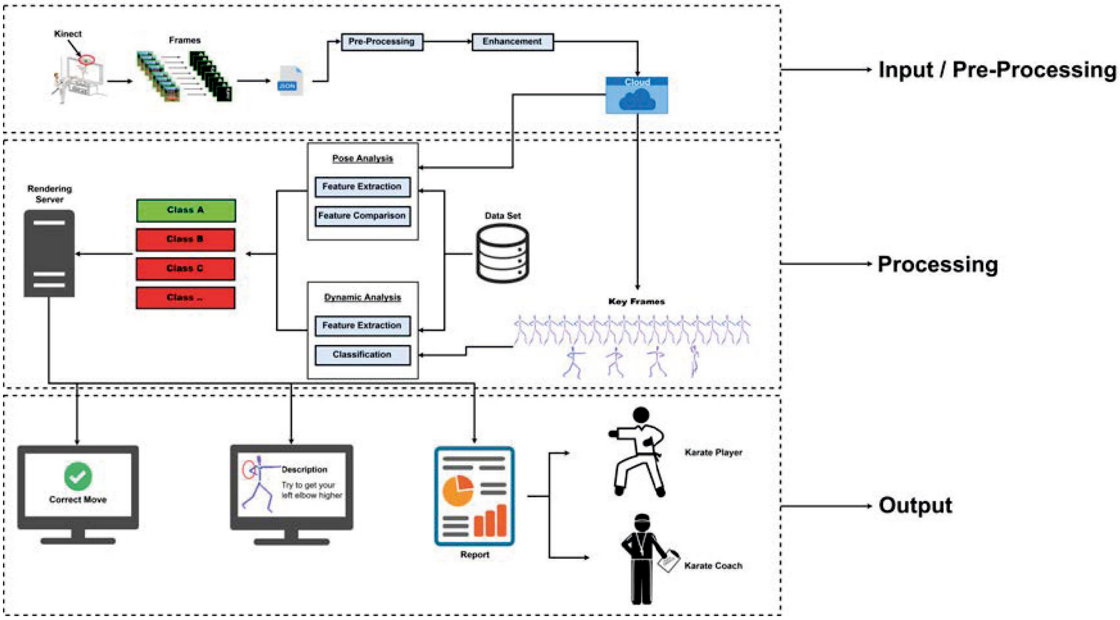

Fig. 2. System Overview

3.1. *Input & Pre-processing*

The input of the system is the coordinates of all the body joints in 3D space (X, Y, Z) captured from the Kinect. Before processing any of the data acquired, some pre-processing had to be done. One of the challenges in capturing motion from the Kinect was normalizing the data.

Normalization: The data that is acquired from the Kinect can not be used directly. While capturing the data, there will be different body proportions (Height, Scale, etc.) or a dominant factor in the data. That's why this approach used Normalization, to make all the features in the data use the same scale. So that no attribute in the data is dominant. The Normalization algorithm proposed in this paper is "Z-score normalization" with EQ. 1

$$X = \frac{Value - \mu}{\sigma} \tag{1}$$

Where "Value" is the data point, μ is the mean value and σ is the standard deviation of the data. If "X" is equals to the mean value of the feature, it will be normalized to zero. If it's below the mean, it will be normalized to a negative number, and if it's above the mean, it will be normalized to a positive number, The "X" value is calculated by the standard deviation. If the un-normalized data had a large standard deviation value, then the normalized values would be closer to zero.

3.2. Processing And Algorithms

After pre-processing the data, Fast-DTW (Fast Dynamic Time Warping) was used to manage the different speeds of the moves taken by the performers using the Kinect and to provide them a real time feedback as accurate as possible. Fast-DTW is an algorithm for measuring similarities between two signals, each signal may have a different speed from the other signals. Fast-DTW is an alignment algorithm which is capable of classifying two different time signals. It's also used in [12], [4], [7], [14], [2], [10], [15]. Fast-DTW could be applied to different types of data, like videos, audio and graphics data. Thus, any data that can be converted to a linear sequence could be analyzed using it. The algorithm was able to recognize the classes of the correct and wrong moves. Fast-DTW could be used with many different distance equations but the "Euclidean Distance" is the one used in this approach as shown in EQ. 2 to compute the distance between the classes.

$$D = \sum_{x_i,y_i}^{n} \sqrt{(x_i - y_i)^2} \tag{2}$$

Where "D" is the distance value, "X" represents the data-set joint position and "Y" represents the performer's joint position. Between each two moves, there is a small gap, this gap would be used to segment each movement and the movement data would be sent to F-DTW algorithm to be classified.

3.3. Output

The last part of the system is the output. Which will be categorized as follows:

Result Screen: This screen will tell the user if they performed the move correctly or not, with a percentage of how much the movement was performed correctly. If the percentage is acceptable and the move was done correctly, the screen will inform the user and display the percentage of the correctness. If the practitioner performed the move in a wrong way, the screen would display to them what they did incorrectly regarding the move and how to perform it correctly with a report.

Report: The second part of the output would be the report. The report will benefit both the student and the coach. Since this report will have a fully-detailed statistics of how accurate the practitioner performed the moves, mistakes and how to improve the performance.

3.4. Data-Set

The data-set is collected by T. Hachaj, M. R. Ogiela, and M. Piekarczyk from the University of Krakow, Poland [11] and was used in [5] and [13]. The data-set includes the moves that are needed in the system, which are: "Age-Uke", "Mae-Geri", "Gedan-Barai", "Soto-Uke" and other multiple Hein-Shodan moves. Their data-set was normalized, so that no attribute can be a dominant factor in the data-set. Moreover, the data-set is divided into two parts, The right way and two common mistakes of each movement. We used 80% of our data-set in training and the rest was for testing.

3.5. Kinect Sensor

The main hardware used in this paper is the Kinect. As mentioned in [19], [16] the Kinect's hardware is composed of an Infrared Emitter to track the body, displaying a basic skeleton and the body's joints using the Microsoft SDK for Kinect.

Furthermore, as stated in Microsoft [16] the Kinect is capable of providing 30 frames per second with a 640 x 480-pixel resolution using its video and depth sensor cameras. The Kinect works by starting the camera and capturing the RGB (red, green and blue) colors of the person to form its image. Then, the monochrome sensor and infrared projector start to receive the rays that were emitted to get the third dimension and form the 3D imagery of the skeleton of the person.

4. Experiments And Results

As mentioned in section 3.4, a data-set has been used from [11] which contains Kata 1 moves including ”Age-Uke”, ”Mae-Geri”, ”Gedan-Barai”, ”Soto-Uke” to evaluate the system. All the joints positions in the movement have been classified. Then the confusion matrix has been calculated. As shown in TABLE 1. The average accuracy that is calculated from the confusion matrix is 90%.

Table 1. Results

Move	Trials	Correct	Wrong	%
Age-Uke	10	10	0	100%
Mae-Geri	10	10	0	100%
Gedan-Berai	10	8	2	80%
Soto-Uke	10	8	2	80%

Afterwards, to prove that the system could run efficiently in real-time, the system was tested with five Karate players from Al Ahly Sporting Club using our temporary data-set collect from Kinect v2. The data-set is divided into three files for each move, one of the files is the move done correctly and the other two files are two different common mistakes for the same move. The temporary data-set contains two moves, which are ”Mae-geri” and ”Age-uke”.

Fig. 3. Real-Life Experiment

As shown in figure 3, The performer is facing the Kinect and doing the move. The system captures the move while he is performing it, stores the data in a file, classifies the data and presents the results in real-time. A Karate coach was present to tell exactly what mistakes were made while performing the move and to discuss the result of the system.

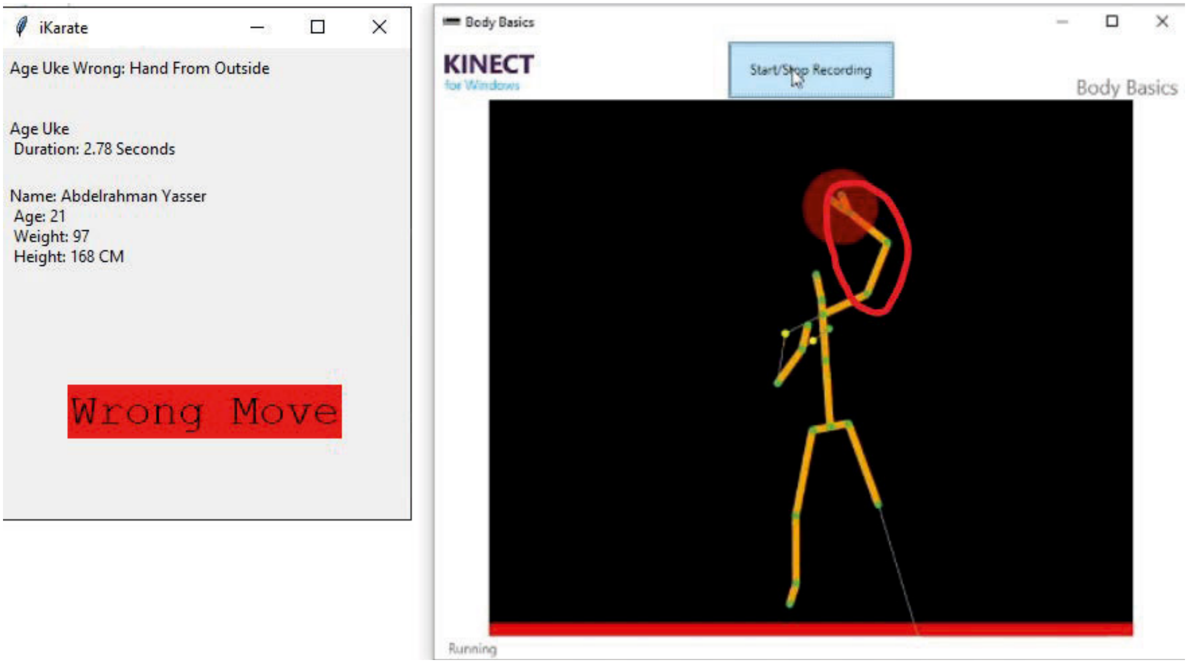

Fig. 4. Real-Life System

As shown in figure 4, the system classified the movement correctly and identified the mistake as the coach stated. Then the system presented the player with the mistake that he had done in the movement.

5. Discussion

The proposed classification method using F-DTW gave us valuable information. The movements that was used and tested in this experiment as mentioned in 3.4 are ”Age-Uke”, ”Mae-Geri”, ”Gedan-Barai”, ”Soto-Uke”, and it was tested with five professional players. F-DTW was able to recognize the full movement and classify it correctly with an average accuracy of 90%. This means that if the input signals were less than the data-set, F-DTW will be able to handle it. While the movement is being analyzed dynamically, it could be detected right before the motions ends if it matches the key-frames.

6. Conclusion and Future Work

It has been concluded from the related work that the Kinect sensor might be better in acquiring the body skeleton and joints. In our pre-processing, ”Z-score normalization” algorithm has been applied in way such that it can accommodate to different body proportions and the dominant factors in the data. F-DTW was used to accommodate to each player’s speed of doing a move. The application achieved an average accuracy of 90% For future work, we intend to collect our own data-set and use better algorithms to enhance the accuracy and the response time. Moreover, outliers should be removed from the data.

References

[1] Alborno, P., Giorgis, N.D., Camurri, A., Puppo, E., 2017. Limbs synchronisation as a measure of movement quality in karate. Proceedings of the 4th International Conference on Movement Computing - MOCO 17 doi:10.1145/3077981.3078027.

[2] Calin, A.D., 2016. Variation of pose and gesture recognition accuracy using two kinect versions. 2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA) doi:10.1109/inista.2016.7571858.

[3] Choubik, Y., Mahmoudi, A., 2016. Machine learning for real time poses classification using kinect skeleton data. 2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV) doi:10.1109/cgiv.2016.66.
[4] Escobedo-Cardenas, E., Camara-Chavez, G., 2015. A robust gesture recognition using hand local data and skeleton trajectory. 2015 IEEE International Conference on Image Processing (ICIP) doi:10.1109/icip.2015.7350998.
[5] Hachaj, T., Koptyra, K., Ogiela, M.R., 2018a. Averaging of motion capture recordings for movements' templates generation. Multimedia Tools and Applications 77, 30353–30380. doi:10.1007/s11042-018-6137-8.
[6] Hachaj, T., Ogiela, M.R., Koptyra, K., 2015a. Effectiveness comparison of kinect and kinect 2 for recognition of oyama karate techniques. 2015 18th International Conference on Network-Based Information Systems doi:10.1109/nbis.2015.51.
[7] Hachaj, T., Ogiela, M.R., Koptyra, K., 2015b. Human actions modelling and recognition in low-dimensional feature space. 2015 10th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA) doi:10.1109/bwcca.2015.15.
[8] Hachaj, T., Ogiela, M.R., Koptyra, K., 2015c. Learning from annotated video: An initial study based on oyama karate tournament recordings. 2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC) doi:10.1109/3pgcic.2015.16.
[9] Hachaj, T., Ogiela, M.R., Piekarczyk, M., 2013. Dependence of kinect sensors number and position on gestures recognition with gesture description language semantic classifier. Proceedings of the 2013 Federated Conference on Computer Science and Information Systems .
[10] Hachaj, T., Ogiela, M.R., Piekarczyk, M., 2014. Real-time recognition of selected karate techniques using gdl approach. Advances in Intelligent Systems and Computing Image Processing and Communications Challenges 5 , 99–106doi:10.1007/978-3-319-01622-1_12.
[11] Hachaj, T., Ogiela, M.R., Piekarczyk, M., 2017a. The open online repository of karate motion capture data: A tool for scientists and sport educators. 2017 IEEE Symposium Series on Computational Intelligence (SSCI) doi:10.1109/ssci.2017.8285270.
[12] Hachaj, T., Ogiela, M.R., Piekarczyk, M., Koptyra, K., 2017b. Advanced human motion analysis and visualization: Comparison of mawashi-geri kick of two elite karate athletes. 2017 IEEE Symposium Series on Computational Intelligence (SSCI) doi:10.1109/ssci.2017.8285269.
[13] Hachaj, T., Piekarczyk, M., Ogiela, M., 2017c. Human actions analysis: Templates generation, matching and visualization applied to motion capture of highly-skilled karate athletes. Sensors 17, 2590. doi:10.3390/s17112590.
[14] Hachaj, T., Piekarczyk, M., Ogiela, M.R., 2018b. How repetitive are karate kicks performed by skilled practitioners? Proceedings of the 2018 10th International Conference on Computer and Automation Engineering - ICCAE 2018 doi:10.1145/3192975.3192989.
[15] Janaki, V.M.S., Babu, S., Sreekanth, S.S., 2013. Real time recognition of 3d gestures in mobile devices. 2013 IEEE Recent Advances in Intelligent Computational Systems (RAICS) doi:10.1109/raics.2013.6745463.
[16] Miles, R.S., 2012. Start here!: learn the Kinect API. Microsoft.
[17] Tanaka, K., 2017. 3d action reconstruction using virtual player to assist karate training. 2017 IEEE Virtual Reality (VR) doi:10.1109/vr.2017.7892343.
[18] Thanh, N.T., Tuyen, N.D., Dung, L., Cong, P.T., 2017. Implementation of technical data analysis of skeleton extracted from camera kinect in grading movements of vietnamese martial arts. 2017 International Conference on Advanced Technologies for Communications (ATC) doi:10.1109/atc.2017.8167625.
[19] Zhang, Z., 2012. Microsoft kinect sensor and its effect. IEEE Multimedia 19, 4–10. doi:10.1109/mmul.2012.24.