



# Salsa dance learning evaluation and motion analysis in gamified virtual reality environment

Simon Senecal<sup>1</sup>  · Niels A. Nijdam<sup>1</sup> · Andreas Aristidou<sup>2,3</sup> · Nadia Magnenat-Thalmann<sup>1</sup>

Received: 2 July 2019 / Revised: 10 April 2020 / Accepted: 5 June 2020 /  
Published online: 23 June 2020  
© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

Learning couple dance such as salsa is challenging as it requires to understand and assimilate all the dance skills (guidance, rhythm, style) correctly. Salsa is traditionally learned by attending a dancing class with a teacher and practice with a partner, the difficulty to access such classes though, and the variability of dance environment can impact the learning process. Understanding how people learn using a virtual reality platform could bring interesting knowledge in motion analysis and can be the first step toward a complementary learning system at home. In this paper, we propose an interactive learning application in the form of a virtual reality game, that aims to help the user to improve its salsa dancing skills. The application was designed upon previous literature and expert discussion and has different components that simulate salsa dance: A virtual partner with interactive control to dance with, visual and haptic feedback, and a game mechanic with dance tasks. This application is tested on a two-class panel of 20 regular and 20 non-dancers, and their learning is evaluated and analyzed through the extraction of Musical Motion Features and the Laban Motion Analysis system. Both motion analysis frameworks were compared prior and after training and show a convergence of the profile of non-dancer toward the profile of regular dancers, which validates the learning process. The work presented here has profound implications for future studies of motion analysis, couple dance learning, and human-human interaction.

**Keywords** Motion analysis · Salsa dance · Virtual reality · Human computer interaction

## 1 Introduction

The analysis and investigation of the effects and intricacies of social dances are ample and find contributions in many sociological, cultural, and psychological areas. This comes at no

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11042-020-09192-y>) contains supplementary material, which is available to authorized users.

---

✉ Simon Senecal  
senecal@miralab.ch

Extended author information available on the last page of the article.

surprise, as social dances already exist for centuries and are embedded in many cultures, ethnic groups and are often related to a social and/or religious context [33]. More particularly, couple dance is a specific declination of social dance acted in pair, traditionally with one man and one woman in a mechanical interaction that allows making some complicated moves, with each of the partners having a specific role (the man is leading, and the woman is following). This type of dance can be found widely across the world, as Forro in Brazil, Tango in Argentina, Fox-trot in the USA, or Valse viennoise in Austria. In more recent studies, the attention to social couple dances is also found in the fields of bio-mechanics, Human Robotic Interaction (HRI), and Human Computer Interaction (HCI), examining the features and its application in the digital domain. Within the latter context, we focus on the predominantly cognitive connection between the dancers while performing a social couple dance. The human to human interaction with full-body movements are coordinated and fine-tuned upon each other, and in most cases attuned to the music, which dictates the rhythm and the “way” a dance is carried out (e.g., slow vs. energetic). Another aspect of the interaction is the “lead” and “follow” roles, which refer to the impulse and response pattern during the dance and the connectivity between the couple. The vastly dynamic and interactive situations of social couple dances bring a plethora of parameters, derived from the physical and cognitive interaction, the musical interpretation and listening (e.g. body “drive”), and represents a tremendous challenge to comprehend and analyze this intricate and interdependent set of parameters. Salsa is a very social couple dance that is popular around the world and whose learning has its challenges:

- Learning in (large) collective classes, which is less effective in spotting errors on individual students.
- The need to practice with another partner on location, meaning the risk of inadequate facilities and/or not having a partner to practice with (either by lack of dance partners or due to personal time schedules).
- Other parameters can influence the study, such as mood, stress, fatigue, and other external social factors.
- Time and location constraints due to other obligations (e.g. studies, work).

Besides, when the student is reaching a similar skill level as its teacher, the student may oppose the advises given by the teacher as to what is “correct”. The status of an *expert* in social dance can be a source of confusion as there is no official diploma state validated but rather a public recognition of skills by pairs. In many cases, the learning process can be less effective, halted, or reconsidered, depending on the relationship between student and teacher. The use of virtual reality exercises have been proved relevant for training in a range of essential jobs (army, pilots, firefighter, etc.) and shows a real improvement of the learner skills, allowing being in the complicated situation as in real life. Giving the complexity of salsa dance, virtual reality is an excellent alternative option for learning dance since it provides the required mechanical interaction between the user and the virtual character, and allows for tracking the full-body movements over an area similar to the one needed for dancing. The main objective of this paper is to demonstrate that we can guide and help users to improve their salsa dancing skills through a Virtual Reality (VR) game that simulates salsa practice. In previous work [40], we showed that six criteria are important for learning salsa; *Rhythm*, *Guidance*, *Fluidity*, *Sharing*, *Styling* and *Musicality*. In this work, we focus on the evaluation of three main skills, which are the *Guidance*, *Rhythm*, and *Style*. In that manner, we have designed a VR application that facilitates a virtual partner, in an interactive environment, and simulates dancing in a couple. Each user wears a VR headset

with hand controllers and performs along with a virtual partner. The motion of the users is recorded using an optical motion capture system, and their movements are linked to the virtual avatar using Inverse Kinematics. The user goes through a series of exercises, and the system returns an overall score to motivate the user to compete against others. We performed an extensive analysis of the recorded exercises, and evaluated the learning skills and progress of the users at different learning stages with regard to the aforementioned important criteria; the analysis was conducted using a number of Music-related Motion Features (MMF) and Laban Motion Analysis (LMA) features. Results demonstrate the improvement in dancing qualities of the *non dancers* that tend to converge to the qualities of the *regular dancers*. Figure 1, shows a visual illustration of our VR environment, where a user interacts with the virtual environment.

The main contributions of this paper are itemized below:

- A VR environment that guides and helps users to practice and improve their dancing skills through dance gamification, and more specifically, via interaction with a virtual avatar. This application also provides seamless motion capture that can be used for further processing and studies.
- A motion analysis that evaluates the influence of our application on the dance skills of users, in terms of three main criteria: the guidance, rhythm, and style. We extract, evaluate, and validate the important MMF and LMA features using a two-class dataset of *regular* and *non dancers*, while their movement is synchronized with music.



**Fig. 1** Our learning salsa gamified VR application. The user wears a virtual reality headset and interacts, in real-time, with a virtual partner using hand controllers. The images of the user in the real-world (left side) and the dancer in the virtual environment (right side) are blended

## 2 Related work

The human motion, during a dance, often carries emotion and is connected with the whole cognitive-motor and psychological system. It has been investigated through multiple scientific studies, including dance motion generation [21, 41], synchronization to music [10, 43], emotion recognition and stylization [8], and represents many challenges for learning [34]. Besides the benefit of social dances for health as improving balance and cognition for elderly [25–27], its interactive aspect has been touched upon by the HRI domain. For example, where through sensors detection, the user's movements transcribed into an intermediary data set to generate poetry [12, 13]. Human to human interaction has also been explored via a setup of patches [42] and scene ranking [47] in the context of an animated character. Another example is the use of robots acquiring the knowledge and skills to perform a dance [30]. However, the research is limited to single instances of a dancer, thus not taking into account the simultaneous act of dancing. The interaction between performers themselves has been studied in the psychological domain [29, 46]. The interaction between the public and the performers has also been investigated by [44]. More in our focus, a set of studies would evaluate the dance performance using various methods, as Kinect [2].

Extracting the motion features from continuous movement is a crucial element for describing, evaluating, and understanding dance and movement in general. For instance, locomotion has been studied with gait analysis and classification with extreme machine learning and leg joint angles data [31, 37]. Studies on everyday actions [18] are proposing a set of features inspired by psychology and physiology to characterized behaviors and the subsequent emotion involved. More specifically, the use of LMA-based features has proved to work well in different situations, such as motion retrieval, indexing, and comparison [4, 6], and is therefore ideal to be used as a base to build a machine learning classifier, as demonstrated for theatre emotional expression [38] or evaluating the performer's emotion using LMA features [3]. Other studies focused on a specific motion feature, for example, the fluidity of the movement that is a critical dance parameter investigated in [32]. In this particular study, it is proposed to see how fluidity can help to describe and to classify dance performance through interdisciplinary research including biomechanics, psychology, and experiments with choreographers and dancers, and they propose a definition that takes specifically the minimum energy dissipation when looking at the human body as a kinematic chain. Another work [1], elaborated upon the expressive qualities, such as rigidity, fluidity, and impulsiveness, to investigate intra-personal synchronization for full-body movement classification. In our previous study [39, 40], we propose a set of motion features that take into account the particular context of salsa dance: motion synchronized with music and interaction with other partners.

Learning is an essential aspect of enhancing dance performance and the use of virtual reality for it is bringing immersion, visualization, and interactivity that shows promising results [11, 22]. Important studies in the field of human visual appearance [9] provide several advises on a good virtual human representation for better interaction. Even a commercial application has been already proposed to learn Salsa dance in virtual reality with a coach [14]. The first study about Forro dance [16], evaluates how the user can learn and improve his dance skills through repetitive training, monitored by his smartphone. The proposed evaluation features computed with the user's motion data from the smartphone's Inertial Measurement Unit (IMU) sensor and the music data: First the "*Rhythm Beats Per Minute (BPM): We calculate the average beats per minute.*", then the "*Rhythm consistency: we calculate the coefficient of variation of the student's BPM across the full dancing exercise*". This study brings exciting insights into characterizing Forro dance learning and

relevant dance features, but the restricted data source (only one point IMU source) is an obvious limitation. In a recent study, a VR interactive simulation of salsa dance using Hidden Markov model to predict the virtual partner dance behavior has been developed [28]. Although the kind of “Top-down” approach, the introduction of jump transitions is making sense as going towards the structure of Salsa dance as learned in classes (based on cycles of 8 beats). This study has good feedback from users regarding the naturalness of the motion and the dance-following feeling. It would be interesting to understand which specific motion features produced by the Markov model enables such perception by the users. Another work is to develop a dance game based on motion capture technology [15], addressing the issue of user’s performance real-time estimation to determine what a virtual dance partner should display as interactive motion. The real-time prediction was based on body parts indexing in conjunction with flexible matching to estimate the completion of motion and reject unwanted motions. A method to control a real-time virtual character using a motion capture system is also proposed [24]. In their method, the character’s motion extracted from a database and pre-processed using a two-layer method. A Markov process is used in the first layer, and a clustering technique is used in the second layer. Finally, a framework is developed [20] for synthesizing the motion of a virtual character in response to the actions performed by a user-controlled character in real-time.

In comparison to the previously mentioned approaches, our work is based on simulating a Salsa dance environment in VR with the focus on user experience and dance skills learning. Indeed, gamification is an interesting process of improving the engagement of user for learning system [19]. We aim at providing the most convincing level of Salsa simulation such that we induce performance improvements. Alternatively to studies that take into account only the basic steps or style elements, we are considering the whole behavior of each partner and their relationship to the music. To analyze and validate the application, comparative motion analysis is required. This analysis is done using, on the one hand, the well-known LMA features, that have shown their accuracy in depicting style in dance within a lot of papers, and on a second hand the MMF, that are a new proposition dedicated to interrelated music and dance motion.

## 3 Design

### 3.1 Overview

Our objective is to develop an interactive dance learning system that is able to improve the dance skills of the engaged users. To achieve that, we propose a framework constituted in three components that fulfill the following technical requirements: a VR salsa simulator, a gamified learning system and motion recording for further analysis. The VR salsa simulator recreates the condition of salsa dance from the leader role side, involving: (a) visual contact and viewing of the engaging partner, (b) natural and physical interaction, (c) an adequate music to dance with the virtual partner, having the ability to guide it into dance, and (d) finally, enough space to allow freedom of movement. This educational and gamification activity ensures the development of dance skills through pedagogical training: it embeds a series of exercises that are easy to understand and start with, it has repetitions, based on timed hand gesture and full body movements, different musical tempos for a dynamic training, and a final score that is accessible at the end of the session to keep up the motivation and engagement. During all exercises, the full body motion is recorded at high frame rate

to allow real time or post processing motion analysis. Figure 2 illustrates an example of a person testing our VR environment.

### 3.2 Salsa simulator

The first step of our work is the design of a VR application based on real salsa practice. For that, we based our work on the observation of real body movements during dance. An important point is the role of each partner. There are one leader and one follower. Both are dancing on the rhythm independently, but the leader will influence the follower motion via his hands, chest or other “connection” tools, and the follower will “listen” this indication and change its dance pattern accordingly. In our game, the user will have the role of a leader, and a virtual partner will be the follower. Similarly to real dance scenarios, our virtual partner behavior can be structured into two animation layer working in parallel: moving the body and feet on the tempo of the music, and reacting to the user *guidance*. The latter reaction has to be natural regarding the user stimulus. Inverse Kinematics (IK) is thus used as it allows to animate the full body (the end-effectors, such as the hands, feet, and head) with time and position constraints. A good and reliable VR setup is necessary to ensure good immersion. We used the HTC Vive, as our VR system, since it possess very high-fidelity and wide space tracking, enough to cover the needed space when dancing Salsa, and it allows the use of additional tracked markers.

#### 3.2.1 Virtual partner model and music-synchronised dance animation

A visually pleasing model, but still a little bit cartoony, is chosen among commercial solutions for the Virtual partner appearance, so to engage the user for interaction. A layer of



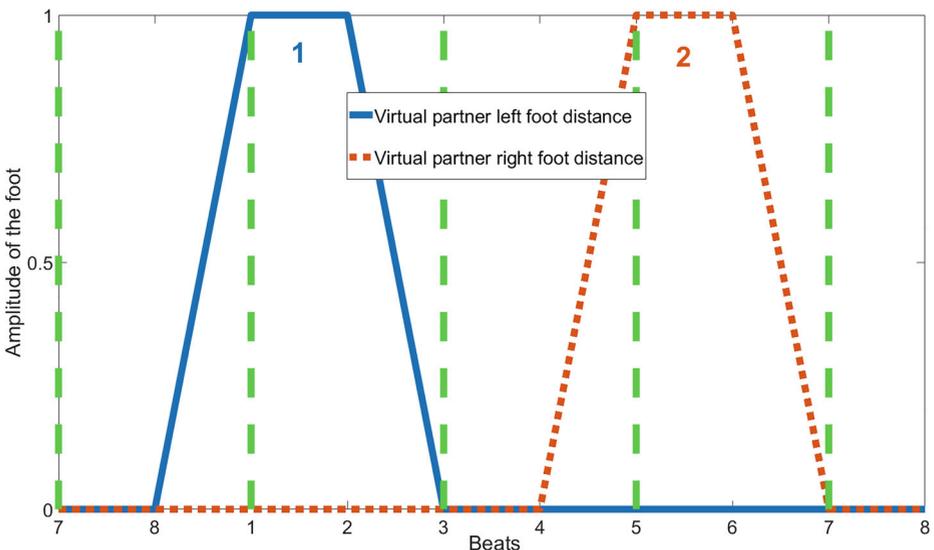
Fig. 2 Example of a person testing the game

inverse kinematics with physical constraints (bending of the upper body and other limbs) is added to the rigged model, allowing to manipulate the end-effectors with ease, achieving constituent motion. The knowledge of basic salsa step's motion in space comes from a previous study [39, 40], from which we extract a motion profile for each foot, as illustrated in Fig. 3. This motion profile serves as a base to set the position in space of the IK targets corresponding to the right and left foot of the Virtual Partner (VP).. The time length of the motion profile is proportional to the music tempo, ensuring the virtual partner always dance "in rhythm". Additionally to this, we move the root at the half distance of the foot position, such that the upper body is always straight and kept balanced. The result is an entirely natural motion that is totally in adequation with the basic salsa steps theoretical description.

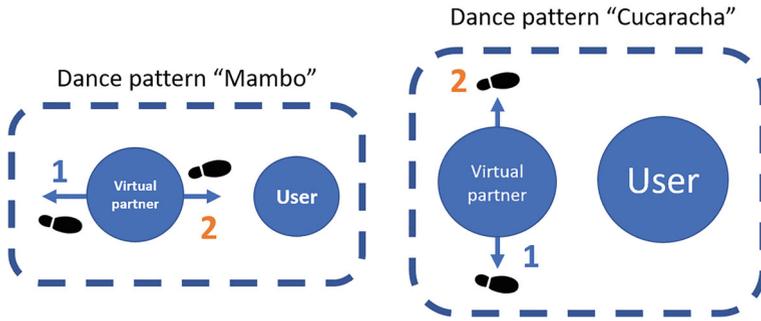
The direction of the basic step using this motion profile can be divided into two main directions, giving us two dance patterns: A forward-backward motion called "*Mambo*" and a right-left motion called "*Cucaracha*", visualized in Fig. 4. The user can follow the steps of the VP in order to catch the music tempo. A drawing of footsteps is placed in front of the VP to help the user be rightly positioned.

### 3.2.2 User interaction: guiding the virtual partner

To simulate the feeling of *guidance*, the user can control the transition of the VP dance pattern via interactive gesture and timing. To give the feeling to hold hands as in Salsa, the hands of the VP are placed near the user's hands in real-time (as an IK position constraint), and the remaining arm is animated through IK, as in the case of manipulating a rag-doll. The correct user hand gesture required to control the transition is detected through the computation of forces. The IK system computes the push force applied from each hand to the respective VP shoulders. Then we extract a forward force (whether the user is pushing or

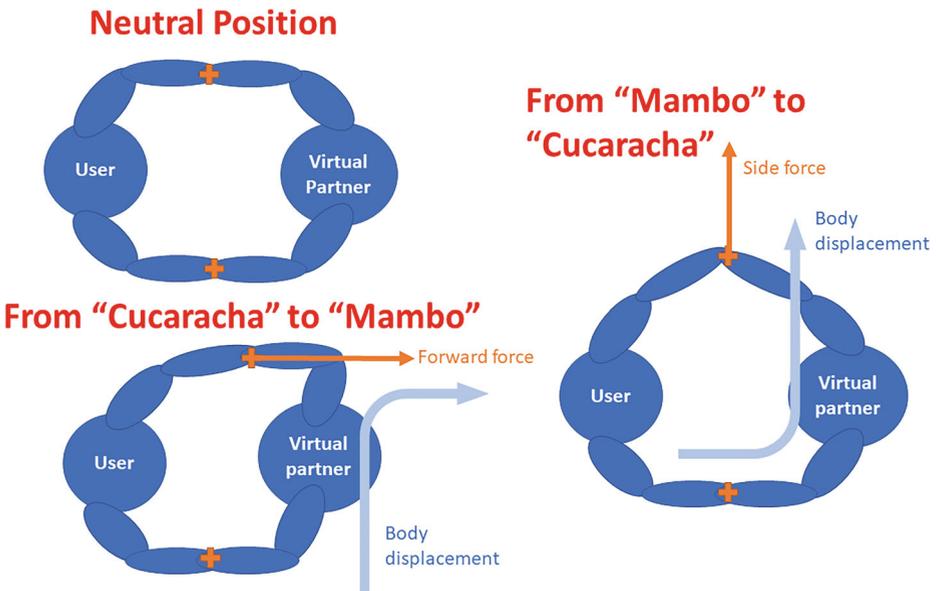


**Fig. 3** The distance for right and left foot of the virtual partner related to a neutral position. This motion profile is repeated every 8 beat of a music. In the application, the curves were smoothed to achieve better naturalness



**Fig. 4** Our two salsa dance pattern. On the left side for “Mambo”, the VP will step backward during the beat 1 and 3, then moves forward during the beat 5 and 7. On the right side for “Cucaracha”, the VP will move his right foot to his right then his left foot to his left alternately

pulling the VP’s arm in front) and a side force (whether the user is pushing or pulling VP’s arms on the sides) with the dot product. This information is calculated in real-time and allows us to know how much force the user is producing on the VP, and in which direction. This analysis gives us two important information: the time the force is applied and the direction of the force (Sides or front). A valid gesture for transition is considered if: The direction of force is perpendicular to the direction of the current dance pattern, and if the force occurs between the beat 7 and 8 (in a similar manner to [39]). The results give the user the feeling of guiding the VP, as illustrated in Fig. 5.



**Fig. 5** Detail of the required gestures to control the virtual partner on its transition to the two possible dance pattern. To guide the virtual partner from a “Cucaracha” motion to a “Mambo” motion, the user have to push the virtual partner between the beat 7 and 8 with his left hand. To do the reverse transition, the user have to pull the virtual partner on his left at the same time

### 3.2.3 Software design

The overall VR application is developed under Unity3D game engine, including all necessary plugins to work with our VR device. When our VR application starts, an initialisation phase waits for the user inputs e.g., the name, to automatically label the saved motion data. In the meantime, the IK animation is activated, allowing to manipulate the virtual partner via holding hands to get familiar with the environment. Then when the user start the training, a countdown is provided and the virtual partner dance animation is triggered, as well as its transition system and the music, all at the exact same time. Finally at the end of the training, the application displays briefly the final score and goes back to the initialisation phase.

### 3.3 Learning and gamification

The main focus of our implementation is to provide the essentials to users to develop two main dance skills: *rhythm* and *guidance* through pedagogical and fun exercises. We set up in our VR application a series of repetitive exercises containing two dance tasks. The tasks consist of the user to move his feet on the music and to guide with his hands the VP to change its basic dance pattern every two-cycles of 8 beats (two simultaneous attentions are needed). There are eight exercises of different tempos in order to vary the difficulty of the task and keep the training dynamic, with a short pause in between them. A feedback, in the form of a final score, is then computed, based on the number of successful guidance attempts compare to a reference number, and provided at the end of the session as reward. Between the first and the last exercise (that are at the same tempo), the user is expected to show an improvement in terms of *guidance*, *style*, and *rhythm*. The gamified aspect of this application is important for the user engagement with a focus on the usability, playability and fun.

### 3.4 Motion data recording

A post-process motion analysis allow to evaluate the ability of learning system to improve dance skills and subsequently, the relevance of our design. The movements of the user are captured via the default VR setup (hands, and head), and additional tracking markers that are placed on the hips and feet. We then get a pose representation in this context of six points. The coordinates of each point are recorded during the training session at a high frame rate (100 frames per second) to ensure high quality and high speed analysis of all kinematic components. This pose representation is giving us enough information for meaningful motion analysis.

## 4 Experiment

One way to show that our VR platform helps users to improve their skills is by computing their MMF and LMA features on the early stage of the training, and then comparing them with the corresponding features at the end of the training. To test our application and evaluate its ability to help users to improve their salsa learning, we conduct experiments using two dancer categories with different experience:

- *Non dancer*: people that never have any class nor experience in salsa dance,
- *Regular dancers*: people that do take class of salsa and have at least one year of practice.

We expect that the performance of the *non dancers*, at the end of the experiment, will converge towards the *regular dancer*'s, indicating an improvement in their learning skills. For each user, the objective is to go through a series of eight exercises. In each exercise, salsa music is played, and the VP moves in synchronization. The aim is to follow the music and guide the VP to change its dance pattern every two iterations. The time of each exercise is about 60 intense seconds during which they are constantly making physical effort to keep the rhythm of the music and the guidance task. The criteria for evaluation are the same for each exercise, with minor variation in difficulty to keep the training dynamic.

The tempo varies in order to stimulate the user but is the same at the beginning and the end of the training for consistent analysis. A summary of the exercises performed by each user is listed in Table 1.

We invited 40 people to participate, half of the participants were *regular dancers* and half *non dancers*. Note that data acquisition is challenging, mainly because it requires the participants to physically participate in the experiment, in our lab, and use our devices. Nevertheless, as shown in Section 6 a training sample of 40 people show a clear learning trend, and suffice to validate this direction.

The setup is not as light as the simplest VR devices, but is light enough so that the participant can move freely (this is also due to the wireless system used). After a short tutorial preparation, each participant went through 8 exercises and got a final score. This score is based on their success to accomplish the given aim (changing dance pattern every 2 cycles) and serve mainly as a motivation for the user to compete against others. With 40 users over 8 exercises, the resulting database represents 320 sequences of motion capture, which are recorded each as 4500 frames of 6 points-skeleton.

## 5 Motion analysis

In this work, we used two well-known motion analysis system to evaluate the movement of the participants, the Musical Motion Features, and the Laban Movement Analysis system.

**Table 1** Summary of the exercises constituting the application

Exercise	Tempo of the music	Remarks
Exo 1	180 bpm	Serve as tutorial for people to get into it
Exo 2	180 bpm	same tempo
Exo 3	180 bpm	same tempo
Exo 4	160 bpm	tempo slower
Exo 5	200 bpm	tempo faster
Exo 6	140 bpm	slowest tempo, easiest for non dancer
Exo 7	220 bpm	fastest tempo, very difficult for non dancer
Exo 8	180 bpm	Final exercise with same tempo as initial

## 5.1 Musical motion features - MMF

Salsa is a specific type of dance in which movements are highly correlated with the music and the other partner. To take that into consideration, we have previously proposed the MMF framework [39, 40], which contains the relevant motion features. MMF indicates excellent performance in classifying motion data with regard to three essential salsa dance skills: rhythm, guidance, and style. In our previous study, we proposed (following dance experts' suggestions) six criteria. However, only three of them were investigated, mainly because the remaining three require complex analysis, and each one a full study on their own. Similarly, in this study we used the same three criteria, which provide though the essentials for developing an accurate prototype for analysis and evaluation of the learning performance of our participants. This framework has been used to distinguish beginner from expert dancers, and was validated through a user study (participants are separated based on their dance level: Beginner, Intermediate and Expert) from a huge amount of motion data (26 couple dancing over 10 songs of 120 minutes). These MMF features carry information relative to dance skills and are therefore a sort of interface between low-level and high level data. Here, the goal of our analysis is to evaluate the performance of one person dancing with a virtual partner that has a predefined behavior.

We consider only a subset of the proposed MMFs, given that features concerning the VP will not vary. We are using sixteen measurements  $\mu_j$  that belongs to five feature categories, extrapolated from three dance skills, that are shown in Table 2. All measurements are observed on a temporal window of given frames corresponding to 8 beats. Previous experiments, e.g., [39], show that 25 frames per second are the best to extract meaningful results. Thus, we downsampled the initial frame rate (100Hz) to 25 frames per second (fps) without

**Table 2** Subset of the musical motion features in our case of virtual reality

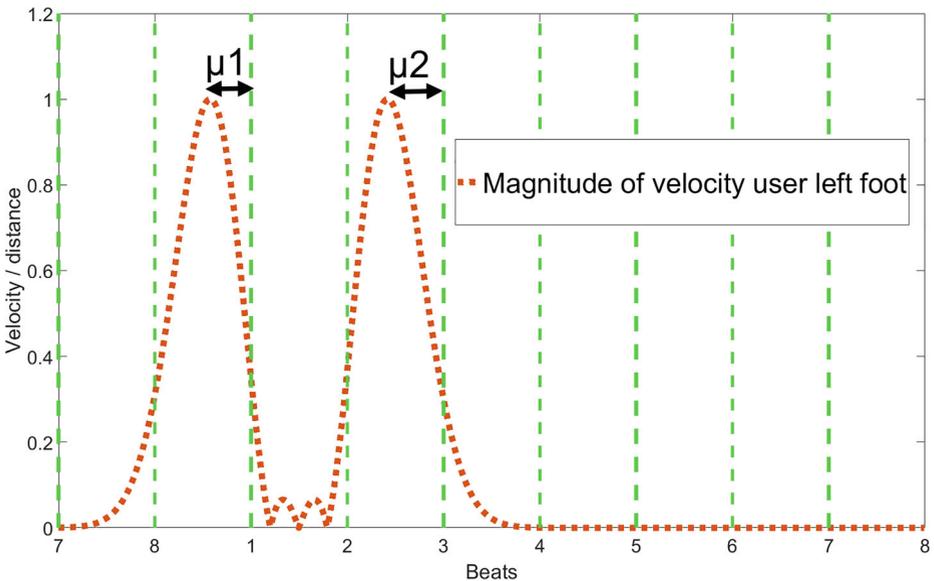
	MMF	Measurements details
Step	$\mu_1$	Temporal accuracy of left foot over beat 1
	$\mu_2$	Temporal accuracy of left foot over beat 3
Accuracy (Rhythm)	$\mu_3$	Temporal accuracy of right foot over beat 5
	$\mu_4$	Temporal accuracy of right foot over beat 7
Rhythm difference between partners (Rhythm)	$\mu_5$	Temporal difference user left foot /VP right foot over beat 1
	$\mu_6$	Temporal difference user left foot /VP right foot over beat 3
	$\mu_7$	Temporal difference user right foot /VP left foot over beat 5
	$\mu_8$	Temporal difference user right foot /VP left foot over beat 7
Foot Correlation (Guidance)	$\mu_9$	Correlation coefficient beat 1 to 3 user left foot / VP right foot
	$\mu_{10}$	Correlation coefficient beat 5 to 7 user right foot /VP left foot
Area (Styling)	$\mu_{11}$	Displacement of the left foot over 8 beats
	$\mu_{12}$	Displacement of the right foot over 8 beats
	$\mu_{13}$	Net velocity change of the left foot over 8 beats
	$\mu_{14}$	Net velocity change of the right foot over 8 beats
Hands Movements (Styling)	$\mu_{15}$	Mean distance left hand to hips over 8 beats
	$\mu_{16}$	Mean distance right hand to hips over 8 beats

loss of the temporal information (see Forbes and Fiume [17]). Finally, each measurement of  $\mu_j$  is normalized between 0 and 1.

### 5.1.1 Dance skill: rhythm

**Step accuracy ( $\mu_1 - \mu_4$ )** One of the essential features when learning salsa dance is rhythm and the ability of the user to follow and be synchronized with the music beats. In that manner, we consider the velocity magnitude over 8 musical beats for each foot. For example, when dancing the “Mambo” pattern, two peaks occur that indicate a movement of the foot on the music. The first peak corresponds to a step forward (beat 1), and the second peak to a step back to the neutral position (beat 3). The same occurs for beats 5 and 7. Given the temporal location of each musical beat, we can compute the step accuracy for each beat as the difference between the musical beat and the user’s foot motion time. Thus, via filtering and peak detection, we can evaluate the temporal location of each of the user’s steps, and compare them to the musical beats, once it is extracted from the music, as shown in Fig. 6. The result is 16 measurements that are extracted through a sliding window of width proportional to the music tempo. The beat 1 is detected by hand at the beginning of each song to ensure the good temporal accuracy of each sliding window.

**Rhythm difference between partners ( $\mu_5 - \mu_8$ )** These features have been placed as a Rhythm skill since the partner motion, in our application, is predefined and therefore acts as a tempo reference. During the dance, the foot motion of the user and the VP are in opposition. Then, similarly to the aforementioned Step Accuracy feature, we detect the temporal location of the user’s beats and compare each of them to those from the VP. Values toward zero are considered as good synchronizations.



**Fig. 6** Velocity peak that occurs during the basic step for the user, and the measurement of step accuracy

### 5.1.2 Dance skill: guidance

**Correlation between foot movements ( $\mu_9 - \mu_{10}$ )** Computing the 2D correlation coefficient of the 8 beats velocity's magnitude between the user and the VP foot motions gives insights about the synchronization of the couple, given that their respective moving feet are supposed to move oppositely and simultaneously (the left foot of the user in the same time as right foot of the VP).

### 5.1.3 Dance skill: styling

**Area ( $\mu_{11} - \mu_{14}$ )** During a cycle of 8 beats, the displacement of the feet is measured by the integration of the velocity over time. In addition, the net velocity change is measured by the integration of the velocity's derivative. These values are computed for each foot, and provide insightful information on the dynamic of the stepping action.

**Hands movements ( $\mu_{15} - \mu_{16}$ )** The dynamic of hands movements provide intuitions and helps in characterizing the styling aspect of salsa. It is computed by taking the mean distance between left / right hands and hips over 8 beats.

## 5.2 Laban movement analysis - LMA

Analyzing human motion is particularly challenging, especially when the goal is to evaluate the learning skills with parameterized geometry and style control. In order to identify and evaluate the learning skills of our platform, we learn motion characteristics based on the LMA principles [23], drawing from the framework described in Aristidou et al. [7]. This framework was strategically designed to capture the diversity of stylistic and geometric characteristics of a set of dancing motions [3], and has been used to analyze and compare folkloric dances [6]. In contrast, the goal of our analysis is to learn features that are characteristic of learning skills among performers with different experiences in dancing.

In this work, we define, as local spatiotemporal descriptors, one-dimensional arrays that encode the LMA-derived features, from selected key joints. We have considered 29 low-level spatiotemporally varying features ( $f^i$ ) of the human body, which were chosen according to the four LMA components (Body, Effort, Shape, Space). For each feature the *minimum*, *maximum*, *mean* and *standard deviation* values were computed, resulting in 114 different feature measurements ( $\phi_j$ ). These measurements are taken by observing each feature over a short temporal-window around a given frame (a 30-frames right anchored sliding window, at 25 frames per second) through each motion sequence, with step 20 frames (10 frames overlap). These feature measurements are after that normalized so as their values range between 0 to 1.

Thereafter, and similarly to [8], we select those features that are **consistent** among the same group of performers (*regular dancers* Vs *non dancers*), and **effective** across the two different groups. This allows us to make a meaningful mapping from the low-level feature space of the underlying motion into the learning skills. To achieve this, we consider in our analysis the mean and standard deviation of the sample values for each feature, for both classes. We define as **effective** and **consistent** features those that their standard deviation is small for motions of the same group ( $< 10\%$  of the value), and the mean values between the two classes have a significant difference ( $> 20\%$ ). Since the movements in our dataset are strictly structured, and the variation in motion is limited, not all LMA features are important in separating the two classes. Based on our LMA feature analysis, we concluded that only

twenty LMA feature measurements are useful for separating the two classes, which are listed in Table 3.

## 6 Results and discussion

Two complementary methods are used to describe the learning effect of the game. In terms of guidance and rhythm (including synchronization), we used the MMF features, and in terms of the movement style (including effort, volume, and space) the LMA features. To evaluate the skills' improvement in learning salsa, we compare the values of the corresponding MMF and LMA features for the second and the last exercises. Note that, we chose not to use the first exercise since it is acted as a training step for the dancers to get familiar with the VR environment.

### 6.1 MMF study

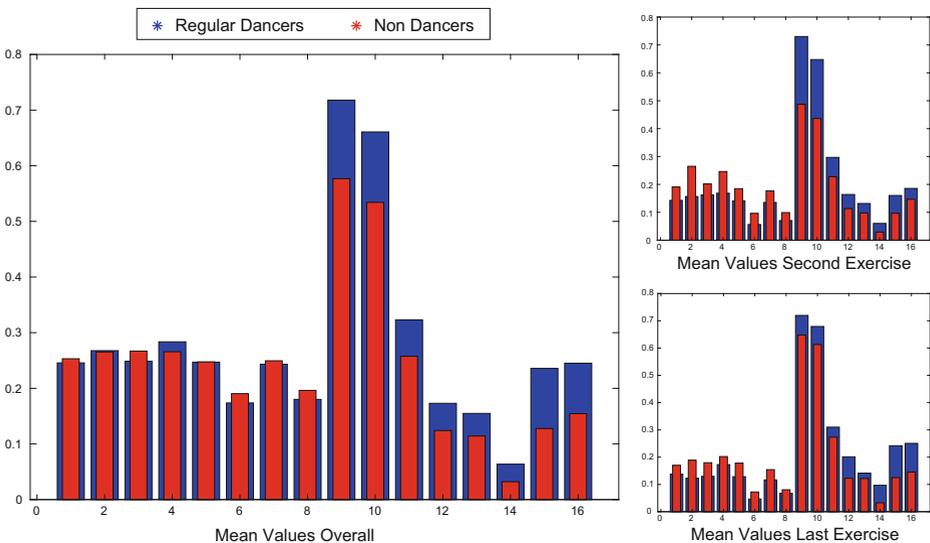
For each performer and exercise, we extract one-dimensional arrays (the windows of MMF measurements using a sliding window of width proportional to the music's tempo), and represent each performance by the mean value of all these local descriptors. Our target is to evaluate the performances of the two categories (*regular dancers* vs. *non dancers*) over time, and observe potential changes in the quality of dance after training.

**Table 3** The consistent and effective LMA-derived feature measurements used for separating the two classes, which provide insights about the learning skills of the participants. The feature measurements indexing follows the numbering of [7]

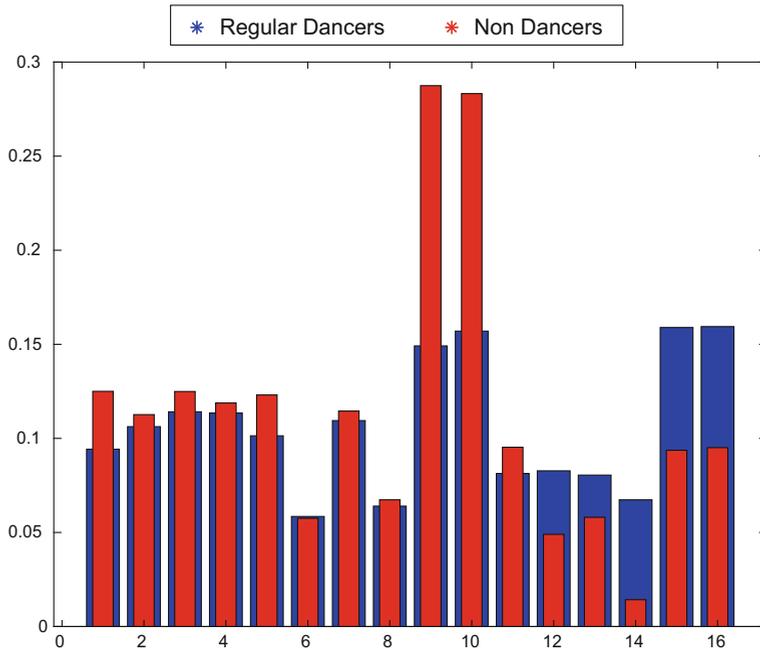
A/A	LMA	Description
1.	$\phi_{11}$	Left hand-shoulder distance (std)
2.	$\phi_{12}$	Left hand-shoulder distance (mean)
3.	$\phi_{15}$	Right hand-shoulder distance (std)
4.	$\phi_{16}$	Right hand-shoulder distance (mean)
5.	$\phi_{19}$	Hands distance (std)
6.	$\phi_{20}$	Hands distance (mean)
7.	$\phi_{47}$	Gait size (std)
8.	$\phi_{57}$	Left hand velocity (max)
9.	$\phi_{59}$	Left hand velocity (mean)
10.	$\phi_{60}$	Right hand velocity (max)
11.	$\phi_{62}$	Right hand velocity (mean)
12.	$\phi_{65}$	Left foot velocity (mean)
13.	$\phi_{68}$	Right foot velocity (mean)
14.	$\phi_{69}$	Pelvis acceleration (max)
15.	$\phi_{71}$	Left hand acceleration (max)
16.	$\phi_{73}$	Right hand acceleration (max)
17.	$\phi_{75}$	Left foot acceleration (max)
18.	$\phi_{77}$	Right foot acceleration (max)
19.	$\phi_{104}$	Torso height (mean)
20.	$\phi_{114}$	Cumulative distribution (mean)

Figure 7 left shows the mean values of the MMF measurements  $\mu_j$  of the performers for the two classes for all exercises, while on Fig. 7 right shows the mean values of the performers for the second (top) and the last exercises (bottom). It can be easily observed that the mean of the MMF measurements for the *regular dancers* have larger values than those of the *non dancers* in regards to the MMF styling and guidance skills. This is in line with our expectations since *regular dancers*, due to their long-time experience, have better guidance than the *non dancers*, and put more effort into dancing, making wider steps and moving their hands more intensely. Another important observation is the significant improvement in the guidance feature for the *non dancers* when comparing the beginning and the end exercises of the training, as well as the notable decrease in their rhythmic error (hence increase their rhythmic accuracy). These two observations indicate an advancement in the performance of the *non dancers*, which supports our claims that our system helps users to improve their salsa learning ability and skills. It is also important to note that *regular dancers* have slightly improved their performance (their MMF features stays relatively the same), reducing their rhythmic error. This indicates that their dance behavior has not changed much during and after the training, which was an expected behavior since they already know the basic salsa steps. Most of the improvements in the *regular dancers* performance seem to be attributed to the familiarization of users with the system.

Another remarkable notice, as shown in Fig. 8, is that the standard deviation (std) of the MMF measurements for non-dancer are much larger than those for the *regular dancers* regarding guidance. This indicates that the movements and guidance skills varied a lot within *non dancers*. This can be justified by the fact that *non dancers*, as non-experienced in salsa moves, have a different sensibility and synchronization of their body movements to the music. In contrast, *regular dancers*' movements have smaller variation since they have prior experience in leading a salsa dance scenario, and control better their body movements and gestures.

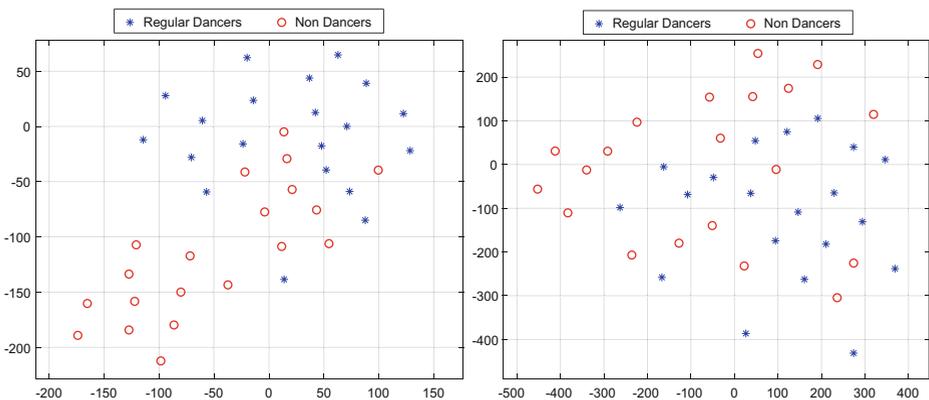


**Fig. 7** The mean values over all exercises for the MMF-derived feature measurements  $\mu_j$  is shown on the left. On the right top is the mean values of the second exercise, and at the bottom is the mean values of the last exercise



**Fig. 8** The standard deviation of the MMF-derived feature measurements  $\mu_j$  for the *regular dancers* (blue) and *non dancers* (red)

To visualize the differences between the two classes, we portray the high dimensional arrays that represent the performance of each participant into a 2-dimensional space using the t-Distributed Stochastic Neighbor Embedding (t-SNE) [45]. We used t-SNE for dimensionality reduction, rather than the Multi-Dimensional Scaling (MDS) [36], since it is particularly well suited for the visualization of high-dimensional datasets such as ours. Figure 9 shows the 2D embedding of the two classes, *regular dancers* and *non dancers*. The most significant observation is that the two classes can be separated at the beginning of the



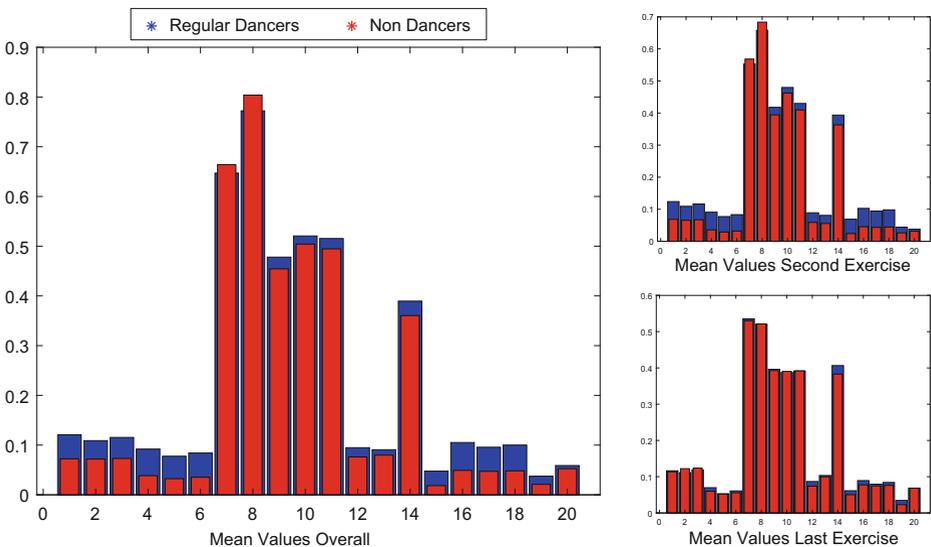
**Fig. 9** t-SNE dimension reduction of the MMFs for the second exercise (left) and the last exercise (right)

training, but as the performers gain more experiences and training (e.g., in the last exercise), the two classes are mixed up. Assuming that *regular dancers* have good learning skills, this is a good indication that the overall guidance and rhythm profiles of the users have been improved, and are converging toward a more homogeneous one, thus validating the learning effect of our training.

### 6.2 LMA study

To evaluate the learning skills and the improvement of the performers in terms of the style/LMA analysis, we performed the following analysis. For each performer, and different learning stages (exercises), we extracted the one-dimensional arrays (the windows of LMA-derived features measurements using a sliding window), and represent each performance by the mean value of all these local spatiotemporal descriptors. In this direction, we aim to conclude to some useful information, e.g., study how the learning skills for each performer or group of performers change over time and observe the differences in the style for users with different dance experiences.

During our motion analysis, we noticed some important observations regarding the two classes (*regular dancers* Vs. *non dancers*). First, the mean of the LMA feature measurements for the *regular dancers* have larger values than those of the *non dancers*, especially at the early exercises of the exercise. That means that the users with regular dance experience put more effort to perform the task than the *non dancers*. Figure 10 shows the mean values of the LMA-derived feature measurements  $\phi_j$  of the performers for the two classes for all exercises (left), and the mean values of the performers on the right for the second (top) and the last exercises (bottom). It can be clearly observed that the two classes are easily distinguishable for the early exercises, but as we move forward to the latest exercises, these differences are getting smaller. Another important observation is that the standard deviation

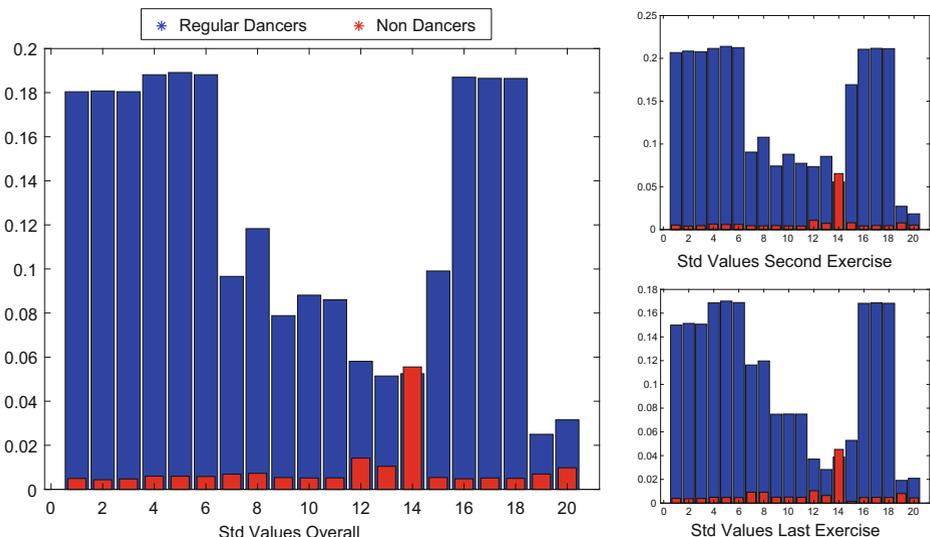


**Fig. 10** The mean values for the LMA-derived feature measurements  $\phi_j$ . The mean values of the performers for all exercises is shown on the left. On the right top is the mean values of the second exercise, and at the bottom is the mean values of the last exercise

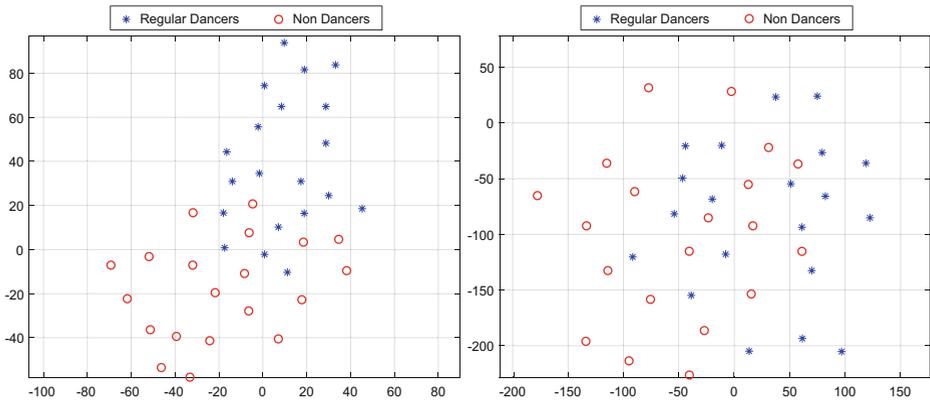
(std) of the LMA feature measurements for the *regular dancers* is larger than those of the *non dancers* (refer to Fig. 11). This indicates that the movements of the *regular dancers* are more variant, while the *non dancers* movements are more compact. One should expect that professional dancers will be more consistent in their movements, and *non dancers* will have larger variation. However, there are many reasons for this peculiarity in the dancers' motion measurements. Unlike non-dancers who put the minimum required effort to do the experiment, and only perform the absolutely basic steps required by the VR application, dancers tend to put more effort on their movements, since each dancer has its own individual dancing style/improvisation/accent, that may be different from others, resulting in larger variation in their LMA feature measurements. In addition, since the dancers who participated in our experiments have no experience with VR environments, while the non-dancers have, we believe that previous VR experiences have a substantial impact on the performance of the participants.

Also, we have studied the effect of our system on the personal style of the dancers. As illustrated in Fig. 11, the std of the LMA features for the *non dancers* remains unchanged over time, since *non dancers* usually oversimplify their movements to only those steps that are required by the system. In contrast, the std of the LMA features for the *regular dancers* seems to converge in later exercises, isolating their personal style, the stylistic nuance of their movement, and their improvisation; std in the last exercise has declined by 20% compared to the second exercise. This indicates that, in a similar way to the case of real teachers, users are getting familiar with the VR environment and accumulate the style of the system (teachers).

Similarly to Section 6.1, we visualize the differences between the two classes, using t-SNE. Figure 12 illustrates the 2D embedding of the two classes, *regular dancers* and *non dancers*. It can be observed that the two classes can be separated, at least for the early exercises, but as users become more familiar with the VR environment, and its tasks, they



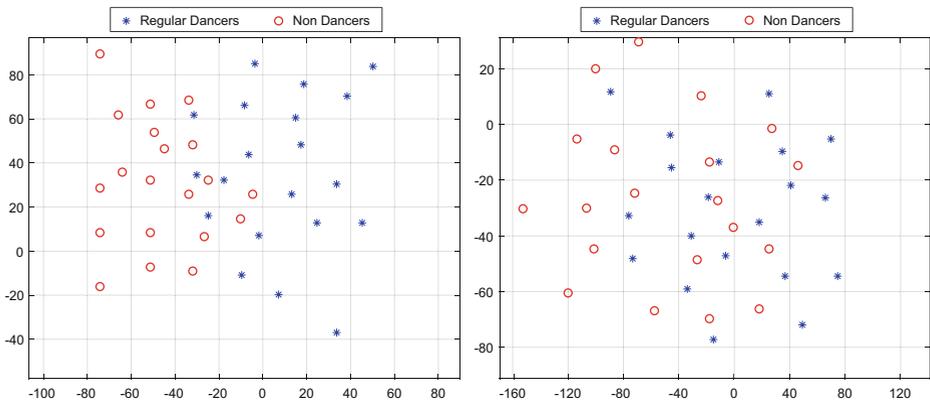
**Fig. 11** The average standard deviation of the LMA features for the *regular dancers* (blue) and *non dancers* (red) over all exercises (left image). On the right, top image, is the standard deviation of the LMA features for the second exercise, and at the bottom, for the last exercise



**Fig. 12** The 2D embedding of the two classes using the mean LMA-derived arrays for the second (left) and final exercises (right). It can be observed that, for the early exercises of the experiment, the two classes can be separated based on the performers learning skills, but as we move forward to later exercises, the dancing skills of the *non dancers* are converging to those of the *regular dancers*

are mixed (it is more difficult to be separated). Figure 12 shows the 2D embedding for the second (left) and last exercises (right).

In addition to the LMA analysis, we evaluated the stylistic behavior (signature) of the movement of the participants, and how it evolves over time. More specifically, we extracted the LMA-derived arrays for all the performances, and similarly to Aristidou et al. [5], we represented each performance by the distribution of its LMA-derived arrays. We positioned all these arrays into a  $d$ -dimensional space ( $d = 10$ ), using Multi-Dimensional Scaling [36], clustered them in this space using  $K$ -means ( $K = 100$ ), and then computed the normalized histogram of the frequency of these arrays for each performance (similar to the concept of bag-of-words). Thus, each performance is succinctly characterized by the distribution of its LMA-derived arrays; stylistically similar performances have a resemblance distribution, while stylistically dissimilar performances have a different distribution. The distance



**Fig. 13** The 2D embedding using t-SNE, when the two classes are represented by the distribution of their LMA-derived arrays [5]. Again, it can be observed that the two classes can be partly separated in the early exercises (left), but as we move forward to the latest exercises (right), the two classes are mixed together. This indicates that our dance VR application helps the participants to improve their dancing stylistic behavior

between these LMA-derived arrays was computed using the Earth Mover's Distance (EMD) metric [35]; note that, EMD performs better than the Euclidean distance, or the Pearson Correlation Coefficient that was originally used in [3]. Again, we applied t-SNE for dimensionality reduction, and the 2D embedding of the two classes for the second and last exercises is illustrated in Fig. 13. Again, it can be observed that the two classes are separable in the early exercises, but tend to converge and be inseparable at the latest exercises.

## 7 Conclusions and future work

We have designed a VR application that simulates salsa dance practice. In our VR environment, the user interacts with a virtual partner via hand to hand contact using controllers and can control the salsa dance pattern's transitions similarly to real dance situation. A six points skeleton of the user is motion captured to provide enough data for analyzing the enforced performance. As a validation, we made an experiment that consists of a series of 8 exercises with different tempos in which the user leads the movements of the virtual partner with specific gestures at given times, as in real life salsa scenarios. We acquired the motion of 40 participants divided into two groups of people from different dance experience, the *non dancers* and the *regular dancers*. The performance was evaluated using MMF and LMA features, which show a clear difference before and after training using our dance VR environment, and significance to classify people upon their learning profile. The results demonstrate an overall improvement of the dance skills for the *non dancers*, and a more uniform profile, that is converging towards the *regular dancers* profile after training.

Our method has some limitations. First, the gesture and timing required to trigger the dance pattern transition felt not enough natural for some users, as there is more complex mechanical interaction to be taken into account. Secondly, the learning duration of our training was too small for some users that shows an understanding of the VR technology. By having longer learning sessions, we expect that the users will feel more comfortable and familiar with the application. In future work, we aim to extend the learning study for a more extended period, e.g., one month with two training sessions per week, to evaluate a more significant impact in terms of performance improvement. Moreover, the diversity of users dance profile was quite broad, and thus, it was challenging to come up with definite conclusions. For example, some of the *non dancers* participants have some minor dance experience or extensive experience with virtual reality applications, and that was not taken into account in our analysis and classification. We want to investigate a more extensive diversity of dancers, that could be categorized based on their experience, e.g., expert dancers, regular dancers, amateur dancers, and non-dancers. Other information, such as previous experiences with virtual reality platforms and applications, age, gender, will also be taken into consideration. Finally, for future work, we also foresee to provide many real-time hints, such as audio clues or the presence of a virtual teacher, to help users to assimilate the given tasks better and improve their skills. The mechanical interaction with the virtual partner can be improved with a more complex vibration-feedback system. From the two motion used in this study, more salsa movements can be investigated as turns and spins. Finally, we look forward to investigate the remaining criteria, as reported in [39], which are more challenging.

**Acknowledgements** This work is co-financed by the European project MINGEI. It has also been partly supported by the project that has received funding from the European Union's Horizon 2020 Research and

Innovation Programme under Grant Agreement No 739578 (RISE-Call: H2020-WIDESPREAD -01-2016-2017- TeamingPhase2) and the Government of the Republic of Cyprus through the Directorate General for European Programmes, Coordination and Development.

## References

- Alborno P, Piana S, Mancini M, Niewiadomski R, Volpe G, Camurri A (2016) Analysis of intrapersonal synchronization in full-body movements displaying different expressive qualities. In: Proceedings of the international working conference on advanced visual interfaces, AVI '16. ACM, New York, pp 136–143. <https://doi.org/10.1145/2909132.2909262>
- Alexiadis DS, Kelly P, Daras P, O'Connor NE, Boubekeur T, Moussa MB (2011) Evaluating a dancer's performance using kinect-based skeleton tracking. In: Proceedings of the 19th ACM international conference on multimedia - MM '11, p 659. <https://doi.org/10.1145/2072298.2072412>. <https://core.ac.uk/download/pdf/11310464.pdf>. <http://dl.acm.org/citation.cfm?doid=2072298.2072412>
- Aristidou A, Charalambous P, Chrysanthou Y (2015) Emotion analysis and classification: understanding the performers' emotions using the LMA entities. *Comput Graph Forum* 34(6):262–276. <https://doi.org/10.1111/cgf.12598>
- Aristidou A, Chrysanthou Y (2013) Motion indexing of different emotional states using lma components. In: SIGGRAPH Asia 2013 technical briefs, SA '13. ACM, New York, pp 21:1–21:4. <https://doi.org/10.1145/2542355.2542381>
- Aristidou A, Cohen-Or D, Hodgins JK, Chrysanthou Y, Shamir A (2018) Deep motifs and motion signatures. *ACM Trans Graph* 37(6):187:1–187:13. <https://doi.org/10.1145/3272127.3275038>
- Aristidou A, Stavrakis E, Charalambous P, Chrysanthou Y, Loizidou-Himona S (2015) Folk dance evaluation using laban movement analysis. *J Comput Cult Herit* 8(4):20:1–20:19. <https://doi.org/10.1145/2755566>
- Aristidou A, Stavrakis E, Papaefthimiou M, Papagiannakis G, Chrysanthou Y (2018) Style-based motion analysis for dance composition. *Vis Comput* 34(12):1725–1737. <https://doi.org/10.1007/s00371-017-1452-z>
- Aristidou A, Zeng Q, Stavrakis E, Yin K, or Daniel C, Chrysanthou Y, Chen B (2017) Emotion control of unstructured dance movements. In: Proceedings of the ACM SIGGRAPH / eurographics symposium on computer animation, SCA '17. ACM, New York, pp 9:1–9:10. <https://doi.org/10.1145/3099564.3099566>
- Bastanfard A, Takahashi H, Nakajima M (2004) Toward e-appearance of human face and hair by age, expression and rejuvenation. In: Proceedings of the 2004 international conference on cyberworlds, CW '04. IEEE Computer Society, USA, pp 306–311. <https://doi.org/10.1109/CW.2004.65>
- Bellini R, Kleiman Y, Cohen-Or D (2018) Dance to the beat: enhancing dancing performance in video. *Computational Visual Media* 4:197–208. <https://doi.org/10.1007/s41095-018-0115-y>
- Chan JC, Leung H, Tang JK, Komura T (2011) A virtual reality dance training system using motion capture technology. *IEEE Trans Learn Technol* 4(2):187–195. <https://doi.org/10.1109/TLT.2010.27>
- Cuykendall S, Soutar-Rau E, Schiphorst T (2016) POEME: a poetry engine powered by your movement. In: Proceedings of the TEI '16: tenth international conference on tangible, embedded, and embodied interaction, pp 635–640. <https://doi.org/10.1145/2839462.2856339>. <http://dl.acm.org/citation.cfm?doid=2839462.2856339>. <http://doi.acm.org/10.1145/2839462.2856339>
- Cuykendall S, Soutar-Rau E, Schiphorst T, Dipaola S (2016) If words could dance: moving from body to data through kinesthetic evaluation. In: Proceedings of the 2016 ACM conference on designing interactive systems - DIS '16, pp 234–238. <https://doi.org/10.1145/2901790.2901822>. <http://dl.acm.org/citation.cfm?doid=2901790.2901822>
- DanceVirtual: (2018). <http://salsa.dance-virtual.com/>
- Deng L, Leung H, Gu N, Yang Y (2011) Real-time mocap dance recognition for an interactive dancing game. In: Computer animation and virtual worlds, vol 22, pp 229–237. <https://doi.org/10.1002/cav.397>
- dos Santos A, Yacef K, Martinez-Maldonado R (2017) Let's dance: how to build a user model for dance students using wearable technology. In: Proceedings of the 25th conference on user modeling, adaptation and personalization. ACM Press, New York, pp 183–191. <https://doi.org/10.1145/3079628.3079673>. <http://dl.acm.org/citation.cfm?doid=3079628.3079673>
- Forbes K, Fiume E (2005) An efficient search algorithm for motion data using weighted PCA. In: Proceedings of the ACM SIGGRAPH/eurographics symposium on computer animation, SCA '05, pp 67–76

18. Fourati N, Pelachaud C (2015) Relevant body cues for the classification of emotional body expression in daily actions. In: Proceedings of the sixth international conference on affective computing and intelligent interaction (ACII2015), pp 267–273
19. Hajarian M, Bastanfard A, Mohammadzadeh J, Khalilian M (2019) A personalized gamification method for increasing user engagement in social networks. *Social Network Analysis and Mining* 9:1–14
20. Ho ESL, Chan JCP, Komura T, Leung H (2013) Interactive partner control in close interactions for real-time applications. *ACM Transactions on Multimedia Computing, Communications, and Applications* 9(3):1–19. <https://doi.org/10.1145/2487268.2487274>
21. Kim TH, Park SI, Shin SY (2003) Rhythmic-motion synthesis based on motion-beat analysis. *ACM Trans Graph* 22(3):392–401. <https://doi.org/10.1145/882262.882283>
22. Kyan M, Sun G, Li H, Zhong L, Muneesawang P, Dong N, Elder B, Guan L (2015) An approach to ballet dance training through MS Kinect and visualization in a CAVE virtual reality environment. *ACM Transactions on Intelligent Systems and Technology* 6(2):1–37. <https://doi.org/10.1145/2735951>. <http://dl.acm.org/citation.cfm?doid=2753829.2735951>
23. Laban R, Ullmann L (2011) *The mastery of movement*, 4th edn. Dance Books Ltd, Binsted
24. Lee J, Chai J, Reitsma PSA, Hodgins JK, Pollard NS (2002) Interactive control of avatars animated with human motion data. In: Proceedings of the 29th annual conference on computer graphics and interactive techniques - SIGGRAPH '02, p 491. <https://doi.org/10.1145/566570.566607>. <http://portal.acm.org/citation.cfm?doid=566570.566607>
25. Merom D, Cumming R, Mathieu E, Anstey KJ, Rissel C, Simpson JM, Morton RL, Cerin E, Sherrington C, Lord SR (2013) Can social dancing prevent falls in older adults? A protocol of the Dance, Aging, Cognition, Economics (DAnCE) fall prevention randomised controlled trial. *BMC Public Health* 13(1):477. <https://doi.org/10.1186/1471-2458-13-477>. <http://www.ncbi.nlm.nih.gov/pubmed/23675705>. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3691670>. <http://bmcpubmedcentral.biomedcentral.com/articles/10.1186/1471-2458-13-477>
26. Merom D, Grunseit A, Eramudugolla R, Jefferis B, Mcneill J, Anstey KJ (2016) Cognitive benefits of social dancing and walking in old age: the dancing mind randomized controlled trial. *Frontiers in Aging Neuroscience* 8:26. <https://doi.org/10.3389/fnagi.2016.00026>. <http://www.ncbi.nlm.nih.gov/pubmed/26941640>. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4761858>
27. Merom D, Mathieu E, Cerin E, Morton RL, Simpson JM, Rissel C, Anstey KJ, Sherrington C, Lord SR, Cumming RG (2016) Social dancing and incidence of falls in older adults: a cluster randomised controlled trial. *PLOS Medicine* 13(8):e1002112. <https://doi.org/10.1371/journal.pmed.1002112>. <http://www.ncbi.nlm.nih.gov/pubmed/27575534>. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5004860>
28. Mousas C (2018) Performance-driven dance motion control of a virtual partner character. In: 25th IEEE conference on virtual reality and 3D user interfaces, VR 2018 - Proceedings, pp 57–64. <https://doi.org/10.1109/VR.2018.8446498>
29. Ozcimder K, Dey B, Lazier RJ, Trueman D, Leonard NE (2016) Investigating group behavior in dance: an evolutionary dynamics approach. In: Proceedings of the American control conference, vol 2016-July. IEEE, pp 6465–6470. <https://doi.org/10.1109/ACC.2016.7526687>. <http://ieeexplore.ieee.org/document/7526687/>
30. Paez Granados DF, Kinugawa J, Hirata Y, Kosuge K (2016) Guiding human motions in physical human – robot interaction through COM motion control of a dance teaching robot. In: IEEE international conference on humanoid robots (humanoids). IEEE, pp 279–285. <https://doi.org/10.1109/HUMANOIDS.2016.7803289>. <http://ieeexplore.ieee.org/document/7803289/>
31. Patil P, Kumar K, Gaud N, Semwal VB (2019) Clinical human gait classification: extreme learning machine approach. In: 2019 1st international conference on advances in science, engineering and robotics technology (ICASERT), pp 1–6
32. Piana S (2016) Movement fluidity analysis based on performance and perception. In: CHI extended abstracts on human factors in computing systems, pp 1629–1636. <https://doi.org/10.1145/2851581.2892478>. <http://dl.acm.org/citation.cfm?doid=2851581.2892478>
33. Powers RSU (2020) Brief histories of social dance. [https://socialdance.stanford.edu/Syllabi/dance\\_histories.htm](https://socialdance.stanford.edu/Syllabi/dance_histories.htm)
34. Raheb KE, Katifori V, Rc A (2016) HCI challenges in dance education. *ICST Transactions on Ambient Systems* 3(9):6–10. <https://doi.org/10.4108/eai.23-8-2016.151642>. <http://eudl.eu/doi/10.4108/eai.23-8-2016.151642>
35. Rubner Y, Tomasi C, Guibas LJ (2000) The earth mover's distance as a metric for image retrieval. *Int J Comput Vision* 40(2):99–121
36. Seber GAF (2008) *Multivariate observations*. Wiley, New York

37. Semwal VB, Mondal K, Nandi GC (2015) Robust and accurate feature selection for humanoid push recovery and classification: deep learning approach. *Neural Comput Applic* 28:565–574
38. Senecal S, Cuel L, Aristidou A, Magnenat-Thalmann N (2016) Continuous body emotion recognition system during theater performances. *Comput Animat Virtual Worlds* 27(3-4):311–320. <https://doi.org/10.1002/cav.1714>
39. Senecal S, Nijdam N, Thalmann N (2019) Classification of salsa dance level using music and interaction based motion features. In: GRAPP 2019 - international conference on computer graphics theory and applications, pp 100–109. <https://doi.org/10.5220/0007399701000109>
40. Senecal S, Nijdam NA, Thalmann NM (2018) Motion analysis and classification of salsa dance using music-related motion features. In: Proceedings of the 11th annual international conference on motion, interaction, and games, MIG '18. ACM, New York, pp 11:1–11:10. <https://doi.org/10.1145/3274247.3274514>
41. Shiratori T, Nakazawa A, Ikeuchi K (2006) Dancing-to-music character animation. *Computer Graphics Forum* 25(3):449–458. <https://doi.org/10.1111/j.1467-8659.2006.00964.x>
42. Shum HP, Komura T, Shiraishi M, Yamazaki S (2008) Interaction patches for multi-character animation. *ACM Transactions on Graphics (TOG)* 27(5):114
43. Tang T, Jia J, Mao H (2018) Dance with melody: an lstm-autoencoder approach to music-oriented dance synthesis. In: Proceedings of the 26th ACM international conference on multimedia, MM '18. ACM, New York, pp 1598–1606. <https://doi.org/10.1145/3240508.3240526>
44. Theodorou L, Healey PGT, Smeraldi F (2016) Exploring audience behaviour during contemporary dance performances. In: Proceedings of the 3rd international symposium on movement and computing - MOCO '16. ACM Press, New York, pp 1–7. <https://doi.org/10.1145/2948910.2948928>. <http://dl.acm.org/citation.cfm?doid=2948910.2948928>
45. van der Maaten L, Hinton G (2008) Visualizing data using t-SNE. *J Mach Learn Res* 9:2579–2605
46. Whyatt CP, Torres EB (2017) The social-dance. In: Proceedings of the 4th international conference on movement computing - MOCO '17. ACM Press, New York, pp 1–8. <https://doi.org/10.1145/3077981.3078055>. <http://dl.acm.org/citation.cfm?doid=3077981.3078055>
47. Won J, Lee K, O'Sullivan C, Hodgins JK, Lee J (2014) Generating and ranking diverse multi-character interactions. *ACM Transactions on Graphics (TOG)* 33(6):219

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Simon Senecal<sup>1</sup>  · Niels A. Nijdam<sup>1</sup> · Andreas Aristidou<sup>2,3</sup> · Nadia Magnenat-Thalmann<sup>1</sup>

Niels A. Nijdam  
nijdam@unige.ch

Andreas Aristidou  
a.aristidou@ieee.org

Nadia Magnenat-Thalmann  
thalmann@miralab.ch

<sup>1</sup> University of Geneva, Geneva, Switzerland

<sup>2</sup> University of Cyprus, Nicosia, Cyprus

<sup>3</sup> RISE Research Center, Nicosia, Cyprus