A theoretical framework of extrinsic feedback based-automated evaluation system for martial arts

Wan Mohd Rizhan Wan Idris1,2,*, Ahmad Rafi3, Azman Bidin2, Azrul Amri Jamal1

1Faculty of Informatics & Computing, Universiti Sultan Zainal Abidin (UniSZA), Besut Campus, 22200, Terengganu, Malaysia
2Faculty of Creative Multimedia, Multimedia University (MMU), Cyberjaya Campus, 63000 Cyberjaya, Selangor, Malaysia
*Corresponding author E-mail: wmrizhan@unisza.edu.my

Abstract

Martial arts (MAs) are considered as a preserved heritage primarily due to the fact that it promotes certain level of identities of a culture. MA refers to the art of combat and self-defense which normally combines offensive and defensive techniques. Technology advancements have made motion capture (MoCap) to be widely used in MA to capture and evaluate human performance. Nevertheless, researches on extrinsic feedbacks (EFs) of MA through the developed evaluation system are scarce. Furthermore, there is no complete framework of evaluation system suggested for MA. This paper presents the theoretical framework of EF-based automated evaluation system in the context of traditional local MA. The framework contains three modules including MoCap, recognition and evaluation. The MoCap module tracks human body accurately in order to generate skeleton, tune focused target, and record human movements. Recognition module develops a script of motion for templates and classification purposes using Reverse-Gesture Description Language (R-GDL) and GDL respectively. Evaluation module produces the extrinsic feedback in terms of pattern and score for the performed movements. This theoretical framework will be used in the development of the digital tool to measure the accuracy and effectiveness of motions performed by one of the traditional local MAs.

Keywords: Martial Arts, Motion Capture, Gesture Description Language, Reverse-Gesture Description Language.

1. Introduction

A country’s cultural heritage provides a foundation for the uniqueness of the nation and its people [1]. Nonetheless, with the advancement of technology, many cultural values have been forgotten [1]. Tradition and cultural values might be lost in the near future if there is no proper steps taken to preserve them. Martial arts (MA) can be considered as a preserved heritage primarily due to fact that it represents certain level of identities and cultures. In general term, MA refers to the art of combat and self-defense [2]. It normally combines offensive and defensive techniques [3]. In Asia, MAs may promote identities of certain races and cultures [4]. There are various styles of MAs practised in the world for examples ESDO, Aikido, Wing Tsun, Judo, Jujutsu, Karate, Taekwondo and Kendo (Asian MAs [5]). Tai-chi-chuan, pa-koua and Shing-hi-chuan (Chinese MAs [6]), Capoeira (Brazil [7]), Mixed Martial Arts (MMA) [8]. In Southeast Asia, traditional Malay MA is a branch of knowledge (physical and spiritual) used for self-protection systematically [4]. Each MA exists with a distinct history, philosophy and set of techniques [2]. Capturing and analysing human motion or frequently called human motion capture is a rapidly growing research area due to the large number of potential application and its inherent complexity. Apart from its usage in biomechanics research, the source of motion data is important for computer animation as well as education, training, and sports [9]. Motion capture (MoCap) is the process of recording a live movement event to obtain a single 3D representation of the performance by translating it into usable mathematical terms using tracking a number of key points in space over time [10]. In other words, MoCap refers to the technology for translating a live movement into a digital performance. Using MoCap technology, the results of the effectiveness and performance in MA techniques can be seen through the feedbacks given by the developed system. The feedbacks can be divided into 2 types: intrinsic feedback; and extrinsic feedback. Intrinsic feedback refers to information that comes through a person’s own sensory-perceptual or nervous system such as muscle spindles, joint receptors, lumbar muscle and lumbar proprioception [11]. Extrinsic feedback (EF) refers to information comes from outside sources such as teacher, therapist, and videotape [12]. Nevertheless, research on EF of the MAs using MoCap technologies are scarce. Furthermore, there is no complete framework of evaluation system suggested for MAs. Feedbacks given by the evaluators such as teacher, referee and jury to evaluate the MA techniques are ambiguous [13], lack of fairness and objectivity [14]. EFs using teacher’s verbal and videotape require the presence of teachers or instructors to detect and reprimand one-by-one of the errors via video. Therefore, in this paper, a theoretical framework of EF-based automated evaluation system is presented and discussed in detailed to evaluate the effectiveness of performing techniques of a traditional local MA called Seni Silat Cekak Malaysia (SSCM).
2. Materials and methods

The traditional MAs practiced among Malay people are known as “Silat”. Silat involves physically and mentally movements to defend the practitioners systematically [15]. There are approximately 439 styles of traditional Malay MAs registered in National Silat Federation (PESAKA) [16]. The main purpose of PESAKA is to set up a merged MA organisation in Malaysia. PESAKA was established on September 24, 1983. Four Malaysia level organizations acknowledged as the Founder Bodies in PESAKA are Seni Silat Lincah Malaysia, Silat Seni Gayong Malaysia, Seni Gayung Fatani Malaysia, and Seni Silat Cekak Malaysia with 16 other state level organisations [16]. In this study, Seni Silat Cekak Malaysia (SSCM) has been selected to be the case study. SSCM is an established traditional Malay MA and one of the founder bodies of PESAKA. Furthermore, SSCM has been registered as co-curriculum subject in most of Institutions of Higher Learning (IHLs) in Malaysia. SSCM is a genuine traditional Malay combat-oriented art covering 99% defending techniques and 1% attacking technique [15]. This shows that SSCM focuses on defensive aspects and symbolizes Malay self-defense movements [15]. In SSCM, components of movements can be classified into several methods namely Kaedah (Fend off technique), Buah Asas (Continuation of fend off technique before taking down the opponent), Buah Jatuh (Takedown technique), Buah Potong (Clinch techniques that teach the practitioner to escape even when in close proximity with the opponent) and Buah Serang (Strike technique) [15].

In order to ensure the success of this study, preliminary study has been done by conducting acquaintance meeting, procedures and approval, and interview. The outputs from these approaches are very important for the experiment and proposed framework.

2.1. Acquaintance meeting

An acquaintance meeting with an influential individual in SSCM association (PSSCM) has been held in order to obtain the overview of PSSCM. The influential individual is the chairman of Centre of Education in PSSCM and at the same time acted as a SSCM supervisor and teacher in Islamic Science University of Malaysia (USIM) and Sepang Vocational College (KVS). Information regarding history, mission and vision, administration, taught subjects, and members of the association has been obtained during the meeting. Additionally, the meeting allowed findings from the literature study to be validated. Several potential informants were identified for this study including the informants for pilot study.

2.2. Procedures and approval

Several procedures with PSSCM have been followed in order to gain approval to implement survey, observation, and experimental study of SSCM. Starting with the acquaintance meeting between the author and influential individual in the PSSCM, several person-in-charge in the association have been identified including the SSCM secretary, teachers and trainers. In order to obtain permission to do the research on SSCM, a letter of application has been delivered to the PSSCM together with author’s verification letter of the research issued by Multimedia University for support, recommendation, and reference. The application letter has been distributed to ask permissions for conducting research along with SSCM teachers, trainers and trainees in exercise class and for implementing studies survey, observation and experiment in the exercise class using the proposed method in the study. The association has stated the permission to conduct the research on SSCM by replying the approval letter. This permission allows the author to plan the next agendas for this research including conducting interview and experiment for selected teacher.

2.3. Interview

After obtaining the approval from PSSCM, the author has approached one of the experienced SSCM teachers through an interview. The SSCM teacher is also a SSCM supervisor in University of Sultan Zainal Abidin (UniSZA) since 2005. The purpose of the interview is to gain information about techniques of SSCM and to deliver information about the experiment to be implemented. Tools of audio recording and notes have been used in this interview for author’s reference, support and further tasks.

In learning SSCM in IHLs, the trainees/students will be tested before going to higher levels of techniques. Two important tests are called Ujian Peringkat Asas (Basic Level Test) and Ujian Peringkat Jatuh (Takedown Level Test) where all the techniques in Buah Asas and Buah Jatuh are tested respectively. Students will be able to continue to Buah Potong once they passed the Ujian Peringkat Jatuh [17]. Since the author has no knowledge and experience in SSCM, technique selection is done by the SSCM teacher. Each level has 21 techniques, and the techniques can be classified into 4 clusters based on Kaedah A, Kaedah B, Kaedah C and Kaedah D. The SSCM teacher has agreed and decided to choose one technique from each cluster for the preliminary study. The chosen techniques are: Buah Pertama, Buah Kuntau Jatuh, Buah Klias Hadapan, and Buah Hengpok from Kaedah A, B, C, and D respectively. These selected techniques are relatively easier than other techniques in its cluster.

2.4. Experiment

For each cluster, the SSCM teacher will be advised to execute the selected technique from Buah Jatuh. This is because techniques in Buah Jatuh is an extension from Buah Asas.

Each execution of the selected technique needs to be repeated 10 times to obtain the right motions using the proposed method. This repetition is to ensure that the captured postures generated by R-GDL method are accurate. Therefore, there will be 40 executions for all selected techniques. Planning and scheduling the experiments are crucial to obtain good data collection for analysis and experiments to be done.

Due to practicality and portability factors, a standard Microsoft Kinect for Windows Version 2 and Kinect SDK 2.0 have been chosen to capture, track and record executions of the selected techniques. In term of practicality, participants involved in this experiment do not need to wear any marker which can limit their movements and performance [18]. The proposed tool used is based on marker-less technique. Furthermore, the participants do not need to attend in static studio for motion capture since the chosen tool is easier to carry. The experiments are conducted indoor. During the experiment, the performers are positioned at the distance 2.65m from the device since the distance are suitable to track the whole body even if some techniques executed require the performers to step forward or backward [19].

2.5. Proposed Framework

A theoretical framework of MAs evaluation system has been proposed in this paper to evaluate MAs movements based on extrinsic feedback. As illustrated in Figure 1, the framework can be divided into 3 important modules consisting of Motion Capture module, Motion Recognition module and Motion Evaluation module.

2.5.1. Motion capture

Motion Capture module mainly acts as the first module in the proposed conceptual framework. The function of this module is to accurately capture and track human bodies to generate skeleton, find focused target, and record motions in the repository. This module consists of Skeleton Tracking model, Focused Target model and Skeleton Recording model. This module becomes a reference for respectively motion template model and motion classification model in Motion Recognition module.
The purpose of the Skeleton Tracking model is to generate data based on the human bodies tracked or captured by the device. The device can produce infrared, colour, depth, and skeleton data through its sensor. The skeleton data has been chosen because the data is more accurate for evaluation of joints positions. Microsoft Kinect For Windows Version 2 can detect 25 joints in each human body, they are: SpineBase, SpineMid, Neck, Head, ShoulderLeft, ElbowLeft, WristLeft, HandLeft, ShoulderRight, ElbowRight, WristRight, HandRight, HipLeft, KneeLeft, AnkleLeft, FootLeft, HipRight, KneeRight, AnkleRight, FootRight, SpineShoulder, HandTipLeft, ThumbLeft, HandTipRight and ThumbRight [19].

• Focused target model
By default, the sensor may recognise up to 6 human bodies simultaneously. Thus, the events involving partners such as MAs can take advantage through this sensor’s capabilities. In real situation of MAs, the performers are normally surrounded by other people such as teachers, trainers, judges or other trainees as well as audience. Standard tracking method in this situation however, may cause confusion in determining the targeted performers for the evaluation. This is because all bodies are tracked using the same methods or codes by the system. Therefore, Focused Target model is proposed to only track the person being evaluated. This model is based on method of the first person to be seen by the sensor. Since the sensor can track 6 bodies, they are assigned in an array of the Body. Whenever one of the bodies in the array is tracked, the system records the index of the body as a class level variable. The system then use the index recorded to access the array for generating the skeleton data and ignore all other bodies.

• Skeleton recording model
Skeleton Recording model is used to record or save the skeleton data tracked by the sensor into repository. The data saved is then used in other main modules such as Motion Recognition and Motion Evaluation modules. Continuous motions are stored in a file containing frame list. A second of recording may contain 25 to 30 frames in which each frame contains of a full of human body skeleton [19]. In the file, data of time period, clip edges, hand left confidence, hand left state, hand right confidence, hand right state, restriction and tracking status of body, lean, lean tracking state, joints' orientation and position as well as tracking status are stored along with the frames list (Figure 2).

The tracking status is important to obtain percentages of tracked and inferred joints during the experiments. The percentages can be used to measure the correlations between the tracking status and the results of classification. Through the correlations, information about tracking status affecting whether good or poor results of classification can be obtained. The expectation of the tracking status, if the recordings may track 80% joints and above, this percentage exhibits high accuracy of the movements and effectiveness of the system. The results of classification are also expected to be high due to the system can track more than 80% joints and postures of the body during the execution of the techniques. If the recordings track below than 80%, it shows low accuracy of the movements. The low accuracy of the movements is expected to affect the results of classification (either for other expert or students).

2.5.2. Motion recognition
Motion Recognition module is purported to interpret motion data captured using Motion Capture module for creating motions template and classification. This module consists of Motion Template model and Motion Classification model. Both models in this module are used to support Motion Evaluation module. Common approaches to motion recognition are using statistical methods, neural networks and fuzzy sets, optimal path finding, semantic methods and finite state machines and natural interface [20]. These approaches however have limitations in terms of requiring very large training and validation sets which might be unintuitive for a skilled system user because the techniques have to be manually tuned and difficult to add new gesture to be recognised without additional intensive training of the classifier [21].

In this study, Motion Recognition module utilises Gesture Description Language (GDL) and Reverse-GDL (R-GDL) created by Dr. Tomasz Hachaj (as illustrated in figure 2). The advantages of GDL/R-GDL are as follows [22]:

• It is capable of classifying human body movements in real time.
• It can classify not only simple, real life gestures but also complicated movements like Karate techniques.
• It does not require large training dataset. Gestures are defined by user in GDL script. User can utilise as many body features as he or she needs in each rule definition.
• Gestures are split into key frames that appears in some order under given time restriction.
• The input data for classifier is set of body joints that arrive from tracking software in real-time (approximately with frequency 30 Hz).

• Motion template model
Motion Template model is used to develop a script of motions as a template for classification and matching purposes using R-GDL. This model utilises teachers’ motions data recorded previously in the Motion Capture module to be visualised, segmented and scripted. R-GDL method has been selected because of its capability to classify data from unsegmented (continuous) recording of teachers’ motions that contains multiple actions in order to generate the scripts [23].

R-GDL can generate syntactic descriptions which are easily understood by a computer programmer or an expert in the field. Using heuristic GDL, unwanted behaviour of classifier can be observed because the heuristic GDL may distinguish between key frames that belong to different classes and solve the false positive error without necessity of defining the new feature set and prolonged re-training of whole classifier [24].

In R-GDL, appropriate features for important limbs/joints have to be created. Since SSMC involves simultaneous movements in executing the techniques, the movements of both right and left hands and both right and left legs should be required and considered as important limbs/joints. The hands movement part may consist of wrist, elbow and shoulder joints while the legs movement part consists of hip, knee and ankle joints. For accurate measurement of both parts movements, joint angles function is used to measure the angles of right and left elbows, right and left shoulders, right and left knees, right and left hips, between wrists and between ankles. Therefore, the features for these parts movements are suggested as Figure 3.

The number of cluster in the Clusters Count part will be filled based on the number of important steps in the executed techniques mentioned or advised by the SSMC teacher. This is because, the SSMC teacher is the expert in this field. Based on Hachaj & Ogiea
epsilon parameter is set to 10 in considered cases because all features use angles factor (angles in GDL are calculated in degrees from range [0, 180]). This value is about 10% of overall possible range of angle values.

The quantisation time is also set to \( t_0 = 0.5 \) sec because the data capturing hardware used for data acquisition is a multimedia device that might be highly inaccurate [23]. GDL 1.1 version will be chosen since this version is the latest and enhanced. The new GDL script containing Rules are generated below the written Features in the writing script part in the Main interface. The new GDL script will be saved automatically in the Debug folder as new *.gdl.gdl. This new GDL script will be used as reference or template for classification purpose. The generated rules in the GDL script will possibly be sorted manually if required.

GDLs files contain of initial features definition, epsilon values for all features defined, centres of clusters computed as average value of all elements assigned to them (all features have the same weight), spatial dimensionality of clusters computed as standard deviation value of all elements assigned to them (all features have the same weight) and key frames definition (each key frame corresponds to the single cluster). Single key frame has the following pattern of rule definition [23].

![Interfaces of GDL and R-GDL approaches](image)

- **Motion Classification Model**

Motion Classification model functions to real-time classify the performers’ (trainees’/students’) motion pattern and score based on the GDLs stored in the database. This model utilises GDL approach created by Dr. Tomasz Hachaj. GDL is a classifier capable for real-time human body static poses and body gestures recognition. This approach uses an automated reasoning module to perform forward chaining reasoning for feature extraction.

**2.5.3. Motion evaluation**

Motion Evaluation module produces the extrinsic feedback in forms of motion pattern and scores. Motion pattern recognition can be interpreted using the GDL approach in the motion classification model. Based on Hoon Choi & Jong Joo [14], the score can be produced through the sum of items A, B, C and D: A: Whether or not the movements are in order, B: Accuracy of the motion, C: Speed and balance and D: Harmony (Strength and smooth, Tempo, Rhythm). In this research, the score will be generated based on the traditional local MA approach. In SSCM, there is no show or performance. Therefore, the harmony item is excluded because this item is to evaluate shows or performance.

**3. Data collection and analysis**

Data collection is implemented after the experiments have been done. The data collected will contain 40 recordings of selected techniques using Motion Capture module. The recordings consume SKL file formats which are skeleton data to store the movements of selected techniques in computer. These file formats later will be used in R-GDL and GDL methods respectively for motions template and classification purposes.

In this study, R-GDL approach is used to generate motion templates using SSCM teachers’ movements while GDL method is purposed to classify the trainees’ movements based on the created motion templates. The outputs of R-GDL approach are in forms of GDL scripts containing information about the procedures of executing the selected techniques using Rules-based method. Meanwhile, extrinsic feedbacks in forms of pattern recognition and score are the outputs of GDL.

Analysis on data collected can be done towards both R-GDL and GDL approaches as the followings:

- To validate the effectiveness of the generated GDLs by comparing the results generated for motion templates in forms of GDL files using R-GDL approach with the stored teachers’ motions again. This analysis is important to obtain accurate motions since this data will be referred as important templates for classification and preservation purposes.
- To investigate the feasibility of Cluster numbers suggested by the SSCM teachers toward the captured postures for each selected technique. This analysis explores the capability of the utilised approaches to capture the desired and important postures for each selected technique.
- To analyze the trainees’ skills and experience in executing the selected techniques. This analysis can be done real-time by comparing the trainees’ performance or movements with the GDLs using GDL approach.

**4. Expected results**

Using the proposed framework for MA evaluation system, trainers and trainees can test their knowledge and skills of MAs techniques. This framework is proposed for PSSCM to be a preparation tool to evaluate the trainees’ performance before they come to the real tests which are Basic Level Test and Drop Level Test. Motion pattern and scores generated through the system allow the trainees to identify their levels of mastering the techniques. Such evaluations may encourage the trainees to improve their techniques Data of teacher’s motions collected for motion templates can be stored digitally for reference and inheritance in the future. Therefore, the data is believed promising the authenticity and originality of the MAs techniques and at the same time preserving this kind of valuable cultural heritage.
FEATURE angle[ShoulderRight.xyz[0] - ElbowRight.xyz[0], WristRight.xyz[0] - ElbowRight.xyz[0]] AS RightElbow
FEATURE angle[ShoulderLeft.xyz[0] - ElbowLeft.xyz[0], WristLeft.xyz[0] - ElbowLeft.xyz[0]] AS LeftElbow
FEATURE angle[SpineShoulder.xyz[0] - ShoulderRight.xyz[0], ElbowRight.xyz[0] - ShoulderRight.xyz[0]] AS RightShoulder
FEATURE angle[SpineShoulder.xyz[0] - ShoulderLeft.xyz[0], ElbowLeft.xyz[0] - ShoulderLeft.xyz[0]] AS LeftShoulder
FEATURE angle[ShoulderRight.xyz[0] - ElbowRight.xyz[0], ShoulderLeft.xyz[0] - ElbowLeft.xyz[0]] AS BetweenWrist
FEATURE angle[HipRight.xyz[0] - KneeRight.xyz[0], AnkleRight.xyz[0] - KneeRight.xyz[0]] AS RightKnee
FEATURE angle[HipLeft.xyz[0] - KneeLeft.xyz[0], AnkleLeft.xyz[0] - KneeLeft.xyz[0]] AS LeftKnee
FEATURE angle[SpineBase.xyz[0] - HipRight.xyz[0], KneeRight.xyz[0] - HipRight.xyz[0]] AS RightHip
FEATURE angle[SpineBase.xyz[0] - HipLeft.xyz[0], KneeLeft.xyz[0] - HipLeft.xyz[0]] AS LeftHip
FEATURE angle[HipRight.xyz[0] - KneeLeft.xyz[0], HipLeft.xyz[0] - KneeLeft.xyz[0]] AS BetweenAnkles

5. Conclusion

A theoretical framework of extrinsic feedback using automated evaluation system is presented in this paper to evaluate the effectiveness of performing techniques of a traditional local martial art called SSCM. To ensure the success of this study, preliminary study has been done earlier including acquaintance meeting, procedures and approval and interview. In the framework, due to the practicality and portability factors, MoCap module utilises Microsoft Kinect for Window Version 2 and Kinect SDK 2.0 respectively as a tool and a runtime platform to capture, track and record the executions of selected techniques. Recognition module for motion template in the framework utilises R-GDL approach for visualizing, segmenting and scripting the teachers’ motion data. The data is stored in forms of GDLs files for classification and preservation purposes. Recognition module for motion classification meanwhile applies GDL approach for classifying the trainees’ motion based on the GDLs files stored. The output in this model will be used in evaluation module to produce the extrinsic feedback in forms of motion pattern and scores for the performance executed. Therefore, the framework proposed for the MAs evaluation system can measure the accuracy and effectiveness of performed motions of one of the traditional local MAs and is believed to guarantee the authenticity and originality of the MAs techniques and at the same time to preserve this kind of valuable cultural heritage.

Acknowledgements

The authors gratefully acknowledge the Ministry of Higher Education (MoHE) and Universiti Sultan Zainal Abidin (UniSZA) for the financial support through RAGS Project No. RAGS/1/2015/ICT04/UNISZA/03/1. Thanks to Multimedia University (MMU) for providing facilities and technical assistance for this research. Special thanks to Persatuan Seni Silat Cekak Malaysia (PSSCM) and its teachers for the kind support and cooperation for this research.

References

[17] Cairin Anwar Bin Mohamed (June 20, 2017), Personal Interview, Faculty of Economy & Management Science, UniSZA, Gong Badak Campus.
[23] Tomasz Hachaj, Marek R. Ogieła, Full body movements recognition – unsupervised learning approach with heuristic R-GDL method,