ANALYSIS OF THE PROBLEM OF MACEDONIAN FOLK DANCE RECOGNITION

Darko Martinovikj, Marija Mihova and Mile Jovanov

Ss. Cyril and Methodius University,
Faculty of Computer Science and Engineering
Skopje, R. Macedonia

ABSTRACT

Human experts in dancing and even the enthusiasts can easily conclude if someone is following the sequence of movements of a dance according to the music. But, if the task is given to the computer, it wouldn't be an easy task, because the computer can only understand and work with numbers.

In this paper we analyse the existing solutions for dance recognition and expand the idea toward building a system for dance recognition of traditional Macedonian folk dances. Different approaches for the used hardware, algorithms and methods are described and compared with Macedonian folk dance characteristics.

Keywords: Macedonian folk dance characteristics, Dance Recognition, Kinect sensor, gesture Recognition

I. INTRODUCTION

A dance is a series of motions expressed by the body, which match the speed and rhythm of the played music [9].

Every country has a specific dance style, emerged from the culture and tradition of the country and from a particular ethnic group of people. The type of the performed body movements mostly depends on the type of the dance. For example, in some dances the legs are more dominant and moved, compared to the arms.

Compared to the freedom of movement, there are two types of dances: dances where specific sequence of movements must be followed and where the dancer has the freedom of making the choreography.

As far as participation is concerned, dances can be performed individually, in pairs or as a participation of a group.

A. Characteristics of the traditional Macedonian folk dance

Macedonian folk dances are composed of fast-paced movements where the movement of the legs is dominant compared to the other parts of the body. The dancers are aligned next to each other, holding by their hands. All of the dancers are dancing synchronously, i.e. each dancer performs the same steps in the same time with the other dancers.

The leader of the dancing group determines the trajectory of the dancers' movement, and often stands out, showing great dancing skills in the rhythm with the music.

The basic and most used movements include moving left, moving right, jumping, hopping, dropping to knees, and other similar movements involving the legs. The movement of the arms is minimal and in most cases it doesn't affect how good the dancing is.

The rhythm of the traditional Macedonian folk dances is unique and complex. Common metrics are 7/8 and 2/4. Other metrics are also found 5/8, 9/16, 4/4 etc.

There are also slightly variations in some of the movements of the dances. For example, depending of the gender, males tend to raise the right leg, while the females just put the right leg forward. Also, there can be differences in length of the made steps and speed of the movement execution, depending on the tempo of the music.

B. Dance recognition

The automatic recognition of a dance type is not an easy job. To be able to give some information about the dancing, the computer needs some visual input data for the user's tracked movements.

There are some systems created for this manner. Most of them are built for a specific type of dances, for example tango, polka, rock'n'roll etc. and the motion tracking is done using an appropriate equipment [4,3]. Our goal is to build a system for recognition of some of the traditional Macedonian folk dances.

The usage of this system will be huge: it can be used for learning to dance Macedonian folk dances, to check if someone is following the sequence of movements of a dance, how good he is dancing, creating a choreography etc.

In this paper we analyse the existing solutions for dance recognition and expand the idea toward building a system for dance recognition of a traditional Macedonian folk dance.

II. Analysis of existing systems for dance recognition based on the freedom of movement

Techno and disco dances, that are popular in discotheques, are dances that don't have strict sequence of movements. The dancers have a freedom of making the choreography during the dancing, making movements that are familiar to the type of dance.

So a system for this purpose needs to synchronize with the tempo of the music and classify the recognised movements (gestures).

As explained in [3], a system of this type has been built, to provide additional feedback to the DJs, VJs or other stakeholders for better overall experience. So by checking how well the dancers are dancing to the specific music, predictions about the next played music can be made. This can contribute in better user experience on the dance floors.

Also in [4], a virtual dancer has been introduced which follows the dancer's movements and moves accordingly with the beat of the music.

On the other hand, systems with predefined dance choreography have been built, where the dancer has to follow that strict sequence of pre-choreographed movements in the same moment of time. This type of system is very popular and today has shown great success in gaming industry. Good examples of this type of system are Harmonix' Dance Central [10] and Ubisoft' Just Dance [11], where a dancer repeats the motion posed by an animated character in the game.

Macedonian folk dances have a strict sequence of movements (though there can be slight variations in some movements as we mentioned before) that are played according to the rhythm of the music and all the dancers are dancing synchronously. So, the second type of systems is closer to the problem we want to solve, but we should note that some parts of the first type of system can be used for song analysis, like getting the song tempo.

Classical dances, like tango, waltz etc. follows a strict sequence of movements, and two or more people are dancing synchronously. So this is sufficient indicator that solutions for dance recognition for this type of dance can be used in our case. However, there are also some differences. In classical dances dancers aren't following a fixed trajectory, while in Macedonian folk dances the leader of the dancing group determines the trajectory. Also, other body part movements despite of the legs must be tracked in classical dances, to be able to give some information about the dancing.

III. Analysis of existing systems for dance recognition based on the used hardware

There are number of different approaches for capturing the visual input data of the user's movements.

One approach is to use a standard video-camera of web camera. In [4] it is shown that the visual data is gathered using a video-camera and processed using a computer vision software. From the stream the user's silhouette is extracted, center of mass is determined and the silhouette pixels are distributed in 16 equal-area radial bins for activity detection.

If the user's body orientation is always towards the camera, then this approach can give satisfactory results. But in Macedonian folk dances the trajectory is determined by the leader of the dancing group and the trajectory it's not a straight parallel line always oriented toward the camera, so determining only the pose of the user is not enough.

Another approach for capturing the visual input data is by using sensing floors or sensing shoes. These sensors are used for feet movement. In [1] a system called PogoBoard has been built and for feet detection the users need to have metal plates placed under their shoes. After that a pressure sensor is used for velocity estimation when the shoe touches the board. In [8] an interactive shoes were designed with pressure sensors on the heel and toes for detecting lower body motion.

This approach is appropriate for Macedonian folk dances recognition where the movement of the legs affects mostly the type of dance. But these sensors aren't so cheap to be used for a crowd of people. Also not every person is confortable for wearing a sensor on them.

Another approach that is getting more and more popular today for gesture recognition is the Kinect sensor. Kinect sensor is an input device for motion sensing and speech recognition, developed by Microsoft [12]. What made kinect so popular, compared to the other existing sensors for motion tracking, is the low price, availability to use with traditional computer hardware and existence of developers' tools for kinect application development.

The sensor is like a camera but apart from the colour data stream, using an infrared sensor, it can generate the depth data stream. From this stream the skeleton data can be generated. The tracked skeleton joints of the user's body are shown in Fig. 1. Using this sensor and the skeleton data, a gesture classifier was created in [5] which greatly eases the problem of dance recognition.

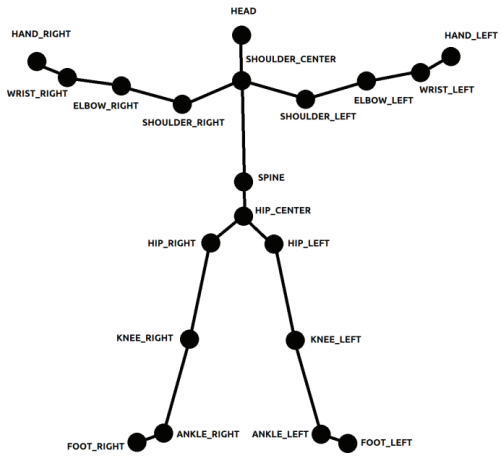

Figure 1: Tracked skeleton joints of the user's body

Kinect is also appropriate for Macedonian folk dances recognition because it can track the ankles and knees positions for improved dance recognition. The advantages for using the sensor is the cheap price, the dancers do not need to wear any sensors on them and the usage with traditional computer hardware. The disadvantages are the limitation of the maximal distance from the sensor which is 4 meters and the maximal number of tracked skeletons which is 2. But more kinect sensors can be added in a room and they can work in parallel. One interesting conclusion about using a kinect sensor for Macedonian folk dance recognition it can handle tracking of multiple dancers in the field of view of the sensor. So, as the dancers are moving in a trajectory, and are arranged one next to each other, when one dancer leaves the field of view, the next one is tracked. Also, only the legs joints needs to be tracked, as the legs affect mostly of the type of dance. These joints are marked in red in Fig. 2.

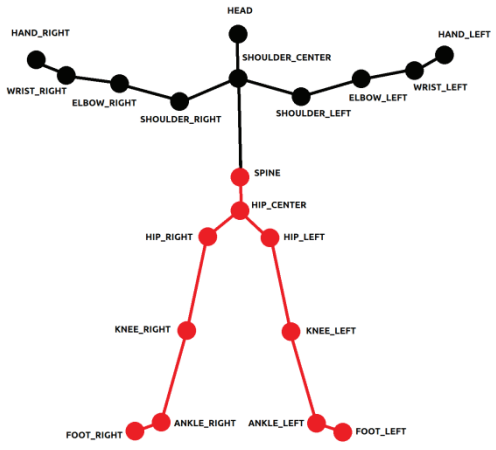

Figure 2: Skeleton joints that should be tracked of the user's body (marked in red)

IV. Analysis of existing systems for dance recognition based on the used methods and algorithms

Many algorithms and methods have been proposed for gesture recognition, starting from the use of Dynamic Time Warping (DTW) [6] to Hidden Markov Models (HMM) [2].

Dynamic Time Warping is an algorithm used for measuring similarity of two data sequences [6]. The two sequences are warped non-linearly in the dimension of time. For time-spatial data, like data of the dancer's movements, this method is used to cope for different speeds and time-deformations between these two data sequences. One data sequence represents the pre-stored gesture sequence of the dance and the other sequence is the executed gesture sequence of the dancer which needs to be evaluated.

The standard DTW method is dynamic programming algorithm which calculates the cost measure as the distance between two elements in the sequences. Then based on the cost, a classification technique must be used.

In [3] a modified DTW, called FastDTW [7] has been used for sequence alignment and a threshold clustering approach for classification. Every dancer's movement sequence is compared to the existing label classes and the class with the nearest DTW cost is chosen. Then using a threshold value for the label class the sequence is mapped to that class or a new class is created.

In [5] a gesture classifier has been created. A kinect sensor is used for getting the skeleton data of the dancer. To reduce the number of data for processing, the data is mapped in a smaller set of features. A spherical coordinate system is constructed and every skeleton joint is represented with two angles. In that way an angular skeleton representation is achieved. For the classification a cascaded correlation-based max-likelihood multivariate classifier has been created which uses the gesture data of a professional dancer and a database of gesture instances.

The DTW method seems usable method for solving the problem for Macedonian folk dance recognition. As the method is distance calculation of two sequences, a good reference sequence must be chosen. Also there can be slight variations in the movements in Macedonian folk dances, so the classification technique must be carefully chosen for less error detection rate.

Hidden Markov Model is statistical model which is used for sequential data. The model has hidden states, i.e. states that are not visible to the observer. Every state can emit one output symbol, which is dependent from the current state of the model. With the evaluation of the sequence of outputs, information about the states can be obtained. Successful usage of HMM are found in speech, handwriting and gesture recognition [2].

The model is based on sample-training. So the parameters of the model are calculated and improved as new output sequences are added to the model.

In [4] a left-to-right HMM for simple gesture detection has been created. Detection of the simple gestures is done by checking the joint's positions of the dancer as the gesture is executed. A dance is represented as a pattern of simple gestures, so for dance recognition another more complex HMM is used. For detecting the most-probable sequence of states the Viterbi algorithm [2] has been used.

The HMM also look like a good model for the Macedonian folk dance recognition. As the sequence of gestures for every Macedonian folk dance is known, the corresponding states in the HMM are also known. The transition and emission probabilities can be estimated by model sample-training.

V. CONCLUSION

In this paper we have analysed the existing solutions for dance recognition and specified how they can be used toward building a system for dance recognition of traditional Macedonian folk dances.

As the legs affect mostly the type of dance, sensors and methods for leg's movement detection should be used. Also, by analysing the dancer's movement data some features can be extracted that are specific for the Macedonian folk dances.

As a future work, we are planning to create a model for dance recognition of Macedonian folk dances based on the analysis of these solutions.

ACKNOWLEDGEMENT

The work has been supported by the FCSE.

REFERENCES

[1] E. Johnstone, "A MIDI Foot Controller - ThePodoBoard," *International Computer Music Conference*, San Francisco, 1991.

[2] G. A. Fink, MarkovModels for Pattern Recognition - From Theory to Applications", 2008

[3] H. Pohl and A. Hadjakos "Dance pattern recognition using dynamic time warping", SMC Conference 2010

[4] I. S. Cerdán, A. Nijholt and D Reidsma, "Dance Movement Patterns Recognition" *CHI '05 Computer Science Final Report, University of Twente, The Netherlands, 2007*

[5] M. Raptis, D. Kirovski and H. Hoppe "Real-time classification of dance gestures from skeleton animation" *SCA '11 Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp. 147–156, New York, 2011

[6] M. Müller, "Information Retrieval for Music and Motion", 2007

[7] S. Salvador and P. Chan "Toward Accurate Dynamic Time Warping in Linear Time and Space" *KDD Workshop on Mining Temporal and Sequential Data, Seattle, 2004*

[8] Y. Kim, D. Jung, S. Park, J. Chi, T. Kim, and S. Lee, "The Shadow Dancer: A New Dance Interface with Interactive Shoes," *International Conference on Cyberworlds*, IEEE Computer Society, 2008.

[9] Dance Wiki Page [Online] Available: http://en.wikipedia.org/wiki/Dance

[10] Dance Central Page [Online] Available: http:// www.dancecentral.com/

[11] Just Dance Page [Online] Available: http://just-dance-thegame.ubi.com/jd-portal/en-us/home/index.aspx

[12] Kinect [Online] Available: http://www.microsoft.com/en-us/kinectforwindows/