Folk Dance Evaluation Using Laban Movement Analysis

ANDREAS ARISTIDOU, EFSTATHIOS STAVRAKIS, PANAYIOTIS CHARALAMBOUS, and YIORGOS CHRYSANTHOU, University of Cyprus
STEPHANIA LOIZIDOU HIMONA, Frederick University

Motion capture (mocap) technology is an efficient method for digitizing art performances, and is becoming increasingly popular in the preservation and dissemination of dance performances. Although technically the captured data can be of very high quality, dancing allows stylistic variations and improvisations that cannot be easily identified. The majority of motion analysis algorithms are based on ad-hoc quantitative metrics, thus do not usually provide insights on style qualities of a performance. In this work, we present a framework based on the principles of Laban Movement Analysis (LMA) that aims to identify style qualities in dance motions. The proposed algorithm uses a feature space that aims to capture the four LMA components (Body, Effort, Shape, Space), and can be subsequently used for motion comparison and evaluation. We have designed and implemented a prototype virtual reality simulator for teaching folk dances in which users can preview dance segments performed by a 3D avatar and repeat them. The user's movements are captured and compared to the folk dance template motions; then, intuitive feedback is provided to the user based on the LMA components. The results demonstrate the effectiveness of our system, opening new horizons for automatic motion and dance evaluation processes.

Categories and Subject Descriptors: I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism—*Animation*

General Terms: Folk Dance Evaluation Using LMA

Additional Key Words and Phrases: Folk dances, Laban Movement Analysis, motion capture, motion comparison, motion evaluation

ACM Reference Format:

Andreas Aristidou, Efstathios Stavrakis, Panayiotis Charalambous, Yiorgos Chrysanthou, and Stephania Loizidou Himona. 2015. Folk dance evaluation using laban movement analysis. ACM J. Comput. Cult. Herit. 8, 4, Article 20 (August 2015), 19 pages.
DOI: http://dx.doi.org/10.1145/2755566

1. INTRODUCTION

Intangible Cultural Heritage (ICH) is an integral part of the cultural identity of any society. ICH encompasses collective knowledge of communities, skills, practices, expressions, and representations that do not have a tangible form. In this work, we focus on devising state-of-the-art methods for digitization,

This work is co-financed by the European Regional Development Fund and the Republic of Cyprus through the Research Promotion Foundation under contract DIDAKTOR/0311/73.
Authors' addresses: A. Aristidou, E. Stavrakis, and Y. Chrysanthou, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus; emails: a.aristidou@ieee.org, {stathis, yiorgos}@cs.ucy.ac.cy; S. L. Himona, Frederick University, 7, Y. Frederickou Street, 1036 Nicosia, Cyprus; email: com.ls@frederick.ac.cy.
Author's current address: P. Charalambous, INRIA Rennes - Bretagne Atlantique, Campus de Beaulieu, 263 Avenue Général Leclerc, 35042 Rennes, France; email: panayiotis.charalambous@inria.fr.

DOI: http://dx.doi.org/10.1145/2755566

analysis, and dissemination of folk dances. We demonstrate our methods using Cypriot folk dancing as a representative use case.

Folk dances are learned informally and passed on from one generation to the next. The main difference between choreographed dances and folk dances is that the latter are often improvisations by nonprofessionals that take place in social events and other daily life activities. Folk dancing is a rather “malleable” form of ICH, as it is modified and adapted over time and across different geographic locations. Although each folk dance has a basic set of steps and postures that dominate, folk dancers will typically modify and often enrich the dance with their personal style. The implication of these stylistic mutations is that there is no single ground truth for a folk dance.

There are various methods one can use to learn dancing. For example, one may use a self-learning approach, for example, by utilizing books and videos or attending a class and learning from an experienced instructor [Kassing and Jay 2003]. Brain research based on dance learning experiments provided evidence that learning to dance is facilitated through both physical and observational learning [Cross et al. 2009]. Irrespective of the learning method, dance students usually learn the choreographic aspects of the dance faster, for example, the basic steps and postures, but take longer to master the dynamics of movement (e.g., flow, weight, and the like)

Motion capture technology has enabled the documentation and preservation of ICH artifacts such as folk dances. However, digitization alone is not sufficient to pass folk dancing to the newer generations. Therefore, interactive virtual reality 3D applications, for example, games [Tang et al. 2011] and dance learning platforms [Magnenat-Thalmann et al. 2008], have emerged as teaching aids for users wishing to learn how to perform these dances. Dance teaching applications usually feature a virtual 3D teacher who first performs a prerecorded expertly executed dance or segment of a dance. The user will then perform this motion physically while being monitored by a motion capture system attached to the application. The motion is then analyzed and compared to the teacher’s motion, and the user is provided with feedback.

One of the main aspects of motion analysis is the understanding of different types of human movements, such as basic human actions (e.g., walking, running, or jumping) and stylistic variations (e.g., emotion, intention, expression, or gender). Stylistic variations, though, are difficult to measure; the movement of the human body is complex and hard to completely describe. An important role in the description and categorization of a dance performance is that played by the intensity and fluidity of each movement, reflecting its *nuance*. The *nuance*, along with the shape, concentration, and energy needed to carry out the action, can provide additional information with regard to the style of the performance. General purpose motion evaluation algorithms are limited in their capacity to acquire the stylistic elements of dance performances (e.g., the emotion, expression, and interaction between the performer and the environment); however, choreographers and movement analysts take into consideration movement characteristics that show the style of the dance, which play an important role in the evaluation of movements. Based on the principles of movement observation science, specifically using Laban Movement Analysis (LMA) [Laban 2011; Maletić 1987] components, we aim to extract the so-called nuance of motion and use it for motion comparison and evaluation purposes. LMA is a multidisciplinary system, incorporating contributions from anatomy, kinesiology, and psychology that draws on Rudolph Laban’s theories to describe, interpret, and document human movements. It is one of the most widely used systems of human movement analysis, and has been used extensively to describe and document dance and choreography over the last century.

In contrast to previous approaches that compare and evaluate dances, our technique uses LMA to qualitatively assess the similarity of two dancing motions. It determines characteristics that a student would find useful for the improvement of one’s skills. For example, we do not report the angular offset of a student’s limbs in comparison to the teacher’s. Instead, our system generates higher-level hints, such as a percentage of correctness in the flow and intensity of the motion inferred from a large set of low-level motion features. This approach of intuitively exposing the quality aspects of the student's motion makes it easier for that student to focus on improving a particular aspect of one's performing skills, for example, overall posture or speed, rather than a specific body part.

2. RELATED WORK

Motion matching or comparing algorithms typically use discrete motion samples that represent body postures to compute an aggregate distance metric between the two postures. In literature, the majority of methods can be grouped into those using (i) the distances between the positions of body joints, (ii) the angular differences between respective joint pairs, and (iii) the velocities of respective joints, or a combination of these methods. Various techniques in the area of indexing, classification, and synthesis search for logical similarities between motions; for instance, Motion Graphs [Kovar and Gleicher 2004] is a data structure widely used to compare motion clips (i.e., using distance metrics between postures) and represent transitions between them for motion synthesis. A variety of different metrics that capture and compare the geometric properties of motion were introduced by Müller et al. [2005] to establish a content-based retrieval method for motion similarity purposes; different techniques have also been proposed for spatial indexing of motion data [Keogh et al. 2004; Krüger et al. 2010]. Moreover, Deng et al. [2009] and Wu et al. [2009] cluster motion on hierarchically structured body segments, whereas Chao et al. [2012] use a set of orthonormal spherical harmonic functions.

In order to achieve a satisfying simulation for complex human body language, an as simple as possible but as complex as necessary description of human motion is required; LMA satisfies these demands. The principles of LMA have been used in computer animation for over a decade. The EMOTE system, introduced by Chi et al. [2000], synthesizes gestures using the LMA effort component for motion parameterization and expression; Zhao and Badler [2005] used the EMOTE results to design a neural network for gesture animation. Hartmann et al. [2006] quantify the expressive content of gesture based on a review of the psychology literature, whereas Torresani et al. [2006] used LMA for learning motion styles. Chen et al. [2011] emphasized Laban's effort quality for movement analysis and evaluation to construct an e-learning system, whereas Wakayama et al. [2010] and Okajima et al. [2012] used a subset of LMA features for motion retrieval. Kapadia et al. [2013] proposed a variety of features based on Laban principles to encode structural, geometric, and dynamic characteristics of motion as keys; these keys are later combined to define queries for motion retrieval. Santos and Dias [2010] presented a tool that uses Laban theories to describe human basic behavior patterns. Masuda et al. [2009] also expressed the bodily emotion of a human-form robot using a small set of Laban's features. Later, the same authors added four basic emotions to arbitrary movements [Masuda et al. 2010]. Recently, Zacharatos et al. [2013] used a set of body motion features based on the LMA effort component, to provide sets of classifiers for emotion recognition in a game scenario. A set of 3D gesture descriptors, based on an LMA model, have been utilized by Truong et al. [2015] as well to recognize the gesture and emotional content of orchestra conductors using a machine-learning framework.

The wide range of existing techniques for general-purpose motion analysis, segmentation, classification, and retrieval may also be applied to motion-captured dances. However, the scientific community has recently focused on explicitly devising methods to cater to dance-oriented applications, such as dance teaching and dancing games, as well as extraction of choreography, dance annotation, comparison, and so forth. When evaluating dancing motions for educational purposes, the teacher's and the student's motions can be qualitatively similar, although they may technically differ. Magnenat-Thalmann et al. [2008] designed a learning framework for folk dances based on motion capture. They treated the concept of dance holistically without discriminating between movement and context. Within the context of this framework, they developed a Web-based 3D environment in which users can visualize folk dances. Alexiadis and Daras [2014] have recently designed a framework for automatic dance performance evaluation that employs motion capture data using marker-less motion capture. The authors represented the human motion data as sequences of pure quaternions and subsequently introduced a set of quaternionic vector-signal processing methodologies for dance motion evaluation and comparison purposes. Tang et al. [2011] implemented a real-time dancing game using a Progressive-Block Matching algorithm. In addition, Chan et al. [2011] presented a similar system, but focused on performing a comprehensive motion analysis of the player's body parts with respect to the taught-motion template. Deng et al. [2011] developed a real-time motion recognition algorithm that is based on a human body partition indexing scheme with flexible matching to determine the end of a move as well as to detect unwanted motion. This work has been furthered by Yang et al. [2013] to provide tools for automatically generating dance lessons that adapt to the skill of the student dancer. Laban theory has also been used for synthesizing dance motion matched to music [Shiratori et al. 2006]. Aristidou and Chrysanthou [2013] used a variety of LMA features that encode characteristics of motion to understand the performer emotions from acted dance performances. Aristidou and Chrysanthou [2014] have provided a brief analysis of how these features change on movements with different emotional states, finding movement similarities between different emotional states. Recently, Aristidou et al. [2014] presented an LMA-based query-by-example motion retrieval method from a folk dance database.

3. MOTION ANALYSIS

In this work, we have developed a novel motion comparison algorithm, which compares the movements of two characters by taking into consideration not only posture matching (meaning the physical geometry of the avatar) but also style. The proposed evaluation extracts the quality characteristics of a dance performance based on LMA. LMA offers a documentation of human motion, divided into four main categories: BODY, EFFORT, SHAPE, and SPACE. In this section, we present a subset of the LMA components and representative features that are indicative to capture the motion properties, and can be used for motion comparison purposes. The proposed LMA features are calculated to be used for motion comparison and evaluation purposes; the key joints used for the description of the proposed LMA features are indicated in Figure 1(b).

3.1 BODY Component

The BODY component primarily develops body and body/space connections. It describes the structural and physical characteristics of the human body and is responsible for describing which body parts are moving, which parts are connected, which parts are influenced by others, what is the sequence of the movement between the body parts, and general statements about body organization. We propose the following features to define the BODY component and address the orchestration of the body parts:

—*Displacement and Orientations*: Different displacements, such as feet to hips distance (f_1), hands to shoulders distance (f_2), right hand to left hand distance (f_3), hands to head distance (f_4), and hands to hips distance (f_5), are used to capture the body connectivity and the relation between body parts of the performer.

—*Pelvis height* (f_6): The distance of the root joint from the ground, in our skeleton the pelvis; this feature is particularly useful for specifying whether the performer kneels, jumps in the air, or falls to the ground.

—*Legs and Body relation to ground* (f_7): This is the distance of the hips to the ground minus the distance of the feet to the hips, which provides a metric for relation of the body's posture and the extention of the legs from the body.

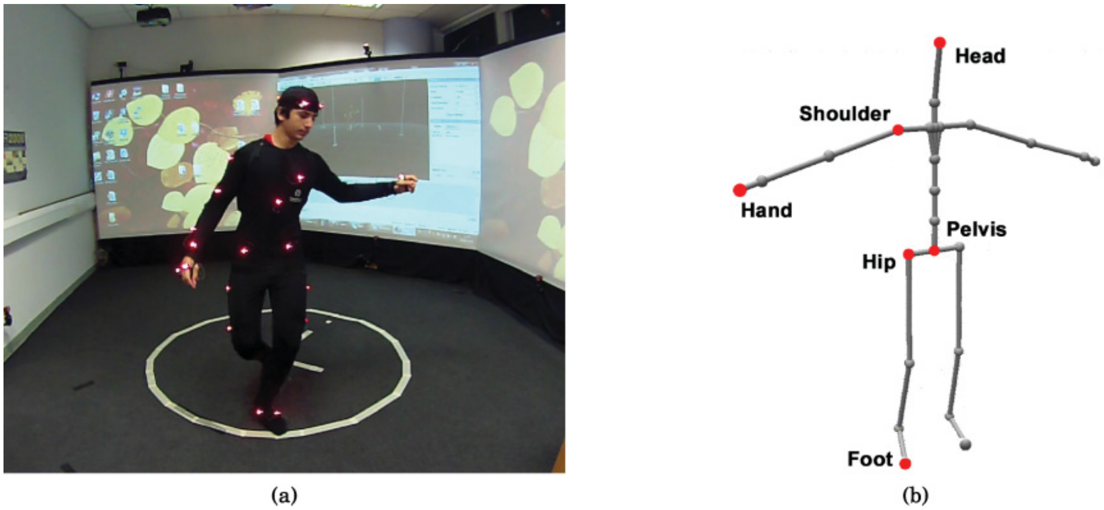


Fig. 1. (a) A dance performer wearing the mocap suit and performing Zeibekiko at our laboratory. (b) Representation of the articulated skeletal structure used to calculate the LMA features. Key joints used in the calculations are clearly indicated.

—*Gait size* (f_8): Gait size is the distance of the right foot to the left. The size of a human gait may be indicative of motion expression, emotion, style, and so forth.

Note that all distances involving joints both from the left and right side of the human skeleton ($f_1 - f_5$, f_7) are calculated symmetrically and averaged. For example, (f_1) feet to hips distance is the average Euclidean distance of the left foot to the left hip and the right foot to the right hip, that is, $f_1 = (\mathrm{d}(L_{foot}, L_{hip}) + \mathrm{d}(R_{foot}, R_{hip}))/2$. Calculating the average of these distances weights equally the two sides of the body and provides a balanced set of metrics to describe the geometry of the performer's body. It would also be possible to introduce features independently for each side of the body, especially for subvolume motion comparison and motion synthesis approaches.

3.2 EFFORT Component

The EFFORT component describes the intention and the dynamic quality of the movement, texture, feeling tone, and how the energy is being used on each motion. It comprises four subcategories—each having two polarities—named EFFORT *factors*:

—*Space* addresses the quality of active attention to the surroundings. It has two polarities: Direct (focused and specific) and Indirect (multi-focused and flexible attention).
—*Weight* is a sensing factor, sensing the physical mass and its relationship with gravity. It is related to movement impact and has two dimensions: Strong (bold, forceful) and Light (delicate, sensitive).
—*Time* is the inner attitude of the body towards the time, not the duration, of the movement. Time polarities are Sudden (has a sense of urgent, staccato, unexpected, isolated) and Sustained (has a quality of stretching the time, legato, leisurely).
—*Flow* is the continuity of the movement; it is related to feelings and progression. The Flow dimensions are Bound (controlled, careful, and restrained movement) and Free (released, outpouring, and fluid movement).

EFFORT changes are generally related to the changes of mood or emotion, and are essential for expressivity. The EFFORT *factors* can be derived as follows:

—*Head orientation* (f_9): The *Space* factor can be derived by studying the attitude and orientation of the body in relation to the direction of the motion. If the character is moving in the same direction as the head orientation, then the movement is classified as Direct, whereas if the orientation of the head does not coincide with the direction of the motion, then this movement is classified as Indirect. To this end, we calculate the angle between the head's orientation and the body path of the performer, which is expressed by the trajectory of the root joint.

—*Deceleration of motion* (f_{10}): The *Weight* factor can be identified by studying how the deceleration of motion varies over time; f_{10} is estimated by calculating the deceleration of the root joint. Peaks in decelerations means a movement with Strong Weight, while no peaks refers to a movement with Light Weight; note that *Weight* is velocity independent.

—*Movement velocity*: The velocity of the performer's movement is indicative of the *Time* factor. It is estimated by calculating the distance covered by the root joint over a time period (f_{11}). In addition, the average velocity of both hands (f_{12}) and both feet (f_{13}) is calculated to distinguish dance movements for which the performer remains at the same position, while the choreography is mainly expressed by changes in body postures.

—*Movement acceleration* ($f_{14} - f_{16}$): The acceleration is another feature for determining the *Time* factor; it is computed by taking the derivative of the aforementioned movement velocities with respect to time; with f_{14} the hips' acceleration, f_{15} the hands' acceleration, and f_{16} the feet's acceleration.

—*Jerk* (f_{17}): A way to extract the *Flow* of each movement is jerk. Jerk is the rate of changes of acceleration or force, calculated by taking the derivative of the acceleration (f_{14}) with respect to time. Bound motion has large discontinuities with high jerk, whereas Free motion has little change in acceleration.

3.3 Shape Component

Shape analyzes the way the body changes shape during movement. It describes the static shapes that the body takes, the relation of the body to itself, the way the body is changing toward some point in space, and the way the torso can change in shape to support movements in the rest of the body. Shape can be captured using the following features:

—*Volume*: The volume of the performer's skeleton (f_{18}) is given by calculating the bounding volume of the five end-effector joints (i.e., head, hands, and feet). In addition, the volume of the skeleton is calculated as the bounding volume of all joints (f_{19}), which enables one to distinguish cases in which hands and/or legs are very close to each other, but the performer's overall volume is still large. In addition, the bodily volume of the performer is subdivided into 4 subvolumes: upper body (f_{20}), lower body (f_{21}), left side (f_{22}), and right side (f_{23}).

—*Torso height* (f_{24}): The distance between the head and root joints indicates whether the performer is crouching, meaning bending the torso; it does not take into account whether the legs are bent, but only whether the torso is kept straight or not.

—*Hands level* (f_{25}): The relation of the hands' position with regards to the body, indicating whether they are moving on the upper level of the body (over the head), the middle level (between the head and the chest), or the low level (below the chest). The hands' orbit level is calculated even if the performer is crouching, kneeling, or jumping.

3.4 Space Component

Space describes movement in relation with the environment, pathways, and lines of spatial tension. Laban classified the principles for movement orientation based on the *body kinesphere* (the space within reach of the body, the mover's own personal sphere) and *body dynamosphere* (the space where the body's actions take place, the general space that is an important part of personal style). SPACE component can be derived using two different features:

—*Distance* (f_{26}): The distance covered over time, which is measured as the length of the projection of the root joint's trajectory to the ground. This prevents vertical translation, for example, jumping, from mistakenly considered as space coverage by the performer.

—*Area* (f_{27}): The total area covered over a time period, which is calculated as the area of the polygon formed by the projection of the root joint to the ground.

Combining f_{26} and f_{27} provides an indication of the relationship of the performer's feelings with regard to the environment. In addition, this combination is a measure of the utilization of the allowable space a performer achieves with one's movements.

4. MOTION COMPARISON

The proposed LMA features can be used to extract information regarding the dance performance, taking into consideration both body variations and style of the performance. In that manner, we are able to evaluate a dance performance and find potential similarities with another, even if the performers' posture geometries have significant differences. In order to assess two performances and find their potential similarities, we have implemented a motion comparison framework.

To extract the proposed LMA features and measure the observations, each motion clip frame is filtered with a 35-frame moving window (clips have 30fps), anchored at the center. We use a window stepping of 1, but this can be increased to speed up computation at the expense of accuracy. In this work, we assume that the clips are already synchronized. A variety of feature measurements were calculated for each of the f_is comprising the features of the LMA components, such as the maximum, minimum, mean and standard deviation, resulting in 87 different feature measurements (ϕ_is). These feature measurements are summarized in Table I.

For each window of a motion clip, a correlation matrix is computed to the respective window of the other clip, which provides an association between the time windows of the two motions. The correlation matrix measures the absolute values of Pearson's linear correlation coefficient [Pearson 1920], (0, no correlation; 1, high correlation). To evaluate the correlation between two performances, each of the four LMA components has been assessed separately for each window, returning a Pearson's linear correlation coefficient for each LMA component; the overall evaluation for a window is a weighted sum of all its LMA components. The weights are user-defined and provide a mechanism to control the importance of each LMA component when comparing motions. For example, the weights used for each of the four LMA components can be set to 25% to weight them equally. The overall correlations computed in each window are then filtered to reduce noise with a 1D Gaussian function with mean $\mu = 0$ and variance $\sigma^2 = 1$. These correlations provide an estimate of the relevance between the windows of the two performances based on the LMA components. Two windows (or frames, as in our case) are considered similar if their overall Pearson's linear correlation coefficient is larger than a user-specified threshold, which we refer to as the *decision threshold*, and is usually set to values higher than 75%.

4.1 Isomap Representation of LMA features

As we mentioned in Section 3, the set of proposed features should be able to capture the motion properties so that meaningful comparisons can be performed. Since our feature space is high dimensional (87 features), it is difficult to visualize the data and understand their properties. Typically, methods

Table I. The Measurements Used in our Implementation

	f_s	Measurement					
		Description	max	min	mean	std	#
Body	f_1	Feet–hip distance	ϕ_1	ϕ_2	ϕ_3	ϕ_4	
	f_2	Hands–shoulder distance	ϕ_5	ϕ_6	ϕ_7	ϕ_8	
	f_3	Hands distance	ϕ_9	ϕ_{10}	ϕ_{11}	ϕ_{12}	
	f_4	Hands–head distance	ϕ_{13}	ϕ_{14}	ϕ_{15}	ϕ_{16}	
	f_5	Hands–hip distance	ϕ_{17}	ϕ_{18}	ϕ_{19}	ϕ_{20}	
	f_6	Hip–ground distance	ϕ_{21}	ϕ_{21}	ϕ_{23}	ϕ_{24}	
	f_7	Hip–ground minus feet–hip distance	ϕ_{25}	ϕ_{26}	ϕ_{27}	ϕ_{28}	
	f_8	Gait size	ϕ_{29}	ϕ_{30}	ϕ_{31}	ϕ_{32}	
Effort	f_9	Head orientation	ϕ_{33}	ϕ_{34}	ϕ_{35}	ϕ_{36}	
	f_{10}	Deceleration peaks					ϕ_{37}
	f_{11}	Hip velocity	ϕ_{38}	ϕ_{39}		ϕ_{40}	
	f_{12}	Hands velocity	ϕ_{41}	ϕ_{42}		ϕ_{43}	
	f_{13}	Feet velocity	ϕ_{44}	ϕ_{45}		ϕ_{46}	
	f_{14}	Hip acceleration	ϕ_{47}			ϕ_{48}	
	f_{15}	Hands acceleration	ϕ_{49}			ϕ_{50}	
	f_{16}	Feet acceleration	ϕ_{51}			ϕ_{52}	
	f_{17}	Jerk	ϕ_{53}			ϕ_{54}	
Shape	f_{18}	Volume (5 joints)	ϕ_{55}	ϕ_{56}	ϕ_{57}	ϕ_{58}	
	f_{19}	Volume (All joints)	ϕ_{59}	ϕ_{60}	ϕ_{61}	ϕ_{62}	
	f_{20}	Volume (upper body)	ϕ_{63}	ϕ_{64}	ϕ_{65}	ϕ_{66}	
	f_{21}	Volume (lower body)	ϕ_{67}	ϕ_{68}	ϕ_{69}	ϕ_{70}	
	f_{22}	Volume (left side)	ϕ_{71}	ϕ_{72}	ϕ_{73}	ϕ_{74}	
	f_{23}	Volume (right side)	ϕ_{75}	ϕ_{76}	ϕ_{77}	ϕ_{78}	
	f_{24}	Torso height	ϕ_{79}	ϕ_{80}	ϕ_{81}	ϕ_{82}	
	f_{25}	Hands level					ϕ_{83} - ϕ_{85}
Space	f_{26}	Total distance					ϕ_{86}
	f_{27}	Total area					ϕ_{87}

that project the data into lower-dimensional spaces (e.g., 2 or 3 dimensions) aim at providing visual representations of the data by preserving interesting structures to aid interpretation.

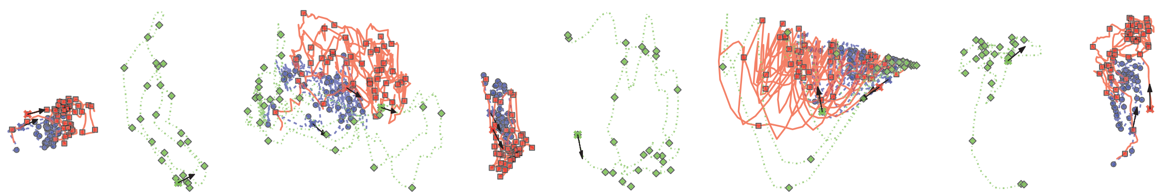
These methods fall into two main categories: linear and nonlinear. One of the most common linear projection methods is Principal Component Analysis (PCA) that projects the data into a new space that is a linear combination of the original features. These methods, however, fail to capture important nonlinear structure in the data. Nonlinear dimensionality reduction approaches assume that the examined data lie on an embedded nonlinear manifold within the high-dimensional space (i.e., the data are artificially high dimensional). In essence, these methods try to find the "intrinsic" variables that really were the cause of the original data. One of the most common approaches in this category is the isomap technique [Tenenbaum et al. 2000] that aims to project the data based on the geodesic distances[1] between neighboring data points.

Figures 2(a) through 2(d) show the projections of two dance performances for different subsets of the LMA features. These performances are of the same dance, which was intentionally performed at

[1]The geodesic distance between two vertices of a graph is the number of edges belonging to the shortest path between them, assuming the graph is connected.

(a) BODY (b) EFFORT (c) SHAPE (d) SPACE (e) All LMA

Fig. 2. Isomap projection of LMA features of the same dance performed at different intensities. Each connected line represents a performance. *X* marks the first set of projected features and the lines indicate temporal relationships between points (i.e., the trajectory of the LMA features). Points are a subset of the data points (every 0.5s, 15 frames). It can be observed that although some of the LMA components have asimilar structure and can be separated (a, b), the overall combination of LMA features provides a more apparent separation (e).



(a) BODY (b) EFFORT (c) SHAPE (d) SPACE (e) All LMA

Fig. 3. Isomap projection of LMA features of three dances. Red and blue represent the same dance performed at different intensities. Green is a Zeibekiko dance, which is apparently very different from the other two. It can be observed in (e) that the isomap projection of all LMA features combined cluster the two dances (red and blue), while separating the third (green).

different intensities. The projection of all LMA components combined is shown in Figure 2(e). The struacture of this combined isomap suggests that the underlying motions share common characteristics that are captured by the chosen LMA feature space. To further verify this, we show the isomap projection of 3 datasets in Figure 3; the two datasets (red and blue) are the same as those of Figure 2, while the third (green) is a rather different dance (Zeibekiko). Again, the isomap projection of some of the LMA components group and separate the projected features, but the combined LMA components seem to group the two first dances together, while setting apart the third.

These graphs indicate that the proposed LMA features capture similarities and dissimilarities between dance performances, and can be used as visual aids (e.g., could be adopted in a game's user interface to provide stylized feedback to players). However, a more complete analysis, such as classification accuracy, could be employed in future work.

5. LMA-BASED DANCE LEARNING PLATFORM

Dancing is largely taught by example, with a teacher performing the movements and the student repeating. Self-learning of dances has been mainly based on educational video material and, more recently, video games. In line with other computer-based dance teaching systems in this section, we present a prototype self-learning dance platform based on our LMA algorithmic framework. The platform takes advantage of high-quality 3D motion capture data of folk dances and uses the motion analysis algorithm presented in Section 3 to provide a set of quality parameters that can be tuned to assess similarity between motions. Furthermore, using the motion comparison algorithm, the platform directly leverages the intuitiveness of the LMA framework to provide user-friendly feedback and

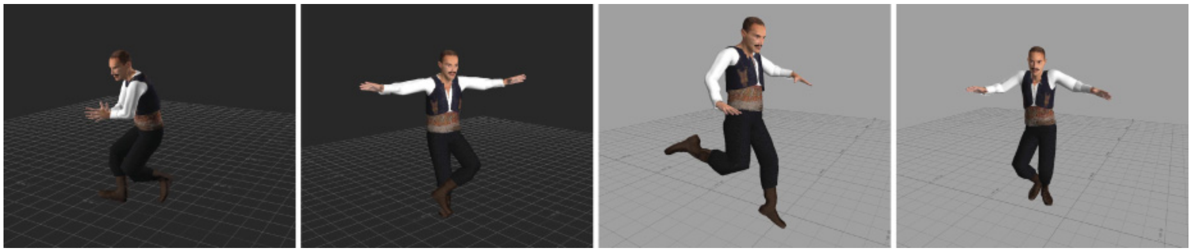


Fig. 4. Sample frames from motion-captured folk dances contributed to the Dance Motion Capture Database. From left to right, it shows Zeibekiko, 1st, 2nd, and 3rd Antikristos, respectively.

parameter control. Note that the dance simulator does not intend to replace traditional dance tuition, but to provide an additional tool for training and education in dance, both at home and at school, using an interactive environment.

5.1 Mocap Folk Dance Data

In parallel to the technical contributions in this work, a considerable effort has been invested in digitizing Cypriot folk dances, as well as acted modern dance performances. The data have been captured using a PhaseSpace's Impulse X2 motion capture system [PhaseSpace 2014], which allows for high-frequency optical tracking of the dance performers (up to $960Hz$). However, the quality of the data is not only due to the technical equipment used. The performers were experienced dancers, the majority of which were active members of cultural organizations and dance schools. Therefore, the motion-captured folk dances document an integral part of Cypriot intangible cultural heritage, which were up to now only documented via text, photographs, and video. These quality and culturally important datasets have been submitted for the enrichment of the Dance Motion Capture Database[2] [University of Cyprus 2014], which has been initialized by Stavrakis et al. [2012], and can be viewed online using the Unity3D Web plug-in in real time. Figure 4 shows snapshots from the folk dances we contributed to the database.

Our datasets are comprised of Biovision Hierarchical (BVH) data files of dance performances that are captured at $480Hz$. The BVH format consists of two parts: the first section details the hierarchy and initial pose of the skeleton, and the second section describes the channel data for each frame, thus the motion section. It is important to note that the BVH skeletons in our dataset are normalized, thus skeleton and joint distances, such as arm span and other displacements, are calculated under the same conditions. Our analysis (Section 3) and comparison (Section 4) algorithms sample these datasets at 30fps, which provides sufficiently good results, but higher frame rates, if required, may unnecessarily increase computation times.

5.2 Dance Learning Platform

The prototype learning platform is built around the concept of students observing a virtual 3D teacher performing dance movements and repeating them. It uses quality motion-captured folk dance data from the database, as described earlier. Motion data represent complex dance choreography, thus can be difficult for beginners to perform all at once. Instead, the motion-captured data are segmented into dance motion primitives, that is, short sequences of distinct movements that usually last between 400 and 900 frames. These motion primitives act as template motions, and can be reassembled into the complete dance.

[2]Dance Motion Capture Database: http://dancedb.cs.ucy.ac.cy/.

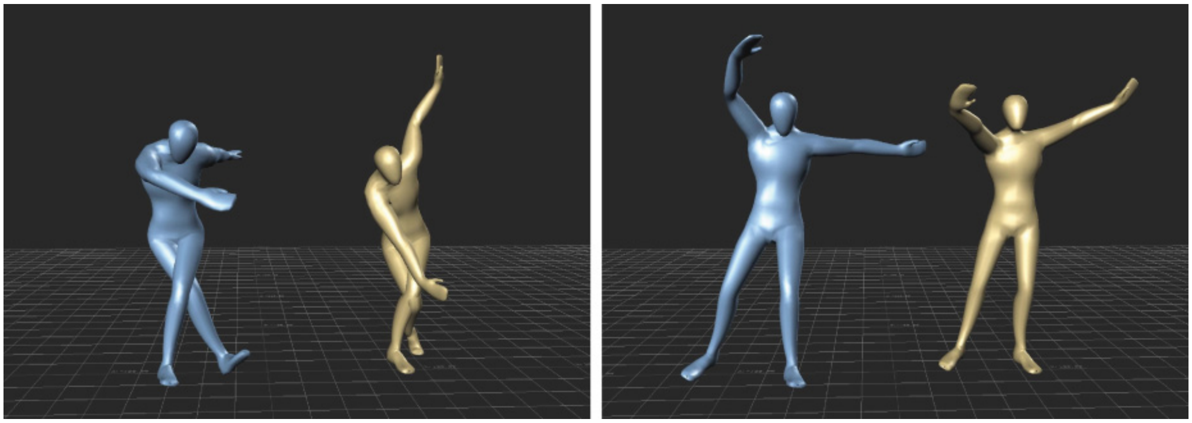


Fig. 5. Snapshots from our experimental data, in which the student (yellow) imitates the teacher's (blue) movements.

During a dance learning session, the user selects the desired dance to learn and a 3D avatar (teacher) selects arbitrary dance motion primitives from the template motions and demonstrates it to the user (student). The user then physically performs the motion, which is captured and passed to the motion analysis subsystem via a full-body motion-capture system. The user's motion is analyzed and compared to the template motion, and an evaluation of the user's performance is generated.

In contrast to other dance learning systems, the user is not explicitly provided with feedback on body parts that have been moving incorrectly. We believe that this type of feedback, although quite helpful, can be daunting to beginners. For example, beginners usually find it easier to learn the body posture (BODY) and steps (SPACE) of a dance, but may find it very difficult to reproduce the flow (EFFORT) and shape qualities (SHAPE) of a dance. Instead, the platform generates an evaluation based on the LMA categories (BODY, EFFORT, SHAPE, SPACE), which exactly point the student to the particular quality characteristic of the performance that needs improvement. This way, our system can be considered as more forgiving toward mistakes that could demoralize the student and play little educational role for that student's skill level, such as an incorrectly bend arm or a slightly misplaced foot.

Furthermore, the learning platform allows the user to modify the sensitivity of the system when comparing the motion of the student to the template motions per LMA category. The four LMA categories are initially equally weighted (25% each). Users can manually adjust the weights to tilt the sensitivity toward one of the LMA components of the dance that they would like to improve on. For instance, users that are comfortable with their body posture may reduce the decision threshold for the BODY and/or increase the threshold of the EFFORT to make the system more sensitive to mistakes in the fluidity of their motion. In addition, the system can be set to adaptively modify the difficulty of achieving a close match of the template motion. This follows the same principles of dynamic difficulty adjustment (DDA) in computer games, with an outlook of focusing the user's attention to aspects of the motion that the user needs to improve on.

6. EXPERIMENTAL RESULTS

This section presents the experimental results of the proposed system. Students were asked to imitate short parts of precaptured dance performances (performed by professional dancers), while their performance was evaluated against the teacher's performance using the proposed LMA-based motion comparison approach. Figure 5 shows two snapshots from our video clips; the teacher (in blue) performs a dance choreography, while the student (in yellow) tries to follow it.

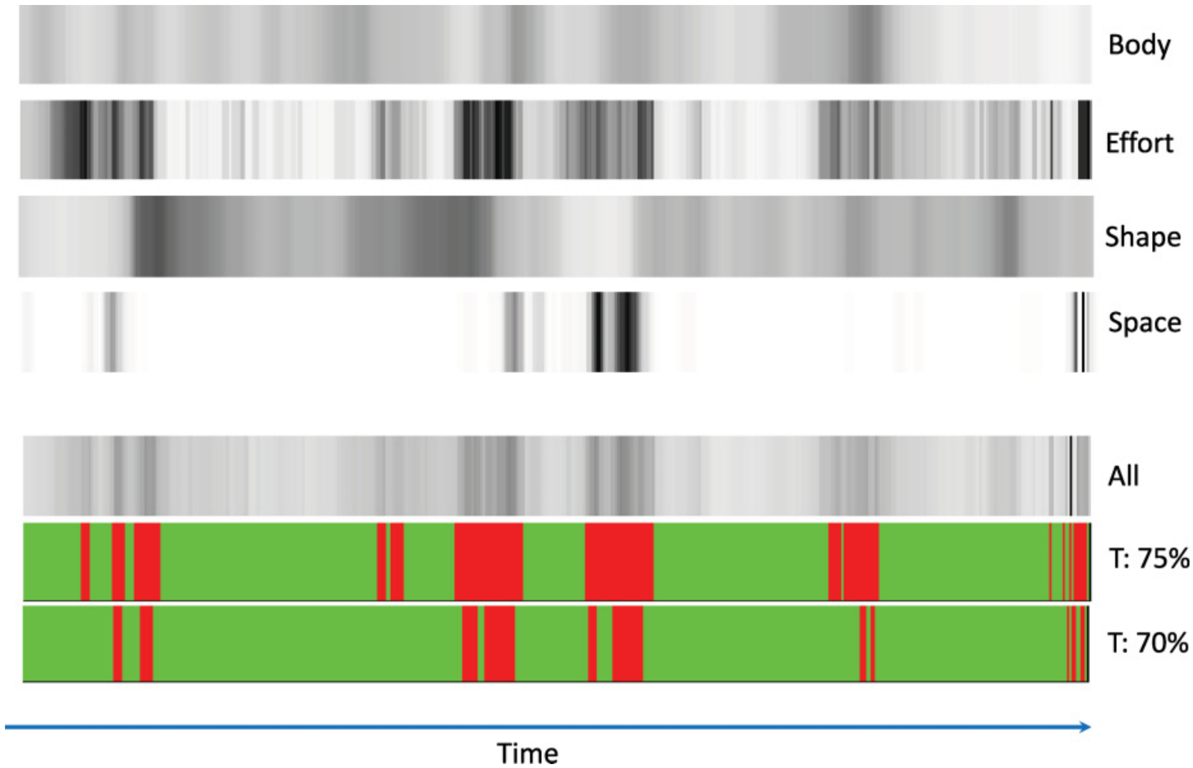


Fig. 6. The correlation between the movements of the teacher and student; the first four bars show the correlation for each LMA component separately, while the next shows the overall correlation taking into consideration all the LMA components. The correlation is presented in gray scale: white indicates a high correlation and black indicates no correlation between the motions. The last two bars show the decision whether the movements under analysis are similar or not, that is, when their correlation is higher than the decision threshold, which is set at 75% and 70%, respectively. Green indicates "pass," while red indicates "fail."

Figure 6 shows the correlation between a student and teacher performance for each LMA component separately (in gray scale, white means high correlation and black means small correlation), as well as the overall correlation when all LMA components are summed. The last two bars show the decision regarding whether these two movements are similar for two cases, when the decision threshold was set at 75% and 70%, respectively; when it is green, the decision is *positive* (above the threshold), while when it is red the decision is *negative* (below the threshold). In addition, Figure 7 presents the same example, indicating the correlation between the student and teacher performances with regard to the Body, Effort, Shape and Space components for each frame. It also states the overall correlation when the weight for each component is set to 25%. For instance, in Figure 7, at frame 250, the Body correlation is 21.4/25, the Effort 24.3/25, the Shape 18.8/25, and Space 24.8/25, while the total correlation is summed up to 89.3%. In contrast, at frame 170, the Body correlation is 15.2/25, the Effort 13.9/25, the Shape 19.1/25, and Space 15.5/25, ending at a total correlation of 63.7%. In order to evaluate the ability of our approach to extract the qualitative characteristics of the movement, we asked a professional dancer to perform the same choreography three times (bachatta dance), but each time with a different intensity (I_1 refers to movement with low intensity, while I_3 to high intensity), as shown in Figure 8. Note that, in all cases, the dance steps can be considered as correct, while the intensity may indicate the dance style. Figure 9 shows the correlation between the performances for each LMA

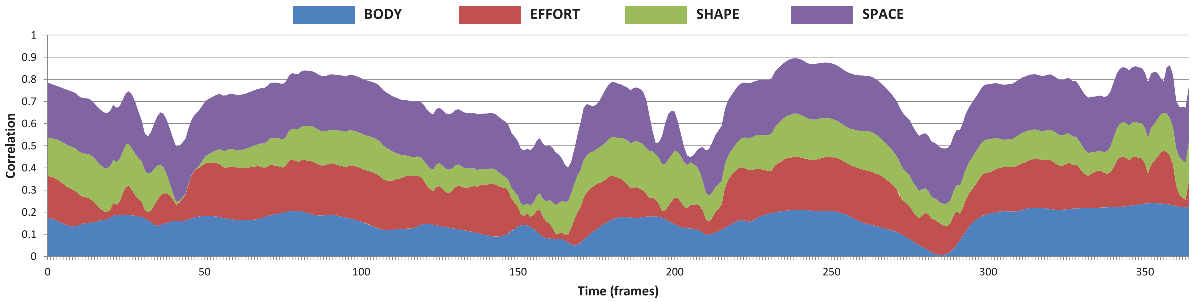


Fig. 7. An example that shows the correlation between the performance of the teacher and student.

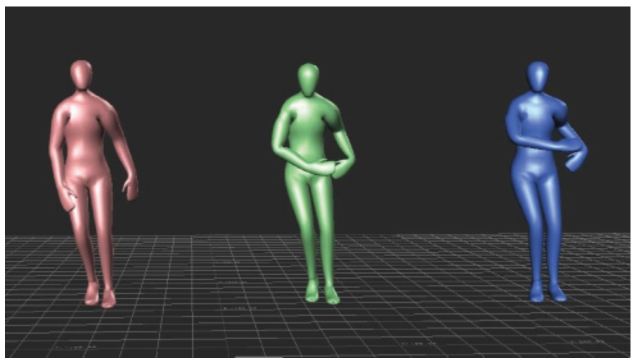


Fig. 8. The dancer performs the same choreography, but each time with different intensity. Starting from the left to the right, the red avatar presents the choreography with intensity I_1, the green with I_2, and the blue with I_3.

component, as well as the overall correlation. In this example, we have observed that the BODY and SHAPE components appear to have high correlation, especially when the I_2 and I_3 performances were compared, unlike the EFFORT and SPACE, which have smaller correlation. This is more obvious when the performances with intensity I_1 and I_3 were evaluated, which has greater variation in their motion intensity.

The dance learning simulator also offers the possibility to choose different weights for each LMA component, in order for the student to focus on individual problems and improve specific skills (based on the LMA components), facilitating the learning of the dance. Figure 10 shows such an example, where the correlation between the performances with intensity I_1 and I_3 have been evaluated, but this time having different weights for each LMA component. For instance, looking at frame 50, it can be observed that when all weights are equal (25% for each LMA component), the correlation is 64%. However, when the weights were set to 50% for the BODY and 16.67% for the other components, the correlation increased to 74.6%. In contrast, when the weights were set to 50% for the EFFORT and 16.67% for the rest, the correlation decreased to 50.8%. Keeping in mind that in the particular set of performances the body movements of the dancer follow the choreography's steps correctly and primarily differ in the intensity of the movements, we conclude that our method can effectively extract the qualitative and stylistic features of the motion.

An important aspect of our motion analysis algorithm is the ability to compare motions based on parts of the skeletal structure of the performers instead of the whole body. More specifically, we have further developed our method so that it can consider the upper and the lower part of the body

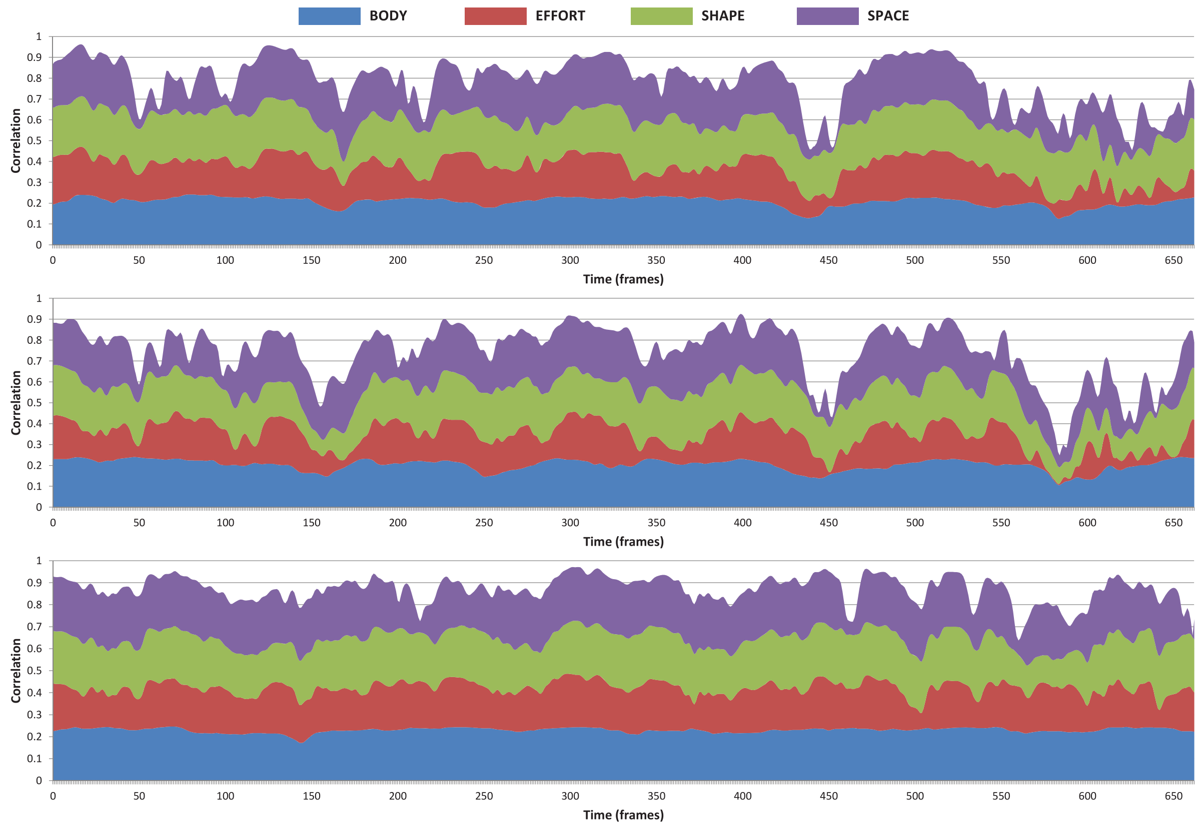


Fig. 9. The correlation between three performances with different intensity (I) when the weight factor for each LMA component is set to 25%: I_1 indicates low intensity, while I_3 indicates high intensityh. Top row shows I_1 compared to I_2, middle row shows I_1 compared to I_3, and bottom row shows I_2 compared to I_3.

separately. This separation is particularly useful in dance learning, as it allows users to focus on a specific body part without having to perform the movements correctly in other parts; thus, the users may concentrate on learning individual elements of the dance, such as the steps (emphasis in the lower part of the body) or the expression (emphasis in upper part of the body). To achieve this, we have applied a subset of the proposed features in two separate parts of the human skeleton: the upper part, which includes the root, the spine, both arms and hands, as well as the head; and the lower part, which contains the root, the hips, both legs and feet. To evaluate our methodology, we used the example in which the dancer performs the same dance but with different intensities. Figure 11 shows the correlation between two performances (I_1 and I_2). At the top row, measurements taking into consideration features of the whole body are shown. The middle row shows measurements obtained by only considering the upper body, and at the bottom row only the lower body is considered. Studying the correlations over time, we observed that the lower body processing results in higher correlation than that of the whole body, while processing the upper body results in lower correlations (e.g., in frame 265, the whole body has a correlation of 82.8%, the upper body correlation is 69.4%, and the lower body correlation is 96.1%). Separating the evaluation into two parts offers the evaluator the ability to assess the dance separately, so that the teacher and student can concentrate on specific points. For example, novices usually focus on learning the steps of the dance, paying little attention to the movements involving

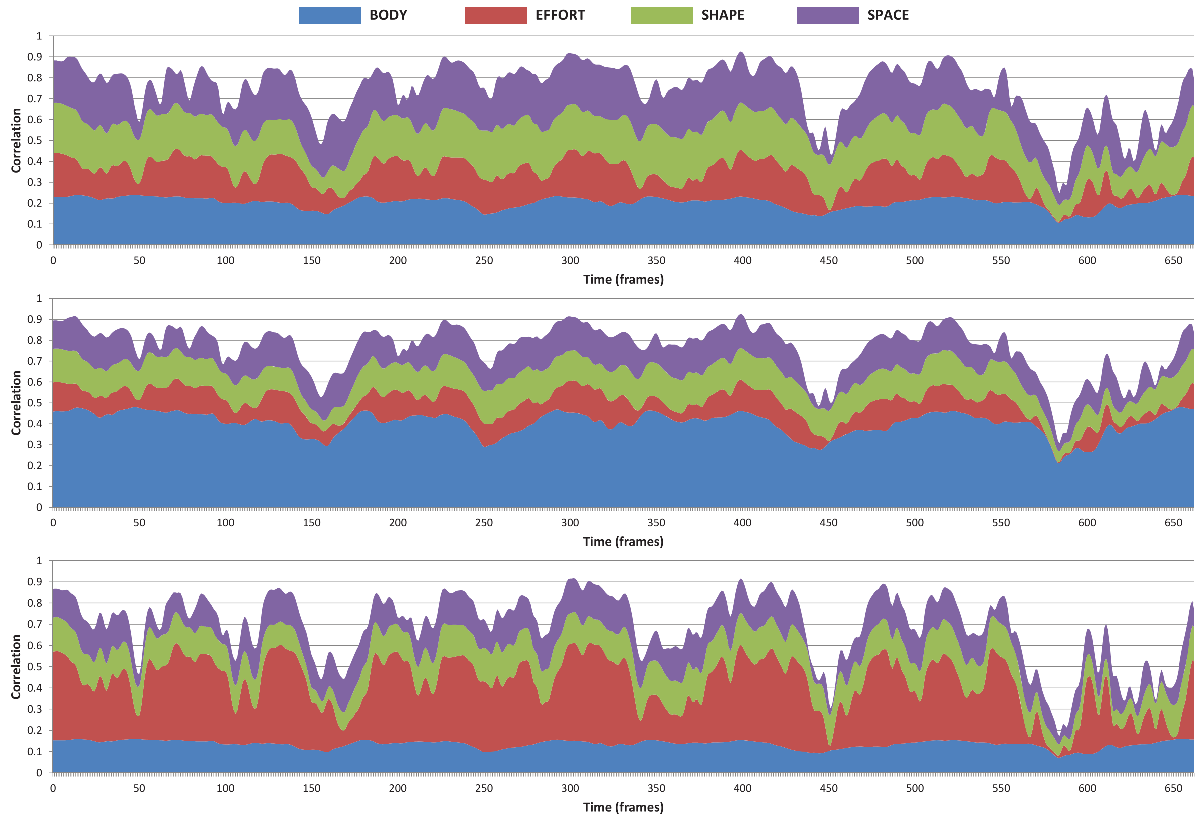


Fig. 10. The correlation when two similar performances with different intensity are compared. Top row shows I_1 compared to I_3 with weights BODY, EFFORT, SHAPE, and SPACE at 25%. Middle row shows I_1 compared to I_3 with weights BODY at 50%, while EFFORT, SHAPE, and SPACE at 16.67%. Bottom row shows I_1 compared to I_3 with weights EFFORT at 50%, while BODY, SHAPE, and SPACE at 16.67%.

their upper body. The usefulness of this selective analysis can also be applicable with inaccurate motion capture data, for example, data from a Microsoft Kinect device that may provide inaccurate or incomplete samples (e.g., for the feet).

The proposed evaluation model allows further customization of the assessment criteria in accordance with the anatomic characteristics of the trainee. Apparently, the trainee is not as fit as the trainer, who is a professional dancer, nor has the same flexibility. For instance, the student may not have the same stretching routine as the teacher, resulting in smaller openings (e.g., of the legs). Using the proposed method, the weight of specific features can be selectively reduced (while others increased) to have less impact on the overall evaluation of motion. In addition, by observing the maximum and minimum values for specific features of the student's and teacher's performance (especially features of the BODY component), we can use a proportional approach that considers the stretching capabilities of the performer. Finally, it is important to note that head orientation (f_9), which offers indications about the immediacy of motion, is not contributing in the evaluation process in cases in which the student is an amateur. In such cases, in which the trainee does not know the steps of the dance and the trainee's head is constantly turned towards the screen, no additional information is offered with regard to the style and quality of the movement, apart from the fact that the head is disoriented.

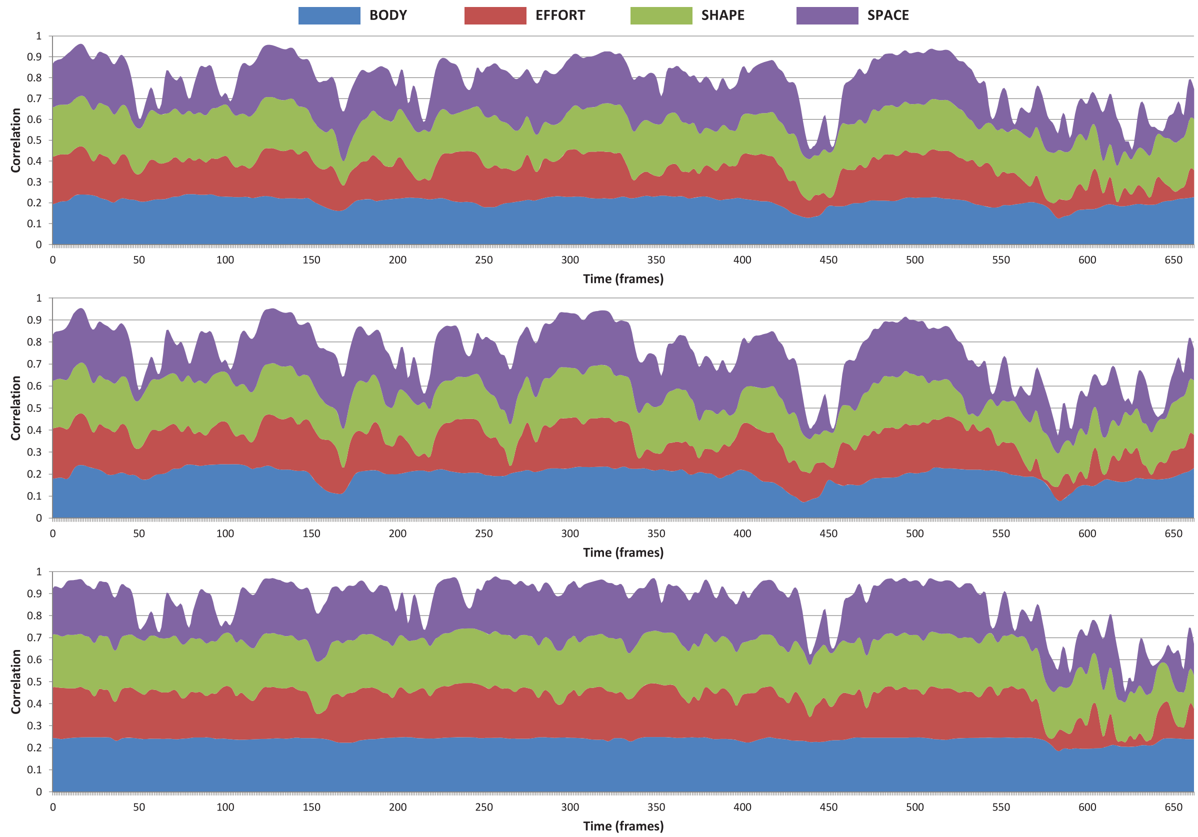


Fig. 11. The correlation when two similar performances with different intensity are compared. Top row shows the correlation when all features were used, middle row when only features of the upper body were used, and bottom row when only features of the lower body were used.

7. CONCLUSIONS AND FUTURE WORK

In this article, we have presented a novel motion comparison algorithm, which compares the movements of two avatars taking into consideration not only posture matching (meaning the physical geometry of the avatar) but also style, including the required effort, shape, and interaction of the performer with the environment. Theories that have been applied in movement analysis and education over the last century have been studied and incorporated to establish algorithms for motion comparison and evaluation. The results demonstrate the effectiveness of our method to extract qualitative and quantitative characteristics of the movement, while dance performances can be evaluated based on the LMA components. Our method is able to find correlations between different dance performances; it also offers the possibility to compare two performances having different weights of influence for each LMA component, giving the opportunity to the trainer, as well as the trainee, to adjust the dance teaching simulator according to one's needs. It may also help the user to identify potential errors in one's performance and improve specific skills. The algorithm may be used to focus only on certain body parts (upper/lower body and left/right side), as well as the whole body.

Although our work was based on real high-quality motion-capture data, this is not a prerequisite for utilizing the proposed motion analysis and evaluation algorithms. It would be possible to use motion-capture data obtained from alternative sources and even commodity hardware setups, such as multiple synchronized Kinects [Kitsikidis et al. 2014]. The only requirement is that the skeletal structure of the performer can be recovered from the motion-capture data, so that the underlying motion can be cast into the LMA feature space for further processing. A limitation of the proposed methodology is that a subset of the features required use of a short time window, resulting in delays in the extraction of the performance characteristics.

We aim to extend the proposed dance teaching simulator to work with the Dance Motion Capture database; in that manner, it will constantly be enriched with new clips and data as soon as they are available. Future work will see the introduction of a large variety of different dances and performances to establish a more complete motion-capture dance library.

In addition, for a real-time dance evaluation system, better motion synchronization and segmentation techniques need to be developed to take into consideration the experience of the user; for instance, different synchronization and evaluation approaches should be considered for amateur or expert dancers since the former needs more time to see and perform, while the latter can do so almost immediately.

The next step is to design enhanced learning tools and processes for teaching and learning dance through understanding and observing one's own movement. The outcome will be a virtual teacher that demonstrates dance through a whole-body interaction environment, giving feedback of the performance to both the trainer and trainee. This learning simulator will aim to help students develop critical skills on movement and enhance their movement literacy (ability to understand and describe their motion).

Finally, while we have focused on introducing qualitative dance comparison methods using LMA, the dance teaching system will have to be formally evaluated with human participants to establish its effectiveness.

REFERENCES

Dimitrios S. Alexiadis and Petros Daras. 2014. Quaternionic signal processing techniques for automatic evaluation of dance performances from MoCap data. *IEEE Transactions on Multimedia* 99, 1–16.

Andreas Aristidou and Yiorgos Chrysanthou. 2013. Motion indexing of different emotional states using LMA components. In *SIGGRAPH Asia 2013 Technical Briefs (SA'13)*. ACM, New York, NY, 21:1–21:4.

Andreas Aristidou and Yiorgos Chrysanthou. 2014. Feature extraction for human motion indexing of acted dance performances. In *Proceedings of the 9th International Conference on Computer Graphics Theory and Applications (GRAPP'14)*. 277–287.

Andreas Aristidou, Efstathios Stavrakis, and Yiorgos Chrysanthou. 2014. LMA-based motion retrieval for folk dance cultural heritage. In *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection*. LNCS, Vol. 8740. Springer International Publishing, Cham, Switzerland, 207–216.

J. C. P. Chan, H. Leung, J. K. T. Tang, and T. Komura. 2011. A virtual reality dance training system using motion capture technology. *IEEE Transactions on Learning Technologies* 4, 2, 187–195.

Min-Wen Chao, Chao-Hung Lin, J. Assa, and Tong-Yee Lee. 2012. Human motion retrieval from hand-drawn sketch. *IEEE Transactions on Visualization and Computer Graphics* 18, 5, 729–740.

J. Chen, W. Lin, K. Tsai, and S. Dai. 2011. Analysis and evaluation of human movement based on Laban movement analysis. *Tamkang Journal of Science and Engineering* 14, 3, 255–264.

Diane Chi, Monica Costa, Liwei Zhao, and Norman Badler. 2000. The EMOTE model for effort and shape. In *Proceedings of SIGGRAPH'00*. ACM, New York, NY, 173–182.

Emily S. Cross, David J. M. Kraemer, Antonia F. de C. Hamilton, William M. Kelley, and Scott T. Grafton. 2009. Sensitivity of the action observation network to physical and observational learning. *Cerebral Cortex (New York, NY)* 19, 2, 315–326.

Liqun Deng, Howard Leung, Naijie Gu, and Yang Yang. 2011. Real-time mocap dance recognition for an interactive dancing game. *Computer Animation and Virtual Worlds* 22, 2–3, 229–237.

Zhigang Deng, Qin Gu, and Qing Li. 2009. Perceptually consistent example-based human motion retrieval. In *Proceedings of I3D'09*. ACM, New York, NY, 191–198.

Björn Hartmann, Maurizio Mancini, and Catherine Pelachaud. 2006. Implementing expressive gesture synthesis for embodied conversational agents. In *Proceedings of GW'05*. Springer, Berlin, 188–199.

Mubbasir Kapadia, I-kao Chiang, Tiju Thomas, Norman I. Badler, and Joseph T. Kider, Jr. 2013. Efficient motion retrieval in large motion databases. In *Proceedings of I3D'13*. ACM, New York, NY, 19–28.

Gayle Kassing and Danielle Mary Jay. 2003. *Dance Teaching Methods and Curriculum Design*. Human Kinetics Publishers, Champaign, IL.

Eamonn Keogh, Themistoklis Palpanas, Victor B. Zordan, Dimitrios Gunopulos, and Marc Cardle. 2004. Indexing large human-motion databases. In *Proceedings of VLDB*. 780–791.

Alexandros Kitsikidis, Kosmas Dimitropoulos, Erdal Yilmaz, Stella Douka, and Nikos Grammalidis. 2014. Multi-sensor technology and fuzzy logic for dancer motion analysis and performance evaluation within a 3D virtual environment. In *Design and Development Methods for Universal Access in Human–Computer Interaction.* LNCS, Vol. 8513. Springer International Publishing, Cham, Switzerland, 379–390.

Lucas Kovar and Michael Gleicher. 2004. Automated extraction and parameterization of motions in large data sets. *Transactions on Graphics* 23, 3, 559–568.

Björn Krüger, Jochen Tautges, Andreas Weber, and Arno Zinke. 2010. Fast local and global similarity searches in large motion capture databases. In *Proceedings of SCA'10*. 1–10.

Rudolf Laban. 2011. *The Mastery of Movement* (4 ed.). Dance Books Ltd., Binsted, Hampshire, UK.

Nadia Magnenat-Thalmann, Dimitrios Protopsaltou, and Evangelia Kavakli. 2008. Learning how to dance using a web 3D platform. In *Advances in Web Based Learning—ICWL 2007*. LNCS, Vol. 4823. Springer, Berlin, 1–12.

Valeri Maletić. 1987. *Body, Space, Expression: The Development of Rudolf Laban's Movement and Dance Concepts*. De Gruyter Mouton, Berlin.

Megumi Masuda, Shohei Kato, and Hidenori Itoh. 2009. Emotion detection from body motion of human form robot based on Laban movement analysis. In *Principles of Practice in Multi-Agent Systems*. LNCS, Vol. 5925. Springer, Berlin, 322–334.

Megumi Masuda, Shohei Kato, and Hidenori Itoh. 2010. Laban-based motion rendering for emotional expression of human form robots. In *Knowledge Management and Acquisition for Smart Systems and Services*. LNCS, Vol. 6232. Springer, Berlin, 49–60.

Meinard Müller, Tido Röder, and Michael Clausen. 2005. Efficient content-based retrieval of motion capture data. *Transactions on Graphics* 24, 3, 677–685.

Seiji Okajima, Yuki Wakayama, and Yoshihiro Okada. 2012. Human motion retrieval system based on LMA features using IEC method. In *Innovations in Intelligent Machines*. Studies in Computational Intelligence. Vol. 376, Springer, Berlin, 117–130.

Karl Pearson. 1920. Notes on the history of correlation. *Biometrika* 13, 1, 25–45.

PhaseSpace 2014. PhaseSpace Inc: Optical Motion Capture Systems. Retrieved June 11, 2015 from http://www.phasespace.com.

Luis Santos and Jorge Dias. 2010. Laban movement analysis towards behavior patterns. In *Emerging Trends in Technological Innovation*. Advances in Information and Communication Technology, Vol. 314. Springer, Berlin, 187–194.

Takaaki Shiratori, Atsushi Nakazawa, and Katsushi Ikeuchi. 2006. Dancing-to-music character animation. *Computer Graphics Forum* 25, 3, 449–458.

Efstathios Stavrakis, Andreas Aristidou, Maria Savva, Stephania Loizidou Himona, and Yiorgos Chrysanthou. 2012. Digitization of Cypriot folk dances. In *Proceedings of the 4th International Conference on Progress in Cultural Heritage Preservation (EuroMed'12)* (October 29–November 3), Vol. 7616. LNCS, Limassol, Cyprus, 404–413.

Jeff K. T. Tang, Jacky C. P. Chan, and Howard Leung. 2011. Interactive dancing game with real-time recognition of continuous dance moves from 3D human motion capture. In *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication (ICUIMC'11)*. ACM, New York, NY, 50:1–50:9.

Joshua B. Tenenbaum, Vin De Silva, and John C. Langford. 2000. A global geometric framework for nonlinear dimensionality reduction. *Science* 290, 5500, 2319–2323.

Lorenzo Torresani, Peggy Hackney, and Christoph Bregler. 2006. Learning motion style synthesis from perceptual observations. In *Proceedings of NIPS'06*. 1393–1400.

Arthur Truong, Hugo Boujut, and Titus Zaharia. 2015. Laban descriptors for gesture recognition and emotional analysis. *The Visual Computer*, 1–16.

University of Cyprus. Dance Motion Capture Database: Retrieved June 11, 2015 from http://dancedb.cs.ucy.ac.cy/.

Yuki Wakayama, Seiji Okajima, Shigeru Takano, and Yoshihiro Okada. 2010. IEC-based motion retrieval system using Laban movement analysis. In *Proceedings of KES'10*. 251–260.

Shuangyuan Wu, Zhaoqi Wang, and Shihong Xia. 2009. Indexing and retrieval of human motion data by a hierarchical tree. In *Proceedings of VRST*. ACM, New York, NY, 207–214.

Yang Yang, Howard Leung, Lihua Yue, and Liqun Deng. 2013. Generating a two-phase lesson for guiding beginners to learn basic dance movements. *Computers & Education* 61, 1–20.

Haris Zacharatos, Christos Gatzoulis, Yiorgos Chrysanthou, and Andreas Aristidou. 2013. Emotion recognition for exergames using Laban movement analysis. In *Proceedings of Motion in Games (MIG'13)*. ACM, New York, NY, 39–44.

Liwei Zhao and Norman I. Badler. 2005. Acquiring and validating motion qualities from live limb gestures. *Graphical Models* 67, 1, 1–16.

Received December 2014; revised March 2015; accepted April 2015