Motion analysis and classification of salsa dance using music-related motion features

Simon Senecal
University of Geneva
Geneva, Switzerland
senecal@miralab.ch

Niels A. Nijdam
University of Geneva
Geneva, Switzerland
niels.nijdam@unige.ch

Nadia Magnenat Thalmann
University of Geneva
Geneva, Switzerland
thalmann@miralab.ch

ABSTRACT

Learning couple dance such as Salsa is a challenge for modern human as it requires to assimilate and understand correctly all the required parameters. In this paper, we propose a set of music-related motion features (MMF) allowing to describe, analyse and classify salsa dancer couple in their respective learning state (beginner, intermediate and expert). These dance qualities have been proposed from a systematic review of papers cross linked with interviews from teacher and professionals in the field of social dance. We investigated how to extract these MMF from musical data and 3D movements of dancers in order to propose a new algorithm to compute them. For the presented study, a motion capture database (SALSA) has been recorded of 26 different couples with varying skill levels dancing on 10 different tempos (260 clips). Each recorded clips contains a basic steps sequence and an extended improvisation sequence during two minutes in total at 120 frame per second. We finally use our proposed algorithm to analyse and classify these 26 couples in three learning levels, which validates some proposed music-related motion features and give insights on others.

CCS CONCEPTS

• **Computing methodologies → Motion capture**; **Motion processing**; *Machine learning approaches*; *Model development and analysis*; Animation;

KEYWORDS

motion analysis, couple dance, motion features

ACM Reference Format:
Simon Senecal, Niels A. Nijdam, and Nadia Magnenat Thalmann. 2018. Motion analysis and classification of salsa dance using music-related motion features. In *MIG '18: Motion, Interaction and Games (MIG '18), November 8–10, 2018, Limassol, Cyprus.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3274247.3274514

1 INTRODUCTION

Social dances have been existing for centuries and can be found, in different forms (duo or otherwise variable group sizes), in almost every ethnic group and culture and pertain to a social and/or religious context[Powers [n. d.]]. The modern social couple dance, such as Latin dances (Salsa, Cha-cha-chá, Bachata) or Western dances (Swing, Jives, West Coast Swing), is gaining more recognition and is developing rapidly[McMains 2015], which is highlighted by the increasing number of international congresses [Congres1 [n. d.]; Congres2 [n. d.]], shows and dance schools, as well as their inclusion as an Olympic category [IDF [n. d.]].

Aside from social sciences and related fields, the study of social couple dances is also interesting in the fields of bio-mechanics, human robot interactions (HRI) and human computer interaction (HCI), as both partners have to dance in unison in an almost mechanical and predominantly cognitive connection. Furthermore, the use of music is very important on the whole synergy of the dance performance, as it dictates for example the rhythm and musicality, which influences the 'way' a dance is carried out.

Social couple dances are a human-human interaction where full-body movements are coordinated and fine-tuned upon each other, and in most cases in adequacy with music. During the dance, one person is leading the dance and the other follows the movements by responding on impulsion, making the relation between dancer very important for the dance. The vastly dynamic and interactive situations of social couple dances brings 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 analyse this intricate and interdependent set of parameters.

To Learn a couple dance, such as Salsa, is a challenge for the modern human as it requires to learn all the different mechanico-cognitivo-interactive parameters from a teacher in mainly collective classes only, which is less effective to spot errors on individual students. In addition other parameters can influence the teacher's teaching, such as personal feelings, fatigue and social pressure. Also at higher experience levels, when the teacher and the student have similar level, the student can reconsider or oppose the advises of his teacher. The status of an expert in social dance can be a source of confusion as there is no universal recognised diplomas but rather a public recognition of skills by pairs. In both cases the learning process can be less effective, halted or reconsidered depending on the relationship between student and teacher. Another challenge during learning is the need to practice with another partner and/or taking some dance classes to progress, meaning they lack access to the facilities and may not have the partner to practice with (either by lack of dance partners or due to personal time schedules). A proposed solution to all these previous challenges would be to set up an artificial intelligence based virtual coach that can analyse the movement of dancer student and provide a positive feedback to

MIG '18, November 8–10, 2018, Limassol, Cyprus

ACM ISBN 978-1-4503-6015-9/18/11…$15.00
https://doi.org/10.1145/3274247.3274514

improve their skills. Therefore an automatic analysis of the different learning states should be developed. In order to automatically analyse and classify a social dancer upon his learning level, the computation of measurable dance motion feature able to distinguish between learning level is required.

Some studies started to identify the rhythm constituency or the performance of solo performers in different dances but do not consider couples as a proper interactive dancing entity. A possible explanation for the absence of extended studies in this direction might be due to the problematic and relative costly acquisition of the data involved (e.g. capture both dancers at a time, requiring an 'expensive' motion capture system).

In this paper, specific music-related motion features (MMF) for dance are proposed and investigated using a database of 3D movements of dancers, synchronised with music. A motion capture system was utilised to acquire first the 3D data (point clouds derived from the motion captured reflective markers) of both dancers, such that we have a certain number of points in 3D space along the time, allowing a precise measure of the dancer's body movement related to each other over the course of time. Then we compute separately the music main feature: the beat (the one used by dancer as a reference of their movement) and its temporal location. As a last step to the process, it is vital to retain synchronisation of audio and 3D motion during the performance recording phase. The proposed motion features are finally tested if they can help to classify the dancer's learning level.

2 STATE OF THE ART

Motion in dance has been investigated through multiple scientific studies. Health studies show the benefit of social dances for balance and cognition for elderly [Merom et al. 2013, 2016a,b]. Moreover, the interactive aspect has been touched upon by the HRI domain, where through inertial measurement unit (IMU) detection the user's movements were transcribed into an intermediary data set to generate poetry [Cuykendall et al. 2016a,b]. Another example is the use of robots acquiring the knowledge and skills to perform a dance [Paez Granados et al. 2016]. 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 [Ozcimder et al. 2016; Whyatt and Torres 2017], even with the audience, that take part in the performing process [Theodorou et al. 2016].

To take into account the uncertainty of observations, the judgement process by a human coach is based on experience, historical knowledge and making assumptions about the state, intentions and methods of the students. It is at this point that bias can appear in decision-making: "fatigue, stress, stakes, prejudices, errors, beliefs, intuitions, the tendency to partiality through ignorance, similarity decision, random correlation belief, great influence of the first time, finding before the evidence, contradictions with unfulfilled beliefs, unjustified emphasis of information interpreted as more egregious. "[Hicks et al. 2004]. These human deficiencies, mostly due to infobesity(i.e. information overload), can be corrected by a virtual coach.

Extracting the motion features from continuous movement is a key element for describing, evaluating and understanding dance and movement in general. The use of Laban motion analysis (LMA)-based motion retrieval and indexing for motion features is a solution that has proved to work well in different situations[Aristidou et al. 2014], and is therefore ideal to be used as a base to build a machine learning classifier, as demonstrated for theatre emotional expression [Senecal et al. 2016] or evaluating the performer's emotion using LMA features [Aristidou et al. 2015]. Some studies focused on a specific motion feature, for example the fluidity of the movement is an important dance parameter investigated in [Piana 2016]. In this particular study, it is proposed to see how fluidity can help describing and classifying dance performance. Through interdisciplinary research including bio-mechanic, psychology and experiments with choreographers and dancers, they propose a definition that takes specifically the minimum energy dissipation when looking at the human body as a kinematic chain. Another work [Alborno et al. 2016], elaborated upon the expressive qualities, such as rigidity, fluidity and impulsiveness, to investigate intra-personal synchronisation for full body movement classification. More recently, several motion features for social dance (Forro) have been proposed, taking the music component into account [dos Santos et al. 2017]. These proposed features are computed with the user's motion data on one hand and the music data (mainly the beat per minutes or BPM) on the other hand. First the "*Rhythm 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 interesting insights on characterising social dancing but the weakness is the accuracy due to the sensor (a simplified IMU for the full body, representing a single point in space).

In comparison to the previous mentioned approaches, our work takes two persons dancing together and defines this as the input entity for analysis, indexing and classification. The work is further set in the context of Salsa social dance. Prior to establishing the input entity, we first reflected upon the most relevant motion features extraction method from literature (BPM rhythm and consistency from [dos Santos et al. 2017]) and reviewed and discussed these through interviews and focus groups of experts in dance (teachers and choreographers). As for the acquisition of the data, a motion capture high precision system was utilised to ensure a maximum accuracy on the movements. Finally we propose a music-related motion feature from the processing of motion and audio file to classify salsa dance.

3 METHODOLOGY

3.1 Field study on criteria improvement

A field study has been conducted in order to improve the motion features from the literature. This study was conducted in Geneva, a dynamic city for social dancing with an official number of 15 active Latin dance schools and hosting international dance congresses; making it a very important central dance area in Switzerland and also in Europe. Experts in social dances are persons with a high-level of expertise and skills, with a subsequent level of reputation, and/or recognised by pairs to be expert as there is no official diploma or formation for social dances (albeit some private schools, and in some countries they do provide a diploma). We therefore define a person as 'an expert' in social dance if it belongs to one of the

Table 1: Retained criteria definitions and relative importance

Proposed criteria	Definition	Relative importance (1 to 10)
Dancing on the rhythm	Being synchronised with the music's tempo.	10
Lead and Follow (Drive)	Being able to guide / follow his / her partner.	7
Fluidity	Being able to move smoothly on the music.	6
Style and Variation	Adding your own variation to the basic movement.	5
Intention and Sharing	Being able to share the moment and enjoy the dance.	7
Musicality	Using your own dance movement with the music's variation.	3

following definitions: *Jury of international competition, Champion of International championships, Invited dancer in international congress, Director of major dance schools* or *Professor of dance.*

Several professors and directors of the Latin dance schools in Geneva have been contacted and invited for an interview. They have been asked about what would be the criteria to teach or evaluate a dance student and indicate per criteria its importance. In addition, a questionnaire was filled out, with additional annotations on which questions were not clear and/or required further explanations. Initially, the questionnaire contained only the motion feature extracted from the literature and was updated with the feedback of the first expert (extending and improving upon further features), then suggested to the next expert and so on. This led to a final list of six features, listed in table 1, ranked by overall importance. Strong importance means that the criteria is essential for dancing, whereas little importance means to be less important (especially at beginner level).

3.2 Motion features algorithm

Three of the six criteria (Rhythm, Driving and Style) from table 1 are analysed along three axis: *(1)* the motion data itself (3D points), *(2)* the relation between motion data and *(3)* music data, and finally the relation between the two dancers. For each component, different parameters have been extracted and tested.

3.2.1 Rhythm. In order to identify the means of how to measure a dancer's rhythm, a good understanding and description of the dance movement itself is required. In salsa, the first fundamental move to learn is the *mambo step*. Mastering this step is essential in learning the dance as it allows both partners to move in synchronisation with each other and the music. It corresponds to a very basic displacement of the feet in a forward-backward motion with an important weight shift from one leg to the other. It must adhere to the eight regular beats from the salsa music. This step is represented visually in figure 1. The depiction of the *mambo step* is from the perspective of the *lead* dancer, commonly a man, whereas the *follower*, commonly a woman, acts in the opposite way (thus mirrored).

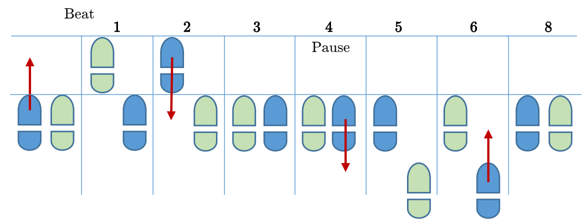

Figure 1: Lead dancer's basic salsa step (Mambo). The red arrow as velocity and the green coloured step for the bodyweight. For the follower dancer, the step is mirrored.

The feet should be moved during a specific beat of the music as highlighted in green and annotated with an red arrow in figure 1. This characteristic is very useful as we can use it to measure the rhythm of a dancer. In an almost mechanical manner the left foot will move to the front and stop on beat *one*. Then, the same left foot will return to its initial position and stop on beat *three*. There is a mirrored behaviour for the right foot on the beat *five* and *seven*. Thus, the dancer 'marks' the beat with alternating feet, while moving in synchronisation with the music. The observation of velocity peaking in relation to the beat is similar for most of the steps in Salsa dance and other social dances. Based on this, we theorise a certain velocity profile that can be expected for the defined basic step and is shown in figure 2.

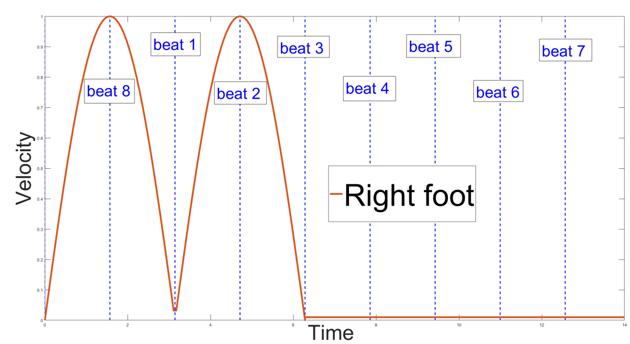

Figure 2: Expected velocity norm over time of the left foot for the basic step. The velocity increases when moving the foot forward and stops on the beat and similar to bring the foot back.

When combining the left and right feet, the theoretical velocity would produce a periodic signal for the full eight beat rythm, illustrated in figure 3.

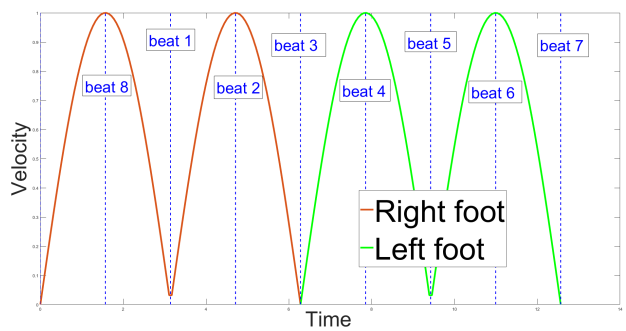

Figure 3: Expected velocity norm over time of both feet when dancing the basic step "mambo". The time window corresponds to 8 beats and will be repeated through the music.

Temporal offset - Rhythm delay. Given the temporal location of the music's beat and the dancer's beat *marking*, we can compute the mean delay of a person by subtracting the temporal location of the detected beat to the nearest theoretical beat location, exemplified in figure 4 as a close-up from figure 3. A perfect dancer would mark the beat exactly on the temporal beat location.

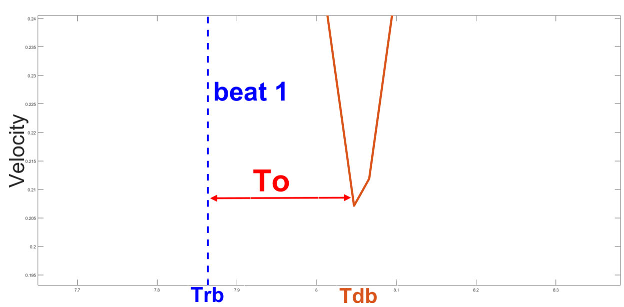

Figure 4: Example for imperfect velocity norm minimum for beat one. The difference between the temporal beat location and the minimum of the velocity norm is the rhythm delay or the temporal offset (*Trb to Tdb*).

The expected velocity of the left foot, figure 4, is set at the minimum when stopping on the beat and increases as the dancer goes to the next one. It is then possible to express a *Temporal offset* (To) of a dancer as the *temporal difference* between the moment the dancer is *marking* the beat (beat number *one* and *three* for the left leg) and the actual musical beat. If Tdb is the temporal location of the dancer's beat and Trb the temporal location of the true musical beat, thus the To, temporal offset for one beat is proposed as:

$$To(t_1) = abs(Tdb - Trb) \tag{1}$$

Note that in the context of couple dance, the beat number *one* is marked by the *lead* dancer with the left leg and by the *follow* dancer with the right leg. It is then possible to apply our proposed algorithm to the *follow* dancer as well, taking the right leg for the Ec computation on the beat *one* and *three*. Once the temporal offset is computed for each beat, we can take the average value and standard deviation as a valuable information corresponding to one song and one couple. The most advanced dancers are expected to have a smaller temporal offset as they are suppose to dance more accurately on the rhythm.

Frequency power spectrum - Rhythm spectrum. Using Fourier transform, on the merged data of both right and left feet, allows us to visualise the main peak of frequencies and resonance frequencies. The main frequencies can be identified as the average tempo to which the dancer is dancing, as shown in figure 5.

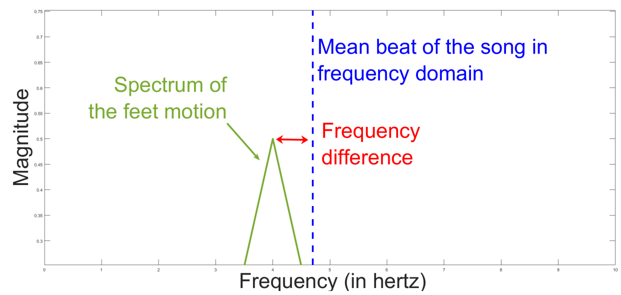

Figure 5: Expected frequency spectrum of feet movement on a specific tempo.

We extract the different spectral location of the main peaks and compare it to the theoretical beat. If Fdb is the temporal location of the dancer's beat and Frb the temporal location of the true musical beat, thus the Fe, Frequency error for one beat is proposed as:

$$Fe(Hz) = abs(Fdb - Frb) \tag{2}$$

It is important to note here that the main frequency at which the people move their feet can be interpreted as the tempo they are dancing on. For our study a conversion from tempo (Beat per Minutes) to frequency (Hz) has been made when using the Fourier transform to the data. It is expected from the most advanced dancer to be precisely on the same main frequency as the theoretical tempo.

3.2.2 Drive - Lead/Follow interaction. In this section, the relationship between the two partners of the dancing couple, the lead and follower, is investigated. Three parameters are proposed: the linear correlation of legs motion, the temporal difference for both dancers when marking the time and the spectral profile correlation.

Linear correlation of legs motion. This parameter corresponds to the correlation between the velocity profile of the follower and the leader during a dance. If $Data\_Man$ is the legs velocity profile of men and $Data\_Woman$ the legs velocity profile of women, then the linear correlation Lc is proposed as the 2d correlation coefficient between these profiles. The linear correlation between the legs motions are expected to be higher for the most advanced dancer as they are suppose to be more synchronised.

Temporal difference man and woman. The beat marking from the follower and leader are not expected to be at the same temporal location, but rather to be around the music's beat (due to natural imprecision and interactive relationship). If Tm is the temporal location of the man's beat and Tw the temporal location of the woman's beat, then the temporal difference Td for one beat is proposed as:

$$Td(t_1) = abs(Tm - Tw) \tag{3}$$

The figure 6 shows a representation of the expected velocity for both feet of both dancers. The two curves are slightly shifted temporally and thus marking the beat differently.

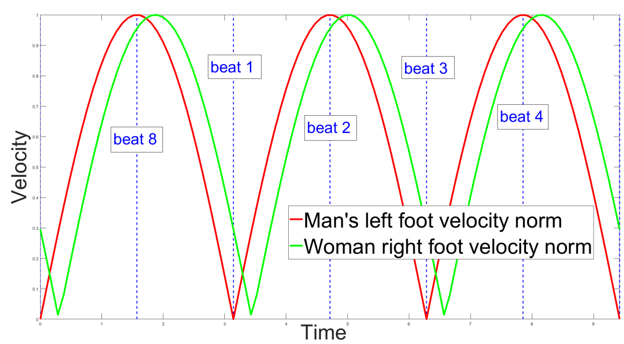

Figure 6: Expected frequency profile of feet movement.

As the man is leading the lady to step on the beat, it is expected to see a smaller time offset between the woman and man for each beat. Even smaller for advanced dancer.

Spectral difference between man and woman. The frequency spectrum from both dancers are expected to be more correlated for advanced dancer. If Fm is the frequential location of the man's main tempo and Fw the frequential location of the woman's tempo, then the Fd, frequential difference is proposed as:

$$Fd(t = beat(1)) = Fm - Fw \tag{4}$$

3.2.3 Style - Variation. Beyond the pure rhythmic features, an important point is the style and variation expressed by the dancers within their dance. Two parameters have been proposed for investigation; The area covered and the quantity of hand movement using the hand to hips distance. These parameters are inspired by the LMA model.

Area covered. To compute the area covered, the integration of both legs velocity profiles is proposed. As V(t) is the velocity norm and T the total time:

$$Area = \int_0^T abs(V(t))dt \tag{5}$$

To ensure data consistency, the same time window should be used for all computations, or the accumulated number divided by the time duration. The more advanced dancers are expected to move more among all tempo variations.

Mean movement quantity hand to hips. This parameter corresponds to the quantity of hand movement, that is considered as a style element in salsa, besides of the leading action of the hand movement. For this point the distance between hand and hips is first computed, then the integral of the derivative (velocity) is taken.

4 EXPERIMENT

Our objective is to measure and validate the proposed music-related motion features as relevant parameter for the classification in the context of couple dance.

4.1 Motion capture data

In order to understand the dancer's motions, we have established a database of motion captures (position and rotation of the body's joints in 3D) of couples dancing the basic salsa moves (SALSA database). A total of twenty six different dancer couples were recorded, of different skill levels (beginner, intermediate and expert), using a set of computer generated music with different beats per minutes (from 100 BPM to 280 BPM, with increments of 20 BPM). The level of dancers have been determined according to their experience: the experts are dance school directors and teachers (Figure 7 shows a couple of experts performing), the beginners started to dance less than six month ago and the intermediate have more than one year and a half of dancing experience (Figure 8 illustrate the recording session with different couples).

The *variable tempo* was determined by performing a BPM study based on commercial salsa songs and music commonly used for teaching in dance schools. The used tempos cover different BPMs from the selection of music which are perceived as the most comfortable to dance to (refined by expert feedback). We ask each couple to perform three basic steps, the *Mambo step*, the *Rumba step* and the *Guapea step*, prior to an improvisation part. A Vicon motion capture system with eight cameras was used for recording (at 120fps), figure 7. The standard template from Vicon for the placement of the markers in each articulation of the body for a total of 52 markers per person was used. For each couple, we asked them to perform a sequence of multiple mambo steps, followed by a sequence of improvisation. Is it important to mention that the numerous occlusion occurring during certain dance moves made the capture (and especially the labelling process) very difficult.

In order to ensure the exact and systematic synchronisation of the music and the captured performance, the music was started simultaneously with the capture through the Vicon software interface.

The result from the capture sessions is a database of 52 people as 26 couples (figure 8), dancing averagely two minutes, representing nearly $26\ couples\ x\ 10\ songs\ x\ 120\ sec\ x\ 120Hz = 3,700,000$ time frames of 104 points. The results have been exported as two fully labelled skeletal entities in C3D formatted files. Each file was segmented for further processing in Matlab into four parts, namely the three main basic step sequences and the improvisation sequence.

4.2 Audio processing

In parallel to the motion capture, it is needed to identify clearly the beat temporal location on the music. Audacity was used to extract the beat temporal location from the audio files in the form of a dual column array containing textual annotations with time-stamps, describing the number of each beat (one to eight) and is shown in figure 9. Thereafter, all the regular beats are marked from the music with the related labels along the duration of the song. Due to the synchronisation, we can directly compare point to point any

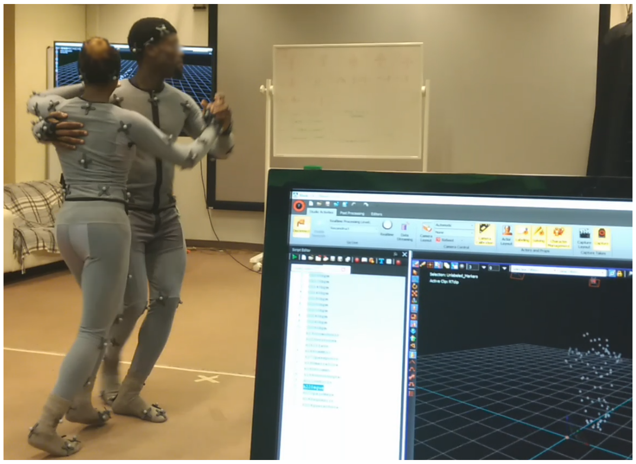

Figure 7: couple performing salsa steps

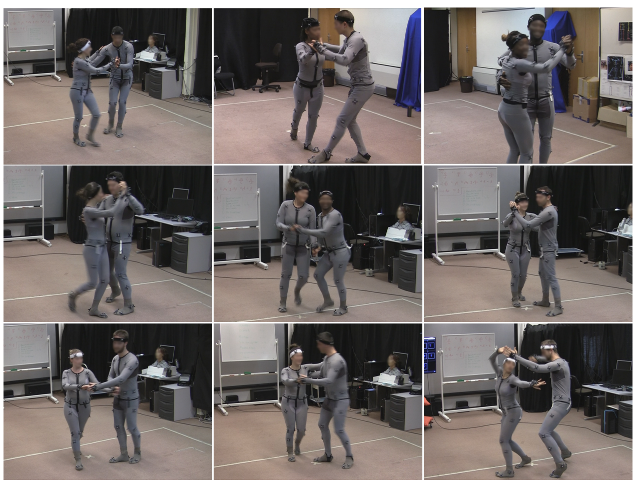

Figure 8: Different couples dancing salsa basic steps

temporal location of a musical event given by the music analysis with the motion data.

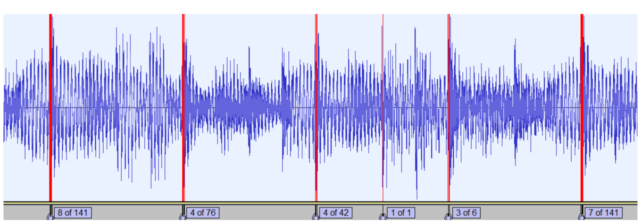

Figure 9: Audio signal (purple) and detected beat (red)

As the audio structure is quite complex, due to the percussive music with a lots of different instruments, we use the fact that the instrument marking the first beat, the "clave", is also the loudest in terms of waveform. Through the use of the clipping detection function it is possible to identify roughly the beat location and then tuned it by hand manually. We improved this step by using a high pass filter in order to get rid of noise artefacts. The hand tuning may result in a loss of accuracy, however the measurement of MMF will be relative and therefore it will suffice for our experiment. The resulting output from this approach is an array containing the timestamps for each beat that is detected. This array is then imported into Matlab for further processing together with the previously mentioned captured data.

4.3 Results

In this section we present the results from the analysis of the two sequences: *(1)*the basic steps and *(2)* the improvisation sequence. The motion features parameters have been computed for each tempo and each couple. Then the median was taken for each group (classified by level of expertise). The median gives an indication of the relative distance among each group, which is even more relevant with higher distances between each group. We also computed the standard deviation of the parameter along each song and also the standard deviation within the groups. Note that the ability to separate dancers into learning level is the objective here, whether if it follows the expectations or not.

4.3.1 Rhythm.

Temporal offset - Rhythm delay. Although the sequence of temporal data for one couple represent the repetition of basic movement *(1)*, an overlay has been made on the first mambo step for visualisation, and comparison with our basic steps model from figure 4. Figure 10 represents the feet of both partners during one mambo step. It is clear that the two curves are indeed placed right before the detected beat and therefore agrees with the theoretical model described in section 3.2. It seems that the follower step is a little delayed as shown in the close up in figure 11, which is explained by the driving parameter (lead and follow action during dance) highlighted by the experts during the survey.

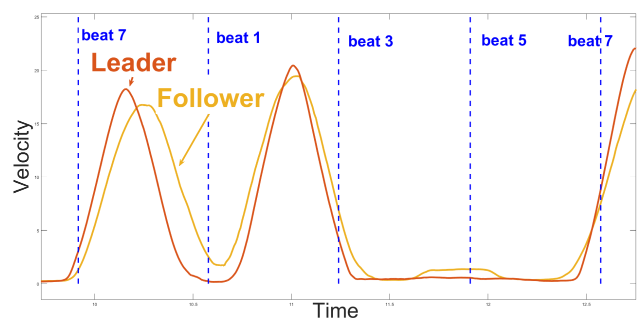

Figure 10: Left foot velocity for the man (Leader - dark red) and right foot velocity for a woman (Follower - light yellow) of a couples during the basic step "Mambo". The blue vertical lines correspond to the detected beat from the audio file.

We then compute the temporal location of minimum velocity (as the dancer marked beat Tdb) and subtract that from the temporal musical beat, Trb to acquire the rhythmic error To of each partner.

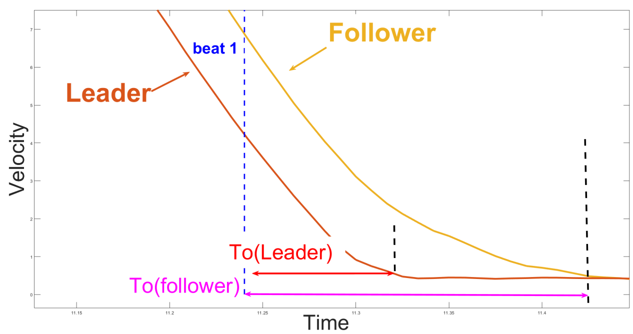

Figure 11: Zoom of the velocities during the basic step "Mambo". The velocity reaches its minimum with a delay regarding the musical beat in blue, and is different for both partners.

Figure 12 shows an overlay of foot velocity from three couples (one from each category expert, intermediate and beginner). Showing a high similarity and thus dancing on the beat. However, their amplitude of movement are not the same. It also shows that the expert couple has a wider velocity span, hence wider movements. The similarity between each velocity of each couple's partner reflects also the presence of this leading / following component during the dance. The beginner couple seems to have huge difficulty to keep up the regularity required by the step. However their movement span is similar to the intermediate couple.

Table 2: Comparison of the median value for each level group of the averaged temporal offset and relative difference between expert and beginner groups

Dance sequence	*Lead dancer*	*Follow dancer*	*Both dancers*
Improvisation	beg > exp > int ; 0.7 %	exp > int > beg ; 0.9 %	exp > beg > int ; 0.2 %
Basic Steps	exp > int > beg ; 1.1 %	exp > int > beg ; 1.7 %	exp > int > exp ; 0.6 %

Table 2 shows the model results with the computation of the mean value for each person, before taking the median among people for each level. The computation has been done with the mean between Lead and follow dancer as well, since a dance style for these dancer could be that the Lead would be in advance and the Follow delayed. The table 2 shows that the mean of the delay parameter does not behave as expected and the small value of the relative difference of the median value indicate that it is not sufficient to distinguish between levels.

In a second time, the standard deviation of the offset along each song for each couple has been computed, and the median has been taken for each level group.

In table 3, a greater difference is shown between expert and beginner, with a median difference around 40% for the basic steps

Table 3: Comparison of the median value for each level group of the standard deviation of the temporal offset and relative difference between expert and beginner groups

Dance sequence	*Lead dancer*	*Follow dancer*	*Both dancers*
Improvisation	beg > exp > int ; 9 %	beg > exp > int ; 6 %	beg > exp > int ; 8 %
Basic Steps	exp > int > beg ; 38 %	exp > int > beg ; 41 %	exp > int > exp ; 38 %

and 8% for the improvisation part. This can be explained by the rhythmic variation introduced by the expert even during the basic steps, as the beginner will be more focused on keeping the rhythm without introducing variations.

Frequency power spectrum - Rhythm spectrum. The observation from the database analysis with spectrum visualisation shows that mostly all dancers present a main frequency peak with some other resonance peaks around, that have frequencies which are a multiple of the main frequency (example of one song for both dancers in figure 13). Taking these resonance frequencies is important as they reflect the use of slower or faster movement in accordance with the rhythm. Within this experiment, the three highest peaks have been considered and compared to the nearest of the song BPM's peaks and its resonance frequencies.

Table 4 shows a substantial difference between expert and beginner, with a median difference around 17% for the male performing the basic steps and 34% for the improvisation part. This can be explained by the rhythmic variation introduce by the expert even during the basic steps, as the beginner will more focused on keeping the rhythm without dancing around.

Table 4: Comparison of the median value for each level group of the standard deviation of the frequency spectrum and relative difference between expert and beginner groups

Dance sequence	*Lead dancer*	*Follow dancer*
Improvisation	exp > int > beg ; 34%	int > beg > exp ; 20%
Basic Steps	beg > int > exp ; 17%	int > exp > beg ; 1%

4.3.2 Drive - Lead/Follow interaction.

Linear correlation of legs motion. The figure 14 is an example of feet motion from the Follower and the Leader during a song. A slight delay is clearly visible, that is then computed across our database.

The parameter is computed for each BPM and visualised in the figure 15 and summarised in table 5. It is showing that the linear correlation among partner is highest during the improvisation part for the expert and smallest during the basic steps part. It is explainable by the highest diversity of advanced dancer movements.

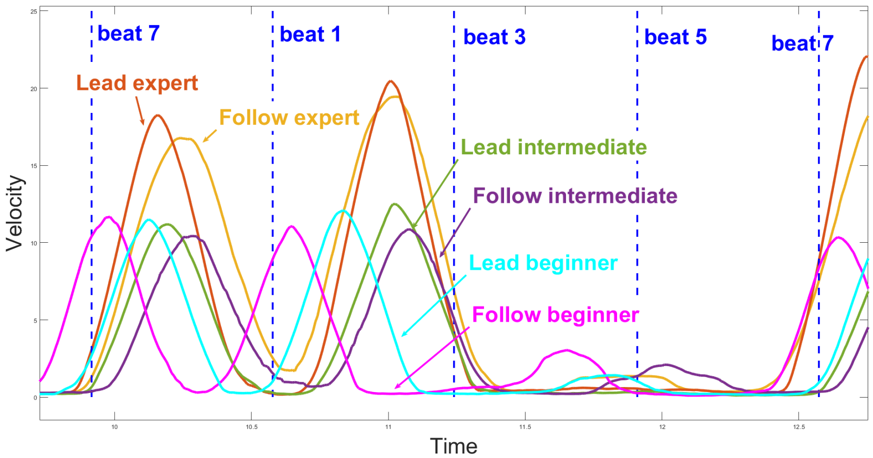

Figure 12: Left foot velocity for the men (Lead - red, green and cyan) and right foot velocity for a women (Follow - yellow, purple and pink) of three couples during the basic step "Mambo". The blue vertical lines corresponds to the detected beat from the audio file.

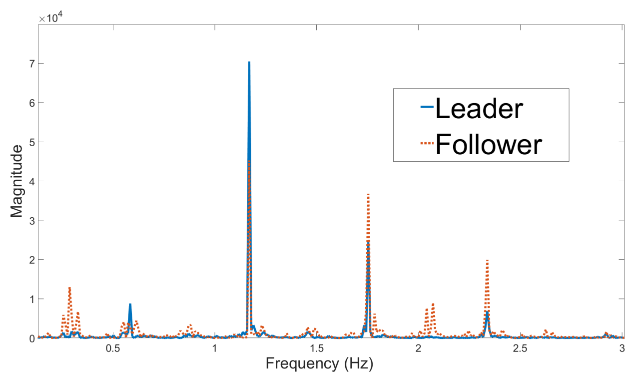

Figure 13: Example of a frequency spectrum from both feet of the Follower and Leader during one dance. A main peak can observed, as well as other peaks that are multiple of the main one. The frequency on the X axis corresponds to different tempos

Table 5: Comparison of the median value for each level group of the linear correlation coefficient and relative difference between expert and beginner groups

Dance sequence	*Median values*
Improvisation	exp > int > beg ; 28%
Basic Steps	beg > int > exp ; 4%

Temporal difference man and woman. The average temporal difference within a couple is shown the table 6.

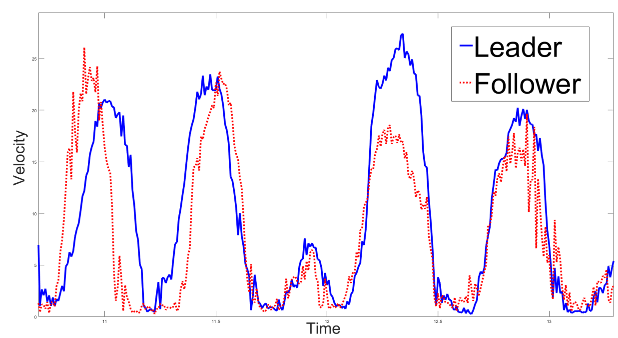

Figure 14: Merged motion of both feet from Follower and Leader during four beats within a song

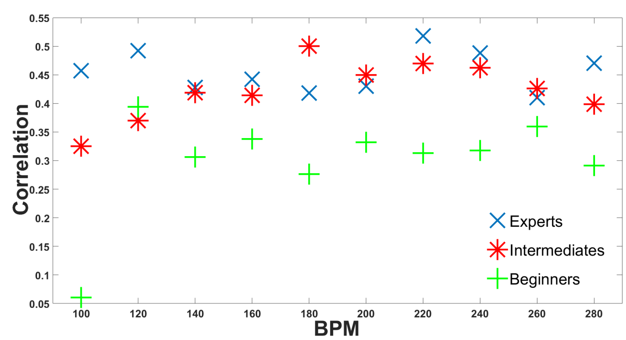

Figure 15: Median value of the linear correlation between the feet motion of the follower and leader across level groups for each song of the Basic sequence

The mean value of this temporal difference is not relevant enough to distinguish between levels of expertise. However the standard deviation shows a significant difference in the improvisation part.

Table 6: Average value of the temporal difference within a dance couple and standard deviation accross all tempos

Dance sequence	*Median values*	*Standard deviation*
Improvisation	exp > int > beg ; 12%	beg > exp > int ; 1%
Basic Steps	exp > beg > int ; 1%	beg > exp > int ; 6%

Spectral difference between man and woman. The correlation between the frequency spectral profile is computed within a couple for each song. A result is visualised in table 7. The relative difference doesn't show a significant difference.

Table 7: Comparison of the correlation of the spectral profile within a couple for each song

Dance sequence	*Median values*
Improvisation	exp > int > beg ; 1%
Basic Steps	beg > int > exp ; 7%

4.3.3 Style - Variation.

Area covered. We compute the integral of the velocity vector of both legs for man and woman separately, with the basic steps sequence and the improvisation sequence. The parameter is computed first for the man during the improvisation sequence and visualised in the figure 16. All results are summarised in the table 8.

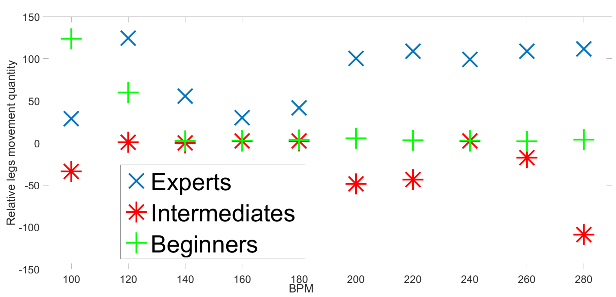

Figure 16: Covered area by both feet for the man during improvisation sequence

Table 8: Comparison of the area covered across each song by group level

Couples	*Man_Median*	*Woman_Median*
Improvisation	exp > beg > int ; 90%	exp > int > beg ; 88%
Basic Steps	exp > int > beg ; 16%	exp > int > beg ; 15%

The area covered shows that the expected order is validated for both man and woman.

Mean distance hand to hips. The computation of the quantity of movement between hand and hips has been computed for the basic steps and summarised in the table 9.

Table 9: Comparison of the mean quantity of movements between hands and hips per level

Couples	*Man*	*Woman*
Basic Steps	beg > exp > int ; 2%	exp > beg > int ; 1%

The results shows no obvious difference between level. A further investigation is needed.

4.3.4 Summary. Table 10 provides a summary of the collected results. Six out of eight parameters have been validated as relevant clues for level classifications during the improvisation part. Note that for some dances, especially those at the slowest tempo (BPM = 100), may show marginal results, due to its extreme characteristic (Salsa dance is therefore not as appropriate and professional dancer would adapt their dance to Cha-cha-cha for example).

Table 10: Summary of the results and validation

Parameter tested	Motion feature relation	Validation
Temporal offset mean	Rhythm	Not Validated
Temporal offset std	Rhythm	Validated
Main frequency error	Rhythm	Not validated
Correlation spectrum	Drive	Not validated
Correlation legs motion	Drive	Validated for improvisation
Couple delay	Drive	Validated for improvisation
Area covered	Style	Validated
Hips Hand motion	Style	Not validated

5 CONCLUSIONS AND FUTURE WORK

From literature and interviews with professionals, a set of six main salsa MMF candidates were identified. Among them, three were modelled and expressed as eight musical-related motion parameters. The basic movement of salsa steps was decomposed theoretically and validated by observation. A database of salsa dance in synchronisation with music was explained and realised. The proposed parameters were computed for each couple on each song, and statistically analysed. The results show that six among eight parameters validates the clustering of space and allows classification, mostly for the improvisation sequence. This is understandable by the fact that improvisation is a more complicated exercises than doing the basic steps. Further investigation will be needed to test the other remaining motion features, as well as a combination of some of them with a coefficient related to their respective importance. Other couple dances can be tested as well.

This study is a first step toward an artificial intelligence based virtual coach that use automatic analysis of learning states for Salsa to improve the dancer's skills. Our proposed music related motion features shows some insight and directions on the aim to describe, analyse and classify social couple dance. This first approach defines a building block for a framework that could be utilised within the other couple dances and the different domains, such as in bio-mechanic studies, human robot interaction or emotional recognition, but also in the domain of virtual reality, avatar systems and generally serious game topics can be of interest and lastly to further our general understanding of motion analysis. The SALSA database and the extracted motion feature can also serve as base for other studies and cultural heritage conservation example.

Future work includes to apply our feature for an automatic recognition algorithm for learning level, but also type of salsa sequence (basic steps or improvisation). Expanding the data set with a categorised set of (Salsa) dance moves. The inclusion of the three remaining music related motion features and tested over the previously mentioned larger accumulated data set. The possibility to include some form of EEG or facial detection study while dancing as to try detect the gaze and *feeling* of both participant.

ACKNOWLEDGMENTS

This project has been partially supported by the European Project Virtual Multimodal Museum (ViMM). The authors would also like to thank Jeremy Patrix for his valuable help in machine learning, the salsa experts for the interesting discussions and all the dancers who performed the salsa dances at our department.

REFERENCES

Paolo Alborno, Stefano Piana, and Antonio Camurri. 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*, Vol. 6455533. ACM Press, New York, New York, USA, 136–143. https://doi.org/10.1145/2909132.2909262

Andreas Aristidou, Panayiotis Charalambous, and Yiorgos Chrysanthou. 2015. Emotion Analysis and Classification: Understanding the Performers' Emotions Using the LMA Entities. *Computer Graphics Forum* 34, 6 (2015), 262–276. https://doi.org/10.1111/cgf.12598

Andreas Aristidou, Efstathios Stavrakis, and Yiorgos Chrysanthou. 2014. LMA-Based Motion Retrieval for Folk Dance Cultural Heritage. *Euromed* (2014), 207–216.

Congres1. [n. d.]. www.salsadancecongresses.com. ([n. d.]). http://www.salsadancecongresses.com/

Congres2. [n. d.]. salsa.wikia.com/wiki/Salsa_congresses. http://salsa.wikia.com/wiki/Salsa{_}congresses

Shannon Cuykendall, Ethan Soutar-Rau, and Thecla Schiphorst. 2016a. POEME: A Poetry Engine Powered by Your Movement. *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction* (2016), 635–640. https://doi.org/10.1145/2839462.2856339

Shannon Cuykendall, Ethan Soutar-Rau, Thecla Schiphorst, and Steve Dipaola. 2016b. If Words Could Dance: Moving from Body to Data through Kinesthetic Evaluation. *Proceedings of the 2016 ACM Conference on Designing Interactive Systems - DIS '16* (2016), 234–238. https://doi.org/10.1145/2901790.2901822

Augusto dos Santos, Kalina Yacef, and Roberto Martinez-Maldonado. 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, New York, USA, 183–191. https://doi.org/10.1145/3079628.3079673

Jeffrey D Hicks, Gregory Myers, Alexander Stoyen, and Qiuming Zhu. 2004. *Bayesian-game modeling of c2 decision making in submarine battle-space situation awareness.* Technical Report. NEBRASKA UNIV AT OMAHA DEPT OF COMPUTER SCIENCE.

IDF. [n. d.]. Olympic certification. ([n. d.]). https://web.archive.org/web/20100626010044/http://www.idsf.net/index.tpl?id=30

Juliet McMains. 2015. *Spinning Mambo into Salsa: Caribbean Dance in Global Commerce.* 108 pages.

Dafna Merom, Robert Cumming, Erin Mathieu, Kaarin J Anstey, Chris Rissel, Judy M Simpson, Rachael L Morton, Ester Cerin, Catherine Sherrington, and Stephen R Lord. 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 (dec 2013), 477. https://doi.org/10.1186/1471-2458-13-477

Dafna Merom, Anne Grunseit, Ranmalee Eramudugolla, Barbara Jefferis, Jade Mcneill, and Kaarin J Anstey. 2016a. Cognitive Benefits of Social Dancing and Walking in Old Age: The Dancing Mind Randomized Controlled Trial. *Frontiers in aging neuroscience* 8 (2016), 26. https://doi.org/10.3389/fnagi.2016.00026

Dafna Merom, Erin Mathieu, Ester Cerin, Rachael L. Morton, Judy M. Simpson, Chris Rissel, Kaarin J. Anstey, Catherine Sherrington, Stephen R. Lord, and Robert G. Cumming. 2016b. Social Dancing and Incidence of Falls in Older Adults: A Cluster Randomised Controlled Trial. *PLOS Medicine* 13, 8 (aug 2016), e1002112. https://doi.org/10.1371/journal.pmed.1002112

Kayhan Ozcimder, Biswadip Dey, Rebecca J. Lazier, Daniel Trueman, and Naomi E. Leonard. 2016. Investigating group behavior in dance: An evolutionary dynamics approach. In *Proceedings of the American Control Conference*, Vol. 2016-July. IEEE, 6465–6470. https://doi.org/10.1109/ACC.2016.7526687

Diego Felipe Paez Granados, Jun Kinugawa, Yasuhisa Hirata, and Kazuhiro Kosuge. 2016. Guiding Human Motions in Physical Human âĂŞ Robot Interaction through COM Motion Control of a Dance Teaching Robot. In *IEEE International Conference on Humanoid Robots (Humanoids).* IEEE, 279–285. https://doi.org/10.1109/HUMANOIDS.2016.7803289

Stefano Piana. 2016. Movement Fluidity Analysis Based on Performance and Perception. *CHI Extended Abstracts on Human Factors in Computing Systems* (2016), 1629–1636. https://doi.org/10.1145/2851581.2892478

Richard (Stanford University) Powers. [n. d.]. Brief Histories of Social Dance. https://socialdance.stanford.edu/Syllabi/dance{_}histories.htm

Simon Senecal, Louis Cuel, Andreas Aristidou, and Nadia Magnenat-Thalmann. 2016. Continuous body emotion recognition system during theater performances. In *Computer Animation and Virtual Worlds*, Vol. 27. Wiley, 311–320. https://doi.org/10.1002/cav.1714

Lida Theodorou, Patrick G. T. Healey, and Fabrizio Smeraldi. 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, New York, USA, 1–7. https://doi.org/10.1145/2948910.2948928

Caroline P. Whyatt and Elizabeth B. Torres. 2017. The social-dance. In *Proceedings of the 4th International Conference on Movement Computing - MOCO '17.* ACM Press, New York, New York, USA, 1–8. https://doi.org/10.1145/3077981.3078055