A Multimodal Dataset for the Analysis of Movement Qualities in Karate Martial Art

Ksenia Kolykhalova*, Antonio Camurri†, Gualtiero Volpe‡, Marcello Sanguineti§,
Enrico Puppo¶ and Radoslaw Niewiadomski‖
DIBRIS, University of Genoa
Genoa, Italy
Email: *ksenia.kolykhalova@dibris.unige.it, †antonio.camurri@unige.it,
‡gualtiero.volpe@unige.it, §marcello.sanguineti@unige.it,
¶enrico.puppo@unige.it, ‖radoslaw.niewiadomski@dibris.unige.it

Abstract—A multimodal dataset is presented, which has been collected for analyzing and measuring the quality of movement performed during sport activities. Martial arts (namely karate) are taken as test-beds for investigation. Karate encompasses pre-defined sequences of movements ("katas") that can be carried out with different qualities, e.g., by performers at different skill levels (highly vs. poorly skilled). The experimental setup and method are described. The dataset is composed of motion capture (MoCap) data, synchronized with video and audio recordings, of several participants with different levels of experience. The raw MoCap data are independent of any particular post-processing algorithm and can be used for other research purposes. In the second part of the paper, a set of measures is proposed to evaluate a kata performance. They are based on the geometrical and kinematic features, such as posture correctness and synchronization between limbs. and were chosen according to karate experts' opinion.

Keywords—*Motion capture, multimodal dataset, karate, movement features*

I. Introduction

Capture and analysis of human motion is a progressing research area, due to the large number of potential application and its inherent complexity. Motion Capture (MoCap) started as an analysis tool in biomechanics research, but has grown importance as a source of motion data for computer animation as well as education, training, and sport. At the same time, there exists the need for developing methods that can objectively assess quality of movement, e.g., to measure sport or dance performance or rehabilitation activities of people with motoric disabilities. The overall quality of a performance depends on many features, such as postural control, coordination, and balance. The main question to be addressed is how such a quality can be measured and possibly automatically computed.

The purpose of this work consists of presenting a dataset that provides an initial reference archive for investigating techniques aimed at evaluating the quality of full-body movements that we plan to perform in future work. Sport is taken as test bed for investigation, as quality of movement is one of the key aspects for a successful sport performance. In more detail, the proposed dataset focuses on katas, i.e., predefined sequences of movements in karate, used for training purposes. For example, in a kata the posture has a great effect on psychological and physical well-being; maintaining the appropriate posture means increasing energy, reducing muscle strain, and preventing repetitive strain injuries. Thus

a kata executed by martial artists with different levels of skill offers (i) enough flexibility in being performed with a different quality of movement and (ii) a reference sequence of specific and pre-defined movements allows comparisons among different performers and qualities. Moreover, various movement features can be extracted from a kata.

For the purpose of this work, 3D MoCap data, synchronized with video and audio recordings, of several kata performances were collected. A collection of features is proposed for future analysis, evaluating different aspects of the performance such as appropriate posture, acceleration of the kicks and punches, and synchronization of limbs movements.

The remaining of the paper is structured as follows. Section II describes the state of the art in motion capture and existing datasets of sport activities. Section III presents the data collection process: experimental setup (participants, equipment) and post-processing of data for further use in analysis. Section IV shows an initial collection of proposed features for analysis of performance quality from MoCap data. These include features for geometrical analysis and kinematic analysis to be extracted from the collected data set. In Section V, conclusions are drawn.

II. State of the Art

A. Analysis of sport activities

Nowadays wide archives of sport activities (e.g. karate, dance) data are used for reseach purposes. For example, karate data is used to investigate the reaction time and anticipation [1], kicking limb coordinative patterns [2], in biomechanical analysis [3] or to develop segmentation techniques [4].

In particular, evaluating and ranking the performance of users based on MoCap data is challenging research question that is recently often addressed. For example, Alexiadis et al. [5] proposed a system that automatically evaluate dance performance against a reference dance i.e., "gold standard" performance. The method is based on motion acquisition via Kinect, human skeleton tracking, and the application of quaternionic signal-processing techniques for a) temporal synchronization, b) spatial aligninment and c) comparison of joint positions and joint velocities of two "dance signals". In a study by Kristan et al. [6] the trajectory-based evaluation of multi-player basketball activity is performed with the use of Bayesian network.

B. Karate datasets in previous studies

Bianco and Tisato [7] proposed an algorithm for karate movement recognition from skeletal motion. It takes as input sequences of skeletal motions that can be derived from either motion capture hardware or consumer-level devices such as Microsoft Kinect. The dataset consists of punches, kicks, and defense karate moves. The items can be single actions or more complex sequences of movements. In the study by Hachaj and Ogiela [8], a multi-Kinect environment was used for recording a dataset consisting of one black belt karate instructor performing four static stances, two blocks, and one kick. The authors proposed a semantic classifier, it is based on forward chaining inferring schema performed on the set of rules that are defined with formal LALR grammar. The set of rules is called "Gesture Description Language" (GDL). The idea of the GDL approach is to code gesture sequences as series of static key frames that appear in a defined order. A study by Witte and colleagues [9] focused on a procedure to visualize movement patterns on the basis of relevant movement angles, in such a way to get a visual impression of the holistic movement. This procedure was applied to a karate front kick, which was executed by five athletes recorded with a VICON system. The authors proposed a method to identify similarities and differences in coordination between the movements of the individual performers. The freely-available Carnegie Mellon University Graphics Lab Motion Capture Database contains MoCap data recordings of eight different katas and front kicks, performed by one individual.[1]

III. Experimental Setup and Method

The data were recorded at the Casa Paganini - InfoMus lab premises. The optical Motion Capture System Qualisys was used, which consists of 9 high-resolution cameras. Recordings were done using the Qualisys system's native software Qualisys Track Manager (QTM) 2.9. Post-processing was applied to the data: labeling the markers, building the models, and cleaning data from noise, occurred by "ghost" or jitter markers. Motion capture data files were exported in TSV format and imported to Matlab for further analysis with use of the MoCap Toolbox [10] and other functions of Matlab.

A. Participants

A total of 7 participants took part in the study (see Table I). All participants received a martial art education in Karate, however they have different levels of experience.

TABLE I.   Participants

#	Description of the participant	Years of practicing	Experience (from 1-low to 5-high)
1	Adult, Male, Karate teacher	>15	5
2	Adult, Male, Black Belt	>10	4
3	Adult, Female, Black Belt	>10	4
4	Teenager, Male	>5	3
5	Adult, Male, Participant of the Championships	>15	5
6	Teenager, Male	>5	3
7	Teenager, Male	>5	3

Participants were asked to perform 2 exercises - katas, each consisting of a specific sequence of movements in karate. They

---

[1] http://mocap.cs.cmu.edu/search.php?subjectnumber=135
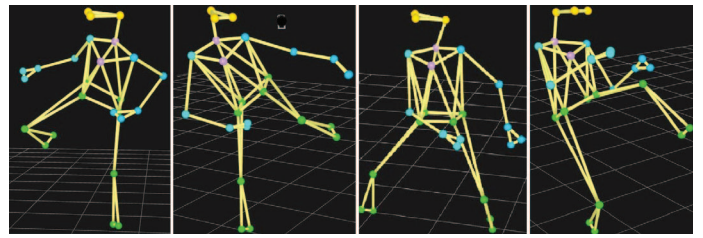

Fig. 1.   Examples of postures from kata Heian Yondan (MoCap data)

could perform them with their own speed and rhythm, not exceeding the specified area of the floor that was captured with cameras.

B. Performance

Martial arts (e.g., karate, taekwondo, etc.) use forms, i.e., katas or patterns in order to help practice certain moves (e.g., kicking combinations) as well as for improving physical conditioning, muscle memory, focus/concentration, and balance. Katas help to practice "offensive" or "defensive" techniques.

For the purpose of this work, two katas were chosen: Heian Yondan and Bassai Dai. The choice was based on discussions with karate experts, who helped to chose the sequences of movements most suitable for investigation. Moreover, the experts indicated essential aspects of each technique. The two katas are different in the movements, speed, rhythm, and difficulty. The first kata focuses on leg techniques, whereas the second one on hand techniques. Heian Yondan is introduced at the green belt skill level as one of the important exercises and consists of 27 karate movements. It is used to practice the varieties of rhythm inherent in karate, the slow increase in tension on the one hand and the sudden release of powerful movements on the other hand. This contrast between slow and fast movements serves to teach rhythm to the beginning students, and helps them to develop the control over the body to avoid a rush. It also helps them to develop their leg movements, and the ability to synchronize leg and arm movements (see Figure 1). Rhythm is very important in this kata.

Bassai Dai is more difficult. It is introduced at the brown belt skill level and often performed for black belt examination. It consists of 38 karate movements (counts). It is characterized by many switching arm blocking combinations (see Figure 2). The most difficult part of Bassai Dai is the hip motion. Fast, strong, and snappy hip rotation are needed to perform a good Bassai Dai [2].
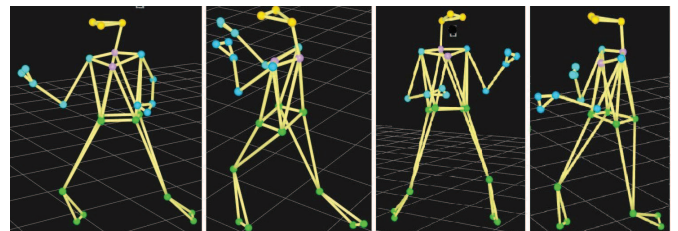

Fig. 2.   Examples of postures from kata Bassai Dai (MoCap data)

---

[2] http://www.tskc.co.uk (JKA Shotokan Kata)

Recordings of these two katas by multiple differently skilled performers make a multi-purpose dataset, which can be used not only for analysis of general aspects (e.g., correctness of the posture, velocity, and so on), but also for more specific ones, such as biomechanics of the punches (for Bassai Dai) or kicks (for Heian Yondan).

C. Recordings

The positions of 25 reflective markers in the 3D space were tracked at a frame rate of 250 Hz. The markers were attached to a black elastic suit on the same joints for each person. The voice sound was recorded with a microphone positioned on the head of the performer. The following audio parameters were set: bitrate - 320 kbit/sec, sampling rate - 48 kHz, two channels (right channel for SMPTE synchronization, left channel for audio). Motion capture and audio recordings were synchronized with a video camera recording of the performances using SMPTE timestamps. The following parameters were set for the video recordings: resolution 1024x576, 50 fps. Participants were recorded individually and were requested to perform 3 trials for each of the two katas, remaining in the center of the capture area indicated by lines on the floor. Therefore, the recorded data consists of 6 recordings for each participant, and of a total number of 42 recordings. The duration of the recordings of kata Bassai Dai is from 71s to 115s, SD = 12.9s. For Heian Yondan it is from 50s to 97s, SD = 13.3s.

D. Post-processing

Automatic identification of trajectories in the QTM software is performed by a module called Automatic Identification of Markers (AIM). In order to create a model of the body skeleton, a manual identification of all the recorded trajectories has to be done. When data is identified, the AIM model can be created and applied to other recordings of the same person. This procedure reduces manual processing time of MoCap recordings. However, 30-35 % of trajectories remains unidentified, and they have to be identified manually. Several problems occurred in the process of identifying trajectories: trajectories overlapping, "ghost" markers, reflections, jittering were eliminated manually. The following list of labels, that was created separately, was used for all recordings (see Table II).

TABLE II.   LABEL LIST

#	Marker Location	Label	#	Marker Location	Label
1	Back Head	BKHD	14	Right Pinkie Finger	RPNK
2	Right Front Head	RFHD	15	Left Pinkie Finger	LPNK
3	Left Front Head	LFHD	16	Right Front Hip	RFHP
4	C7 on the Spine	C7	17	Right Back Hip	RBHP
5	Neck	NCK	18	Left Front Hip	LFHP
6	Right Shoulder	RSHD	19	Left Back hip	LBHP
7	Left Shoulder	LSHD	20	Right Knee	RKNE
8	Right Elbow	RELB	21	Left Knee	LKNE
9	Left Elbow	LELB	22	Right Front Ankle	RFAK
10	Right Wrist	RWRS	23	Right Back Ankle	RBAK
11	Left Wrist	LWRS	24	Left Front Ankle	LFAK
12	Right Index Finger	RIND	25	Left Back Ankle	LBAK
13	Left Index Finger	LIND			

A separate AIM model for each participant was created. For several markers, trajectories were missing because of occlusions and jittering. The most problematic areas were hips, knees, and both hands, due to specific movements in

karate. The missing trajectories were filled in using linear or polynomial interpolation algorithms available in the QTM software. The choice of the interpolation algorithm for the missing trajectories was performed manually for each gap of frames. Polynomial interpolation was preferred for big gaps. Average of the missing trajectories does not exceed 2.5 - 3% for each marker.

IV.   KARATE MOVEMENT FEATURES

Experienced karate martial artists were interviewed to identify the features they deem important for assessing the quality of a kata performance. Based on such experts' knowledge, in this section we propose a collection of features that can be extracted from the recorded movements. The features can be grouped into three sets: geometric features, kinematic features, and features related to synchronization of body movements. Geometric features focus on the correctness of the posture during the performance: shoulders bending and straightness of the back during the whole performance. Kinematic features include maximum acceleration developed during the performance and biomechanical efficiency in isotonic movement as a balance between acceleration and deceleration of the wrist/ankle during a punch/kick. Synchronization features concern intrapersonal synchronization, i.e., whether body joints moments are coordinated. Computation details are described below.

A. Geometric features

1) Angle between shoulders and neck: According to the experts, the performer has to maintain the angle between shoulders (see Figure 3) within a specific range: one has to avoid bending back or moving shoulders forward. However, even if this angle increases, the posture of the participant can still be considered as correct because of punches/kicks. Indeed, when one hand performs a punch and moves to one direction, the other hand is going to the opposite direction moving the shoulder to the back, which makes the angle between shoulders increase. Therefore, this posture is acceptable but the opposite situation (i.e., a decreasing angle) is always considered as an error.

Incorrect postures can be detected by analyzing the changes of the angle between shoulders and spine. The recorded data contain the marker positions in the 3D space. Hence, the x, y, and z coordinates are available for each marker for all the recorded frames. By extracting the positions of markers LSHD, RSHD, NCK, and C7, the values of the angles θ and ϕ can be calculated according to the set of equations 1(see Figure 3).
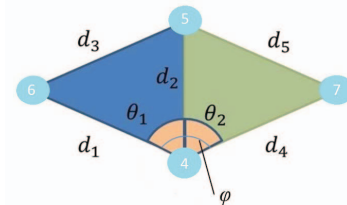

Fig. 3.   The angle between shoulders is computed as the angle ϕ between markers RSHD(6), C7(4), and LSHD(7), NCK(5)
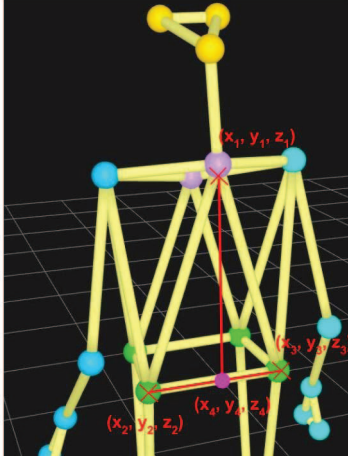
76

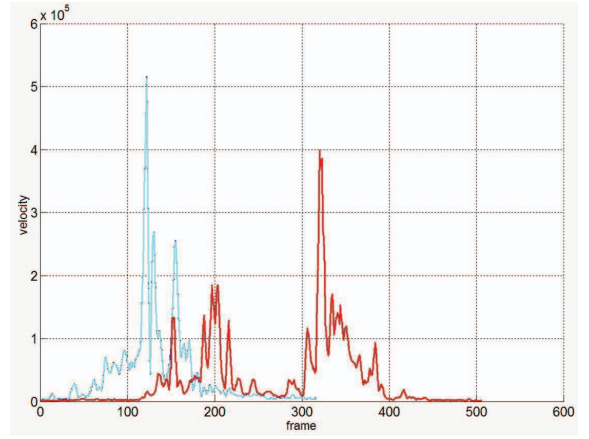Fig. 4. Distance between marker C7 and the mid-point between LBHP and RBHP

Fig. 5. Velocity during a punch. The blue line refers to participant #5 (high experienced), the red to participant #4 (intermediate level of experience)

$$
\begin{aligned}
d_3^2 &= d_1^2 + d_2^2 - 2 \cdot d_1 \cdot d_2 \cdot cos(\theta_1), \\
d_5^2 &= d_4^2 + d_2^2 - 2 \cdot d_4 \cdot d_2 \cdot cos(\theta_2), \\
\theta_1 &= acos(\tfrac{d_1^2 + d_2^2 - d_3^2}{2 \cdot d_1 \cdot d_2}), \\
\theta_2 &= acos(\tfrac{d_4^2 + d_2^2 - d_5^2}{2 \cdot d_4 \cdot d_2}), \\
\phi &= \theta_1 + \theta_2
\end{aligned}
\tag{1}
$$

Next, we find the mean value μ_ϕ of this angle, for each participant over time. Angles that are in interval $[0.9 * \mu_\phi, \mu_\phi]$ are considered as correct; incorrect angles are smaller than $0.9 * \mu_\phi$; acceptable angles are all larger than μ_ϕ. Equation 2 provides the proportion of wrong postures in the performance:

$$
V_1 = \frac{N_{wrong}}{N},
\tag{2}
$$

where N_{wrong} is the number of frames where the angle ϕ between shoulders and neck was incorrect and N is the total number of frames in the performance.

2) Straightness of the back: The back of the performer has to be straight during the whole performance, therefore the distance between head and hips has to be almost constant. When the back is straight, the whole body is aligned, producing significantly more power and having a better balance. In order to measure whether the back is straight, the distance between marker C7 and the mid-point between markers LBHP and RBHP is computed (see Figure 4 and set of Equations 3) by considering the matrices of the x_i, y_i, and z_i, $i = 1..3$ coordinates of the three markers (RBHP, LBHP, and C7) for each frame of a recorded performance.

$$
\begin{aligned}
x_4 &= \frac{(x_2 + x_3)}{2}, \\
y_4 &= \frac{(y_2 + y_3)}{2}, \\
z_4 &= \frac{(z_1 + z_2)}{2}, \\
Dist &= \sqrt{((x_4 - x_1)^2 + (y_4 - y_1)^2 + (z_4 - z_1)^2)},
\end{aligned}
\tag{3}
$$

where x_4, y_4, z_4 are the coordinates of the middle point between RBHP and LBHP. When $Dist$ decreases, the posture of the back is incorrect. Since each individual has a different length of the back, the values of $Dist$ can be compared only when normalized. For each participant the mean μ_{Dist} and the standard deviation σ_{Dist} of all the $Dist$ for each participant over time are computed. In order to compare participants, a measure of how large the standard deviation σ_{Dist} is with respect to mean μ_{Dist} is computed as:

$$
V_2 = \frac{\sigma_{Dist}}{\mu_{Dist}}
\tag{4}
$$

B. Kinematic features

1) Maximum acceleration during the performance: Both power and accuracy are affected by acceleration. The acceleration of the wrist markers is computed using the MoCap toolbox and for each participant, the maximum acceleration in performance is calculated:

$$
V_3 = max(acceleration of marker)
\tag{5}
$$

2) Biomechanical efficiency in punches and kicks: Having a large acceleration value is not sufficient to perform a good punch. The balance between acceleration and deceleration is important, too. According to biomechanics, a punch/kick can be considered as an isotonic movement. In this kind of movements, balance between acceleration and deceleration is very important. In a punch, a big acceleration has to occur in order to increase the momentum, which is responsible for the power of the punch, but velocity has to be slowed down in order to avoid injuries.

Figure 5 shows the peak velocity during a punch, which divides the movement into two parts - acceleration and deceleration. When comparing high experienced with less experienced participants, it can be seen that the movement of participant #5 (with more experience - blue color) displays a more evident peak.

In order to measure how participants maintain balance between acceleration and deceleration, the time spent for both parts of the movement is calculated. This can be done separately for the hand and the leg that is performing an action. Velocity is calculated for the whole segment of a
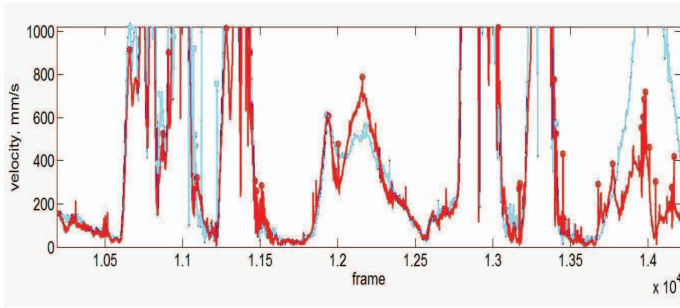
Fig. 6. Example of peaks of velocity (blue refers to the RSHD marker, red to the RWRS marker) for participant #5

punch/kick movement. The maximum velocity is taken as a separation between the first phase - acceleration and the second phase - deceleration, times t_{acc} and t_{dec} are calculated as the time spend for acceleration and for deceleration, respectively. Equation 6 provides a measure of how balanced the action is:

$$V_4 = 1 - abs(t_{acc} - t_{dec}) \qquad (6)$$

In the ideal case $V_4 = 1$ (i.e., time spent for acceleration equal to the time spent for deceleration). In other cases, the closer V_4 to 1, the better balance between acceleration and deceleration is.

C. Synchronization between pairs of joints

According to a previous study, karate experts have more synchronized wrist-shoulder movements, which allow them to generate a strong power while punching [11]. Synchronization of relevant joints pairs such two wrists, or wrist-knee is thus important for assessing movement quality. In order to compute it, velocity is calculated with the MoCap toolbox and peaks of velocities (local maxima) are extracted for both markers. Synchronization of pairs of joints is measured using the event synchronization approach. This was followed in [12] as a method to measure synchronization and time-delay patterns between signals, based on the relative timings of events in time series. Events were defined, e.g., as local maxima and it was proposed therein to count the fraction of event pairs matching in time and how often each time series leads in these matches. An example is shown in Figure 6. The larger the event synchronization index, the more synchronized the movements (see Equation 7).

$$V_5 = I_{synch}(RWRSvs.RSHD) \qquad (7)$$

The analysis of synchronization can be extended to multiple joints following the approach described in [13] for analysis of synchronization in dance performances with multiple participants.

V. CONCLUSION

We presented a multimodal dataset of karate performances including synchronized MoCap, video, and audio recordings. In order to explore several directions for analysis, rather than addressing a single action movements such as a punch or a kick, the work focused on recording complex sequences of movements (katas) that need different hand/leg techniques. 7 participants with different levels of skill were chosen for recording the dataset. We provided examples of measures, which can be used to evaluate the quality of the performance, using this dataset. This is a work in progress. We plan to combine the different measures presented in the paper in a global index. This index will be used to evaluate the overall quality of the performance and will be assessed through perceptive studies with experienced and non-experienced in karate observers.

REFERENCES

[1] S. Mori, Y. Ohtani, and K. Imanaka, "Reaction times and anticipatory skills of karate athletes," *Human movement science*, vol. 21, no. 2, pp. 213–230, 2002.

[2] F. Quinzi, P. Sbriccoli, J. Alderson, A. Di Mario, and V. Camomilla, "Intra-limb coordination in karate kicking: Effect of impacting or not impacting a target," *Human Movement Science*, vol. 33, no. 1, pp. 108–119, 2014.

[3] P. R. Cavanagh and J. Landa, "A biomechanical analysis of the karate chop," *Research Quarterly. American Alliance for Health, Physical Education and Recreation*, vol. 47, no. 4, pp. 610–618, 1976.

[4] J. Barbic, J.-Y. Pan, C. Faloutsos, J. K. Hodgins, and N. Pollard, "Segmenting motion capture data into distinct behaviors," in *In Proceedings of Graphics Interface 2004*, May 2004, pp. 185 – 194.

[5] D. S. Alexiadis, P. Kelly, P. Daras, N. E. OConnor, T. Boubekeur, and M. Moussa, "Evaluating a dancers performance using kinect-based skeleton tracking," in *Proceedings of the 19th ACM International Conference on Multimedia*, ser. MM 11, 2011, pp. 659–662.

[6] M. Kristan and J. Per, "Automatic evaluation of organized basketball activity using bayesian networks," in *Proceedings of the Computer Vision Winter Workshop 2007*, 2007, pp. 1–8.

[7] S. Bianco and F. Tisato, "Karate moves recognition from skeletal motion," in *Proc. SPIE*, vol. 8650, 2013, pp. 86 500K–86 500K–10.

[8] T. Hachaj and M. R. Ogiela, "Recognition of human body poses and gesture sequences with gesture description language," *Journal of medical informatics and technology*, vol. 20/2012, pp. 129–135, oct 2012.

[9] K. Witte, P. Emmermacher, N. Langenbeck, and J. Perl, "Visualized movement patterns and their analysis to classify similarities-demonstrated by the karate kick mae-geri," *Kinesiology*, july 2012.

[10] P. T. Burger, "Mocap toolbox manual," http://www.jyu.fi/music/coe/materials/mocaptoolbox/: University of Jyvaskyla, 2010.

[11] R. Roberts, B. D. P.G. Bain, and M. Husain, "Individual differences in expert motor coordination associated with white matter microstructure in the cerebellum," *Cereb Cortex*, no. 23, p. 22822292, 2013.

[12] R. Quian Quiroga, T. Kreuz, and P. Grassberger, "Event synchronization: A simple and fast method to measure synchronicity and time delay patterns," *Phys. Rev. E*, vol. 66, p. 041904, Oct 2002.

[13] G. Varni, G. Volpe, and G. Mazzarino, "Towards a social retrieval of music content," in *Proc. IEEE International Conference on Social Computing SocialCom'2011*, IEEE, 2011, pp. 1466–1473.