UNIVERSITÉ DE GENÈVE

Archive ouverte UNIGE

https://archive-ouverte.unige.ch

Thèse 2020

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Learning and understanding partner dance through motion analysis in a virtual environment

Senecal, Simon

How to cite

SENECAL, Simon. Learning and understanding partner dance through motion analysis in a virtual environment. 2020. doi: 10.13097/archive-ouverte/unige:142477

This publication URL: https://archive-ouverte.unige.ch//unige:142477

Publication DOI: 10.13097/archive-ouverte/unige:142477

© This document is protected by copyright. Please refer to copyright holder(s) for terms of use.

UNIVERSITÉ DE GENÈVE

FACULTÉ DES SCIENCES

Département d'informatique | Professeur José Rolim

FACULTÉ D'ÉCONOMIE
ET DE MANAGEMENT

Institut de Science de Service Informationnel | Professeur Nadia Magnenat-Thalmann

---

Learning and understanding partner dance through motion analysis in a virtual environment

THÈSE

présentée à la Faculté des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences, mention interdisciplinaire

par

Simon Renaud SENECAL

de

Louviers (France)

Thèse N°5475

GENÈVE

Atelier d'impression Repromail

2020

Acknowledgements

I would like to thank my supervisor Prof. Nadia Magnenat-Thalmann for granting me this unique opportunity at MIRALab. Her guidance, insights and support throughout this PhD were vital during these years, for which I'm very grateful. It allowed me to further my knowledge in new directions and improve my skills in computer science. I also want to thank the members of the jury for their time and effort in reviewing my manuscript; Professor Jose Rolim (University of Geneva), Professor Nadia Magnenat-Thalmann, University of Geneva, Dr Bruno Herbelin, EPFL and Professor Andreas Aristidou, University of Cyprus.

Huge thanks to all my colleagues at MIRALab, past and present, for their collaboration and generally for giving me a great time. In particular, many thanks to Niels Nijdam and Yvain Tisserand for their guidance and support they provided for my topic and collaboration. Thanks to all MIRALab colleagues to make this a memorable and enjoyable stay.

A big thanks to all participants in the dance experiments, to the Salsa experts for the interviews and the Geneva dance schools who kindly supported my work by hosting a session of Virtual reality experiment during parties, as well as Dr Jeremy Patrix for his support.

The presented work was supported by several projects: the European Marie-Curie FP7 ITN-DCH, the European research H2020 ViMM, and the European AAL research Active@Home. These projects enabled me to broaden my topic and further the applicability of my work. I would like to thank all the people for their collaboration and their hospitality during exchanges.

Special thanks to the astronomer Hubert Reeves, who gave me the deep motivation and taste for scientific research from my early years. To my family, with their unconditional support, I thank them.

Simon R. Senecal

Abstract

Partner dance is a physical activity that is social and involves two partners dancing to music. Learning such dance is difficult and presents many challenges, such as finding a partner, learning the right skills, etc. Through the motion analysis of couples as entities with regards to tempos, we aim at developing new ways to understand and characterize partner dance. We are proposing a set of features that can extract high-level information from low-level data, usable enough to help design learning systems.

In the first place, we selected three main dance skills considered important for Salsa dance learning from interviews with experts - namely Rhythm, Guidance, and Styling - and searched for clues in the theoretical description of the movement that could help to describe high-level dance skills in terms of low-level motion measurements. Next, we suggest a set of musical-related motion features (MMF) to characterize Salsa performance (two people interacting and dancing with each other to the music) in terms of learning levels. These features lead to 29 measurements applied to captured motion data and music beat analysis.

Secondly, a database is created with 26 couples performing from three levels of experience (beginners: less than 6 months, intermediates: 1 to 3 years, experts: more than five years - teachers or jury) along with ten music with varying tempos. All MMFs are then extracted from the database and submitted to a machine learning algorithm. The "Random forest" algorithm shows the best results for level classification, up to 90% accuracy. A feature importance study is also conducted to enhance our MMF set by identifying the features that contribute the most to the classification.

Finally, we designed and evaluated an interactive dance learning system that allows a person to learn Guidance and Rhythm skills through a virtual reality Salsa dance environment containing a virtual partner with hand to hand interaction. We installed the experiment in different dance schools and at dance parties to make several dancers of various levels test our system. The overall user motion data was recorded and a learning analysis performed using our Musical Motion Features and the well-known Laban Motion Analysis, showing an improvement of the dance skills of users after training, thus validating the relevance of our design.

More extended studies could be conducted to further validate the learning aspect of our interactive system by defining more precisely the learning groups, and new exercises could be imagined to add new skills to the existing rhythm and guidance. Our proposed framework of MMFs could also be refined by considering the more complex and precise definitions of each of the features, taking more skeleton joints into account than the foot joints for the Rhythm skill. The Guidance

skill can also be redefined in the light of the final design of our learning system, based on gesture recognition on a specific time frame.

Overall, our approach, rooted in the reality of current Salsa dance social performance, shows to be relevant for learning state classification and for improving dance skills with an innovative immersive virtual reality environment.

Résumé

La danse de couple est une activité physique et sociale qui implique deux partenaires en interaction sur la musique. L'apprentissage de la danse de couple est difficile et présente plusieurs défis comme trouver un partenaire, apprendre les bonnes compétences etc. Au travers de l'analyse du mouvement du couple de danse en tant qu'entités en synchronisation avec un tempo musical, nous proposons un set de composantes du mouvement qui permet d'extraire des informations haut-niveau a partir de données bas-niveau, utilisables pour l'élaboration de systèmes d'apprentissage.

Dans un premier temps nous avons interrogé plusieurs experts et identifié trois compétences importantes pour l'apprentissage de la danse Salsa : le Rythme, le Guidage et le Style. Puis nous avons cherché des indices spatio-temporels exploitables dans la description théorique des mouvements de base. Ensuite nous proposons un set de composantes musicale du mouvement (MMFs) capable de nous aider a décrire une performance de Salsa en terme de compétences de danse. Ces composantes sont structurés en 29 mesures faites sur les données 3D d'un corps en mouvement synchronisé au tempo musical.

Dans un second temps, une base de données est créée contenant 26 couples de trois niveaux d'expérience (Débutant : moins de 6 mois d'expérience, intermédiaire : entre 1 et 3 année d'expérience, expert : plus de 5 ans d'expérience, professeurs ou jury de danse.) performant sur 10 musiques ayant 10 tempo différents. Toutes les 29 mesures sont ensuite extraites et soumise a un algorithme de classification statistique. L'algorithme "Random forest" permet les meilleurs résultats de classification avec une précision de 90%. Une études comparative des composantes est aussi réalisée afin de déterminé lesquelles sont les plus contributive a l'algorithme.

Finalement, nous avons designé un système interactif d'apprentissage de danse qui permet a une personne seule de pratiquer les pas de bases de la Salsa et d'intégrer les compétences de rythme et guidage au travers un environnement de réalité virtuelle. Cet environnement permet une interaction tactile avec un partenaire virtuel. Le système a été testé et validé au sein de plusieurs écoles de danse et soirée dansantes. Les données de mouvement du corps des utilisateurs on été enregistrées et une analyse utilisant nos composantes ainsi que l'analyse de Laban montrent une amélioration des compétences de danse après entraînement avec le système, ce qui valide notre design.

Une étude plus longue permettrais de valider encore plus la pertinence de notre système d'apprentissage avec un groupe contrôle et un groupe régulier, en définissant plus précisément les

niveaux d'experience de danse. De nouveaux exercices peuvent êtres ajouté afin d'élargir les compétences apprises. Notre structure de composantes musicale du mouvement peut être amélioré en considérant plus de compétences et une définition plus précise de chaque categorie de composantes . La compétence de guidage peut être redéfini a la lumière de notre système interactif, par exemple en tant que calcul du temps d'un geste de la main droite ou gauche sur le tempo musical.

Globalement, notre approche, ancrés dans la réalité des performance de danse de couple actuellement réalisées au sein de soirées et d'écoles prouve être pertinent pour classifier le niveau d'expérience et améliorer les compétences de danse au sein d'un système innovant d'apprentissage en réalité virtuelle.

List of Publications

List of publication directly related with this thesis:

Peer-reviewed Journals

– S. Senecal, N. Nijdam, A. Aristidou and N. Magnenat-Thalmann, “Salsa dance learning evaluation and motion analysis in gamified virtual reality environment”, *Multimedia tools and application, Springer*, 2020.

– S. Senecal, L. Cuel, A. Aristidou, and N. Magnenat-Thalmann, “Continuous body emotion recognition system during theater performances”, *Computer Animation and Virtual Worlds*, 2016.

Peer-reviewed Conferences

– S. Senecal, N. Nijdam, and N. Magnenat-Thalmann, “Classification of salsa dance level using music and interaction based motion features”, *GRAPP 2019 - International Conference on Computer Graphics Theory and Applications*, 2019.

– S. Senecal, N. Nijdam, and N. Magnenat-Thalmann, “Motion analysis, indexing and classification of salsa dance using music-related motion features”, *ACM conference of motion in games, MIG*, 2018.

List of publications indirectly related to this thesis:

Peer-reviewed Conferences

– C. Coughenour, M. Vincent, M. Kramer, S. Senecal, D. Fritsch, and M. Flores, “Embedding knowledge in 3D data frameworks in cultural heritage”, *25th International CIPA Symposium, Taipei*, p. 103, 2015.

– M. Kramer, C. Coughenour, S. Senecal, “Common Ground for ITN-DCH: a semantically enriched web platform for integrating and disseminating digital cultural heritage research (poster)”, *1st International Conference on Science and Engineering in Arts, Heritage and Archaeology (SEAHA)*, p. 103, 2015.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"Dancing is creating a sculpture that is visible only for a moment." Erol Rozan, writer.

1.1 Partner dance and current challenges

Dance is an art form that generally refers to movements of the body, usually rhythmic, and music, used as a form of expression, social interaction or presented in a spiritual or performance setting. Dance is also used to describe methods of non-verbal communication (ex. body language) between humans or animals (bee dance, patterns of behavior such as a mating dance). In sports, gymnastics, figure skating, and synchronized swimming are dance disciplines, while the Katas of the martial arts are often compared to dances. Dance is part of the most universal and historical fine arts. Body motion expression is here the main communication channel that carries emotion, storytelling, etc. Social dances have existed for centuries and can be found in different forms (a duo or groups of variable sizes), in almost every ethnic group and culture and pertain to a social and/or religious context [111]. The modern social partner dance, such as Latin dances (Salsa, Cha-cha-cha, Bachata) or Western dances (Swing, Jives, West Coast Swing), is gaining more recognition and developing rapidly[82], which is highlighted by the increasing number of international congresses [29, 28], shows and dance schools, as well as their inclusion as an Olympic category [62].

More specifically, partner dance is a physical activity that involves two people, usually, a man and a woman, who dance together following the tempo of the music and moving their bodies according to the rules of the dance ("man" and "woman" in dance roles, that can be exchanged). Last century saw the rise of multiple partner dances [111] originating from different parts of the world like Salsa from Cuba, Forro from Brazil, or Tango from Argentina. An international organization such as the International DanceSport Federation [62] has been recognized by the Olympic committee and presents many worldwide championships.

Dance partners generally stand at a close distance from each other and can be mechanically connected by holding hands or with a more complex "embrace," being within the arms of each other, like a hug. This posture is essential to transmit body movement information to each other. Two roles are well defined: the leader (usually a man) and the follower (usually a woman). Each role is complementary and essential for good performance. The leader has the initiative to decide what is the next dance pattern, and sends the information trough her/his body expression to the follower. The follower has to understand and respond to the received information in an appropriate manner according to the dance. Learning such a partner dance is not an easy task and requires much effort. Around the world, many dance schools provide classes people can attend and then practice their knowledge in dance parties, dance festivals, and other social events.

The human body possess a huge quantity of degree of freedom and is already by itself not easy to understand, even by limiting the level of accuracy used for motion depiction. In partner dance, the movement of two bodies are presented and have intricate mechanical coupling and interaction that is even more difficult to analyze. Among the databases available publicly, the majority of proposed human recorded motion are one body captured. In Salsa dance, the basic motion can be easy to understand, but soon with the complication comes rotations, arms exotic movements which makes it difficult to understand and distinguish, even for a human eye. Many aspects of the motion analysis of partner dance are not fully understood.

To learn a partner dance such as Salsa is a challenge for modern humans as it requires them to learn all the different mechanical-cognitive-interactive parameters from a teacher in mainly collective classes only, which is less effective to spot errors in 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 a similar level, the student can reconsider or oppose the advice of the teacher. The status of an expert in social dance can be a source of confusion as there are no universally recognized 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 a different partner and/or take some dance classes to progress, they might 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). To summarize, here are the **current challenges** for learning and understand partner dance:

- Subjective description of the movement by the teacher
- Quality of the movement vary depending on the teacher's mood or state of mind
- Mimetic learning method
- Practice induce needs for attending classes, have the right schedule
- Partner can be not as good as expected
- Feedback from the teacher is not specific to the person

1.2 Motivations

This thesis intend to expand the practical and experimental knowledge of partner dance motion analysis and learning by proposing a **new framework** to try answering all previously mentioned challenges. This would be to set up a **virtual reality dance simulation** that embed a **virtual coach** which analyzes the movement of the dance student and **provide positive feedback to improve her/his dance skills**. Motion analysis, the treatment of raw 3D motion-captured data (3D coordinate with time), is at the center of dance-related studies, with multimodal approaches that also adds the audio and video aspects. Extracting meaningful information from motion data is the common objective. It implies to transform the low-level data into high-level information that are closer to what human can naturally understand. Therefore an automatic analysis of the different learning states should be developed from motion analysis and dance structure investigation. In order to automatically analyze and classify a social dancer on her/his learning level, the computation of measurable dance motion feature must be able to distinguish between learning levels in terms of dance performance. This framework brings many solutions to the previously stated current challenges:

- Practice at home, alone or with other person, with flexible schedule
- Learn dance skills in an objective, reproducible and efficient way
- Custom feedback personalized for each users

1.3 Related questions and problems

Our proposed framework use case scenario is a virtual reality interactive salsa dance learning system. This system requires two main components:

1.3.1 A simulation of partner dance

To achieve that, we need to simulate the main parameter of a partner dance situation. A virtual reality partner that dances and interact with us is needed. Its role can be Follower or Leader. If we take the Follower role, the virtual partner has to react to the user gestures according to dance rules. This leads to few research questions:

- RQ 1: How to synthesize a salsa dance motion ?

The virtual partner has to be animated such that his body moves accordingly to the dance and in synchronization with the music, by doing specific footsteps. A preliminary investigation is required to define what is a salsa dance motion.

- RQ 2: How to simulate the mechanical coupling between the user and the virtual partner ?

Both partner are close together and their motion are intricate such that we have a mechanical coupling for some time of the dance. Mainly both partner can holds hands-to-hands or be in the "embrace" position. An adequate visual or haptic feedback is crucial to reproduce a real salsa practice.

- RQ 3: How to manage the dance pattern changes depending on the user decision (Leader-Follower roles) ?

The leader will decide what is going to be the next move and guide accordingly the Follower. This information transmission happen at a certain time in a certain manner, that needs to be identified and integrated into the system.

1.3.2 A dance skills motion learning system

A dance learning system means to be able to teach motion skills and evaluate if they have been assimilated by the student to provide feedback. So it is needed to find the most important and relevant dance skills for learning salsa dance, and a way to extract them during a dance performance. Since we have low level motion data, we need to find a solution to convert it to the high level dance skills. Then a learning methods have to be chosen and an evaluation of its efficiency. This leads to these research questions:

- RQ 4: What are the important dance skills ?

Dance skills are related to specific dance, in our case Salsa. When learning salsa dance, the importance is focused on different aspects: guiding the partner, following the music, reproducing footsteps and patterns, etc.

- RQ 5: How to evaluate a partner dance performance in regards of dance skills ?

Raw 3D motion data and musical data is acquired and then the extraction of high level dance skills using motion analysis techniques take place as being part of the dance performance evaluation. We need to identify methods for extracting high level motion information from low level data. The evaluation of correct movements of the users is central to the learning system.

- RQ 6: What are the possible learning strategies and how to evaluate them ?

Learning Salsa is traditionally done by "mimetic" - reproducing the movements of the teacher - and in collective classes. Our framework needs firstly to define a pedagogical structure of the training, for example as multiple exercises that are interactive and fun to experience. Secondly, an evaluation of the training results consistency, that should result in the increase of dance skills.

1.4 Scope and related areas

Partner dance involves the communication through body motion and covers multiple research aspects, being a multidisciplinary subject. In the direction of answering our previously defined research question, multiple research domain can be taken into account:

- Motion analysis (analysing the body motion)
- Animation (synthesing a dance motion)
- Physical modelisation (interaction, haptics, feedback)
- Psychological interaction parameters (mindset when learning, in VR environnement)
- Music (Synchronisation with the movement)
- Learning systems (Pedagogy, teacher feedback, session length and content, evaluation)

1.5 Potential applications

Partner dance understanding leads to analyzing the behavior of two people interacting, and this knowledge can be then used for potential applications in different areas:

- Education (ex: pedagogical application)
- Behaviour recognition (ex: recognition of specific motion pattern)
- Gaming environment (ex: reversely for synthesizing expressive motion in a virtual avatar)
- Medical (ex: as rehabilitation method)
- Psychology (ex: part of social attention)

Motion analysis and classification and indexing of 3D data is a hot topic that is increasingly relevant because of the more widely available motion capture technology. As it is quite a sophisticated form of raw information, segmentation of specific time sequences is a first challenge similar to dance pattern identification. It can be used for behavior recognition for the education systems or psychology studies. Classification of high-level components of dance is useful for emotion recognition, which can be utilized within a medical context or other kind of research.

Some studies started to identify the rhythm constituency or the performance of solo performers in different dances but did 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 relatively costly acquisition of the data involved (e.g., capture both dancers at a time, requiring an 'expensive' motion capture system).

In terms of motion analysis, the body of both partners moving in synchronization with the tempo can be used to develop classification and indexation techniques for cultural preservation. In terms of interaction, the complex mechanism that links leader and follower can help to design new interactive learning or rehabilitative systems via the replacement of the missing partner by a robot or a virtual character. In terms of social analysis, since the focus of the dancer's attention is on the other person, we can see the influence of the other on our "self."

1.6 Contributions

We are proposing a framework that describes an approach to partner dance learning and understanding using a virtual reality simulator. This simulation of salsa dance presents a virtual partner with which you can hold hands and interact similarly as with a real dance partner. A series of exercises are proposed to the user and a specific gesture with precise timing is required during the dance to change the dance pattern, making the person learn guidance and rhythmic skills. A score is visualized at the end of the training as motivation. Our contribution expands the current knowledge in motion analysis in the case of high level motion descriptor for music-synchronised motion in couple. It also brings new methods for developing a virtual reality interactive learning platform and dance simulation. The main contribution are the following:

- A database of Salsa performance by two partners dancing in synchronization with music. Twenty-six couples participated and provided 18 exploitable dance sequences on various tempos.
- A set of audio-related motion features that allows characterizing salsa dance in terms of skills which are important for learning, validated by machine learning.
- An interactive learning system using virtual reality and haptic feedback to teach people guidance and rhythm with a virtual partner, validated by a user study.

First, we did an international survey and a discussion with experts from the field in order to understand what are the relevant high-level skills that are essential for learning. The result is a set of criteria validated by experts as the fundamentals of dance. Secondly, we propose to interpret these previous criteria as audio-related motion features, which we can measure on motion capture data. We made a database of 26 couples dancing salsa sequences to different tempos, with music. The classification by machine learning validated our assumptions. Finally, we use a subset of the previously validated criteria in an interactive dance learning system to teach two fundamental skills within a virtual environment with a virtual partner: Rhythm and Guidance.

1.7 Manuscript Organization

- Chapter 2 *Related Work*
  The chapter is divided into several sections that summarize the state of the art. Starting with some contextual notes, it expands with an extensive tour of related work classified thematically.
- Chapter 3 *Salsa motion modelling: building the musical motion features*
  In this chapter, we explain in details the important dance skills required for learning Salsa and our extraction of corresponding motion features.
- Chapter 4 *Musical motion feature evaluation*
  In this chapter, we build a database of salsa performance and validate our proposed motion features in their ability to distinguish salsa experience levels.
- Chapter 5 *Virtual reality salsa dance learning and motion analysis*
  In this chapter, we expand and propose a new set of motion descriptors in the specific case of virtual reality. We explain the design of an interactive dance learning system and validate its relevance with a user study.
- Chapter 6 *Conclusion*
  The final conclusion of our findings on the presented research, the proposed architecture, the established implementation, and conducted validation during several stages of its design and implementation.

Chapter 2

Related Work

"Dancing is the loftiest, the most moving, the most beautiful of the arts, because it is no mere translation or abstraction from life; it is life itself." Havelock Ellis, writer.

2.1 Context

We are presenting here a literature review that covers multiple domain to get enough information for answering the research questions and realizing our framework. We structure this chapter into six distinct categories from the paper's research domain to have a broad panorama around our subject. Please note that the proposed categories can be discussed for some papers as they are quite spread out and cover multiple thematic. Section 2.2 presents work that is classifying dance motion into high-level space, which can be emotion behavior, motion qualities, dance skills, performance score, etc. Dance motion is in most the case highly repetitive, so Section 2.3 focuses on methods that are classifying performance in terms of dance pattern, that are fundamental building components of many dances such as traditional or ballroom. Section 2.4 contain works on interactive systems, in which the dance experience is enriched with additional visualization, including virtual reality setup. In section 2.5, there is few examples of usage of motion descriptor and features to synthesize animation, with applications on retrieval and indexation. Lastly, section 2.6 condenses work on interactive learning systems that contains motion analysis components and also a learning aspect characterized by the type of feedback provided to the learner. The best ideas are compiled and discussed for each categories and a summary of the selected methods relevant to our research question is presented as conclusion.

2.2 Classification of performance into high level space

This section is a collection of research focused on the motion analysis of performance or other physical activity into high level classification space such as emotion of aesthetic result. The use of intermediate motion features and statistical tools is almost systematic.

Low cost motion capture devices show relevance to evaluate dance, as presented by Alexiadis et al. [2, 3]. In this work, they are extracting motion features from skeleton recording from the database Huawei 3DLife/EMC2. From 13 amateur dancers, their motion is scored and compared with two professionals actings as ground truth via three parameters: joint position, joint velocities, and 3D flow error that are indices of signal similarity. The reproducibility of movements acts here as a validation since professional dancers are supposed to perform similarly, which is validated by the experiment. Each user has been rated by the professional in terms of three criteria: Choreography, Musical timing, and Body balance, in this order of importance. Results show a strong correlation between the ranking resulting from the experiment and the professional annotation.

Laban Motion Analysis has been recently tested for motion analysis and classification of dance. A set of LMA-inspired motion feature, divided in 86 measurements of the four categories of LMA (Effort, Shape, Space, and body) has been developed to evaluate the emotional qualities of dance movement by Aristidou et al. [6, 5, 4, 9] for contemporary dance, by mapping an input motion to a bidimensional emotional space inspired by Psychological models, visualized in the table 2.1. Using random forest to learn the LMA profile for each emotion, the recognition results shows 87% success.

	Features		Measurements			
	f^i	Description	f^i_{max}	f^i_{min}	f^i_{σ}	f^i_{μ}
Body	f^1	Feet-hip distance	ϕ_1	ϕ_2	ϕ_3	ϕ_4
	f^2	Hands-shoulder distance	ϕ_5	ϕ_6	ϕ_7	ϕ_8
	f^3	Hands distance	ϕ_9	ϕ_{10}	ϕ_{11}	ϕ_{12}
	f^4	Hands-head distance	ϕ_{13}	ϕ_{14}	ϕ_{15}	ϕ_{16}
	f^5	Hands-hip distance	ϕ_{17}	ϕ_{18}	ϕ_{19}	ϕ_{20}
	f^6	Hip-ground distance	ϕ_{21}	ϕ_{22}	ϕ_{23}	ϕ_{24}
	f^7	Hip-ground minus feet-hip	ϕ_{25}	ϕ_{26}	ϕ_{27}	ϕ_{28}
	f^8	Centroid-ground distance	ϕ_{29}	ϕ_{30}	ϕ_{31}	ϕ_{32}
	f^9	Gait size	ϕ_{33}	ϕ_{34}	ϕ_{35}	ϕ_{36}
Effort	f^{10}	Head orientation	ϕ_{37}		ϕ_{38}	ϕ_{39}
	f^{11}	Deceleration peaks				ϕ_{40}
	f^{12}	Pelvis velocity	ϕ_{41}		ϕ_{42}	ϕ_{43}
	f^{13}	Hands velocity	ϕ_{44}		ϕ_{45}	ϕ_{46}
	f^{14}	Feet velocity	ϕ_{47}		ϕ_{48}	ϕ_{49}
	f^{15}	Pelvis acceleration	ϕ_{50}		ϕ_{51}	
	f^{16}	Hands acceleration	ϕ_{52}		ϕ_{53}	
	f^{17}	Feet acceleration	ϕ_{54}		ϕ_{55}	
	f^{18}	Jerk	ϕ_{56}		ϕ_{57}	
Shape	f^{19}	Volume (5 joints)	ϕ_{58}	ϕ_{59}	ϕ_{60}	ϕ_{61}
	f^{20}	Volume (upper body)	ϕ_{62}	ϕ_{63}	ϕ_{64}	ϕ_{65}
	f^{21}	Volume (lower body)	ϕ_{66}	ϕ_{67}	ϕ_{68}	ϕ_{69}
	f^{22}	Volume (left side)	ϕ_{70}	ϕ_{71}	ϕ_{72}	ϕ_{73}
	f^{23}	Volume (right side)	ϕ_{74}	ϕ_{75}	ϕ_{76}	ϕ_{77}
	f^{24}	Torso height	ϕ_{78}	ϕ_{79}	ϕ_{80}	ϕ_{81}
	f^{25}	Hands level				ϕ_{82}- ϕ_{84}
Space	f^{26}	Total distance				ϕ_{85}
	f^{27}	Total area				ϕ_{86}
	f^{28}	Total volume				ϕ_{87}

Figure 2.1 – Proposed LMA features from Aristidou et al. [7]

Recently, the concept of motion words and motion signatures has been introduced by the same author [8] as a compact representation of motion sequences, in conjunction with bag-of-words similar description. This concept shows good results for segmentation, retrieval, recognition, and synthesis of motion.

Bernadet et al. [16] did an expert-rated empirical analysis of the LMA system. Asking Certified Movement Analyst to annotated pre-made videos containing basic movement "Knocking a door" and "Showing directions" with variations of the LMA dimension, using a questionnaire, they highlight the weak result of LMA evaluation over the statistical indices Krippendorff α.

A conceptual framework is proposed by Camurri et al. [21] for analyzing the expressive qualities of dance movements, in the context of the European project H2020 "DANCE," illustrated in the figure 2.2. Based on four levels of abstraction, they assume different temporal and spatial scales, and that the overall framework is from an observer perspective. On the first layer, there are the physical signals, meaning the instantaneous raw data from the sensors(3D position of motion capture, Accelerations, for example). On the second layer, the raw data is processed in a collection of features characterizing motion locally in time (Quantity of motion, Symmetry, for example). The third layer operates in a longer time window to extract more structural aspects of the movement. (Lightness, Fluidity, Coordination for instance). The last layer, "Communication of expressive qualities," is about the perception of the qualities of the movements by an external observer (Emotion, Hesitation, Groove, for example). At this level, machine learning is often employed. Within the same framework, Alborno et al. [1] elaborated on the concept of intra-personal synchronization, which carries the motion feature fluidity, impulsivity, and rigidity, for coordination assessment. From the analysis of the speed of arms and hands, they find out that impulsive movements are the most synchronized, as opposed to fluid and rigid ones.

Niewiadomski et al. uses wearable sensors to detect Lightness and Fragility within dances motion [94] (pop, contemporary, classic). Based on a database of professional dancers, the extracted data from IMU sensors, EMG devices, video cameras, and a microphone was segmented using annotation an input into a Naive Baiyes classifier that achieved 77% recognition.

The same author analyzed movement qualities in karate's Kata demonstration [93]. Based on the same multi-layer framework started by the project DANCE [21], they focused on dedicated motion feature for the case study: Biomechanical efficiency, which corresponds to the minimization of muscular energy (minimum jerk and two-third law), Shape, which refers to the correct static posture and Intrapersonal coordination, which is about time relationship between the movements of different parts of the body. A number of Karate experts were asked to judge the qualities of video samples from motion capture, which helped to design and tune the overall system. In total, 16 measurements were computed on the segments defined by the Katas section, and a global movement quality score was proposed for each performance. The comparison with expert rating is calculated using the Pearson correlation coefficient and is up to 84%.

Working toward an application that would help the expression of children with autism, Piana et al. [109] proposed a method to automatically recognize 6 emotions using motion data and features like postures, kinematic and geometries, that are put into an SVM classifier. They could achieve an accuracy of 61%, which is close to the comparative user recognition rate with a low-cost device, the Kinect. The fluidity of the movement is a critical dance parameter investigated by the same author, [107]. In this particular study, it is proposed to see how fluidity can help in describing and classifying dance performance. Through interdisciplinary research, including biomechanics, 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. On a complimentary paper, they developed around the main guidelines of a multimodal repository for the analysis of expressive movements qualities in dance [110], namely, in this case, fluidity, impulsivity, and rigidity. They also present in a more recent work a computational model for automated recognition of emotions from body movements [108].

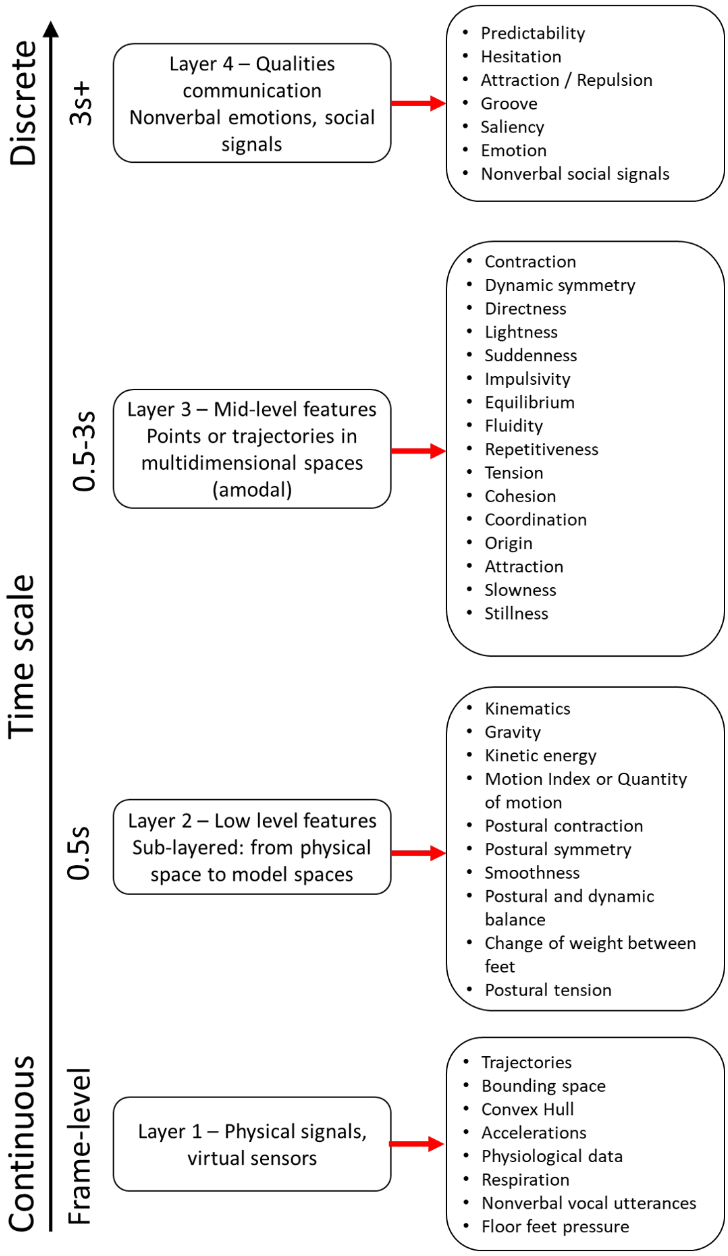


Figure 2.2 – Multilayered framework proposed by Camurri et al. [21]

Following the same technique as in previous paper, they introduce a Sparse Data representation that helps to deal with intra-personal noise (variation of expression for the same movement by the same performer). An automatic system is settled in a real-time application for stimulating expression of autistic children.

Ferguson et al. use Dynamic time warping to compare two contemporary dancer performance along one dimension (Z-axis or the height direction) [46]. By recording a contemporary dance piece with and without music, they could show that time difference induces Lapsing (high rate of changes) in the majority and scaling (gradual changes).

Fourati et al. made a study to identify relevant body cues for the classification of emotional body expression [51, 48, 50, 49]. They first proposed as motion descriptors a set of 114 bodily motion cues structured in anatomy, directional information, and postures. Then they classify the emotion expressed within simple daily action such as Walk, Sit down, Lift, Throw, etc. that were emphasized each by performers along 8 emotions(Joy, Anger, Sadness and so on) with the help of a Random Forest algorithm, with an accuracy rate of about 67% to 97%. Then they compare the ranking of the different motion descriptor in terms of the recognized motion, which give a qualitative hierarchy of the descriptors. Furthermore, a cross-comparison allowed them to reduce the amount of most relevant body cues to 11 features: 2 acceleration features (relative elbow and right hand extension), 5 posture features (backward and forward torso, openness of feet, downward torso ad head flexion.), 3 speed features (absolute, relative elbows and hands extension) and 1 standard deviation feature(backward movement of upper body part). Most of these findings apply to emotion expression during Walk motion, which can be useful for Salsa dance as the basic move is a walk.

Chantas et al. proposed a Multi-Entity Bayesian Network to classify Salsa dance and Tsamiko dance performances regarding two parameters: the execution accuracy or Proficiency Level Assessment and the Synchronisation assessment [24]. For Salsa dance, they based their description in terms of 22 annotated steps. Results show a 13% difference for execution accuracy and a 5% difference for synchronization, on average, among a pool of 30 experts and 30 beginners.

Morioka et al. proposed a method to synthesize music that is suitable for a dance motion data [89]. They achieved this by computing the first 14 motion factors that characterize the movement locally (Linear, Strong, Balanced, etc.) and then use PCA to reduce to a set of 5 motion descriptors (Power, Speed, Stability, Spread, and Regularity). From there, they have a matrix that makes the correspondence with seven emotions (happy, lonely, sharp, for example). On the other hand, they are computing features similarly from music data such as Harmony, Key, Tempo, etc.. A mapping is experimentally tuned. A final user study validates the relevance of the method with a correlation coefficient of above 90% for all emotion except "Sharp" that reached 11% max.

In a study from Cuykendall et al. [32, 33], the motion of persons was recorded using IMU sensors (Accelerometer, Magnetometer, and Gyroscope) and transcribed into poetic verses in order to characterize the complex interaction between words and movements. From experiment and use of the Language of Dance system (based on Labanotation) that links speech and movements, they build up an according vocabulary of motion features that were evaluated trough Gaussian Naive Bayes algorithm at various accuracy value that gets the maximum using the statistic "range" reaching 97% accuracy and 28% of log loss.

Kim et al. uses extreme machine learning to classify K-pop dance motion from Kinect data in the context of plagiarism detection [66, 68, 64, 65]. In order to use the most meaningful and reduce the amount of information, they took angles from the detected skeletal data and reduce the dimension with PCA. Then a comparison is made between the ELM classifier, SVM, and KNN classifiers. The proposed technique outperforms others by a 2% difference, going up to

94.5% accuracy on a database of 400 dance motion.

Oczimder et al. developed a communication model for the leader-follower interaction transition in Salsa represented by link diagram from Knot theory [97]. Physical primitives moves are described by Alexander polynomial function manipulation.

Ozcimder et al. proposed a formal model of salsa dance representation [99]. A sequence of dance is defined as a concatenation of motion primitives that correspond to partner movements over an 8-beat musical window. The dance is seen as a transition model that is using a topological link diagram representing the hand-to-hand configuration at the beginning and at the end of each motion primitive. In the case of beginner learning salsa, they establish four movements primitive and two metrics for evaluating the artistic merit: the energy, that is the distance covered by the dancer, and the phrase complexity, that is, the frequencies of the same motion primitives. An experiment is done with experienced dancers performing four dance sequences containing each 20 motion primitives. An audience of 15 judges were marking each performance in terms of the two metrics as well as an AI judge that recognized the dance primitives using the transition model and compute the same metrics. Results show a correlation of 81% between the two evaluations, validating the use of this formal model for recognition and artistic merit estimation.

Also, in computer vision, there is a classification problem for dance [124]. Samanta et al. tackle to classify classic Indian Dance using pose descriptor based on a histogram of oriented optical flow. Using a database of Youtube video, they feed the processed data into a Support Vector Machine with intersection Kernel. Accuracy of 86.67% is achieved, outperforming Bag-of-words model.

Bailleul et al. made a study on finding good metrics to classify salsa step sequences based on their artistic merit [13]. Using four dance steps, a couple performed ten choreography with different patterns. Twenty judges ranked the performances, and results show that the energy required to move is the most closely correlated with human judgment.

Shiratori et al. propose a method to produce animations of dance synchronized to music tempo that imitates dancers skills [137, 136]. A set of four related motion and musical features are proposed to analyze first a reference database and serve to establish the motion graph needed for the synthesis, for which trajectory that presents the best correlation between music and motion features is selected. A list of assumptions is taken: In the reference database, the rhythm and intensity of the dancer's motion are assumed to be synchronized to that of the music. The first motion feature is the "Keyframe" component that represents whether hands stop moving or not and is related to the rhythm music component. It is computed for all frames as the local minima of velocity. The second motion feature is the "Intensity" component that represents how wild the dance motion is. It is correlated with the music intensity feature and is based on the velocity value of the hands. The first music feature is rhythm and includes beat tracking and degree of chord changes. The second musical feature, music intensity, is computed with the spectral power of the melody line increment. The overall technique has been tested on the CMU motion capture database applied to three music piece before being evaluated visually.

Tang et al. propose a music-oriented dance choreography synthesis method using a long short-term memory (LSTM)-autoencoder [144]. Their database contains 94 minutes of professional dancers motion-captured over four types of dance (Waltz, Tango, Cha-cha-cha, and Rumba). The performers tagged their own motion as well, and the dance figures were time-stamped. From this database, they extracted three types of features: Acoustic features, motion features, and temporal indexes. Using the library "librosa," they could access five acoustic features characterizing the pitch and strength of the music. Three temporal indexes are related to the rhythm and the beat's location. Twenty-one joints position along the three axes are giving the last 63 motion features. The machine learning system here consists of a recurrent neural network that acts as dimension reduction. A user study has been made with 20 participants that were asked to score the dance samples produced by different models and evaluate how these synthesized motion fit the music. The approval rating was 67% for Waltz, 65% for Tango, 76% for Rumba, and 79% for Cha-cha.

The work of El Raheb et al. is about Dance Information System for archiving traditional dance [42, 43, 44]. Using description logics, they tried to represent dance moves in a way that is both machine-readable and human-understandable to support semantic search and movement analysis. The proposition is DanceOWL that is based on Labanotation concepts due to the high need to organize data and make this knowledge accessible. DanceOWL is a semantic dance move representation model based on rule-based extractions of logical descriptions from existing Labanotation scores. In a second time, they proposed a specific ontology structure based on a multilayer conceptual approach: a Labanotation symbols layer, a Labanotation concept layer, a general movement concept layer, and a specific movement vocabulary layer.

Synthesis and discussion

The literature presented here bring some methods to classify the motion data into high level space or categories. High level space means a characteristic behavior that human can evaluate with his judgment. Examples from this literature are: emotions, level of performance (artistic merit, aesthetic properties, execution accuracy towards an objective) or global motion qualities (ex: "fragility", "lightness"). The evaluation is often done using a reference data, that can be either an expert or annotated data, computed with machine learning.

From the raw motion data, motion features are computed to extract meaningful low-level information that is an essential first step towards high level evaluation. These motion features can be of various complexity and processed on top of each other to increase the abstraction level. We can sort the ones present in our literature into families:

- Kinematics: velocities, acceleration, jerk of body joints. Velocity peaks. Quantity of movement. Energy. Center of mass. Motion smoothness.
- Geometry: Volume defined by key joints, distances between key joints, bounding box, trajectories in space. Distance and area covered by the person. Postural symmetry.
- Postures: whole body joints position in space.

- Other data: footsteps impact, multimodal data (physiological signals)

Among the most structured framework, the Laban Motion Analysis from Aristidou et al. [7] show good results and a large description of the movement based on many motion features that describes distances, volumes and kinematics. However it lacks relationship to any musical synchronization or rhythmic consideration, beside being focused on emotional behavior. Another one is the DANCE framework proposed by Camurri et al. [20] which contains a precise description of motion with various level of abstraction based on their timescale. Various usage of this structure shows good results [107],[93]. However this study focuses more on the quality of the movement from an external perspective than technical dance skills that have to be learned. Finally the approach of Oczimder et al. [98] allows a formal description of the hands-to-hands particularity of salsa dance and its evolution, for evaluating the phrase complexity. This make sense in the context of partner dance and can be added to a more global learning system evaluation.

2.3 Extraction of temporal features and segmentation

This part of the literature review covers the methods for segmenting the motion and music data as well as to extract temporal features linked to dance pattern or rhythm, which is essential for partner dance evaluation.

Dos Santos et al. investigates the assessment of rhythm skill objectively using inertial sensors [125]. They proposed RiMoDe, an algorithm that tracks physical rhythmic abilities designed using 282 independent evaluation from an expert on 94 dance exercises performed by seven dance students. The algorithm uses acceleration peak time and extracts features such as "time between peaks." Results show a major gap between the purely algorithmic approach and how experts evaluate dance rhythm. Six themes are identified as necessary in this context: Synchronicity, Weight transfer, Limbs/Joints, Quality of the movements, Posture, Gaze.

The work of Kitsikidis et al. focused on dance pattern classification using unsupervised machine learning from Kinect skeletal data, tested on Traditional Greek Tsamiko dance [69, 70]. They are using an exemplary-based hidden Markov model that takes each frame as a hidden state. A first segmentation is done for each motion sequence divided into periods based on a monodimensional variation of the waist position. Then the position and rotation of the ankle, knee, foot, and hip are fed into the HMM to determine its parameters for pattern prediction. A test has been made with three experts and nine dance students, with dance patterns annotated manually. The segmentation method results in a 0.78% error rate on average for periods and a 5% max error rate for pattern recognition.

In the topic of multimodal classification, Masurelle et al. used footstep impact detection to help segmenting salsa dance sequence [80], from 8 dancers doing six different steps. Four on-floor piezoelectric transducers gave the information of footstep impact temporary location to segment body motion trajectories, filtered by a one-class SVM. Based on the 3DLife database, they took Hidden Markov Models as a classifier. They introduced sub-trajectories computed from a reduced amount of joint: ankles, knees, and hips, resulting in having a high classification rate of about 74% as F-measure. Another study from Karavarsamis et al. took the same problems and updated it with a new annotation and classification model based on Random Forest, reaching 94% F-measure classification [63]. By developing a dance annotation software for Kinect signal, they could annotate sequences of motion very accurately and filter out noise. Extracting features with PCA show to increase classifier accuracy and control more class imbalance. Finally, they tried to classify 20 dance steps but reached only 45% precision.

Labanotation is a way to write dance motion. A tool for generating Labanotation automatically from motion capture data is proposed by Choensawat et al. [25]. It contains a GUI that enables users to specify which key-frames corresponds to which postures. The system has been tested with experts in Kinetography Laban and shows a questionnaire-based evaluation of averagely 4.5 over 5. However, experts stipulate that the resulting information is still limited.

Davis et al. show a method to extract rhythmic features in videos and music [36]. Using a power spectrum, onset envelope detection, and Tempogram, they could extract beats and apply it to time-warp a video.

Dong et al. [38] use Hilbert-Huang transform and empirical mode decomposition to describe dance motion from Waltz and Salsa into motion primitives, outperforming Short-time Fourier transform and wavelet analysis.

The convolutional deep neural network has been investigated to classify Indian classical dance "Lasya" micro-steps by Faridee et al. [45] using multiple inertial sensors placed over the body. Data was collected from one professional and four learners and annotated by the expert. A sample window with 90% overlap was chosen, and highly correlated features removed using PCA. Results show that the proposed deep neural network outperforms state of the art shallow machine learning method as random forest by 7%.

Ofli et al. proposed a multimodal framework for dance analysis using the correlation between music features and dance figure labels [95]. Assuming a synchronized segmentation between music and dance, they are using the Chroma feature of the music to extract rhythm and hidden Markov model for motion pattern recognition. Experimentation on traditional Turkish dance Kasik allowed the system to learn proper parameters, and the evaluation was on the synthesis of a possible suite to the dance.

Shiratori et al. highlight the importance of dance digitization for cultural heritage preservation [135]. Working with the case of traditional Japanese Dance that includes specific music and motion with pauses, they proposed first to extract the speed of center of mass and hands. Looking at the local minimum of these values define a first batch of segmentation candidates. Then by Onset detection, they refined the candidates with beat tracking. They tested their algorithm on a database of three dances type. Qualitative results show a good match between the resulting segmentation and the "Mai" pauses inherent of the Dance.

Tongpaeng et al. proposed an adaptation of Labanotation to Thai traditional dance using the translation model [146].

Davis et al. proposed a robust function for beat tracking of any song [35]. It consists of using a complex spectral difference in the onset detection function. Results show an excellent performance for Dance (79.1%) and Rock music (68.1%), with more difficulties for Jazz (36.1%), Folk (24.9%), Classical (17.5%) and Choral (2.5%).

The problem of beat tracking during a Salsa song, giving the multiplicity and complexity of musical instruments involved, was tackle by Dong et al. [39]. They propose a method to estimate a Salsa Beat Interval based on Onset position detection from the Spectrogram and Onset strength curve. This method has been developed as Offline or Online, which then computes essential parameters during the 10s at the beginning. The overall system has been tested on 40 music that has also been annotated by experts. The various metric tests show the method outperform other methods of beat tracking, offering up to 85% and 65% accuracy for respectively the Offline and Online method.

Dyaberi et al. propose a method to detect the phrase structure in contemporary dance [41]. Using a small motion capture database, they computed the Mahalanobis distance between frames using Kinetic energy, momentum, and force of the recorded skeleton joints. Between two phrases, namely ABA and Rondo, they reach up to 93% accuracy.

More on the computer vision side, Bellini et al. proposed a video post-processing method that enhances the rhythm of a dancing performance [15]. They detect first the "motion-beat," the moment when there is a motion stop for a few frames in between a significant change in direction using Super-pixel Optical Flow. Then they stretch or contracted the video depending on the matching beat with the music, allowing synchronization. A user study of 26 people was conducted and show that they preferred the enhanced videos for the most part, except one, which means that sometimes keeping the beat is not enough to present a good performance.

In recent years, low cost and consumer-ready motion capture technology came more readily available and has been heartily welcomed also in the research communities. To great benefits of body motion analysis research, where extensive research has been conducted following this recent development of motion capture technology, in terms of detection, segmentation and classification. Several case studies of solo performers ranging from Indian classical dances by Saha et al. [120] recorded with a Microsoft Kinect for neural network classification. The same author worked on recognizing Indian dance postures using skeletal information from a Kinect sensor [121]. Using only a set of 16 joints information from the position of the limbs, they could process the value of the average limb in the space of 4-sided polygon, allowing up to 92.7% recognition accuracy using the SVM algorithm, computed in 4.726 sec, from a database of 7 dancers.

Synthesis and discussion

To understand precisely how people are dancing since partner dance is practiced by doing redundant dance pattern synchronized with music, identifying temporal features from motion and music is a way to help segmenting the data and evaluating the dance performance. From the raw motion data, time-dependent motion features are extracted. We list here the most used temporal motion features from our literature:

- Joints frequency analysis: Spectrogram with main frequencies, FFT, hilbert huang transform.
- Joints trajectories: Waist horizontal motion, Angles from skeletal data, center of mass, acceleration peaks, footstep impact detection.
- Postures: whole body joints position in space.
- Music analysis: onset envelope detection, chroma features.
- Kinematic: energy, momentum and forces.

The work of Dos Santos et al. [125] on Forro is usable for Salsa dance as they are very similar. Their method seems useful as to identify rhythm components within a solo performance, but point out other important aspects to take into account (Synchronicity, Weight transfer, Limbs/Joints, Quality of the movements, Posture, Gaz) for a global dance evaluation. The work of Karavarsamis et al. [63] shows the importance of footsteps temporal location, which make sense as learning partner dance is focused at the beginning on practicing the right steps more than moving the upper body. Other work highlighting the importance of frequency analysis of body motion can be used to extract temporal characteristics of salsa dance performance.

2.4 Interactive visualization system

This part of the literature review presents a collection of existing interactive systems that can help us design our system.

A first study is focused on the Haptic hand interaction between dancers for Discofox dance, by Holldampf et al. [60]. They record the physical profile of what happens during partner dance using a force and torque sensor and recreate the behavior of male (lead) dancer. Three steps were considered: basic step (forward-backward), sidestep, and windmill (rotation). The recreated male behavior involved two components: the force sensor for hands and the translation for steps. A body trajectory is first generated, and then the step size is adapted depending on the measured interaction forces and their direction. The dynamic of the system is expressed as vector fields. Results show an excellent reproduction of the link between force and step size.

In a recent study from Mousas et al. [90], a VR interactive simulation of salsa dance using the Hidden Markov model to predict the virtual partner dance behavior has been developed, as illustrated by the figure 2.3. Each animation that follows the current one is generated based on a hidden markov model and an initial database. As more of "Top-down" approach, the introduction of jump transitions between dance animation is getting towards the structure of Salsa dance as learned in classes (based on cycles of 8 beats), this study was evaluated with a questionnaire of satisfaction regarding the naturalness of the avatar motion and the dance-following feeling. Results show that users perceive the synthesized motion as more natural when applying the smoothness factors. They means that hand contacts are beneficial since the feeling of controlling the partner properly is enhanced It would be interesting to understand which specific motion features produced by the Markov model enables such perception by the users.

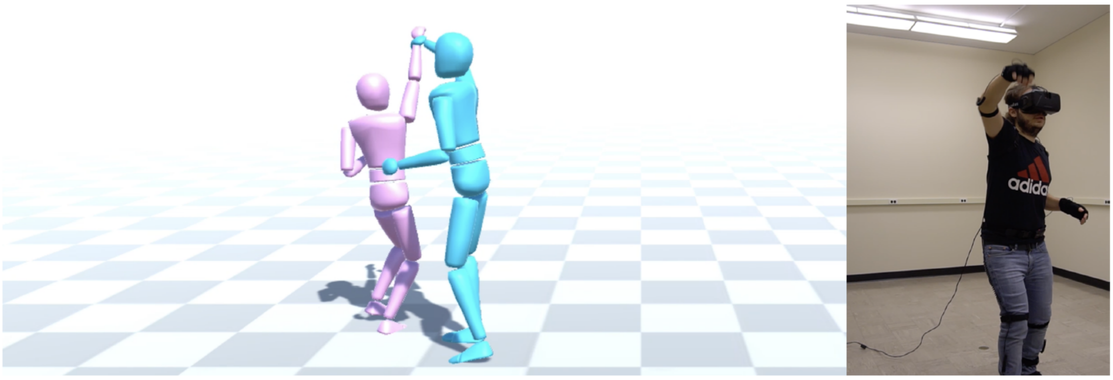


Figure 2.3 – Interactive Salsa dance system by Mousas et al. [90]

A new study from Gentry et al. highlight the importance of guidance for partner dance [52]. In the context of Lindy Hop dance, they asked five experienced couples to state if each of them were listening to the same song, given the guidance they can feel from each other. The follower detected the mixed song condition successfully with a score of 13 out of 16 but not the leaders with a score of 8 out of 16. This brings the idea that the guidance within the couple is robust to a difference in song structure.

Peeters et al. did an experiment at visualizing motion-captured Tango performance from professional dancers [106, 105]. Sensors were put on the feet, hips, shoulders, and hands of both dancers. The visualization was projected on a flat surface near the dancers during a performance. In conclusion, a need for a less obtrusive type of sensors is needed, and various parameters of the visualization like colors according to speed movements, the timing of the decay of visual traces. Also, projecting onto a surface that surrounds the dancers is another suggestion. Tango experts are currently using the whole system for teaching activities.

The first work from Brown et al. [18] investigated the effect of musical feedback from real-time analysis of movement during tango dance in order to deepen the internal experience. They are focusing on a partner dance's important feature called "Connection," which corresponds to the kinetic synchronization between the two partners' bodies. Since the tango embrace position, occlusion problems were too difficult to manage, and wearable inertial sensors were preferred for motion capture. Dancers had three sensors, one behind each ankle and another on the upper back. Step detection by thresholding the jerk of ankle sensor was implemented, and steps off the beat were filtered out. The spatial similarity between mirroring dancer limbs was sensed using cross-covariance. Also, the windowed variance as an activity level measure and its derivative was computed as the suddenness of the movement. As a movement-sound relationship, the onset of two to three-note melodies was defined to be triggered by dancer foot onsets, and reverb tail affected by the length of steps. Perceptual qualities were modeled as a space to map how movements feature lead to melody modification: pointedness/roundedness, orchestral thickness, and busyness/sparseness.

Brenton et al. present an interactive visualization that responds to the free-form movements of a non-professional dancer via interactive machine learning [17]. The idea is to train the system with an embodied, holistic view of movements. Their last prototype uses an Oculus Rift headset and Kinect for tracking, as well as the weighted nearest neighbor algorithm for learning. A dynamic visualization enables us to understand in real-time what the classifier is doing in the shape of stickmen representation. Overall it allows designing interfaces by giving examples of movements instead of analysis.

Heydrich et al. [57] highlighted the importance of Synchrony of physical and audiovisual stimulus for self-identification within a virtual reality environment, by experimenting with people being poked in the back with a stick while seeing different bodies that they can believe it is their own.

In the topic of Virtual reality, the work of Serino et al. [133] proposed to compute as neurophysiologically metric the bounding box of a peripersonal zone around the body where the tactile and audiovisual information has to be highly time-correlated for enhancing immersion, giving clues about the neural mechanism involved in embodiment and self-consciousness.

In this early work, Schiporst et al. use a six degree of freedom motion capture system "Flock of birds" to visualize one movement in real-time with a delay, giving the interactive possibility to see your movements [129]. Tsampounaris et al. established an interactive system to experiment with various visualization of the body while dancing [147]. Different visualization parameters were tested like particle systems, motion trail, smoke, fire, etc. as well as custom-designed bodies that have more less abstract shape. All of these parameters can be activated through an in-app menu by the user, giving the freedom to choose what he likes the most. This experiment tries a new approach to learning dance by exploring your possibilities through interactive visualization that reflects your motion.

Using the Huawei/3DLife Grand Challenge datasets, Gowing et al. proposed a method for multimodal data synchronization and visualization of Salsa performance for augmented reality devices [53]. The synchronization is based on the music data that is segmented from onsets detection of the on-floor piezoelectric sensor and ankle inertial sensors, extracting the Step impacts. Joints positions are extracted from the skeletal data and visualized.

The spatial and temporal coordination of dancers facing each other or being back to back is studied by Shikanai et al. [134]. They recorded the timing and correlation of velocity as for rhythm and knee height and angle for spatial body movements during the asynchronous and non-synchronous task of up and down movements. Results show that timing difference tends to reduce while being face to face, and dancers are able to synchronize with each other regardless of rhythm changes.

An interactive system linking a contemporary dancer and an electromechanical piano were developed by Palacio et al. [103]. The produced artistical piece presents the dancer and the piano as two performers on stage whose bodily movements are mutually interdependent. Assuming that the physical movements of a pianist represent a choreography, they attached inertial sensors to the dancer's body. The interactivity is defined by the raw sensory data as well as high-level features such as Smoothness, Weight, Energy, and Dynamic Symmetry inspired by Rudolf Laban. The musical composition is generated from the previous feature and different algorithm: Swarm simulation, for example, that take distances within an intervallic relationship between notes.

Another work is to developed a dance game based on motion capture technology by Deng et al. [37], 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 movements and reject unwanted motions.

Other works include the process of understanding how interactive dance visualizations can inspire coordinated dance moves between dancers in informal contexts, by Griggio et al. [54]. In their experiment, they captured the motion data of persons using smartphone's inertial sensors. They mapped it to visualization in the shape of "sphere of light," that moves and evolves depending on the user's movements. A questionnaire has been filled to extract the feeling of each participant. Results show that people were interested in interactive visualization while having less social interaction with each other. A set of useful design consideration is proposed as a conclusion: User gets first familiar with the interactive system through a discovery phase where they play. Visual representation of the rhythm of dance help users their contribution to the experience. An evident visual connection between body movement and visual effects is likely to inspire coordination between dancers.

A framework is developed by Ho et al. for synthesizing the motion of a virtual character in response to the actions performed by a user-controlled character in real-time [59]. They are introducing the interaction mesh, a method to enhance virtual character interaction with the digitized user by matching live captured data with close interaction.

A method to control a real-time virtual character using a motion capture system is also proposed by Lee et al. [76]. In their approach, the character's motion extracted from a database and preprocessed using a two-layer method. A Markov process is used in the first layer, and a clustering technique is used in the second layer.

Theodorou et al. have also investigated the interaction between the public and the performers. As the analysis of audience motion pattern during a contemporary dance performance [145]. Facial expressions, hand gestural, and upper body movements were extracted trough infrared camera recording and various methods (optical flow, ELAN, SHORE software). Thirty-three audience members were recorded (30 female and three male). Results show that hands are moving more toward the face as the performance continues; the upper body gets more static. The faces of people are recognized as "Angry," which corresponds to a blank face finally, indicating concentration. Overall, it suggests that stillness may be an indicator of engagement. A correlation is also computed and reveal a synchronization between the dancer body movement and the audience. The conclusion is that the viewer moves very little and have predominantly expressionless faces during the performance itself, in contrast to before and after. Hands seem to play a significant role since they are moving more freely.

Another work from Oczimder et al. investigated group behavior in dance using an evolutionary dynamic model [98]. Their approach is motivated by observations of a group of nineteen dancers during a performance in which they choose a sequence of dance movements from a finite set of allowable movement modules as they perform. Results show evidence that subgroups of dancers performing the same movement module with more excellent representation are aware of their dominance, which in turn influences their switching rates between modules.

Curtis et al. used a low-cost robot toy that can learn dance moves for child therapy [30]. Via tapping on the robot, the user can enable different mode, in which the robot track a yellow piece of tissue and move his head accordingly, recording the motion. Then using audio signal energy to compute the beat in real-time, the toy can reproduce previously learned motion adjusted on a beat produced by a metronome.

Sakashita et al. explore in an early work [123] a new way to interact with a tangible representation of motion: by 3D printing the critical posture of a dance example, people can "touch" the motion and change their point of view similar to virtual reality experience.

Synthesis and discussion

Overall, most of the studies point out the importance of synchronicity between the activity or task and the real-time visual feedback. More particularly, the hands-to-hands simulation specific to partner dance is beneficial and brings more trust from the user, if it induces changes of the virtual partner that seems natural. The work from Mousas et al. [90] uses typical parameters of salsa dance such as 8-beat time window and hands contact in a creative way that is satisfactory for the user. However the motion of the virtual partner is generated and does not reproduce the basic steps that you can learn in class, with specific synchronisation with music.

2.5 Animation

This section explore the literature in the field of Animation about the motion analysis of databases, and the design of dance-based virtual character animation.

Partner dance represents a challenge in terms of animation, partially because of the close ballroom hold - the embrace - that is a very intricate mechanical connection between the two partners. Herbison et al. proposed a series of design advise based on the Labanotation-based developed animation software LINTEL [56]. First observation is five areas of contact: The man left hand holding the woman right hand, The woman left hand resting on the man right upper arm, the man right hand placed on the left shoulder blade on the back of the woman, the woman left elbow rests on the man elbow, that right area of the chest of each partner touches that of the other. The design pieces of advice are: In terms of locomotion, the torsos of the two figures must be moving at the same time. For the foot release, the thigh of the moving foot must first be flexed, and then the knee flexed, in order to avoid body rise. For the hold maintenance, the inter-pelvis spacing should be maintained, and the arm and hand contact is maintained by raising or lowering the man's arm.

In their work, Matthews et al. tackle the ability to add a "Cuban motion" procedurally to any motion file [81]. They define it by a list of how the weight of the body related to posture has to be depending on the tempo of the music. They explain how to do it with XML description.

Okada et al. proposed a new method to add more expression to a dance animation while keeping the timing as original [96]. Considering the difference between "Natural" and "Happy" state, they use Dynamic Time Warping to extract the feature "Accent" and amplitude of motion to obtain the feature "Emphasis." Applying a filter that increases these two features results in slow and rapid changes of movements, and also more translation of the center of gravity. The final motion is more expressive.

On the topic of motion synthesis constituent to a rhythmic constraint, Kim et al. proposed a method to extend an unlabelled motion to a new motion while keeping its rhythmic pattern [67]. They first extract information from the example motion to build a movement transition graph, and then make a new motion using the sound input that is synchronized and validates kinematic constraints. To detect beat motion, they are using zero-crossing of the second derivative of each motion signal (joint orientation) at every frame. Among these values, they select the reference beat estimation with pulse repetition signals. To build up the transition graph, two constraints must be respected: The kinematic continuity, which is the measure of transition smoothness, and the behavioral continuity, which is the degree of satisfying the choreographic pattern. The evaluation was done using Waltz and Tango music style at a tempo of 150 beats per minute. The motion generation could handle up to 2000fps and down to 46fps. The Waltz experiment generated animations of a couple dancing in ZigZag. The Tango experiment produced a crowd of 20 couples dancing, moving with a concept of flow force and interaction forces.

LMA has been used to set up a motion retrieval and indexing system for Folk dance by Aristidou et al. [7], that can be useful for motion databases. LMA based style analysis can also be used to prune transitions of motion graphs, shown by Aristidou et al. [11], improving the coherence of stylistic continuity for motion synthesis.

Avatar to avatar interaction motion synthesis has also been explored via a setup of interaction patches, by Shum et al. [138]. In this work, a scene with multiple animations of virtual characters interacting with each other is generated from a set of high-level pattern description, motion capture, and motion graphs. Multiple interaction animations are simulated and ranked using an inspired web-page ranking algorithm and cost function from desired 3D animation parameters, allowing to generate scenes with many avatars.

In another work, Won et al. proposed a method to generate and rank rapidly and semi-automatically multi-character interaction scenes from high-level graphical description composed of simple clauses and phrases [154]. Using a database of motion capture from fighting scenes, they are using the VisualRanking method and some diversity ranking to define the plausibility of the scene in terms of five metrics. While limited by causal chains, the resulting scene is chosen among 1000 candidates.

In the interest of computer games, Mori et al. developed a technique for automatic dance animation generation from musical annotation [87]. They are using motion networks, similarly to motion graphs and a database of samples containing association between music and motion. Upon musical score, they are choosing a matching posture and interpolate using "spare key motion" to fit the time window.

Important studies in the field of human visual appearance [14] provide several advises on a good virtual human representation for better interaction.

The virtual avatar appearance is discussed by Mori et al. with a study that highlights the risk of repulsion from the spectator at a certain point of similarity with a real human body, called the "uncanny valley" [88]. They noted that the familiarity of spectators to a robot is increasing along with the resemblance of the robot to the natural human body, until a certain point where it drops, yet to grow again. They assumed it might come from the resemblance of the robot too close to a dead body.

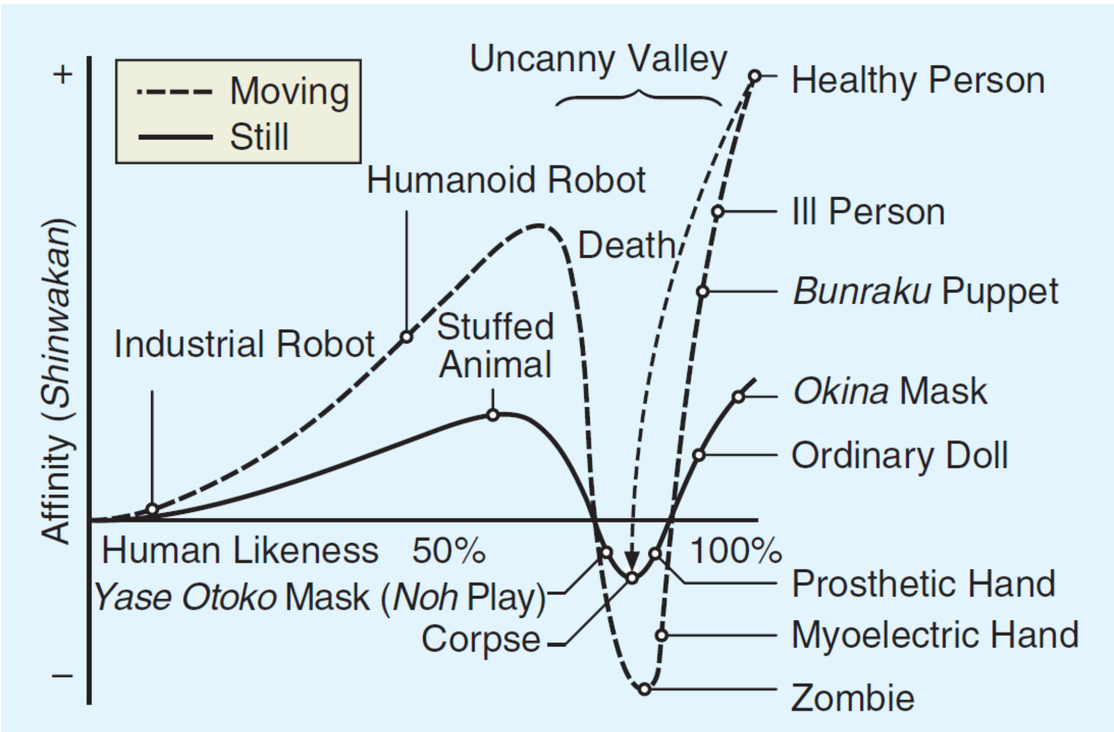


Figure 2.4 – The uncanny valley, affinity versus natural resemblance graph, by Mori et al. [88]

Synthesis and discussion

A set of design consideration can be extracted from the literature concerning the animation of our virtual partner:

- Steady distance between the animated character and the user in regards to torso, also inter-pelvis distance.
- The thigh of the moving foot must first be flexed, and then the knee flexed, in order to avoid body rise.
- The hand contact is maintained by raising or lowering the virtual partner arm.
- Weight distribution within a posture depends on the music.
- A kinematic continuity between the different animations.
- The appearance of the avatar should be avoiding to reach the "uncanny valley"

As within salsa dance, the virtual partner needs to be animated depending on the music, what it is supposed to be dancing and its the interaction with the user.

2.6 Interactive Dance Learning System

This part covers the literature of interactive dance learning system that has been tested and evaluated. It involves the user, a learning method, an evaluation and a feedback system with videos or virtual environment.

To take into account the uncertainty of observations, the judging 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. [58]. These human deficiencies, mostly due to infobesity (i.e., information overload), can be corrected by a virtual coach.

Cuykendall et al. [31] discusses the definition of expertise in dance and propose new categories beyond just the role or years of experience that take into account the complexity of crossed-skills, within a framework of dance enaction. Four categories of experts are proposed: Virtuosic, Expressive, Kinematic, and Somatic, as illustrated by figure 2.5. These categories can be the global objective of the dance student.

El Raheb et al. [112, 114] investigated the methods of dance teaching and learning in relation to interactive technology, emphasizing the point that dance, as a domain of mastering expressive movement, can help to the design of whole-body interaction experiences. They show that learning methodologies can belong to different categories:

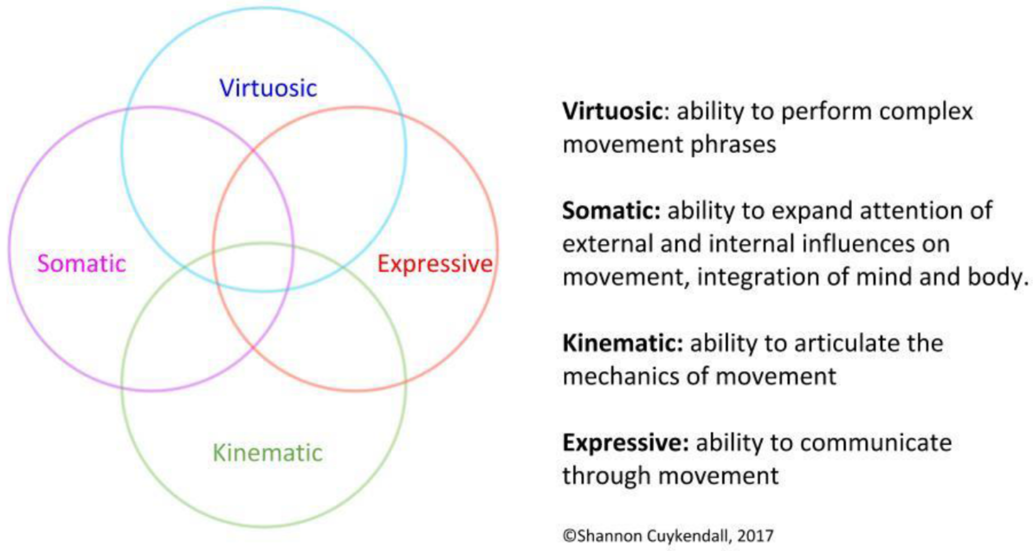


Figure 2.5 – Expertise categories by Cuykendall et al. [31]

- Mimetic method: In this method, the students mainly imitate or copy the teacher's movement or sequence of movements. It is also known as "see and do" approach.
- Traditional method: where the professor makes all the decisions, and the student follows these decisions. The method requires accuracy of performance, and the right/wrong paradigm is strongly applied. Also known as command style teaching.
- Generative method: the teacher gives the student an exercise/phrase/sequence as a starting point to achieve technical and creative goals;
- Reflective method: the student have a movement task to work with, improvising without trying to achieve a specific sequence, and the professor provides feedback. The aim is more on mastering or exploring specific aspects of movement (principles, qualities, etc.), rather than reproducing specific moves.

Based on a literature review, they also propose a structure of the important design parameters of interactive dance learning systems:

- Initialization: How the experience starts? Initiative (Frequency of Intervention/Timing) What is the student asked to do?
- Capturing the Student's movement: How to capture the student's movement? Equipment and setting (devices, hardware, etc.)
- Processing (Movement) Data: What are the objectives for evaluating a student's movement? Movement parameters to evaluate.
- Feedback Phase: Visualization of body and movement. Modalities used (e.g., visualization, sonification, audio, text, speech, other) Continuous vs. Discrete. Correction vs. Reflection.

In a recent paper, the same author investigated the different ways of visualizing dance movement via interaction with virtual avatars, objects, visual effects, and 3D representation, providing an adapted environment for self-reflection learning methods [113].

A study about Forro dance by Dos Santos et al. [128, 127] 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*". Different types of feedback are provided by the developed App, illustrated in figure 2.6.

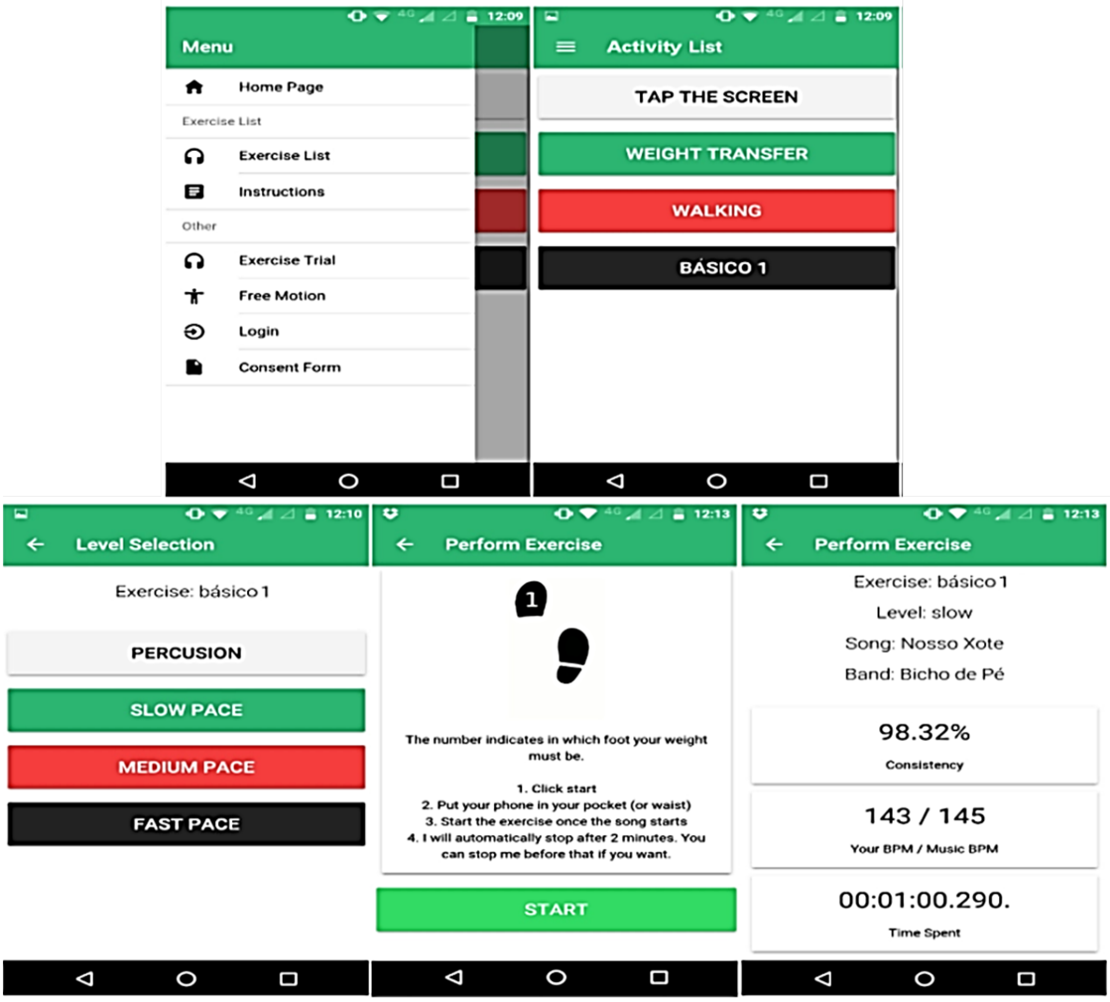


Figure 2.6 – Forro trainer app by Santos et al. [128]. From left to right: menu, activity list, difficulty level list, instructions and performance result for the last exercise.

More recent work from the same author [126] explores the video annotation method as a tool to support dance teaching. Facing the difficulty of teachers to give feedback to all students in classes, they proposed a web-based method to assess student performance using a video annotation tool. The case study concerns Forro dance. The design was developed using 94 videos from real students annotated by six dance teachers. The tool includes comments about previously identified relevant skills: dancing with a consistent rhythm, pausing at the right time, Synchrony with the song, time between movement, correct weight transfer, correct step size, avoiding dance jumping, avoiding stepping strongly, and correct hip movements. The evaluation of the tool was made using 960 1-minutes videos annotated by four teachers. As a result, four only mandatory skills prove to be relevant: rhythm/tempo, pause, weight transfer, and step size. The tool proves to be useful as a post-class feedback tool and can be proposed as a teacher training system or online learning system.

Another dance learning platform using pressure pads under the participants is proposed by Drobny et al. for Slow Waltz [40]. They are using force sensors to detect steps of dance couples and compare their timing to the music one. Feedback is provided in case of rhythm mistakes by increasing linearly up the volume of the percussion inside the music, which is not disturbing. Eight couples tried the system over five songs, with the control group included. The resulting data show better improvement of the dancer skills that have the feedback and a questionnaire submitted confirm their agreement on the help provided. Extra information is that the standard deviation of steps timing is a good value to determine the experience of dancers, and this system cannot identify "reverse" dancing (starting on the third beat instead of the one).

A team with Molina et al. studied the effect of a custom interaction device in dance studio [86]. This device is called "Delay Mirror" records a video stream rendered immediately on a large screen with a delay of few seconds and shows to be useful in a context of dance class, evaluated by qualitative questionnaires after an advanced-level adult ballet course. Out of eight dancers plus an interview of the teacher, results show that the device seems to be particularly useful at supporting the student to better comprehend teacher feedback, by looking at yourself. Other advantages pointed out are the ephemeral quality, which does not imply recordings.

An original device proposed by Tsuchida et al. tackle the problem of formation dancing learning[148, 149]. They propose a support system for performing the formation smoothly using a self-propelled screen (OMNIKIT2010), even if there is no dance partner. Nine dancers were in the experiment and had to dance with and without the device, and also with a simple projection of the missing partner, doing choreography that emphasizes some collective spatial movements. The evaluation was made using a questionnaire and show that the bulkiness and sensibility of the device were not helping in terms of the trajectory where the partner have to be very close. However, it proves to be good at learning the path of the missing partner in space, and the projected screen component did help in terms of feeling to be together.

In the study from Sun et al. [139], a self-learning framework using Laban movement analysis has been developed to facilitate trainees in automatically analyzing their movements (with Microsoft Kinect) and correcting dance technique without an expert. The body-shape components of the LMA analysis were computed using a time warping algorithm, the rhythm assessed via curve fitting, and the effort measured by the standard deviation. Results show that 23 LMA parameters can be used to describe the movement, and the trainee gets 10% improvements by using the framework.

Web 3D platform for dance learning has been investigated by Thalmann et al. [79]. The movement of Folk dance expert are recorded and playback using a web interface, which usability enhance the learning process.

Camurri et al. proposed a system to support the learning of movement qualities in dance [20]. Using a computational model of dynamic symmetry of human movement(from the jerk of hands linear acceleration and kinetic energy of the hand's movements), they offer an interactive sonification exergame based on two inertial sensors that teach symmetry or asymmetry. The auditory feedback provides reward depending on the level of coordination, symmetry, and synchronization achieved with the arm.

A virtual reality dance (Hip-hop and Agogo) training system using optical motion capture has been developed by Chan et al. [23]. Based on mimetic - or the pedagogical method to imitate the teacher movements, the user should reproduce the motion of a virtual teacher and is evaluated automatically using motion matching (to a reference motion from experts), having subsequently different type of feedback: An immediate feedback, a score report and a slow-motion replay. Experimenting with nine subjects training during 15 minutes with the system in comparison with subject with only videos, the results show that the system can successfully guide students to improve their skills.

Hoysniemi et al. made an extensive online study [61] with a questionnaire that has been submitted to 556 people from 22 countries that examines the player background about the game Dance dance revolution. Results show that the game has a positive effect on social life and physical health. More precisely, the study shows that some specific movements such as jumping or crossover are more learned than using right or left foot only since the coordination required. The primary method of teaching is by mimicking other players while others are playing and watching videos.

Kyan et al. use a Kinect to capture ballet dancer as skeletal joints in space and provides feedback as visualization within a CAVE environment [74]. Using an unsupervised method self-organizing map (SSOM) they propose to map specific feature descriptor "gesture trajectories" in a posture space via a predefined gesture database in order to recognize the movements obtaining a high rate (95%). Using mainly angles of the body joints as input, which is relevant to ballet, they employ three types of feedback: A visualization side-by-side and an overlay and a score graph. Results show a slight improvement of the gesture learning by the user, over 4%.

Nakamura et al. [91] present a multimodal presentation method for a dance training system. Watching a video to learn dance induce a lack of two primary information: the translation in space and the timing. To solve this issue, they propose an image display and Vibro-motors on a mobile robot that provides these missing information as the video itself. After two experiments with 16 beginners and evaluated using a motion capture system, the proposed system proves to be useful for dance learning.

Virtual reality simulator that can teach traditional Greek dances has also been developed based on LMA by Aristidou et al. [12, 10]. The analysis would extract a score for each of the 4 LMA component as feedback for the learner, which have high-level indications, at a global point of view of his performance.

Kitsikidis et al. [71] tested a dance learning system with Hidden state Conditional Random Fields classifier as motion analysis and Fuzzy inference system to compare the learner performance to the expert one. Results show an accuracy between 40 and 77% of choreographic reproduction after one session of 4min.

Yang et al. proposed an automated two-phase lesson generator to guide beginner to learn dance move [155], in a way that offers a suitable cognitive load. An experiment was done with 52 college students divided into test group and control group recorded with optical motion capture that went trough both phase: first to learn simple patterns and then assemble these patterns into a choreography, in the case of Latin dance and Hip-hop. The dance patterns were identified using a self-similarity matrix from the motion of a teacher. Their learning time has been compared, and a questionnaire assesses their motivation. Results show a significant difference between the groups in terms of learning time (ANOVA test). The control groups express to have a cognitive overload.

Benefit of social dances for health has been highlighted for balance and cognition by Merom et al. [85, 83, 84]. For instance, locomotion has been studied with gait analysis and classification with extreme machine learning and leg joint angles data [104, 131].

Recent papers by Cisneros, Camurri and Rizzo [26, 27, 22, 115] sets out a framework to explore the creative potential of VR in the context of the EU-founded project WhoLoDancE [153]: Whole body interaction learning for dance education (case study on Greek Folk, Ballet, Contemporary and Flamenco). After making a database of 3D motion capture, a framework for immersive visualization of dance performances using low-end head-mounted devices have been developed. As an experiment, record dancers could see their own performance from a projected virtual space using Hololens. This new visualization encouraged them to rethink differently about they moved and can be exploited in their creative process. From another experiment, choreographer shows that the product of creation using VR tools may be the performance itself, designed to be watched using VR devices. In this study, the situation was modeled into four parts: the virtual dancer, the virtual environment, the virtual performance, and the spectator.

Tang et al. made an interactive dancing game using optical 3D motion capture and block matching algorithm to recognize dance moves in real-time in the case of A-go-go dance (1960s collaborative dance) [143]. Upon tracked performance, a virtual partner is animated accordingly. A user study validates the work with questionnaires. The recognition part is done using a finite state machine and motion posture template, via frame matching cost function (weighted sum of joint ankle difference). Seven subjects tried out the experiment for twenty minutes before filling a questionnaire. Results show that the users gave the note 4.1/5 that agree positively with the virtual dancer interaction and smoothness. Side comments pointed out the importance of getting a score for motivation and the presence of synchronized music.

With the European research project ALIZ-E, Ros et al. designed a robot companion that can establish affective interaction with hospitalized children that have diabetes and obesity through dance teaching lessons [117, 116, 118]. A first experiment leads to a series of design consideration for learning interaction: Measuring pertinent motor-skills that constitutes the exercise as balance, flexibility, layered movement, synchronization, strength, speed, recall, and rhythm. Also, a list of critical points are highlighted: The movement selection, performance evaluation, and feedback. In a second experiment, they explore the creative dance learning called "concept-based learning," where the child has to move their body in a way that corresponds to the concept-word. They implemented this on a robot NAO in a "request-respond" process: first, the robot asks the child to make a concept move. Then if the child fails to respond, the robot shows a move accordingly. A test was performed with 11 children over four days within three sessions. Results show that the children present a significant increase in their engagement and response to the task.

A research project initiated by Kosuge et al. [72, 73] uses a robotic platform whose motion are described by Liu et al. [77, 156], a dance partner robot that can reproduce partner dance interaction with haptics,illustrated in figure 2.7, leads to many discoveries and publications: In a first study concerning Waltz, Takeda, Takahiro and Nakayama proposed to reproduce the following behavior of the female-type partner [142, 141, 140, 92]. They are using a control architecture that is composed of a knowledge base, a step estimator, a step detector, and a motion generator. For each detected step from the user, a transition is estimated using dance rule and human intention. Real-Time analysis of the time series data of force and moment applied by the human is evaluated using a hidden Markov model to include the uncertainty. The motion itself is generated based on recorded trajectories of dance steps and physical interaction. The dance step stride length is adapted in real-time, depending on the input force. In a related study by Sakai et al., they implemented a male-type behavior to the robot [122]. In this case, the step estimation is replaced by a step selection problems that includes collision avoidance with other couples and respect of a so-called line of dance, that is relevant to the case study of Waltz. Another study from Buondonno and Wang [19, 151, 152] focused on the translation part of the robot design by considering Linear Inverted Pendulum model to simulate body dynamic using an additional feature such as considering the swinging foot. Accurate and reactive behavior is obtained using Kalman predictor and model predictive control, handling human intentions and noisy signals. A last work by Paez et al. focused on male-type partner interaction [102, 100]. The center of mass elevation is included as part of the guidance and shows an improvement in understanding the motion direction from 70% to 90%. It also reduces the needed interaction forces by 10N, achieving more human-like leading. A new design for leading type dance robot is also rethink considering the minimum degree of freedom and force required. The final implementation includes three degrees of freedom for the mobile base - translation and four degrees of freedom for the upper body. The minimum force required to reproduce human-like guidance was implemented as 30 Nm for wrists, 40Nm for elbows, and 25Nm for shoulders, based, for example, of the 35N for the leading forward movement. A comparative learning study was proposed by the same author [101] with the robot as a dance teacher, using a progressive teaching method, which adjusts the difficulty based on user performance. The system was tested on 12 volunteers, and the psychological safety parameters were evaluated in comparison to other learning methods also using the robot. Results are in favor of progressive learning, although a forgiveness factor should be integrated with time gaps between practices. Indeed, gamification is an interesting process of improving the engagement of user for learning system [55].

Figure 2.7 – Ballroom dance learning with a male-type robot, by Paez et al. [102]

Synthesis and discussion

The literature presented here show some various method for learning. Mimetic method is the basic one, that consist in reproducing the movement of a teacher. The reflective method allows more freedom to the student and aim to improve its skills by exploring its learning space. Generative method is the development of the student by itself from a starting point. Two-phase method starts by teaching basic blocks of knowledge and then assembling them into a complex pattern, releasing the cognitive load. Concept-based learning propose to the student a concept for which it has to respond appropriately. Progressive method adjust the difficulty of the learning stages depending on the user performance. Repetition is a key aspect of almost all learning methods.

The evaluation method is done using motion analysis and extraction of features that serves for the feedback component, which are often computed to a reference point, the expert performance or the music (for example motion matching or comparison with the music beat).

Feedback is a crucial part of learning and is the cognitive mechanics that will push the student to improve. It is divided into two categories: the real-time feedback that will indicate the need of an immediate improvement and the post-training feedback that allows more self-thought on the performance evaluation. Classically a score or graph of the performance can be provided, as well as slow motion replay with different views or positive textual feedback. More for remote teaching, there is the possibility of annotated videos. In terms of real-time feedback, there is first, interactive visualizations that changes depending on our actions, more less delayed.Then virtual or robotic virtual partner that simulates the trajectory of real dancer. Auditory or vibration located in space that provides some time and spatial indication. Music parameters as the volume varying depending on the performance.

The experiment of Dos Santos et al. [128, 127] present similar aspect to our framework as the objective to improve partner dance skills. However their learning system does not involve any interaction with any sort of partner and so does not allow to enhance these essential skills. Also the motion capture using smartphone, although to be not invasive is limiting the quality of dance evaluation. Drobny et al. for [40] developed a learning method using the percussion track of a music that increases linearly with the rhythmic error of participants. Although very smooth for the participant, this technique needs to prepare the music generation to control each track separately, which is difficult to achieve with commercial music. Also the rhythm is the only dance skills that is improve so far. Finally the pressure pads can have a surface texture that can be disturbing to dancers that are used to smooth dance floor. Takeda, Takahiro and Nakayama proposed to reproduce the following behavior of the female-type partner [142, 141, 140, 92]. The haptics feedback used in robotics shows great relevance for the interaction between partner. An adaptation can be researched to have similar feelings in virtual reality.

2.7 Summary and Conclusion

We have explored the core technologies and research areas that revolve around the domain of partner dance learning and analysis, with a strong focus on interactive learning systems. The most relevant work for motion analysis and interactive learning system are summarized into the tables C.1,C.2, C.3 and C.4, C.5, C.6 respectively, that are placed in the annexes.

One important aspect of our framework lies around the motion analysis of partner dance performance. Through the literature review, we conclude that the motion data is represented by different abstraction level, which we illustrate with the figure 2.8 containing examples.

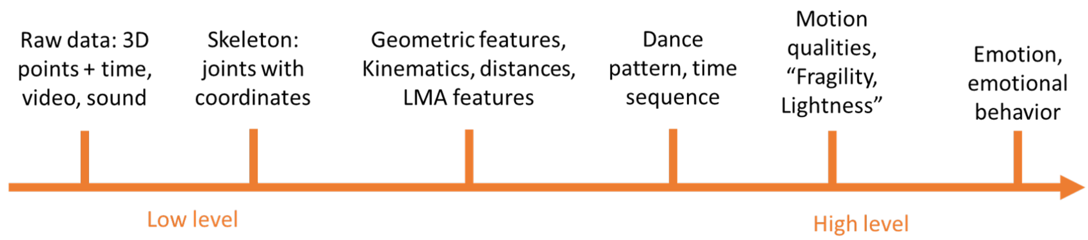


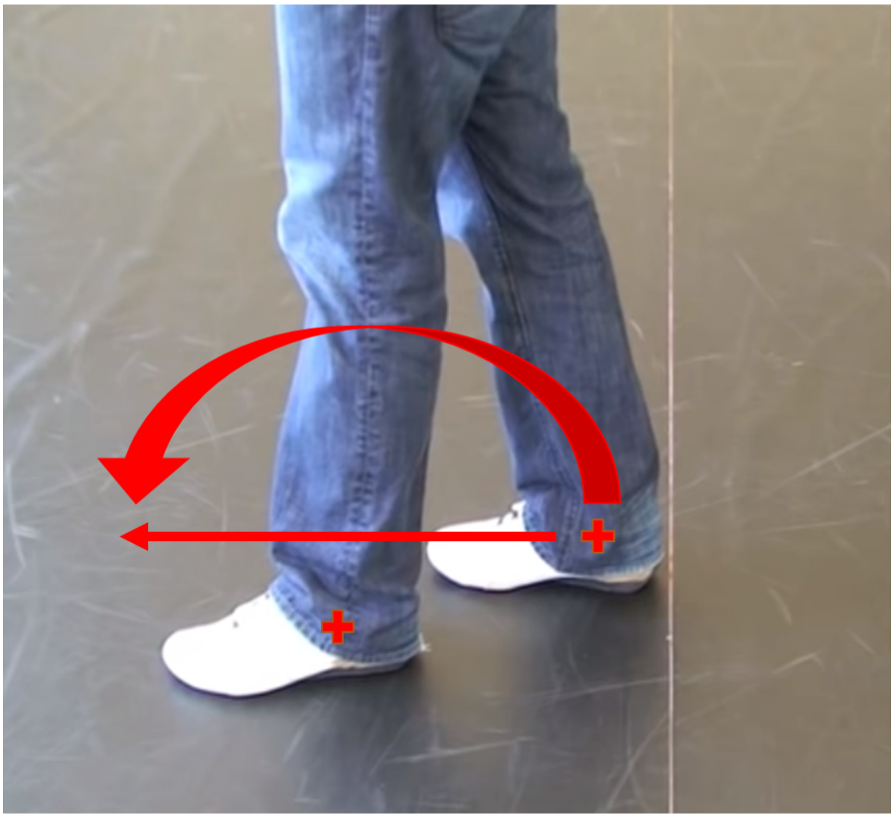
Figure 2.8 – Abstraction level as a dimension of motion data representation. Ranging from low level as 3D points with time to high level emotional state.

The several methods investigated give some clues to answer our research questions. In terms of dance skills (RQ4), the most important ones are the rhythm, capacity to dance in synchronization with the music, and the guidance, being able to lead the partner by giving the right physical information at the right time. More research is needed however to identify other important dance skills by asking experts in the field, like salsa teacher. For the dance evaluation (RQ5), we can use intermediate motion features from geometry, kinematics or anatomic as starting point and add temporal feature such as frequency analysis and trajectories. All motion features have to be relevant to the learning objectives of salsa dance and include the music synchronization as well as the partner interaction. This guidance skills is not really evaluated in the current literature. Our learning method (RQ6) can be reflective, generative and progressive as the virtual environment is a good place for taking the time to explore and practice with this new vision and feelings. It has to include repetition and increasing difficulty. A score can be provided as performance feedback. The virtual partner can be hands to hands (RQ2) with the user and then his upper body motion influenced by this connection. An haptic feedback such as vibration can be introduce to give additional timing or spatial information. Its lower body, especially the feet, have to move accordingly to the music (RQ1) to be a visual reference of rhythm. A specific gesture (RQ3) or motion can be detected in a given time as main information to change the dance pattern. The main limitation encountered in the methods from the literature is the lack of complete motion analysis entirely designed for music-synchronized partner dance performance evaluation.

In comparison to the previously mentioned approaches, our work tends to overcome the limitations by establishing a framework based on taking two persons dancing together with the music defined as the input entity for Analysis, indexing, and classification. The work is further set in the context of Salsa social dance. We propose a set of musical motion features that take into account the body motion of partners and the music rhythm to express dance skills and allow classification of dance performance upon the level of dance (Beginner, Intermediate and Experts). We created a big database of motion-captured Salsa performance in synchronization with music that contains sequences of 104 tracked points in time at 120Hz, allowing to test motion analysis on a bigger scale than seen in the related work. Finally, we built an interactive dance learning system based on virtual reality tools than includes Haptic real-time feedback with a very high degree of immersion, that has been tested at dance school and validated by a user study.

Chapter 3

Salsa motion modeling: building the Musical Motion Features



"Dancing is the poetry of the foot" John Dryden, poet.

3.1 Introduction

From the literature review of the previous chapter, two dance skills appears in experiments: rhythm and guidance. A more complete study is needed to ensure and validate the relevance of learned dance skills from salsa class. Also the subsequent performance evaluation regarding these dance skills requires adapted description of the recorded movements with motion features that are relevant to our context. Based on the several examples from previous work, we can investigate salsa motion to find the most adapted ones. This chapter starts with the identification of salsa dance skills via an interview of experts. Then, a selection of motion clues are discovered from the model of basic foot motion. Finally a set of adapted motion feature is defined and proposed to evaluate a dance performance.

3.2 Identification of the essential dance skills

As a first step, knowing what skill is important for Salsa dance is essential. To be able to analyze Salsa performances, we need to understand what makes a good performance, and so what high-level dance skills are improved when learning.

3.2.1 Salsa dance motion literature

A review of method books for learning salsa has been done to extract the basic pattern and motion profile of Salsa basic steps, as well as important dance knowledge related to dance skills.

3.2.2 Field study

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 recognized 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 following definitions: *Jury of international competition, Champion of International championships, Invited dancer in an 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 dance skills 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 required further explanations, available in the annexes C. Initially, the survey 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.

Table 3.1 – Selected criteria definitions and relative importance on a scale from 1 to 10

Dance skill	Definition	Importance
Rhythm	Being synchronized with the music's tempo. Dancing on the rhythm.	10
Guidance	Being able to lead / follow her/his partner.	7
Fluidity	Being able to move smoothly on the music.	6
Styling	Adding your own variation to the basic movement. Use of more hands gestures	5
Sharing	Being able to share the moment and enjoy the dance.	7
Musicality	Using your own dance movement with the music's variation.	3

3.2.3 Selected dance skills

A summary of the selected dance skills is shown on the table 3.1, ranked by overall importance. It is a fusion between our field study and information from our literature review. Substantial importance means that the criteria are essential for dancing, whereas little importance means to be less important (especially at a beginner level). Some of the proposed features are similar to the ones from literature, such as the "Dancing on the rhythm" and "Fluidity." Others are more specific to couple dancing, "Lead and Follow," and "Intention and Sharing." The final two features, "Style and variation" and "Musicality," are more related to the advanced Latin dancer. We describe here their meaning and the assumptions we make. Most important and prominent, dancing on the *rhythm* is essential for good practice. This means to have your footsteps in synchronization with the music such the two are aligned temporally. As a partner dance, two complementary roles are defined, the leader and the follower. While both partners are dancing simultaneously, the leader has to *guide* the follower who let himself follow indeed. As early beginners will not be sure of themselves and dance a little bit like a robot, the *fluidity* of movements is seen as being in control of her/his body. For more experienced dancers, the *styling* represents the ability to use the upper body part, such as arms, shoulders, and torso in an aesthetic manner. Correct also for other dances, having the pleasure to dance together and *share* the moment is crucial. Finally, playing with the *music* variation and your body moves also adds to the diversity of the dance.

3.3 Motion modeling: Seeking for spatiotemporal clues

3.3.1 Context

Salsa dance can appear to be very complex in terms of body movements. To help defining useful motion descriptors, an understanding of the spatiotemporal structure of salsa dance is needed. We consider the smallest entity of the dance structure, which corresponds to the eight beat window of any Salsa music. During this time, dancers have to move their feet following a specific pattern, more or less complex, depending on their skills. We consider among the most straightforward pattern possible, the one taught in Salsa class at the first lesson, the *Mambo* step. (Or basic step). In a second time, we also consider the *Cucaracha* step, very similar, and other dance patterns, to find common motion clues.

Although the full-body motion is involved during a dance, the feet motion is the most critical part for beginners as they need to master dancing on the rhythm. We consider after that only the movement of the feet. We also don't take complicated moves such as spins into account.

3.3.2 Mambo rhythm structure

The dance pattern (or dance step) *Mambo* consists of a forward-backward movement of the whole body, with the feet moving on the eight musical beats of the song. This pattern is described in figure 3.1.

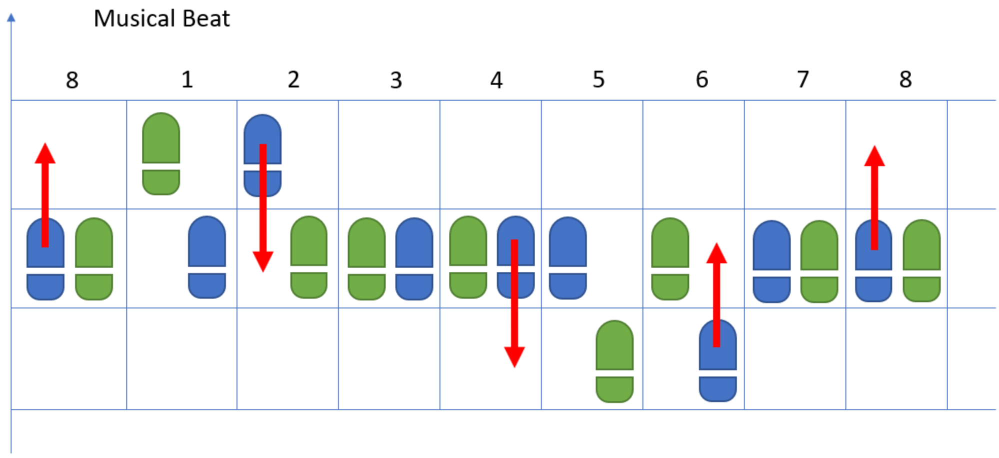


Figure 3.1 – Lead dancer's basic salsa step (*Mambo*). The red arrow when the foot is moving and the green coloured step for the body-weight. For the follower dancer, the step is mirrored. From the neutral position, the dancer steps forward then go back to neutral position, and then steps backward and back to neutral position.

We can observe from the figure 3.1: The right foot is moving between the beat 4 and 5, and also between beat 6 and 7, the left foot moving between the beat 8 and 1, and even between beat 2 and 3. There are only two possible positions in space for each of the feet.

For more accessible analysis, we consider the feet motion within a coordinate system with the X-axis along the forward-backward direction, the Z-axis in the course of both feet, and the Y-axis in the vertical direction of the person, as illustrated in figure 3.2.

The most important and significant movement of the feet is along the X-axis. We can draw a theoretical model of the feet motion based on their description, which we illustrated in figure 3.3. Please note that we represented the absolute value of each motion for practical illustration purposes. In natural movement, this graph would not have sharp discontinuities, which is to keep in mind. The displacement of each foot corresponds to a ramp in terms of distance. The feet move in the Y-axis is similar to this, but with lower amplitude, as the dancer lifts a little bit her/his foot off the ground when moving. The motion along the Z-axis is considered null.

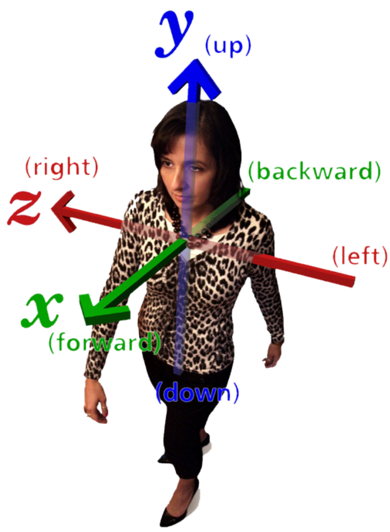


Figure 3.2 – X,Y and Z axis considered for our motion modeling

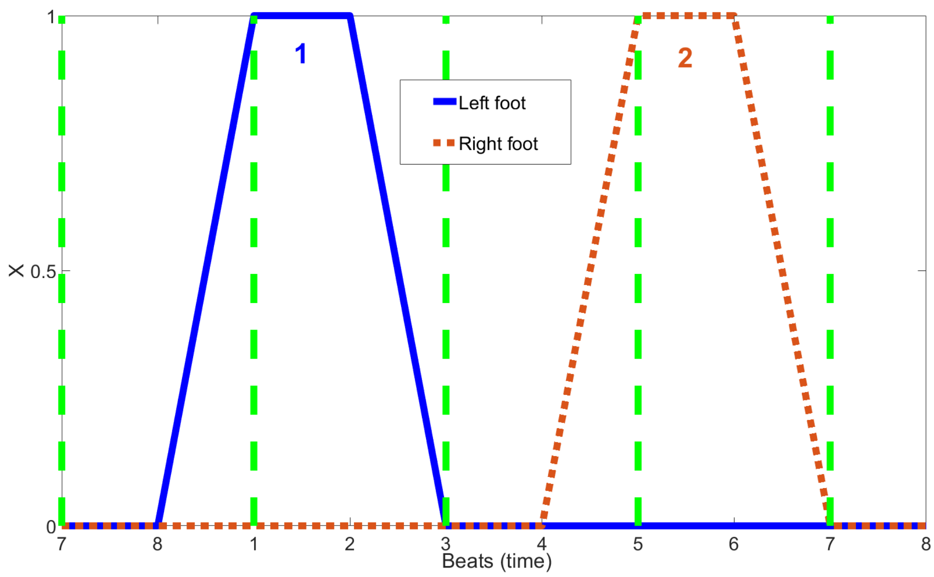


Figure 3.3 – Model of distance norm over time of both feet when dancing the basic step "mambo". The left foot makes two velocity peaks during the beat 8 and 2 whereas the right foot have velocity peaks for the beat 4 and 6.

The velocity profile is shown in figure 3.4. The position changes over time are translated in velocity peaks. We can observe that the peaks are located at a regular distance of a musical beat (by half a beat), making it possible to detect if the person moves her/his feet at the same time as the musical beat.

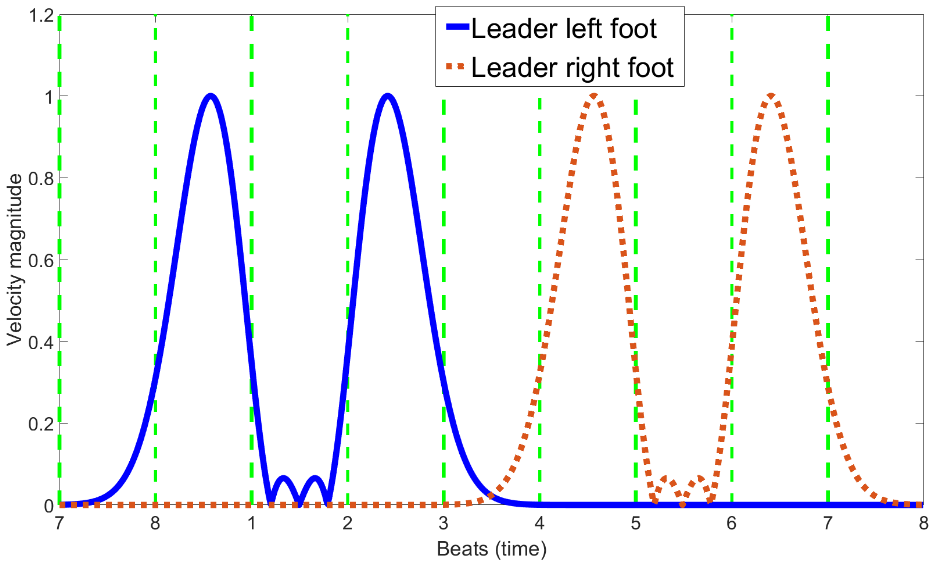


Figure 3.4 – Model of velocity norm over time of both feet when dancing the basic step "mambo". The left foot makes two velocity peaks during the beat 8 and 2 whereas the right foot have velocity peaks for the beat 4 and 6.

3.3.3 Cucaracha rhythm structure

The Cucaracha dance pattern is similar to the Mambo step in the way that the feet have to move in a right-left motion, as illustrated in figure 3.5.

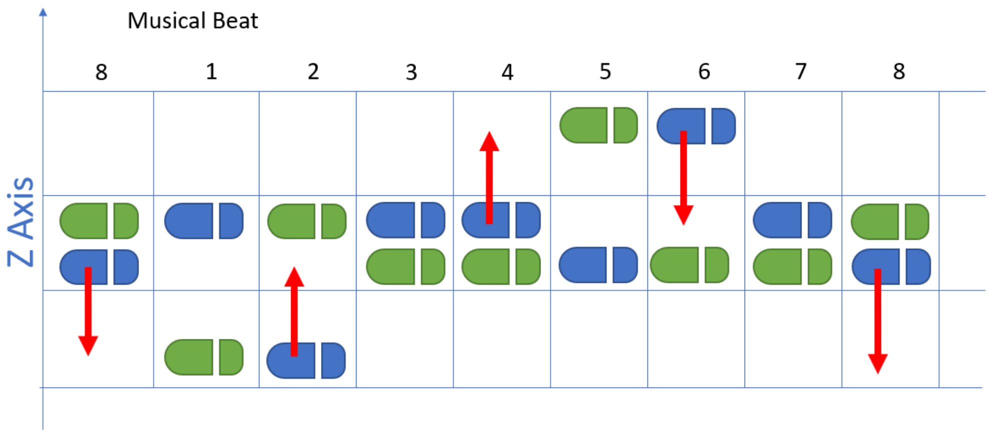


Figure 3.5 – Lead dancer's *Cucaracha* step. The red arrow when the foot is moving and the green coloured step for the body-weight. For the follower dancer, the step is mirrored. From the neutral position, the dancer steps forward then go back to neutral position, and then steps backward and back to neutral position.

Looking at the motion and velocity profile in time, we found the same peaks in the same position. Along this time, the Z-axis.

3.3.4 Selected motion clues

Other steps follow the same observation: The feet velocity profile would show peaks regularly near the musical beats in different directions. We assume here that if we take the velocity norm of each foot, the peaks related to the musical rhythm will be present whatever is the direction and is general for all basic steps. (As long as the alternate of feet is kept during the dance).

Table 3.2 – Selected motion clue definitions

Motion clue	Definition	Details
Velocity norm feet	Velocity norm profile over eight beat and peaks propriety	Refer to the alignment of the dancer moves to the musical beat
Hands Hips Position	Distance between hands and hips	Means if the person moves her/his arms
Power spectrum feet	Fourier transform of the feet merged motion data	The main frequency of the feet, so the rhythm

3.4 Musical Motion Features

In this section, we propose a set of music-related motion features that allows characterizing the motion of a dancer couple in terms of dance skills. Our approach is based on the motion clues from table 3.2 in the form of measurements μ_j that are meaningful in regards to corresponding dance skill from table 3.1. As a first approach, the three simplest skills are investigated here: Rhythm, Guidance, and Styling. The other ones have a more complex and fuzzy definition, which may also involve other domains of expertise (for example, Sharing may include some form of psychological analysis).

3.4.1 Dance skill: Rhythm

This dance skill means if the dancer is moving her/his feet according to the music's tempo. We consider the motion of each dancer (Leader and Follower) within the dance couple separately and assume that a skilled dancer is in perfect synchronization with the music tempo. We present two MMF that is related to this skill: firstly, the rhythmic accuracy of each foot on each musical beats and the average dancer tempo.

3.4.1.1 Feet rhythmic accuracy

The idea here is to measure the accuracy of the dancer's step for each musical beat based on our model. Within a time window of 8 musical beats, while performing a basic step, the feet velocity magnitude shows some peaks corresponding to the movement that is aligned in a regular way close to the musical beats.

As illustrated in the figure 3.6 and based on our model, we can compute the ideal time of the velocity magnitude of the foot as half a beat Δt before the time of the musical beat 1 $t(beat_1)$,

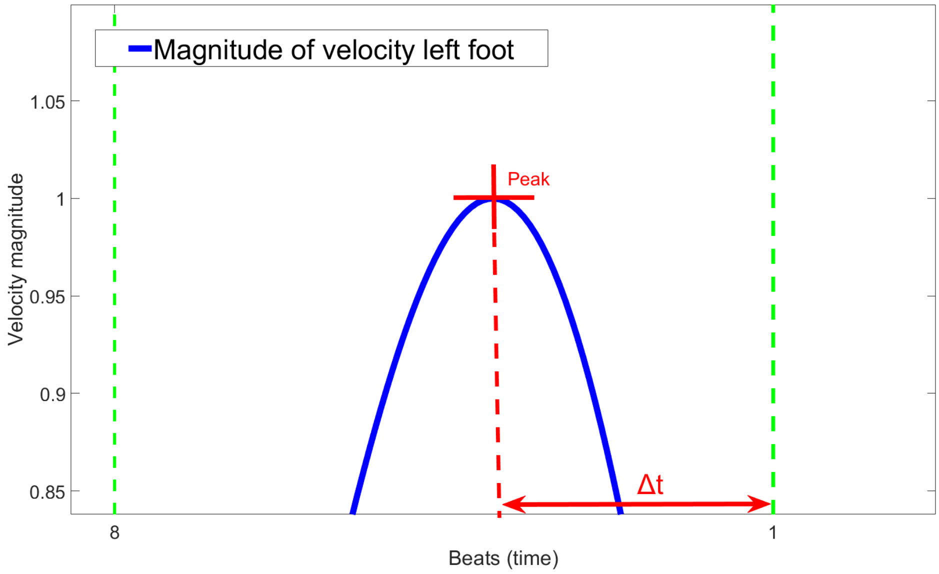


Figure 3.6 – First peak velocity magnitude of the left foot. Based on our model, the velocity magnitude peak occurs Δt before the musical beat one temporal location, where Δt is the length of a half-beat. The time difference between this theoretical ideal time and the real measured one gives us the rhythm accuracy.

in the context of 8 beat window. The rhythm accuracy is then the difference between this pre-computed time and the actual measurement of the peak's time, as written in the equation 3.1.

$$Rhythm\ accuracy(beat_1) = abs[t(velocity\ peak(beat_8\ to\ beat_1)) - t(beat_1) - \Delta t] \tag{3.1}$$

Two peaks occur for the left foot in the first half of the 8beat time window that helps to compute the accuracy for the beat 1 and 3 and two other peaks in the second half from the right foot, which leads to the accuracy for the beat 5 and 7. Until then, we considered only the steps of the Leader. Since the steps are mirrored for the Follower, we interchange the feet to the associated beat. This overall consideration leads to a series of 8 measurements, 4 for each partner, twice per foot indicated in the table 3.3.

Table 3.3 – Feet Rhythmic Accuracy's measurements

$\mu_1 = abs[t(velocity\ peak_{left\ foot\ leader}(beat_8\ to\ beat_1)) - t(beat_1) - \Delta t]$
$\mu_2 = abs[t(velocity\ peak_{left\ foot\ leader}(beat_2\ to\ beat_3)) - t(beat_3) - \Delta t]$
$\mu_3 = abs[t(velocity\ peak_{right\ foot\ leader}(beat_4\ to\ beat_5)) - t(beat_5) - \Delta t]$
$\mu_4 = abs[t(velocity\ peak_{right\ foot\ leader}(beat_6\ to\ beat_7)) - t(beat_7) - \Delta t]$
$\mu_5 = abs[t(velocity\ peak_{right\ foot\ follower}(beat_8\ to\ beat_1)) - t(beat_1) - \Delta t]$
$\mu_6 = abs[t(velocity\ peak_{right\ foot\ follower}(beat_2\ to\ beat_3)) - t(beat_3) - \Delta t]$
$\mu_7 = abs[t(velocity\ peak_{left\ foot\ follower}(beat_4\ to\ beat_5)) - t(beat_5) - \Delta t]$
$\mu_8 = abs[t(velocity\ peak_{left\ foot\ follower}(beat_6\ to\ beat_7)) - t(beat_7) - \Delta t]$

3.4.1.2 Average Tempo Accuracy

Another useful information is the mean tempo people are dancing on. A musical tempo is also expressed as a frequency (for example, 180BPM = 3Hz). Using the merged data from an entire performance of both right and left feet gives us a periodic signal that can be used with a signal processing technique. Using fast Fourier transform on that signal, we can access to a spectrum. The main frequency can be interpreted as the average tempo to which the dancer is dancing, as shown in figure 3.7, and be compared to the music tempo in the frequency domain.

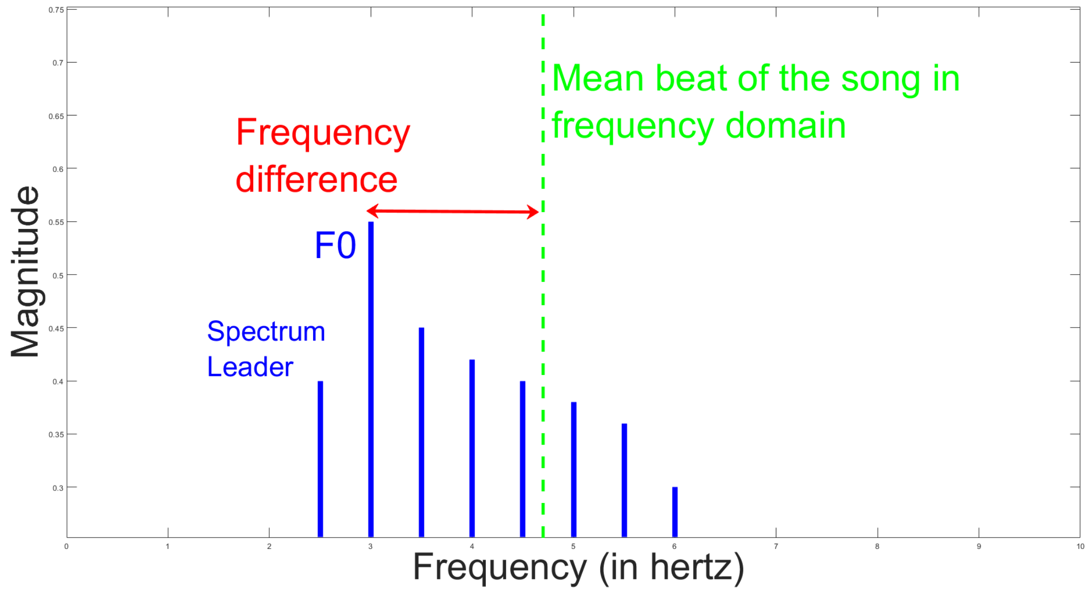


Figure 3.7 – Spectral analysis of the feet velocity magnitude. The main frequency that have the maximum magnitude can be compared to the music's tempo.

If we look at the difference between the tempo and the main frequency from the spectral analysis of the dancer, showed in equation 3.2, we can estimate how far the dancer is from the tempo, giving us the other indication on its overall performance regarding rhythm.

$$Average\ Tempo\ Accuracy(Hz) = abs(main\ frequency\ merged\ velocity\ feet(Hz)\ -\ tempo(Hz)) \tag{3.2}$$

We can measure this feature for each partner. We need to take into account the full performance that could last one song or less. We then merge the data of each foot and take the velocity. Using a fast Fourier transform, we can get the power spectrum and locate the main frequency f_0. Finally, we take the difference between the obtained frequency and the music's tempo, as written in the table 3.4.

Table 3.4 – Average tempo measurements

$\mu_9 = abs(FFT(velocity\ left\ foot_{leader} + velocity\ right\ foot_{leader})_{f0} - tempo(Hz))$
$\mu_{10} = abs(FFT(velocity\ left\ foot_{follower} + velocity\ right\ foot_{follower})_{f0} - tempo(Hz))$

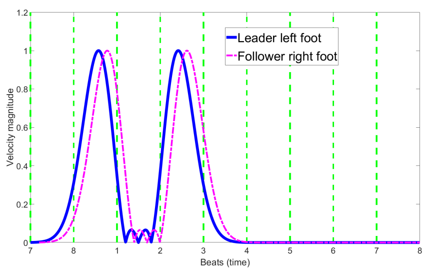
3.4.2 Dance skill: Guidance

This skill focuses more on the relationship between the two partners: the leader and follower. The leader has to guide the follower mechanically through the dance via gestures and anticipate what he is going to do, whereas the follower must respond to this stimulus accordingly. We assume that the Follower motion depends on the guidance from the Leader. Comparing both motion profile, thus gives us an idea of how connected the two partners are. Three features are proposed: the linear correlation of legs motion, the feet rhythmic difference, and the average tempo difference.

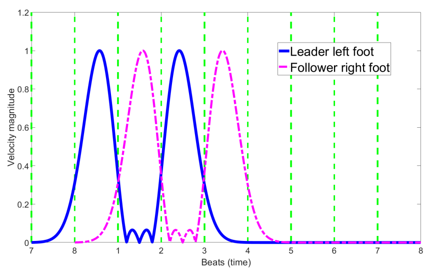
3.4.2.1 Linear correlation of legs motion

A first approach is to compute the linear correlation between the velocity profile of the feet of both dancers during an eight beat time frame window. As they are supposed to dance on the same rhythm, their foot is in perfect synchronization (Illustrated in figure 3.8), and therefore describes the same motion profile. Considering the first beat, the calculation is made between the left foot of the leader and the right foot of the follower, as written in the equation. Given that A is the feet velocity of the follower and B the feet velocity of the leader, $\overline{A}$ and $\overline{B}$ the average of these values, and m,n the points of the data, we can write the equation 3.3.

$$L.C.\ Coefficient = \frac{\sum_m\sum_n(A_{mn}-\overline{A})(B_{mn}-\overline{B})}{\sqrt{(\sum_m\sum_n(A_{mn}-\overline{A})^2)(\sum_m\sum_n(B_{mn}-\overline{B})^2)}} \tag{3.3}$$



(a) High linear correlation of follower and leader motion foot motion profile. They are moving simultaneously.



(b) Low linear correlation of follower and leader motion foot motion profile. The follower motion is delayed.

Figure 3.8 – Linear correlation examples on beat 1 and 3

Having a high correlation between the two partner's feet is a sign of good guidance skills. We compute this linear correlation coefficient for the beat 1 and 3, which corresponds to the motion of left foot for leader and right foot for the follower and for the beat 5 and 7 which corresponds to the motion of the right foot for the leader and the left foot for the follower. This gives us finally two measurements, written in the table 3.5.

Table 3.5 – Linear Correlation's measurements

$\mu_{11} = Corr(Velocity\ Leader\ left\ foot, Velocity\ Follower\ right\ foot)$
$\mu_{12} = Corr(Velocity\ Leader\ right\ foot, Velocity\ Follower\ left\ foot)$

3.4.2.2 Feet rhythmic difference

More precisely, we can estimate the foot time difference on each beat, similarly to section 3.3.1.1. The rhythm of the follower and leader can have some minor differences due to the different role of anticipating and responding to the music, as illustrated in figure 3.9. It can serve as a relevant feature for the guidance skill (also due to natural imprecision). Even though the couple is not dancing perfectly on the rhythm, their synchronization as a result of guidance skill induces that they show the same rhythm accuracy. To measure this, we compare velocity peaks time location between the leader and the follower related feet for the beat 1, 3, 5, and 7 would indicate if they are both synchronized, as written in the equation 3.4.

$$Feet\ rhythmic\ difference(beat_1) = abs[t(velocity_{leader}\ peak(beat_8\ to\ beat_1)) - t(velocity_{follower}\ peak(beat_8\ to\ beat_1))] \quad (3.4)$$

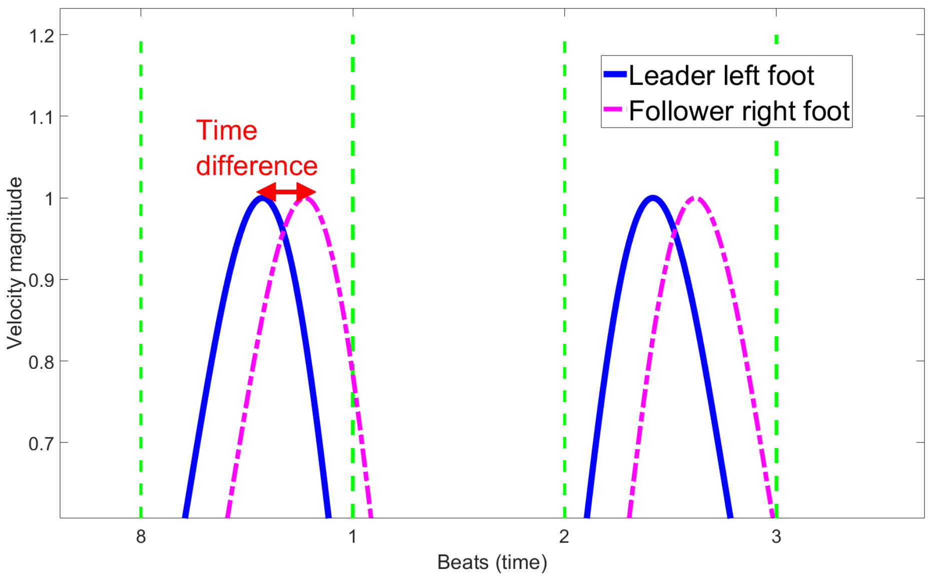


Figure 3.9 – Time difference between partners velocity magnitude of the feet

To measure this feature, we compute the time difference between the peaks of velocity associated with each beat (1, 3, 5, and 7). We assume that a value of 0 means a perfect synchronization of the feet, and therefore a good skill. The measurements are summarized in table 3.6.

Table 3.6 – Feet Rhythmic Difference's measurements

$\mu_{13} = abs[t(velocity\ peak_{left\ foot\ leader}\ (beat_8\ to\ beat_1)) - t(velocity\ peak_{right\ foot\ follower}\ (beat_8\ to\ beat_1))]$
$\mu_{14} = abs[t(velocity\ peak_{left\ foot\ leader}\ (beat_2\ to\ beat_3)) - t(velocity\ peak_{right\ foot\ follower}\ (beat_2\ to\ beat_3))]$
$\mu_{15} = abs[t(velocity\ peak_{right\ foot\ leader}\ (beat_4\ to\ beat_5)) - t(velocity\ peak_{left\ foot\ follower}\ (beat_4\ to\ beat_5))]$
$\mu_{16} = abs[t(velocity\ peak_{right\ foot\ leader}\ (beat_6\ to\ beat_7)) - t(velocity\ peak_{left\ foot\ follower}\ (beat_6\ to\ beat_7))]$

3.4.2.3 Average tempo difference

The mean tempo each partner is dancing, and the difference give a piece of important information about their musical coupling from a global point of view. We can extract the velocity magnitude of the combined feet for each partner and apply a fast Fourier transform to get the main frequency of the main periodic signal, that corresponds to their main movements of the dance (Illustrated in figure 3.10). Comparing the main frequency from both partners allows seeing if they are dancing on the same rhythm, at a global point of view, as written in the equation 3.5.

$$Average\ Tempo\ Accuracy(Hz) = abs(main\ frequency\ merged\ velocity\ feet_{leader}(Hz) - main\ frequency\ merged\ velocity\ feet_{follower}(Hz)) \quad (3.5)$$

This feature has only one measurement, that is the frequency difference between the main frequencies of the leader and the follower, as written in the table 3.7.

Table 3.7 – Average tempo difference's measurement

$\mu_{17} = abs(FFT(velocity\ left\ foot_{leader} + velocity\ right\ foot_{leader})_{f0} - FFT(velocity\ left\ foot_{follower} + velocity\ right\ foot_{follower})_{f0})$

3.4.3 Dance skill: Styling

This dance skill expresses the aesthetic variation that is mastered by more advanced dancers. An important point is the styles and variations expressed by the dancers within their dance. Multiple criteria can be taken into account, as shown in the several successful dance style studies using LMA [68, 7], as the spatial aspect or stylistic variation of the movements. Among them, three features have been proposed for investigation: The net velocity change, the area covered, and the quantity of hand movement.

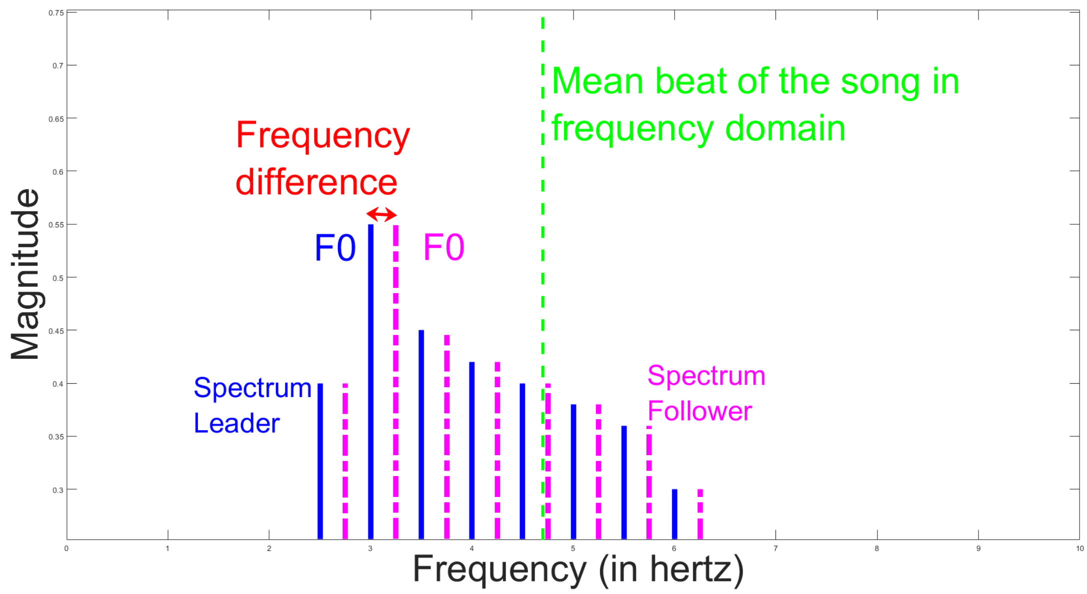


Figure 3.10 – Spectrum of the feet velocity from the leader and the Follower. The main peak corresponding to their dance tempo can be compared.

3.4.3.1 Net Velocity change

The net velocity change brings information about the effort put while dancing. Assuming that dancers are trying to follow the rhythm, having more effort is an indicator of good styling skills. We can compute it by taking the integral of the acceleration of the feet, as illustrated in figure 3.11, and written in the equation 3.6.

$$Net\ Velocity\ Change = \int Acceleration\ dt \tag{3.6}$$

As measurements, we can compute the feature for each foot and each partner, which gives us four measurements, written in the table 3.8.

Table 3.8 – Net Velocity change's measurements

$\mu_{18} = \int Acceleration_{left\ foot\ leader}\ dt$
$\mu_{19} = \int Acceleration_{right\ foot\ leader}\ dt$
$\mu_{20} = \int Acceleration_{left\ foot\ follower}\ dt$
$\mu_{21} = \int Acceleration_{right\ foot\ follower}\ dt$

3.4.3.2 Area covered

At a fast tempo, doing wide steps while dancing on the rhythm means good effort management. As well at a slow tempo, maintaining small steps is also a sign of good effort management. The area covered by the dancer within a time range can help differentiate the level of expertise, as part of the style component (Illustration in figure 3.12). To compute it, the integration of the

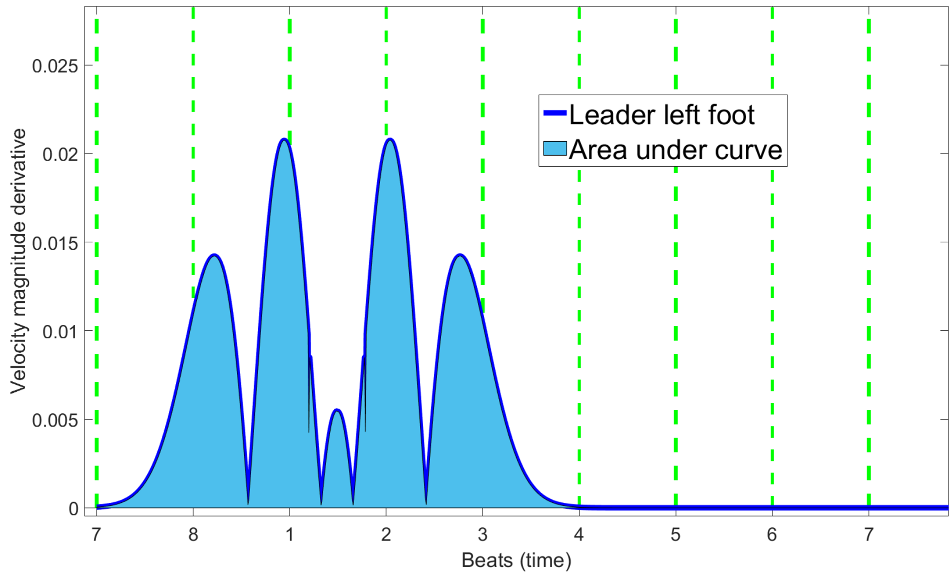


Figure 3.11 – Area under the curve of the acceleration magnitude for the dancer foot. This give an insight on the effort put in movements while dancing.

derivative of velocity is taken, as written in the equation 3.7.

$$Area\ Covered = \int Velocity\ dt \tag{3.7}$$

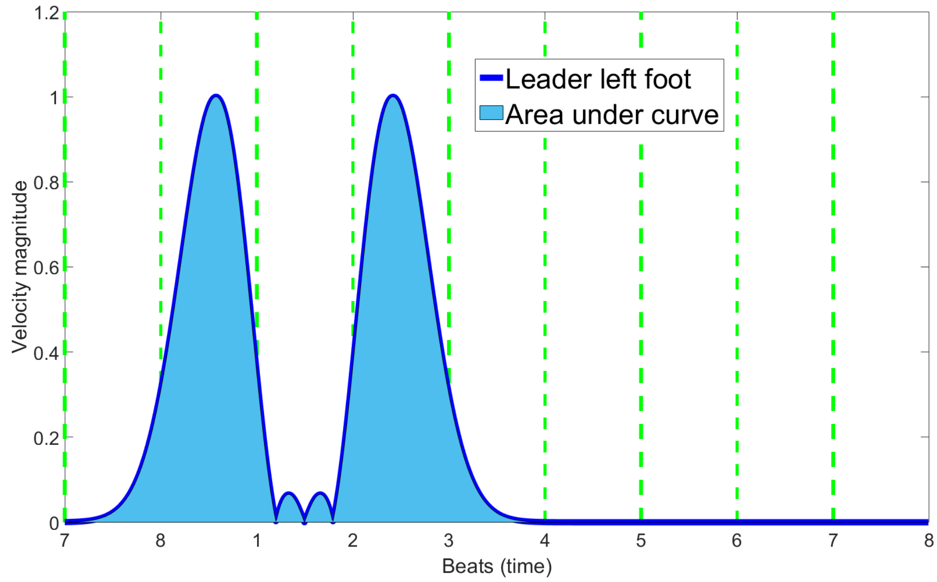


Figure 3.12 – Area under the curve of the velocity magnitude for the dancer foot. This give an insight on the area covered by the dancer while dancer.

For measurements, we can compute the features for each foot and each partner, which gives us four measurements, written in table 3.9.

Table 3.9 – Area Covered's measurements

$\mu_{22} = \int Velocity_{left\ foot\ leader}\ dt$
$\mu_{23} = \int Velocity_{right\ foot\ leader}\ dt$
$\mu_{24} = \int Velocity_{left\ foot\ follower}\ dt$
$\mu_{25} = \int Velocity_{right\ foot\ follower}\ dt$

3.4.3.3 Hands movement

Oppositely to the feet that are strongly connected to the tempo, hands movements, and specifically freehand movements (When dancer don't hold hands for guidance) are part of styling aesthetic as illustrated in the figure 3.13.

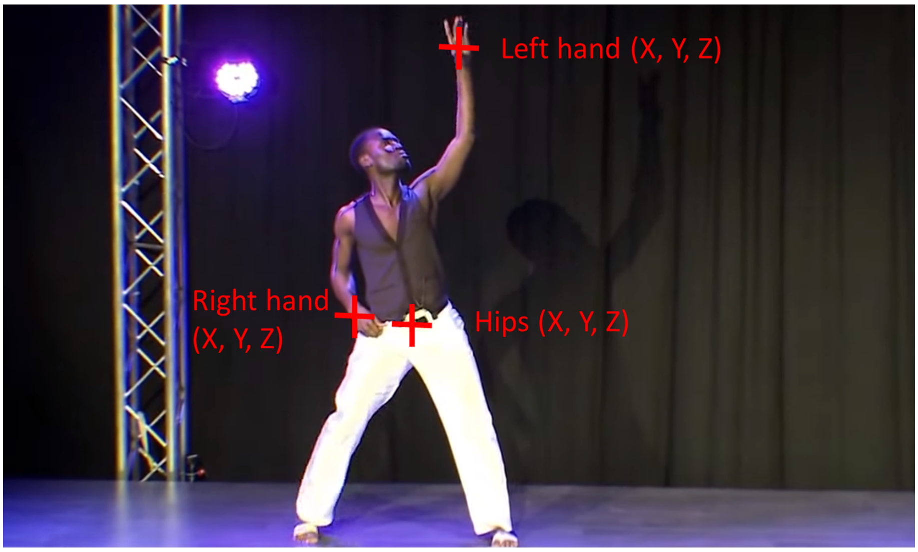


Figure 3.13 – Example of professional salsa dancers having an aesthetic pose. Courtesy of Kouame Dancefloor.

During salsa dance, the movement of the upper limbs is also vital for styling, additionally to the guidance action. More precisely, the hand can be lifted high for aesthetic effect. This is taken into account through the proposed feature of hands movement. As in the position of hands-to-hands, the height of hands is mostly steady; looking at the hips-to-hand distance allows for tracking the free-hands movements. For this feature, the average Euclidean distance between hand and hips during the dance is considered. The 3D location of the hips and both hands are taken, and then the gap between hips and each hand is computed, as shown in equations 3.8.

$$Handsmovement = Average_{8beat}\left(\sqrt{(X_{hand} - X_{hips})^2 + (Y_{hand} - Y_{hips})^2 + (Z_{hand} - Z_{hips})^2}\right) \quad (3.8)$$

For measurements, we can compute the features for each hand and each partner, which gives us four measurements, written in the table 3.10. We are computing here the distance between two points, but since it is averaged on a certain time window and since the basic position of dance is by holding hands, any motion that is different would mean a movement from the hand, therefore a way for use to calculate hands movement.

Table 3.10 – Hands movement's measurements

$\mu_{26} = (\sqrt{(X_{Rhand} - X_{hips})^2 + (Y_{Rhand} - Y_{hips})^2 + (Z_{Rhand} - Z_{hips})^2})_{Leader}$
$\mu_{27} = (\sqrt{(X_{Lhand} - X_{hips})^2 + (Y_{Lhand} - Y_{hips})^2 + (Z_{Lhand} - Z_{hips})^2})_{Leader}$
$\mu_{28} = (\sqrt{(X_{Rhand} - X_{hips})^2 + (Y_{Rhand} - Y_{hips})^2 + (Z_{Rhand} - Z_{hips})^2})_{Follower}$
$\mu_{29} = (\sqrt{(X_{Lhand} - X_{hips})^2 + (Y_{Lhand} - Y_{hips})^2 + (Z_{Lhand} - Z_{hips})^2})_{Follower}$

3.5 Summary of the MMF

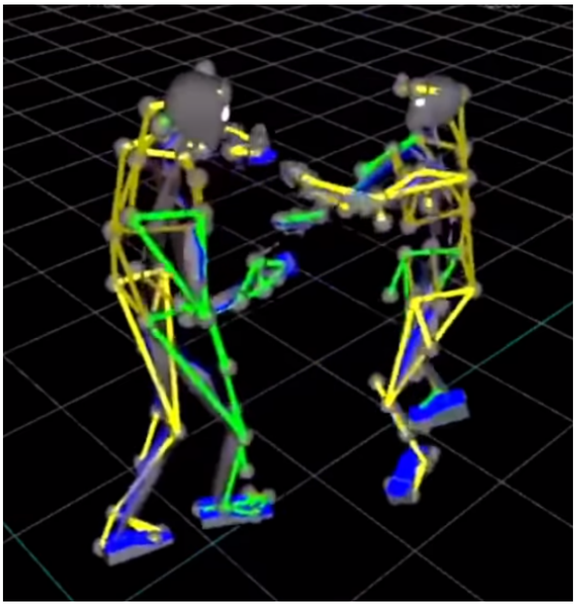
In total, eight features are proposed: two features related to rhythm, three related to guidance, and three related to style elements, for a sum of twenty nine measurements μ_j. A summary is available in the table 3.11. These features allow to characterize a salsa dance performance in terms of its main dance skills components, that are computed from the motion data of both partner in relation with the music.

Table 3.11 – Summary of the proposed features, μ_j is the feature measurement number associated to the MMF and dancing skill.

μ_j	*Skill*	*MMF*	Definition
μ_1	Rhythm	Feet rhythmic accuracy	$Abs[t(vel.\ peak_{L.F.\ leader}(beat_1)) - t(beat_1) - \Delta t]$
μ_2	Rhythm	Feet rhythmic accuracy	$Abs[t(vel.\ peak_{L.F.\ leader}(beat_3)) - t(beat_3) - \Delta t]$
μ_3	Rhythm	Feet rhythmic accuracy	$Abs[t(vel.\ peak_{R.F.\ leader}(beat_5)) - t(beat_5) - \Delta t]$
μ_4	Rhythm	Feet rhythmic accuracy	$Abs[t(vel.\ peak_{R.F.\ leader}(beat_7)) - t(beat_7) - \Delta t]$
μ_5	Rhythm	Feet rhythmic accuracy	$Abs[t(vel.\ peak_{R.F.\ follower}(beat_1)) - t(beat_1) - \Delta t]$
μ_6	Rhythm	Feet rhythmic accuracy	$Abs[t(vel.\ peak_{R.F.\ follower}(beat_3)) - t(beat_3) - \Delta t]$
μ_7	Rhythm	Feet rhythmic accuracy	$Abs[t(vel.\ peak_{L.F.\ follower}(beat_5)) - t(beat_5) - \Delta t]$
μ_8	Rhythm	Feet rhythmic accuracy	$Abs[t(vel.\ peak_{L.F.\ follower}(beat_7)) - t(beat_7) - \Delta t]$
μ_9	Rhythm	Average tempo	$Abs(FFT(vel._{L.F.\ leader} + vel._{R.F.\ leader})(f_0) - tempo(Hz))$
μ_{10}	Rhythm	Average tempo	$Abs(FFT(vel._{L.F.\ follower} + vel._{R.F.\ follower})(f_0) - tempo(Hz))$
μ_{11}	Guidance	Linear Correlation	$Corr(vel._{L.F.\ leader}, vel._{R.F.\ follower})$
μ_{12}	Guidance	Linear Correlation	$Corr(vel._{R.F.\ leader}, vel._{L.F.\ follower})$
μ_{13}	Guidance	Feet Rhythmic difference	$Abs[t(vel.\ peak_{L.F.\ leader}\ (beat_1)) - t(vel.\ peak_{R.F.\ follower}\ (beat_1))]$
μ_{14}	Guidance	Feet Rhythmic difference	$Abs[t(vel.\ peak_{L.F.\ leader}\ (beat_3)) - t(vel.\ peak_{R.F.\ follower}\ (beat_3))]$
μ_{15}	Guidance	Feet Rhythmic difference	$Abs[t(vel.\ peak_{R.F.\ leader}\ (beat_5)) - t(vel.\ peak_{L.F.\ follower}\ (beat_5))]$
μ_{16}	Guidance	Feet Rhythmic difference	$Abs[t(vel.\ peak_{R.F.\ leader}\ (beat_7)) - t(vel.\ peak_{L.F.follower}\ (beat_7))]$
μ_{17}	Guidance	Average tempo difference	$Abs(FFT(vel._{L.F.\ leader} + vel._{R.F.\ leader})(f_0) - FFT(vel._{L.F.\ follower} + vel._{R.F.\ follower})(f_0))$
μ_{18}	Styling	Net Velocity change	$\int Acceleration_{left\ foot\ leader}\ dt^2$
μ_{19}	Styling	Net Velocity change	$\int Acceleration_{right\ foot\ leader}\ dt^2$
μ_{20}	Styling	Net Velocity change	$\int Acceleration_{left\ foot\ follower}\ dt^2$
μ_{21}	Styling	Net Velocity change	$\int Acceleration_{right\ foot\ follower}\ dt^2$
μ_{22}	Styling	Area covered	$\int Velocity_{left\ foot\ leader}\ dt$
μ_{23}	Styling	Area covered	$\int Velocity_{right\ foot\ leader}\ dt$
μ_{24}	Styling	Area covered	$\int Velocity_{left\ foot\ follower}\ dt$
μ_{25}	Styling	Area covered	$\int Velocity_{right\ foot\ follower}\ dt$
μ_{26}	Styling	Hands movement	$(Euclidean\ distance(Right\ hand, Hips))_{Leader}$
μ_{27}	Styling	Hands movement	$(Euclidean\ distance(Left\ hand, Hips)_{Leader}$
μ_{28}	Styling	Hands movement	$(Euclidean\ distance(Right\ hand, Hips)_{follower}$
μ_{29}	Styling	Hands movement	$(Euclidean\ distance(Left\ hand, Hips)_{follower}$

Chapter 4

Musical Motion features evaluation



"To be fond of dancing was a certain step towards falling in love." Jane Austen, writer.

4.1 Introduction

This chapter is focused on the validation of the proposed musical motion features and their ability to classify salsa dance performance into their level of learning (beginner to experts). Firstly, Since the lack of public databases that fulfill our needs, we setup a database of salsa dance performance synchronized with music with several couples having various learning level dancing on various tempos. Secondly we evaluate the proposed musical motion feature using our database and machine learning algorithm.

4.2 Music-synchronized motion database - SALSA

In order to evaluate our musical motion features and allow further investigation on salsa performance, we constructed a database of motion captures (position and rotation of the body's joints in 3D) of couples dancing the basic salsa move that is synchronized with music. Some publicly available databases, such as the one from Carnegie Mellon University [150] does contain samples of Salsa dance but in insufficient quantity and without the parameters that are our interest (varying tempo, music synchronization, specific steps).

4.2.1 Overview

To test the relevance of our proposed MMF measurements to characterize the motion of dance couples in terms of dance skills, we designed and made a database of recorded couples from different dance mastering levels that are performing basic Salsa moves. We assume that the dance skills are more developed with the level of experience of the dancers. We hypothesize that the MMFs are able to distinguish the three categories.

4.2.2 Design

To construct the database, multiple parameters have to be taken into account, summarized in the table 4.1.

Table 4.1 – Parameters of our database

Parameters	Proposed values
Type of music	Commercial music and computer-generated
Tempo of the music	Ranging from 100bpm to 280bpm
Types of salsa steps	Mambo, Cucaracha, Guapea
Sequence of dance	Basic steps sequence, turns and improvisation
Category of participants	Expert, Intermediate and Beginner
Number of participants	Ideally 30 couples, 10 of each categories
Motion capture technology	Vicon system, 52 markers per person via optical tracking

Commercial music induces more energy to the participants, whereas computer-generated music does feel more awkward to them. However, since its regularity, computer-generated music allows for tracking precisely the beats temporal location over time and access the musical information needed for the computation of some of the MMF measurements. The *variable tempo* was determined by performing a tempo study based on commercial salsa songs and music commonly used for teaching in dance schools as well as playlists of notorious salsa deejay. Centered around 180 BPM that is perceived as the most comfortable to dance to (refined by expert feedback), proposing a range of dance from 100 (Among the slowest Salsa songs) to 280 BPM (Among the fastest ones) by steps of 20 would reveal how the participant behave in different situations. We assume that for extreme tempos, the more experienced dancers would show better dance skills. Following our modeling phase, we ask participants to perform basic steps, turns, and a sequence of improvisation to test if our modeling still stands in a more complicated situation. Each couple has to perform three basic steps, the *Mambo step*, the *Cucaracha step*, and the *Guapea step* (mirrored Mambo step for the leader), before an improvisation part. The full sequence is constituted in four mambos, four Cucaracha, four Guapea, two cross-body motion, and finally, a sequence of improvisation. The level of dancers have been determined according to their experience: the experts are dance school directors and teachers (As illustrated in the figure 4.1), beginners started to dance less than six months ago, and intermediates have more than one year and a half of dancing experience. This selection should show a wide span of dance skill mastering.

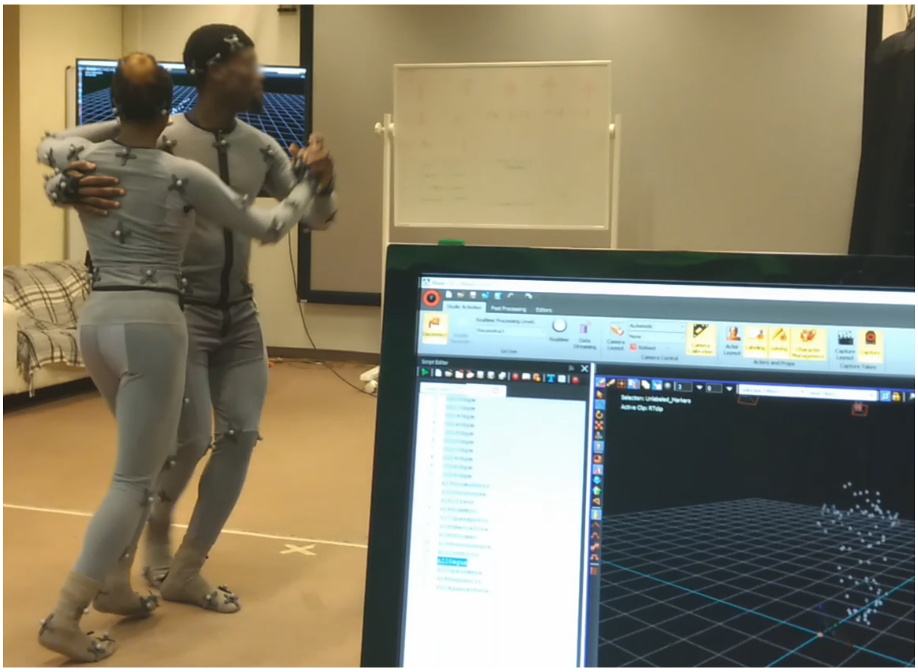


Figure 4.1 – Couple of experts performing our defined sequence

Ten couples in each category are considered. Since they are performing each of them 20 times, it would make a database of 600 sequences, big enough for statistical analysis. A Vicon motion capture system with eight cameras was used for recording (at 120fps). Especially at high tempos, the dancers will move fastly, hence the need for high-speed motion capture. The standard template from Vicon for the placement of the markers was used in each articulation of the body for a total of 52 markers per person.

4.2.3 Experiment

Twenty-six couples were recorded: 9 experts, 9 beginners, and 8 intermediates (Figure 4.2 illustrates the recording session with different couples).

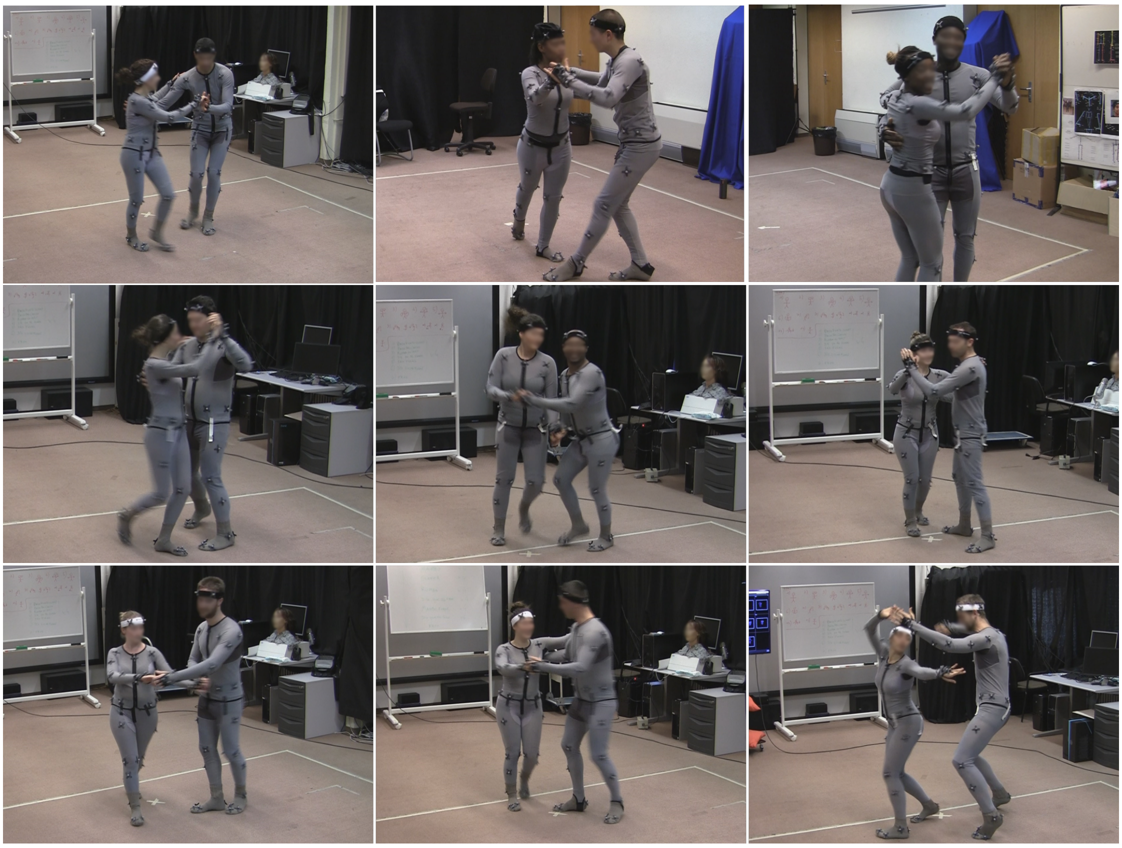


Figure 4.2 – Different couples dancing salsa basic steps

Each of the couples was asked to perform a sequence of dance, including the basic steps and an improvisation part. The commercial songs used for the recording are listed in the table 4.2. The computer-generated ones ranged from 100 to 280 BPM by steps of 20, with the same pattern.

Table 4.2 – Commercial music in their recording order

BPM	Song title
190	Ave maria lola - Sonora Carruseles
155	Velenosa - Alexander Abreu
172	Pasaporte - Alexander Abreu
212	Vehicle - Roberto Roena
166	Cambio del pie - Marc Anthony
220	Agua - Los Van Van
130	Como Duele - Maiyembe
254	Guarachera - Celia Cruz

It is essential to mention that the numerous occlusion occurring during certain dance moves (mainly the closed position) made the capture challenging (and especially the labeling process).

To ensure the exact and systematic synchronization of the music and the captured performance, the music was started simultaneously with the capture through the Vicon software interface. Each performer had to wear a special motion capture suit with markers on them, a headband, and special socks that could hold the markers. A short training was given to them to get familiar with the system, although the situation and the particular state of the ground (carpet) was not giving them help. The order of the danced songs was starting from the most comfortable tempo (180 BPM) and alternate progressively to the extreme, following the order in table 4.2. For each couple, the overall operation lasted two hours.

4.2.4 Post-processing

4.2.4.1 Motion processing

The raw data of the motion capture have to be processed into several steps. The first and most important one is labeling. For each marker tracked in space, the system has to recognize which one is which one and associate to it a label. Problems such as loose of the markers during some time can confuse the labeling, and the results are less accurate. On our data, the occlusion of some markers, especially at the upper part of the body, made the labeling task difficult and highlight why there is not so much database of people dancing together. Secondly, a segmentation of the sequences is done to separate the improvisation sequence to the basic steps sequence. Finally, two skeletons are extracted from the labeled points, named, and the data exported as .C3D (or .FBX) file.

4.2.4.2 Audio Processing

In parallel to the motion capture, it is needed to identify 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 4.3. After that, all the regular beats are marked from the music with the related labels along with the duration of the song. Due to the synchronization, we can directly compare point to point any temporal location of a musical event given by the music analysis with the motion data.

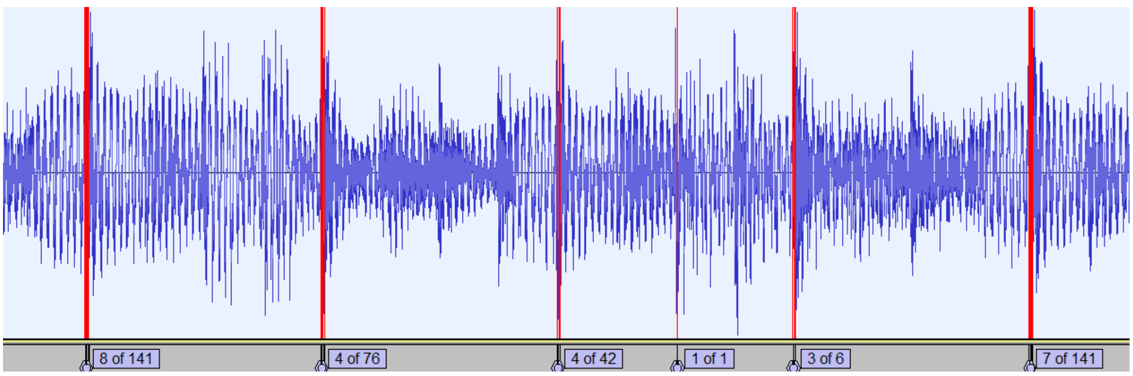


Figure 4.3 – Audio signal (purple) and detected beat (red)

As the audio structure is quite complex, due to the percussive music with a lot of different instruments, we use the fact that the instrument marking the first beat, the 'clave,' is also the loudest in terms of the 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 to get rid of noise artifacts. 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 time-stamps for each beat that is detected. This array is then imported into Matlab for further processing, together with the previously mentioned captured data.

4.2.5 Results

The result from the capture sessions is a database of 52 people as 26 couples, dancing averagely two minutes per songs, representing nearly 26 $couples\ x$ 10 $songs\ x$ 120 $sec\ x$ $120Hz = 3,700,000$ time frames of 104 points, in the shape of 260 motion sequences. Each sequence has been exported as two fully labeled skeletal entities in *C3D* formatted files, in synchronization with each music file and segmented for further processing in Matlab into two parts, namely the basic steps sequences and the improvisation sequence, leading to a grand final total of motion sequences of 520. In parallel, each music file has been post-processed to extract the timestamp of each steady musical beat (beat 1, 3, 5, and 7) along with the music, with a time dimension that matches the motion file one.

4.2.6 Qualitative validation

Although the full recorded motion sequences correspond to a repetition of the basic salsa pattern mambo, we choose an arbitrary time window of 8 beats from the data of an Expert (Tempo at 180 Bpm) for visualization. To see if our models were correct, we overlayed the foot motion velocity of leader and follower on the same time window. The expert's motion is illustrated in figure 4.4. This particular time window does represent the general motion graphs we obtain in our motion database. We can see that the curve of the follower seems delayed a little bit compare to the leader's one, which makes sense considering the guidance information that passes from one partner to another. The height of the peaks, which translates to the width of the step, is similar and also makes sense as partner dancing together would have the same step length. A zoom from the figure 4.4 is shown on the figure 4.5. Here the difference between the decreasing gradient of both motion curves can be explained as the time difference of both feet touching the ground.

Figure 4.6 shows an overlay of the foot velocity from three couples (one from each category: expert, intermediate, and beginner). The expert couple has more prominent peaks, meaning wider velocity span, hence broader movements. The beginner couple has smaller peaks height and also further peak location between leader and follower, which means less synchronization.

In figure 4.7, we can visualize the spectrogram of an expert couple leader and follower foot

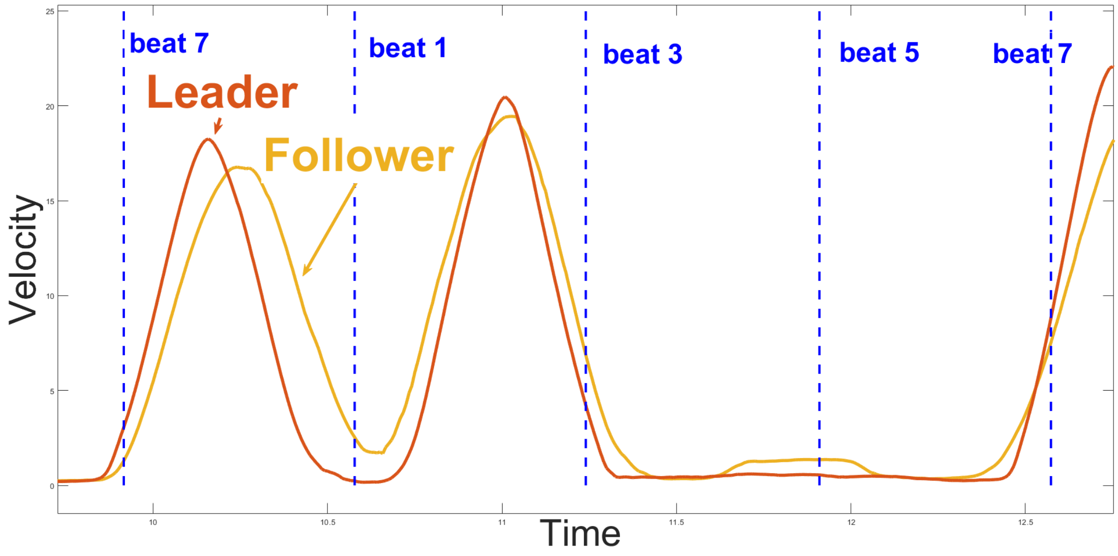


Figure 4.4 – Left foot velocity for the man (Leader - dark red) and right foot velocity for a woman (Follower - light yellow) of a couples during the basic step "Mambo". The blue vertical lines correspond to the detected beat from the audio file.

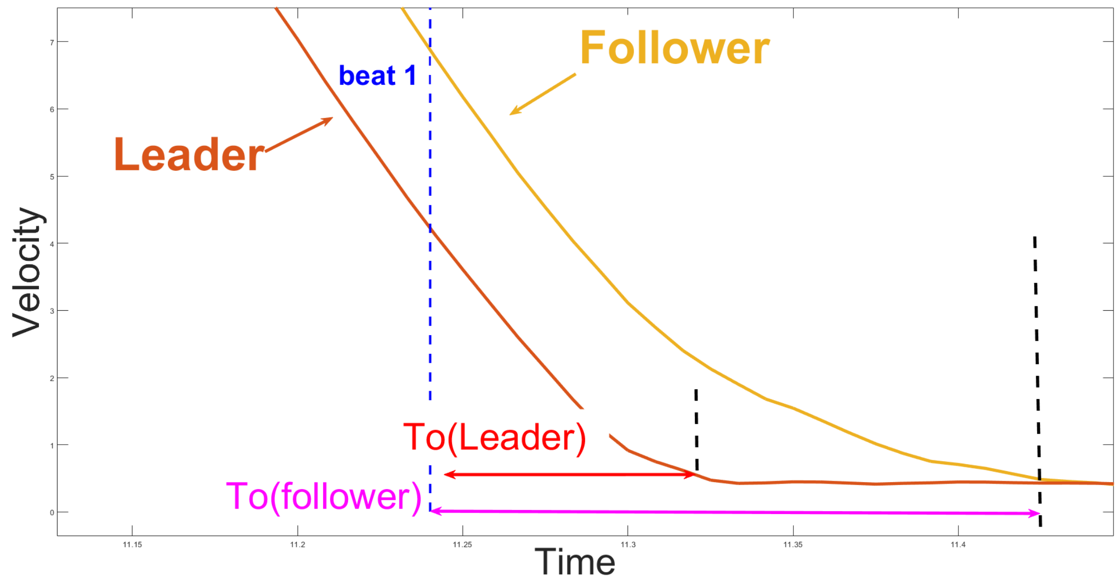


Figure 4.5 – Zoom of the velocities during the basic step "Mambo". The velocity reaches its minimum with a delay regarding the musical beat in blue, and is different for both partners.

velocity that is representative of what we can find within the database. We can see a certain number of peaks being the same for both partners, and that the highest peak is also the same. This corresponds to the main rhythm to which both partners are dancing.

As a last remark, we can see on the figure 4.8 the foot velocity of leader and follower during an arbitrary time window. The curve is not smoothed here and illustrate the needed processing steps to extract usable data. However, the main characteristic, follower being here delayed compared

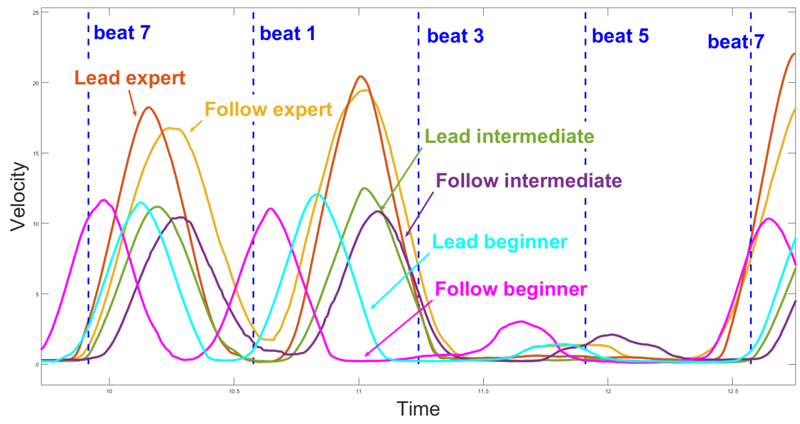


Figure 4.6 – Left foot velocity for the men (Lead - red, green and cyan) and right foot velocity for a women (Follow - yellow, purple and pink) of three couples during the basic step "Mambo". The blue vertical lines corresponds to the detected beat from the audio file.

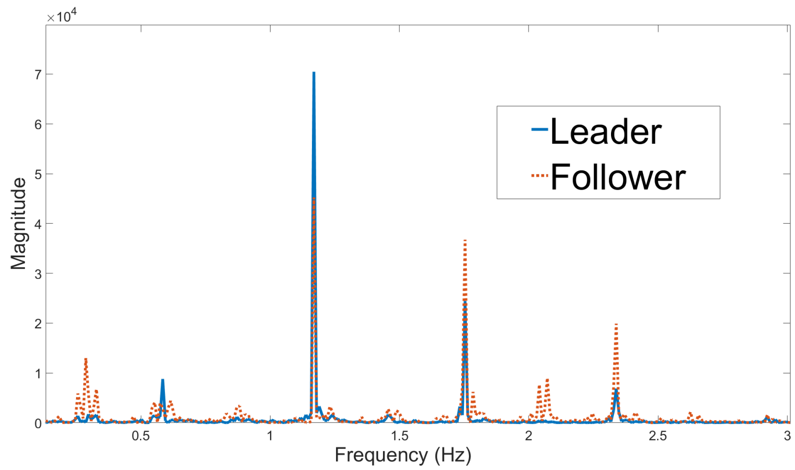


Figure 4.7 – Example of a frequency spectrum from both feet of the Follower and Leader during one dance. The main peak can be observed, as well as other peaks that are multiple of the main one. The frequency on the X-axis corresponds to different tempos

to the leader, is visible.

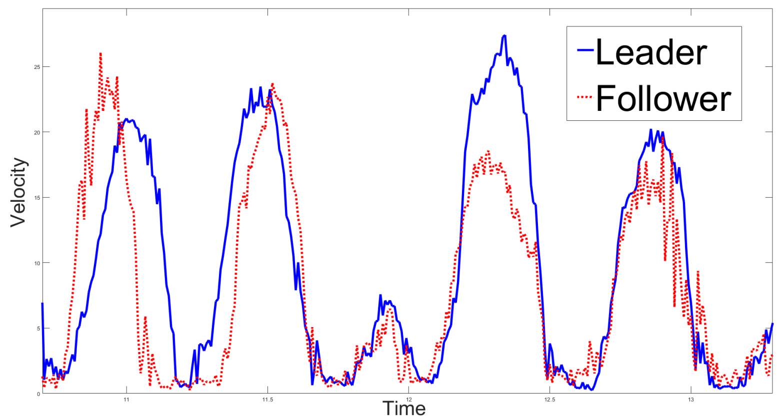


Figure 4.8 – Merged motion of both feet from Follower and Leader during four beats within a song

4.3 Evaluation of the MMF

4.3.1 Features extraction

Each motion file (recording) that corresponds to one song and one couple has been first segmented into two sequences for which we do the computation separately: the basic steps sequence and the improvisation sequence. Then, the three global features have been extracted (μ_9, μ_{10} and μ_{17}) for the entire sequence. For the remaining features, we are computing them using a sliding window of width equal to the time length of 8beats of the current song and moving on every eight beats. This gives us finally a vector of 29 values μ_j for each time window within each sequence. This choice ensures to compute our feature measurements on a period that makes sense with the music and, therefore, the dance requirements. The total number of computed feature measurements vector are:

- Basic steps sequence - 3780 samples.
- Improvisation sequence - 2650 samples.

An illustration of our processing pipeline is visualized in figure 4.9.

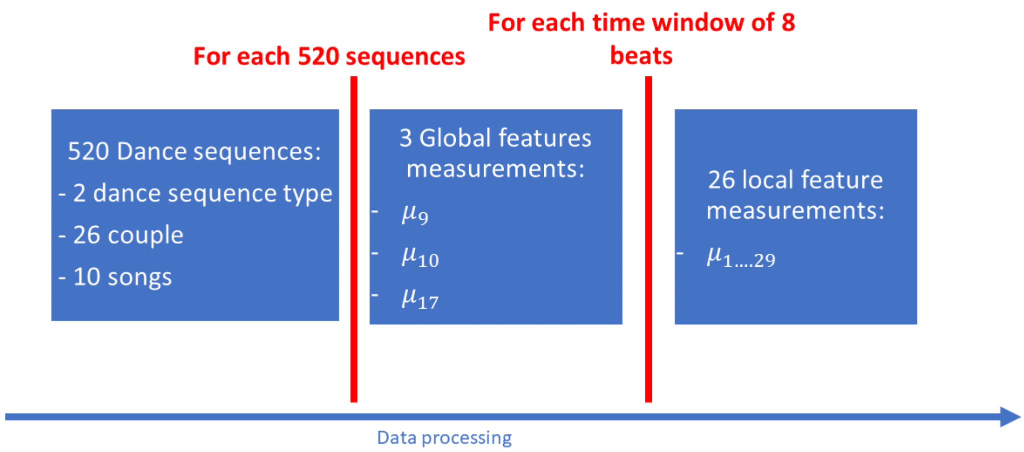


Figure 4.9 – Pipeline of the feature measurement μ_{1-29} extraction

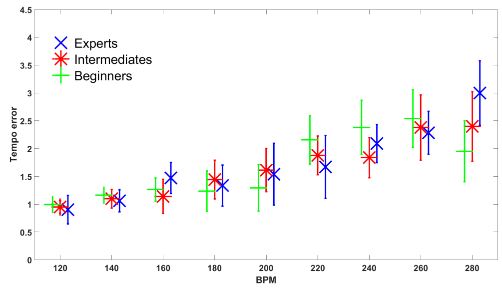
4.3.2 First data investigation: Tempo analysis & MMF visualisation

For this first data analysis, we compare the average value of the feature measurements per song and per category of dancers. The median was taken for each packet of samples for each song (each sequence of motion). The median indicates the relative distance among each song, which is even more relevant with higher gaps between each song. Then the average was taken per category (classified by level of expertise) for each tempo. We also computed the standard deviation within the groups. We aim at identifying the differences between groups of dancers regarding different tempos. All graph in results are available in the annexes.

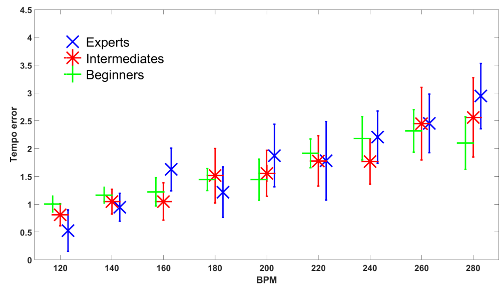
4.3.2.1 Dance Skill: Rhythm

Feet rhythmic accuracy μ_{1-8} This feature is measured by the time difference between foot velocity peak and beat location. The compiled results tempo per tempo are shown in the figure C.1 and C.2. We notice that the average value and the standard deviation are decreasing along with the tempo, on a steep slope from 120 to 180 Bpm, and then more light slope from 180 to 280 Bpm. This can be explained by dancers more comfortable to practice at a higher tempo. The standard deviation is reducing toward high tempo, more significantly for the beat 1 and 3. For the beat 5 and 7, the standard deviations stay constant. Since the first beat is the one playing reference for the rhythm, it makes sense that it is the most mastered. No major difference spotted between category here. It may be due to measurement imprecision.

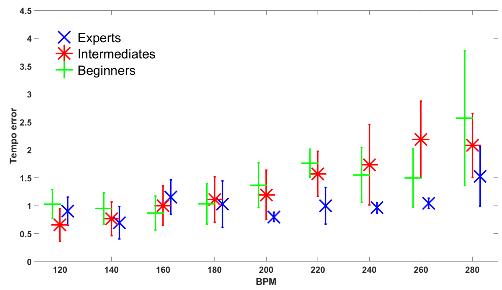
Average tempo accuracy μ_{9-10} This feature corresponds to the time difference between the mean frequency peak from the foot velocity spectrum and the mean beat of the song, which measures the average frequency of foot motion, so the average rhythm, shown in figure 4.10. Globally, the value increases with the tempo, as well as the standard deviation, meaning that maintaining the proper rhythm of reference is more difficult at a high tempo. We also notice that for the follower, a clear difference is visible depending on the category. Experts are more likely to have constituent rhythm than beginners.



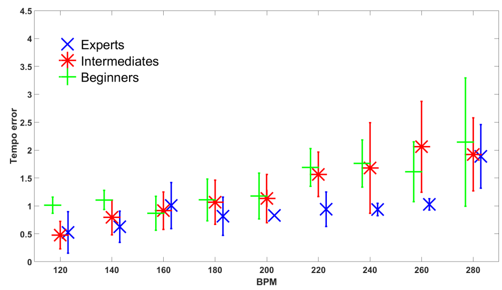
(a) Leader main foot velocity spectrum frequency for the basic sequence (μ_9)



(b) Leader main foot velocity spectrum frequency for the improvisation sequence (μ_9)



(c) Follower main foot velocity spectrum frequency for the basic sequence (μ_{10})



(d) Follower main foot velocity spectrum frequency for the improvisation sequence (μ_{10})

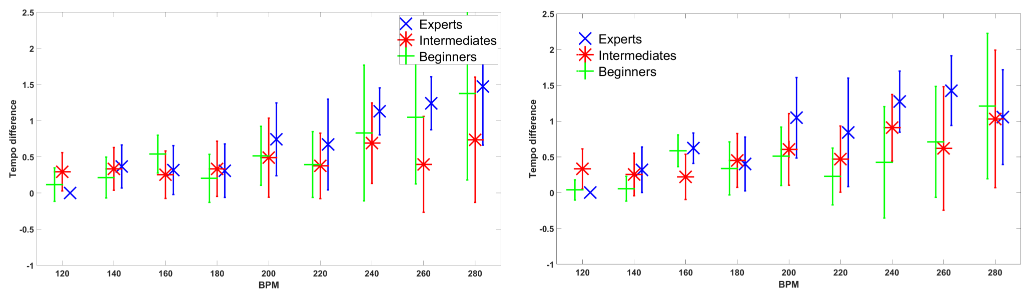
Figure 4.10 – Average values and standard deviation over each tempo of the main foot velocity spectrum frequency, as the average tempo, μ_{9-10}, averaged per category.

4.3.2.2 Dance skill: Guidance

Linear correlation of legs motion μ_{11-12} This feature is the linear correlation coefficient of foot velocity between each partner, visualized in the figure C.4. The standard deviation is very high, although a clear difference is visible in terms of average value between categories. The expert seems to have a smaller value than beginners.

Feet rhythmic difference μ_{13-16} This feature corresponds to the time difference between the velocity peaks of the foot motion and the musical beat, illustrated in figure C.5. We can see that the value is decreasing as the tempo is higher. This means that both dancers are more correlated as the tempo gets higher, and so their rhythm skill is converging. Also, here we can see two main phases: one between 120 and 180Bpm and another between 180 and 280Bpm.

Average tempo difference μ_{17} This feature is computed by taking the frequency difference between the main peak of the foot velocity spectrum from both partners, illustrated in figure 4.11. Here the value is increasing with the tempo. Also, the expert category seems to have higher frequency differences than other categories, meaning that there are more rhythmic differences between leader and follower. The standard deviation is increasing very much as the tempo gets higher, which means that users have difficulty to keep precisely the good tempo.



(a) Partners frequency difference between the main peak of the foot velocity spectrum for the basic steps sequence (μ_{17})

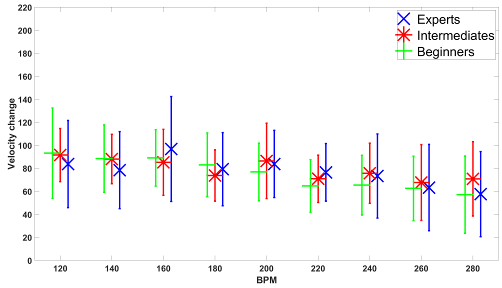
(b) Partners frequency difference between the main peak of the foot velocity spectrum for the improvisation sequence (μ_{17})

Figure 4.11 – Average values and standard deviation per category of foot velocity main spectrum frequency difference between the main peaks of leader and follower μ_{17}

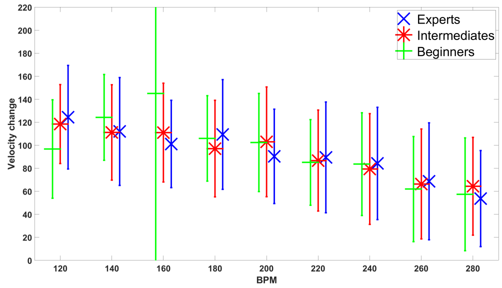
4.3.2.3 Dance skill: Styling

Net velocity change μ_{18-21} This feature is computed as the area under the foot acceleration curve under the time window range and illustrated in figure 4.12. Here we can globally see that the value is decreasing towards high tempos. Also, for μ_{21}, a clear difference is visible between categories in terms of average value and standard deviation. Experts can make more complex and stylized moves, which can explain this result.

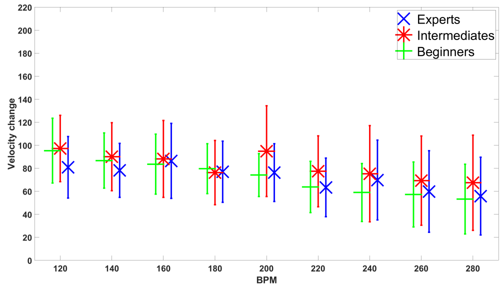
Area covered μ_{22-25} This feature is computed as the area under the foot velocity curve under the time window range and illustrated in figure C.8. The average value and standard deviation are smaller for the basic sequence, which makes sense as this sequence demands more control



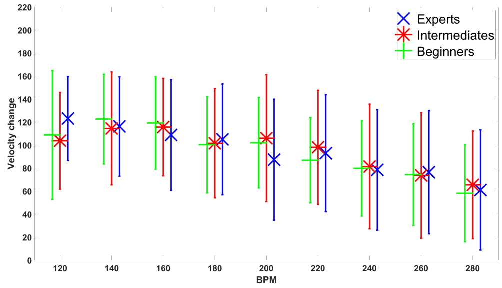
(a) Leader net left foot velocity change for the basic sequence (μ_{18})



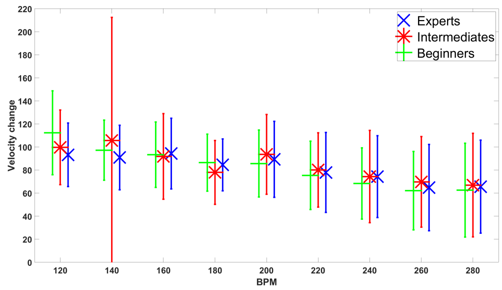
(b) Leader net left foot velocity change for the improvisation sequence (μ_{18})



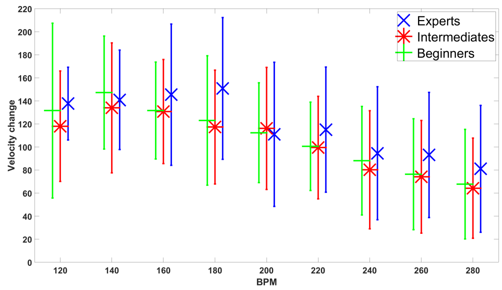
(c) Leader net right foot velocity change for the basic sequence (μ_{19})



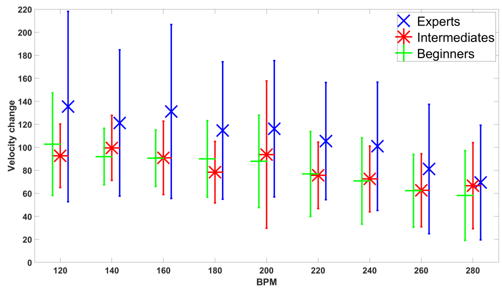
(d) Leader net right foot velocity change for the improvisation sequence (μ_{19})



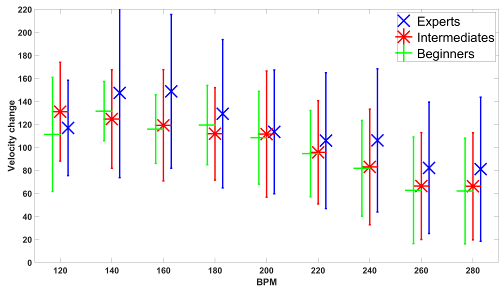
(e) Follower net right foot velocity change for the basic sequence (μ_{20})



(f) Follower net right foot velocity change for the improvisation sequence (μ_{20})



(g) Follower net left foot velocity change for the basic sequence (μ_{21})



(h) Follower net left foot velocity change for the improvisation sequence (μ_{21})

Figure 4.12 – Average values and standard deviation per category of the feet net velocity change μ_{18-21}

and precision from the dancer. Globally the value decreases towards high tempo. This follows the idea of making small movements to save energy as the rhythm increased. Experts seem to have a slightly bigger area covered than beginners, which can be explained by the most significant skills and energy-saving they have.

Mean distance hand to hips μ_{26-29} This feature is computed as the mean distance from the left and right hands to hips under the time window range and illustrated in figure C.9. The value and standard deviation are bigger for the improvisation sequence. Oppositely to many previous features, the value seems constant, but the standard deviation is increasing with tempo. This can be explained by the excitation of speed music to express more emotion with the upper body. For μ_{28-29}, a clear difference is visible along with categories.

4.3.2.4 Summary and conclusion

The differentiation between categories of dancers is not showing on all proposed features. Since it is a per-feature view, another study that takes the combination of features is needed. 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 appropriated and professional dancers would adapt their dance to Cha-cha-cha for example). Few main trends are noticed:

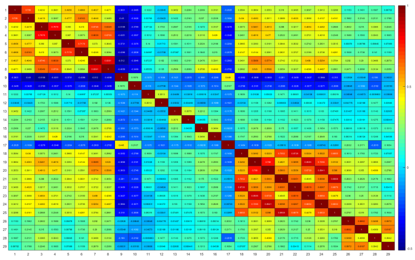
- Value and standard deviation increasing with tempo ($\mu_{9,10,17}$)
- Value and standard deviation decreasing with tempo ($\mu_{1-8,13-16,18-25}$)
- Difference between categories ($\mu_{10,11,12,15,16,21,28,29}$)

4.3.3 Second investigation: Correlation matrix among features

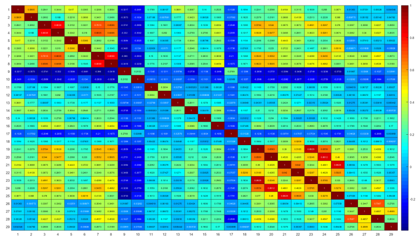
Visualizing the cross-correlation between all features per category allows seeing whether some features are more or less related depending on the experience level, giving insights on separating these categories. Quite evident for some features as they are closely related, it is still interesting to observe it globally across all features, which is shown in figure 4.13. We notice first packets of highly correlated features, especially for the basic step sequence, knowingly the μ_{18-25} ensemble that shows the most correlation for beginners. Characterizing Guidance, and more precisely, the synchronization of both partners, these results point out the variation induced naturally by more experienced dancer, oppositely to what we could expect. Also, the μ_{26-29} group seems very correlated, especially for experts and then intermediates. Here it seems in line with expectations of the more experienced dancer using more the upper body part. μ_{1-8} seems very correlated for the intermediate category. The three global feature $\mu_{9,10,17}$ are less correlated with other features, which make sense as they keep the same value for each sequence.

4.3.4 Third data investigation: Salsa performance classification

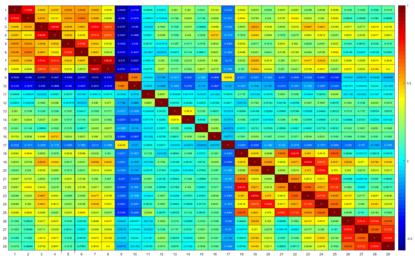
In this section we present the classification method and results from the extraction of three sequences: *(1)* the basic steps, *(2)* the improvisation sequence and *(3)* the combination of both. For each sequence, the data processing and segmentation produce a vector of 27 features (numbers) with a variable sample size between one thousand to eight thousand. This data is introduced into different multiclass classifiers to try distinguishing between firstly the three learning levels (beginner, intermediate, and expert) and in a second time between beginner and expert only.



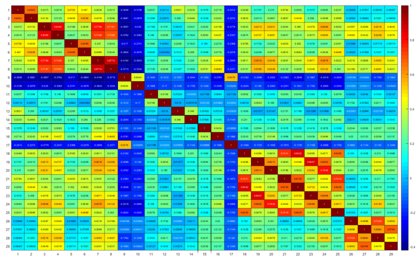
(a) Cross correlation for the beginner category during the basic steps sequence



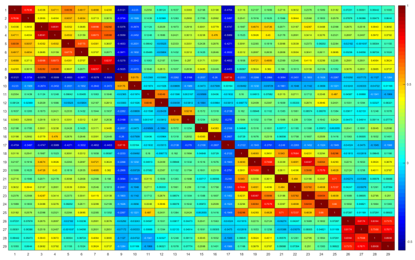
(b) Cross correlation for the beginner category during the improvisation sequence



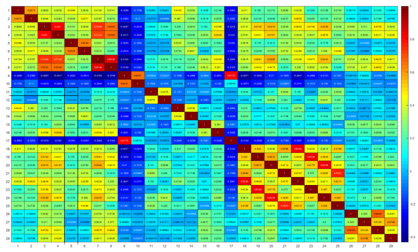
(c) Cross correlation for the intermediate category during the basic steps sequence



(d) Cross correlation for the intermediate category during the improvisation sequence



(e) Cross correlation for the expert category during the basic steps sequence



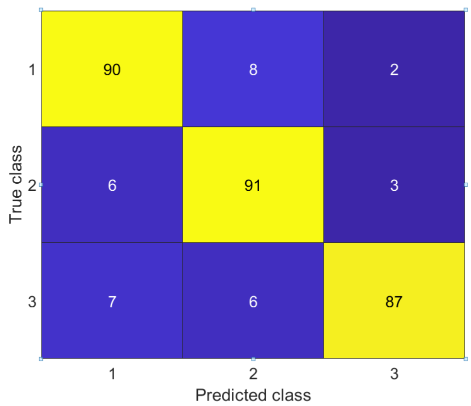
(f) Cross correlation for the expert category during the improvisation sequence

Figure 4.13 – Cross-correlation matrices of the 29 feature measurements μj among the whole database

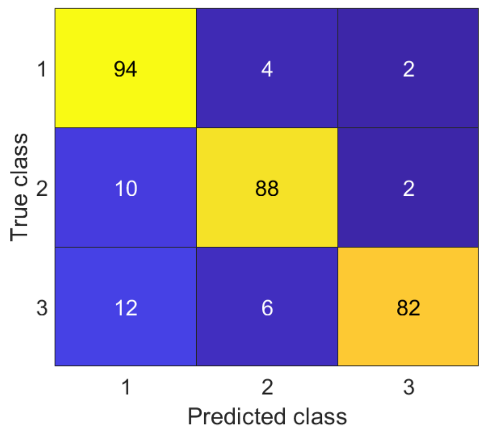
4.3.4.1 Classification results

In this section, we are evaluating the proposed MMF to distinguish dancer categories in terms of a combination of 29 measurements. All data from all sequences have been inserted successively into machine learning classifiers with the 29 features as inputs and the dancer's level as target (a 1-dim vector with the number 1,2,3 that corresponds to the three levels). Three of the most popular of the families of multiclass classifiers have been tested, namely k-nearest neighbours (KNN) algorithm, weighted Support Vector Machine (SVM) with city block metric, and Random Forest (RF) algorithm. The classification has been tested for three levels of dance, giving a total of 9 results. The figure 4.14 shows the confusion matrices for the classifier showing the best results (Random forest).

The table 4.3 provides a summary of the collected statistical results among all classifier. The



(a) Basic steps sequence (percentage)



(b) Improvisation sequence (percentage)

Figure 4.14 – Confusion matrices

evaluation of classifiers is made through the following statistic: Recall (R), the proportion of motion parts of specific level which have been identified to be the correct level, Precision (P) the proportion of the motion classified as a specific level, whose true class label was that level and Accuracy (A) the global proportion of data classified correctly. Given TP is *true positive*, TN *true negative*, FP *false positive* and FN *false negative*, the definition is:

- *Recall = TP/(TP +FN)*
- *Precision = TP/(TP +FP)*
- *Accuracy = (TP +TN)/(TP +TN +FP +FN)*

Table 4.3 – Summary of the resulting statistics from all classifiers: K-nearest-neighbor, Support Vector Machine and Random Forest. R is recall, P precision and A accuracy. Each class represents the ability to classify the validation data into the three categories: Beginners, Intermediate and Experts.

	Three level classification (beginner - intermediate - expert)								
	KNN			SVM			RF		
Seq.	*R*	*P*	*A*	*R*	*P*	*A*	*R*	*P*	*A*
Basics	76.10	78.07	77.54	80.12	81.91	80.10	91.26	87.48	**90.04**
Impro.	69.13	70.09	70.02	65.75	64.29	65.12	83.75	84.63	**85.30**

The classifier that shows the best result is a random forest algorithm with a number of 100 learners that shows a maximum of 90% accuracy for basic sequence between the three dancer categories. In all cases, the improvisation sequence presents a lower classification accuracy, indicating that the diversity of produced movement increases the difficulty of the classification task. Indeed a combination of basic and specific movements such as spin and other rotation as well as subtle rhythm variation would bring more noise to the data, and so is constituent with lower accuracy. Other good results are produced by the SVM algorithm with Cubic kernel, using one Vs. one strategy. In third are the KNN, nearest neighbor algorithm with distance metric as 'Cityblock,' 10 nearest neighbors, and inverted squared weight distance function.

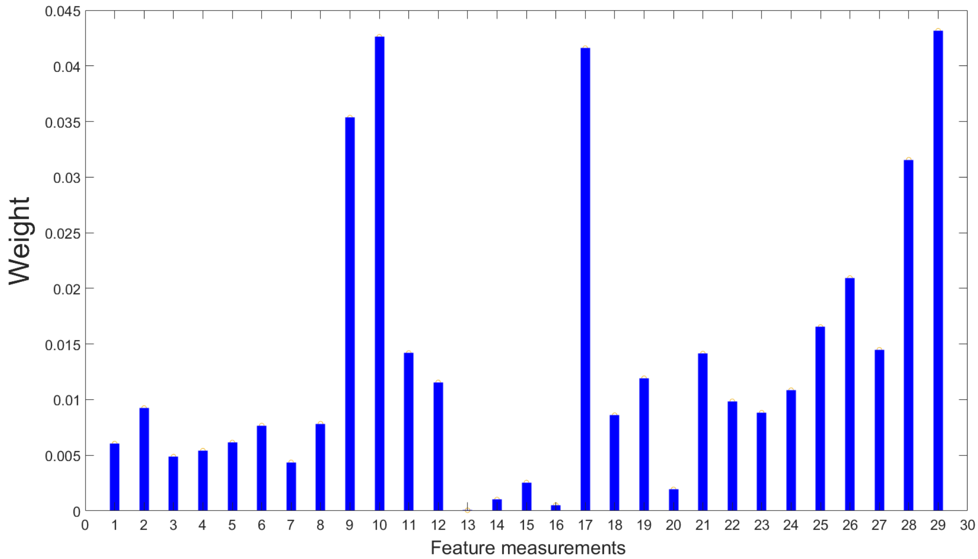


Figure 4.15 – Importance of the different features measurements μj upon the basic steps sequence

This level of 90% accuracy of classification for dance level is quite high compare to previous research [49, 24, 94] and prove the possibility to classify dance couple in categories of experience, regarding motion features that correspond to dance skills. The attribution of each couple to their according category is the most discussable parameter here, and mainly the Intermediate group is difficult to define precisely as learning is a continuous process. Classifying dancers in terms of performance can be from the aesthetic point of view, technical performance, or others. In our case, from the learning point of view, we defined these MMFs that are supposed to help to describe dancers that are progressing. The good accuracy of our results emphasizes the relevance of our approach.

4.3.4.2 Features importance

The importance of each feature has been investigated through the computation of their relative weight for the classification, that is shown from the RF classifier on the figure 4.15 and figure 4.16 for the basic steps sequence and then improvisation sequence respectively.

The features from the styling skill, 18 to 29, are playing a more important role in the basic step sequence. It is constituent to the fact to have more upper body movements and steps variation during the improvisation sequence, adding more noise. The last features of upper body movements 26 to 29 show great importance to a separate category. We can explain it by the more significant and broader upper movement expert are doing, which constituent to the order of learned dance skills defined with the interviewed experts. Clearly, the feature 13 to 16 concerning guidance are not contributing to classification and can be reconsidered for future optimization. The three global features, 9,10 and 17, are participating a lot in both sequences. It validates to highlight the interest of using spectrum analysis to extract rhythm information.

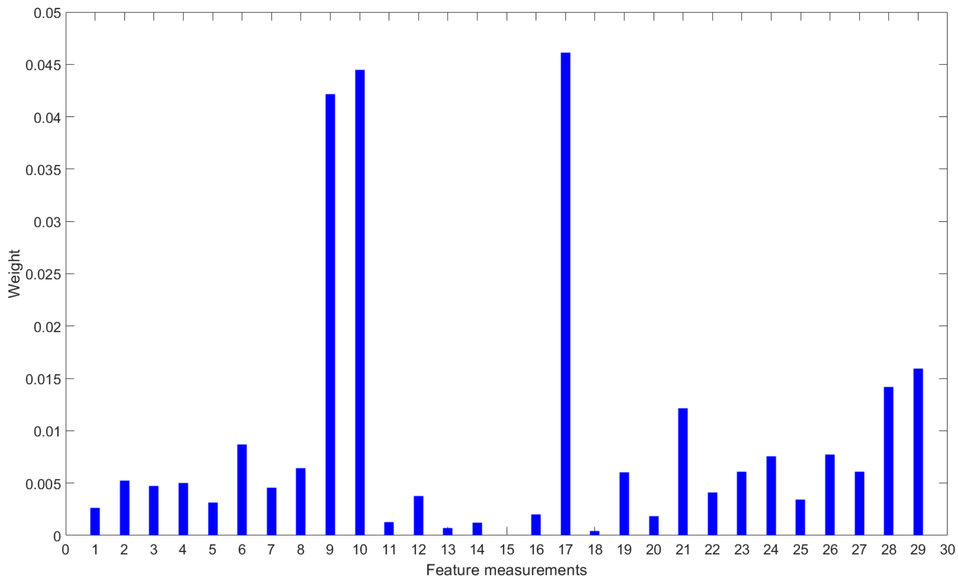


Figure 4.16 – Importance of the different features measurements μj upon the improvisation sequence

4.4 Conclusion

From literature and interviews with professionals, a set of 6 dance skills with varying importance were identified. Among them, three were interpreted and expressed as 29 interactions based on musical-related motion parameters (MMF). A database of salsa dance in synchronization with music was realized. The proposed parameters were computed for each couple on each song within a sliding window and inserted into classifiers. The results show an accuracy of up to 90% and mostly above 75%, validating most of the parameters and our sliding window method. A latter analyze of the feature importance shows that 23 features out of 29 are relevant for learning level classification, allowing to have a complimentary evaluation of salsa dancer during couple performance.

This study is the first step toward a virtual coach that uses automatic analysis of learning states for Salsa to improve the dancer's skills. Our proposed music and interaction-based motion features show some success in classifying social couple dance performance. This first approach defines a building block for a framework that could be utilized within the other couple dances and the different domains, such as in robot interaction, emotional recognition or bio-mechanic studies, but also in the field 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 a base for other studies and cultural heritage conservation examples.

Future work includes improving the classification via the optimization and fine-tuning of the classifier as well as studying feature importance and weight. Implementation of a neural network can be also interesting for classification and real-time application as well as reducing the number of features able to discriminate the different levels. 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 electroencephalogram or facial detection study while dancing as to try to detect the emotional states of both participants. A learning study can be developed using the proposed features to investigate how they can have a real impact on the improvement of social dance learning.

Chapter 5

Virtual reality salsa dance learning and motion analysis

"When you dance with a partner you are close and the dance is very suggestive, but it is not personal. Close is what the music inspire you to become. The embrace looks personal, but what we are actually embracing is the music." Carlos Gavito, dancer.

5.1 Introduction

Learning partner dance represents many challenges, firstly because of the accessibility to dance class, dance partner, and good quality teaching. Even a good teacher, depending on her/his current mindset and emotional phase, can have varying pedagogical efficiency and does not constantly provide good quality lessons. Most of the time, partner dance classes are assumed collectively. Being in groups or classes also undoubtedly means that the teacher cannot focus all the time on all the students and thus provides valuable feedback. Secondly, because of the vast dynamic and interactive parameters to assimilate, such as dance pattern, synchronization with music, and guidance derived from the physical and cognitive of fundamental partnership interaction. This musical interpretation and listening/following (e.g., body guidance) represent a tremendous challenge to comprehend and analyze this intricate and interdependent set of parameters. We can summarize these challenges here:

- 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 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).

Additional difficulties may arise along the learning process 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 a *expert* in social dance can be a source of confusion as there is no official diploma state validated (There is diploma provided by organization or schools though) 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 can provide the required mechanical interaction between the user and the virtual character. It allows for tracking the full-body movements over an area similar to the one needed for dancing, giving the possibility for real-time feedback. The main objective of this chapter is to demonstrate that we can guide and help users to improve their salsa dancing skills through a VR game that simulates salsa practice. In previous chapters, 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

Figure 5.1 – 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.

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 in relation to the aforementioned important criteria; the analysis was conducted using a number of Music-related Motion Features (MMF) and 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 5.1, shows a visual illustration of our VR environment, where a user interacts with the virtual environment.

The main contributions of this chapter 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.

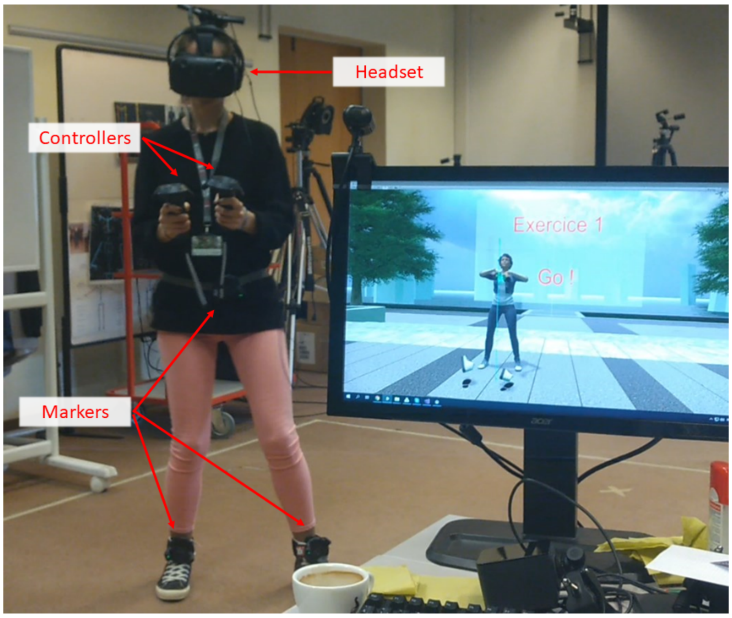


Figure 5.2 – Example of a person testing the game

5.2 Design of the interactive dance learning system

5.2.1 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 f 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 5.2 illustrates an example of a person testing our VR environment.

5.2.2 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.

5.2.2.1 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 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 the previous chapters and experiments, from which we extract a motion profile for each foot, as illustrated in figure 5.3. 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 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 figure 5.4. 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.

5.2.2.2 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 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

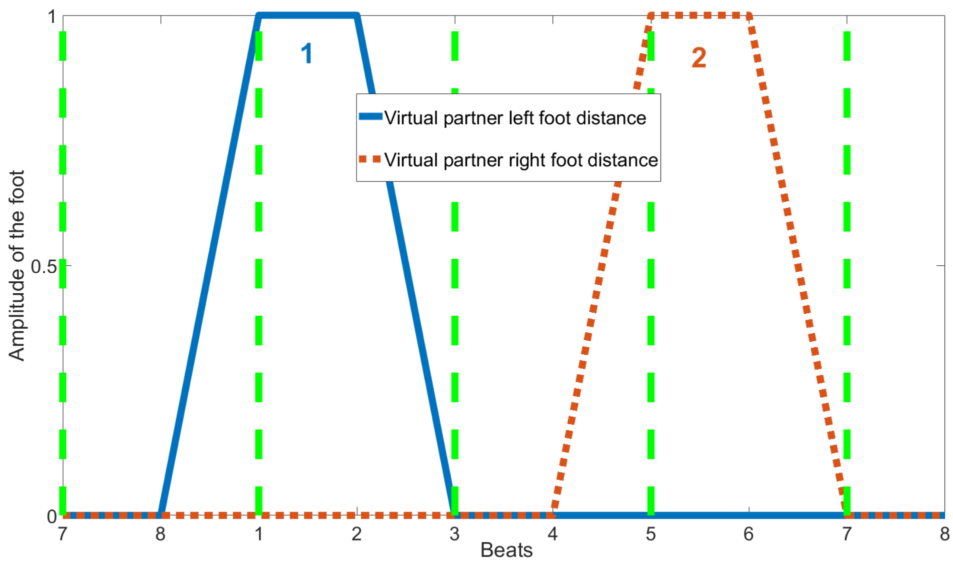


Figure 5.3 – The distance for the right and left foot of the virtual partner related to a neutral position. This motion profile is repeated every 8 beats of music. In the application, the curves were smoothed to achieve better naturalness.

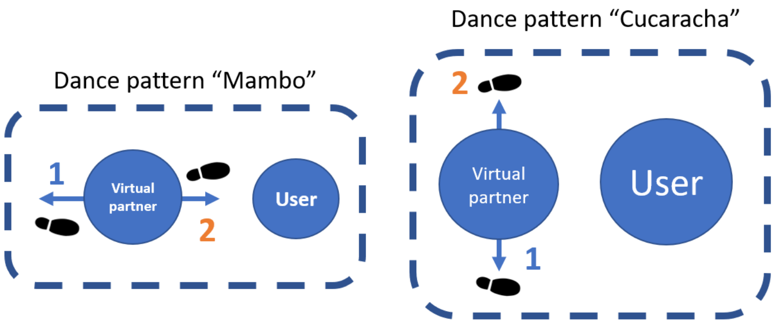


Figure 5.4 – Our two salsa dance pattern. On the left side for "Mambo", the VP will step backward during the beat 1 and 3, then moves forward during the beat 5 and 7. On the right side for "Cucaracha", the VP will move her/his right foot to her/his right then her/his left foot to her/his left alternately.

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 [132]). The results give the user the feeling of guiding the VP, as illustrated in figure 5.5.

5.2.2.3 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

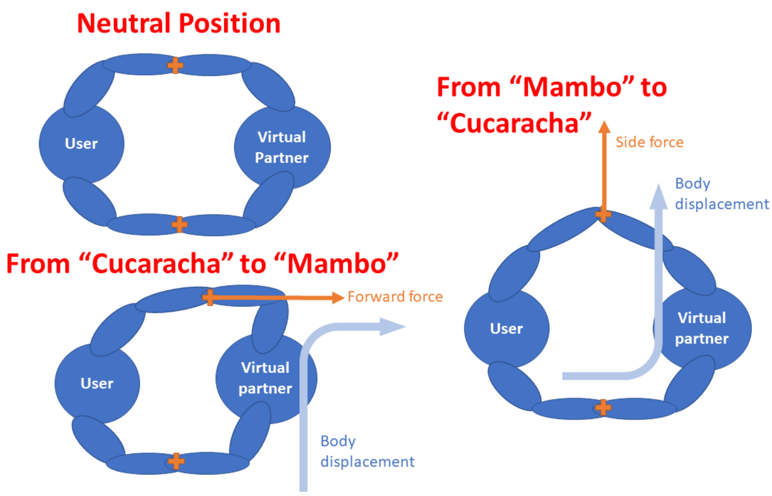


Figure 5.5 – Detail of the required gestures to control the virtual partner on its transition to the two possible dance patterns. To guide the virtual partner from a "Cucaracha" motion to a "Mambo" motion, the user has to push the virtual partner between the beat 7 and 8 with her/his left hand. To do the reverse transition, the user has to pull the virtual partner on her/his left at the same time.

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.

5.2.3 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.

5.2.4 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.

5.3 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 a 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 exercises 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 5.1.

Table 5.1 – 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

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 5.4.2.1 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 in accomplishing the given aim (changing dance pattern every two cycles) and serves 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.

5.4 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.

5.4.1 Virtual reality Musical Motion Features - VrMMF

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, 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 expertsâ€™ 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 $_j$ that belongs to five feature categories, extrapolated from three dance skills, that are shown in Table 5.2. All measurements are observed on a temporal window of given frames corresponding to 8 beats. Experiments in previous chapters 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 loss of the temporal information (see Forbes and Fiume [47]). Finally, each measurement of $_j$ is normalized between 0 and 1.

5.4.1.1 Dance skill: Rhythm

Step accuracy (μ_1 - μ_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 eight musical beats for each foot. Two peaks occur that

Table 5.2 – Subset of the Musical Motion Features in our case of Virtual Reality

	MMF	Measurements details
Step Accuracy (Rhythm)	μ_1	Temporal accuracy of left foot over beat 1
	μ_2	Temporal accuracy of left foot over beat 3
	μ_3	Temporal accuracy of right foot over beat 5
	μ_4	Temporal accuracy of right foot over beat 7
Rhythm difference between partners (Rhythm)	μ_5	Temporal difference user left foot /VP right foot over beat 1
	μ_6	Temporal difference user left foot /VP right foot over beat 3
	μ_7	Temporal difference user right foot /VP left foot over beat 5
	μ_8	Temporal difference user right foot /VP left foot over beat 7
Foot Correlation (Guidance)	μ_9	Correlation coefficient beat 1 to 3 user left foot / VP right foot
	μ_{10}	Correlation coefficient beat 5 to 7 user right foot /VP left foot
Area (Styling)	μ_{11}	Displacement of the left foot over 8 beats
	μ_{12}	Displacement of the right foot over 8 beats
	μ_{13}	Net velocity change of the left foot over 8 beats
	μ_{14}	Net velocity change of the right foot over 8 beats
Hands Movements (Styling)	μ_{15}	Mean distance left hand to hips over 8 beats
	μ_{16}	Mean distance right hand to hips over 8 beats

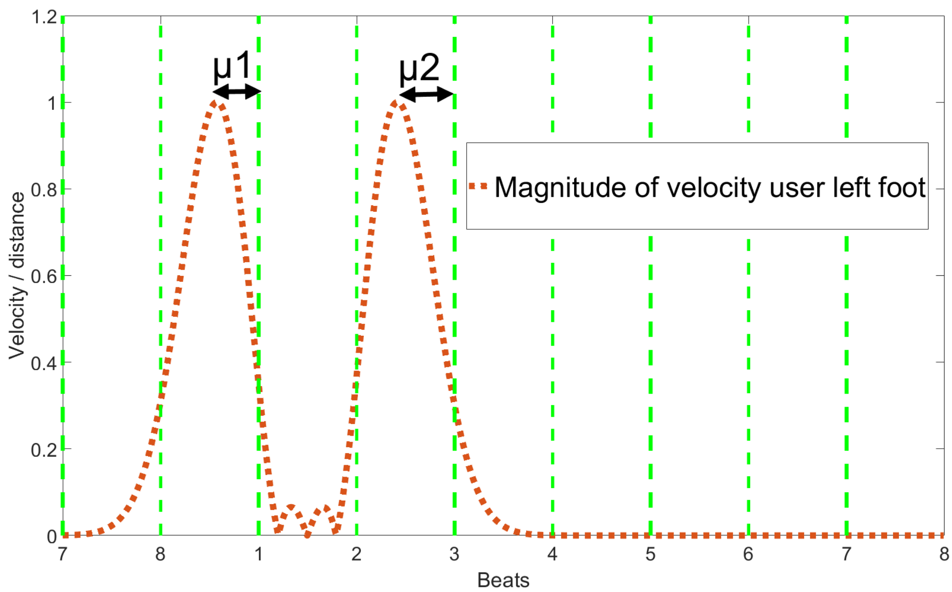


Figure 5.6 – Velocity peak that occurs during the basic step for the user, and the measurement of step accuracy.

indicate a straight movement of the foot on the music. For example, when dancing the "Mambo" pattern, the first peak should correspond 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 (that has to be done in preparation for each music), we can compute the step accuracy for each beat as the difference between the ideal musical beat and the user's performed beat. 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 Figure 5.6. The result is 16 measurements that are extracted through a sliding window of width proportional to the music tempo. The beat one is detected by hand at the beginning of each song to ensure the good temporal accuracy of each sliding window.

Rhythm difference between partners (μ_5 - μ_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.

5.4.1.2 Dance skill: Guidance

Correlation between foot movements (μ_9 - μ_{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).

5.4.1.3 Dance skill: Styling

Area (μ_{11} - μ_{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 (μ_{15} - μ_{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.

5.4.2 Laban Movement Analysis - LMA

5.4.2.1 Preliminary work on Laban Motion Analysis

In order to compare our MMF with other framework, we developed in parrallel an approach to high level classification of motion using LMA system. The study of human multimodal emotional behavior can lead to the creation of virtual characters with different moods, emotional states, or personalities. In order to create social-emotional agents and allow them to interact, it is necessary to understand the emotion expressed through the various communication channels, such as body motion, facial expression, and voice. A number of studies on emotion recognition, across different research domains, have been published. The majority of the first studies were focused on facial expression. Inspired by the work of psychologists and behaviorists such as Ekman during the 1970s, systems that allow discrete recognition of emotion were developed and even integrated into some applications like with interactive robots. More recently, emotion recognition from body motion was possible owing to the recent advances in motion capture systems. Using visual pattern recognition at the origin, new techniques using statistical analysis and machine learning give the possibility to consider the dynamics of the body and not only static postures or gestures. Their main difficulty is to overcome the fact that a large number of parameters may influence the recognition as the classification method and the quality and quantity of the data. However, a very high recognition rate can be obtained from specific sets of data. In order to achieve a good simulation for the body language, a simple but efficient body movement description system is needed: the Laban Movement Analysis (LMA) system fulfills these requirements. The LMA system draws on Rudolph Laban's theories and allows to describe, interpret, and document human movements. It is a multidisciplinary system, incorporating contributions from psychology, kinesiology, and anatomy. It is one of the most used languages to analyze human movement and has been regularly used to document and describe choreographies and dance over the last century. The relation between movement and emotion has been studied extensively in psychology. Several emotion models were developed over the last 30 years. Still, one of the most used in computer applications is the continuous Russell Circumplex Model (RCM), which maps the different emotions following two dimensions: the pleasure/displeasure value and the activation/deactivation value. Coming from a long comparative work and experimentation, its classification allows the emotion to be placed on a 2D diagram for visualization. Taking into consideration the broad range of classification criteria for emotion recognition, it is difficult to imagine a system that can adapt to any possible situation, and for different performers. Indeed, looking at the body movements, the meaning may vary significantly regarding the characteristic of the person itself: internal moods, social context, stress, culture, energy, and so on. Among other possibilities of body expression situation, theater performance is a good compromise between natural and clear expression, which is closely related to emotion. In this work, we aim at recognizing the emotional state of theater actors by mapping their motion, concerning the expressed emotion, onto the Russell bi-dimensional space. This can be considered as the first step over a complete emotion recognition system. We integrate neural networks, because of their ability to approximate multivariate functions, to form the mapping. The neural networks are trained using the attribution of emotion to performance, and the corresponding measured LMA features. It is then possible to have a correspondence with LMA features and coordinates of the emotion that vary continuously, allowing us to study and visualize emotion's dynamic.

Methodology Theater performances are a compelling case of interaction where one of the two protagonists is rather passive (the public) and the other one very active, proficient on the communication. During theater performances, actors use a variety of ways to communicate emotion and story with the public. Body expression is, therefore, of high importance and plays a vital role. The interaction aspect, although it can be placed under consideration as the actor, does not have real feedback from the public. Still, the action of performing will make the actor emphasize on its most expressive capabilities. Notice that to have a powerful expressivity doesn't mean to express something of high intensity (as "Very happy") but rather to be easily understood. It is essential to understand here the subtle difference between the internal emotional state of a person

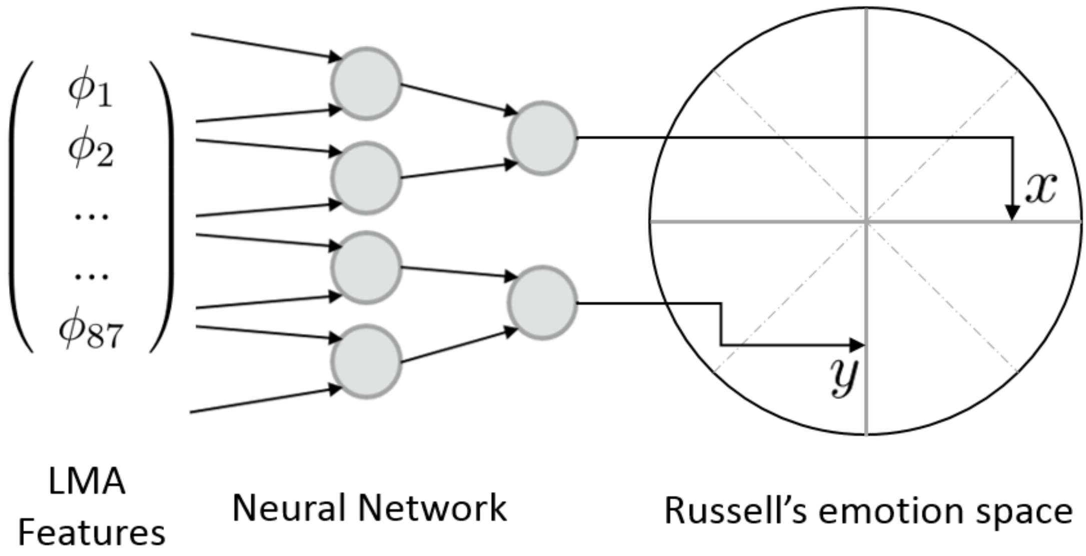


Figure 5.7 – Pipeline of mapping motion to Russell's space. This system allows continuous visualization of motion in terms of activation and pleasure. From 87 LMA features extracted from motion, we have 2 dimensional coordinates of emotion

(what he/she is thinking) and the expressed emotion that is seen externally. The difference can be misleading, and our study is focused only on the second case.

We firstly record motion capture data of people expressing emotions corresponding to Russel's classification. We chose eight emotions that are distributed along with the four quadrants of the model: Happy, Excited, Afraid, Annoyed, Sad, Bored, Tired, Relaxed. Since we want a context highlighting interactive behavior, we chose theater expression as it is a good case of emotional communication. We invited several actors to perform motion sequences while they express only one emotion at a time. Using the Aristidou's method, we extracted the LMA features by averaging measures on a sliding window of 35 frames (on data captured at 30fps). Then, to map the input motion data to an emotional space, we need a multivariable function that outputs a dual of values (Russel's coordinates) from a high number entries (86 LMA values). To do this, we introduced a neural network to approximate the coordinates of the input motion onto the RCM, managing at the same time with a large number of the LMA variables. The proposed pipeline is illustrated in figure 5.7.

Experiment

Database of theater movement In our experiments, we used a Kinect for Windows V2 as a motion capture device. It's a mobile, fast, and efficient device that tracks the skeleton data directly with internal processing. Although the low resolution in space and time compare to professional motion capture systems, the discreet characteristic presented some advantages concerning the naturalness aspect of recorded motion. Data acquisition problems came mostly from the fact that the Kinect is a depth camera and not a 3D sensing system. As a result, Kinect faces occlusion problems because the recorded person rotates itself, sometimes facing the camera or not. A

Figure 5.8 – Professional actor performing theater's expression of emotion and Kinect for windows v2.

person presenting her/his side is more challenging to detect. It leads to some requirements for the actor's performance. Operating at 30 fps with a resolution of 512Ã—424 pixels for the depth camera, although non-communicated, the accuracy for skeleton body tracking is sufficient to recognize basic gestures and small spatial displacement. We used a smoothing function during the motion capture for the joint tracking to have better measurements and reduce the subsequent problems. It is important though to note the variability of the data due to the several factors that influence the performance, such as emotional state and current moods, actor's skills and experience, actor's personality and idiosyncrasies, cultural references, current environment, the reaction of public, etc. The most difficult aspect of our study is that the expression of emotion remains anyway subjective. Thus, we have tried to reduce the potential influence of external factors that affect the quality of the motion during the capturing procedure. Since the Kinect device is lightweight and noninvasive, compared to the traditional mocap system, the performers were more confident about themselves and could behave more naturally.

The only technical requirement is to face in a relatively short angle the device. To avoid a face-to-cameras situation, we ask the remaining actors to face the performer as there was a real public. Ten different actors from a professional theater school were playing eight different emotional behavior (figure 5.8). The performers were professional actors, one male, and nine females, while their age ranges between 18 and 27-year old. The actors were asked to perform an emotional state for 40-50sec, improvised from their range of acting repertoire. Most of the time, they ask other actors to launch their behavior with a dialogue that triggered an emotional response. In total, we end up with 80 performances, with approximately 53 min (96,000 frames) of motion; the data set is sufficiently large, having in mind that in the experiments we do not use individual theater motion, but the entire sequence, where the performer acts during the 40s without repeating her/his/her movements. To avoid corrupting the data, the information communicated to the performers was minimal and did not contain the analysis criteria. We stored and used the data as a BVH format, mapped to the same 3D character for uniform processing and analysis.

Laban Motion analysis Human motion analysis is particularly challenging, especially when stylistic characteristics are of high importance. The difficulty is even more pronounced when motion is used to describe and classify human emotions or behaviors. In this work, we utilize a human full-body motion analysis that is based on the LMA principles, aiming to identify those factors that describe the movement signature of the performer.

LMA is a language for interpreting, describing, visualizing, and notating human movement; it offers clear documentation of the human motion, and it is divided into four main components:

- BODY, which describes the structural and physical characteristics of the human body;
- EFFORT, which describes the intention and the dynamic quality of the movement, the texture, the feeling tone, and how the energy is being used on each motion;
- SHAPE, which analyzes the way the body changes shape during movement; and
- SPACE, which describes the movements in relation to the environment.

The EFFORT component, which is generally related to the changes of mood or emotion, is further divided into four sub-categories, each having two polarities named EFFORT factors:

- Space addresses the quality of active attention to the surroundings. It has two polarities, direct (focused and specific) and indirect (multi-focused and flexible attention).
- Weight is a sensing factor, sensing the physical mass and its relationship with Gravity. It is related to the movement impact and has two dimensions: strong (bold, forceful) and light (delicate, sensitive).
- Time is the inner attitude of the body toward the time, not the duration of the movement. Time polarities are sudden (has a sense of urgency, staccato, unexpected, and isolated) and sustained (has a quality of stretching the time, legato, and leisurely).
- Flow is the continuity of the movement; it is related to the feelings and progression. The flow dimensions are bound (controlled, careful, and restrained movement) and free (released, outpouring, and fluid movement).

In order to extract movement characteristics that discriminate human behaviors with regard to emotion, we utilize the LMA framework described by Aristidou et al. [7] for full-body motion analysis. We used a 35-frame sliding window with a 1-frame step (our motion data are sampled at 30 fps) to extract the LMA features. The features and their respective measurements are summarized in Table 5.9, that originated from previous work of Aristidou et al. [4].

Results

Training We use a neural network that has 86 inputs and 2 outputs with a total of 10 hidden layers. 70% of the data was used for training using the Levenberg-Marquardt method, and 15% for validation. The 86 measurements proposed by Aristidou et al.[4] are used as input values, whereas the output values are the x,y coordinates of the RCM diagram, indicating the Pleasure

	Features		Measurements			
	f^i	Description	f^i_{max}	f^i_{min}	f^i_{σ}	f^i_{μ}
BODY	f^1	Feet-hip distance	ϕ_1	ϕ_2	ϕ_3	ϕ_4
	f^2	Hands-shoulder distance	ϕ_5	ϕ_6	ϕ_7	ϕ_8
	f^3	Hands distance	ϕ_9	ϕ_{10}	ϕ_{11}	ϕ_{12}
	f^4	Hands-head distance	ϕ_{13}	ϕ_{14}	ϕ_{15}	ϕ_{16}
	f^5	Hands-hip distance	ϕ_{17}	ϕ_{18}	ϕ_{19}	ϕ_{20}
	f^6	Hip-ground distance	ϕ_{21}	ϕ_{22}	ϕ_{23}	ϕ_{24}
	f^7	Hip-ground minus feet-hip	ϕ_{25}	ϕ_{26}	ϕ_{27}	ϕ_{28}
	f^8	Centroid-ground distance	ϕ_{29}	ϕ_{30}	ϕ_{31}	ϕ_{32}
	f^9	Gait size	ϕ_{33}	ϕ_{34}	ϕ_{35}	ϕ_{36}
EFFORT	f^{10}	Head orientation	ϕ_{37}		ϕ_{38}	ϕ_{39}
	f^{11}	Deceleration peaks				ϕ_{40}
	f^{12}	Pelvis velocity	ϕ_{41}		ϕ_{42}	ϕ_{43}
	f^{13}	Hands velocity	ϕ_{44}		ϕ_{45}	ϕ_{46}
	f^{14}	Feet velocity	ϕ_{47}		ϕ_{48}	ϕ_{49}
	f^{15}	Pelvis acceleration	ϕ_{50}		ϕ_{51}	
	f^{16}	Hands acceleration	ϕ_{52}		ϕ_{53}	
	f^{17}	Feet acceleration	ϕ_{54}		ϕ_{55}	
	f^{18}	Jerk	ϕ_{56}		ϕ_{57}	
SHAPE	f^{19}	Volume (5 joints)	ϕ_{58}	ϕ_{59}	ϕ_{60}	ϕ_{61}
	f^{20}	Volume (upper body)	ϕ_{62}	ϕ_{63}	ϕ_{64}	ϕ_{65}
	f^{21}	Volume (lower body)	ϕ_{66}	ϕ_{67}	ϕ_{68}	ϕ_{69}
	f^{22}	Volume (left side)	ϕ_{70}	ϕ_{71}	ϕ_{72}	ϕ_{73}
	f^{23}	Volume (right side)	ϕ_{74}	ϕ_{75}	ϕ_{76}	ϕ_{77}
	f^{24}	Torso height	ϕ_{78}	ϕ_{79}	ϕ_{80}	ϕ_{81}
	f^{25}	Hands level				ϕ_{82}- ϕ_{84}
SPACE	f^{26}	Total distance				ϕ_{85}
	f^{27}	Total area				ϕ_{86}
	f^{28}	Total volume				ϕ_{87}

Figure 5.9 – The features and measurements used to extract the movement characteristics based on the LMA components. Each feature is decomposed up to 4 measurements: maximum (fmax), minimum (fmin), standard deviation (fsigma), average (fmu), computed on a 35 frames time window.

and Activation correspondence of the emotion. The remaining 15% data served as test data on which we could use the trained network. All inputs are not normalized, but they have the same factor. We choose 10 layers as it shows the best results. It took 20 iterations to have a sufficiently trained network. After data processing and training of the neural network, we were able to extract a mathematical function that takes 86 values and give 2 values. Note that this function is similar to the dimensional reduction method, such as Principal Component Analysis. We try this function first with the 15% remaining test data and secondly with another motion sequence, where the actor has been asked to change the expressed feeling between a number of different emotions within the same performance.

Recognition results Recognition results We can see in figure 5.10 the result of the trained neural network. Each color corresponds to one emotion, coming from all actors. The dispersion of the clouds highlights the diversity of movement produced by different people. However, most of the points for each cloud are contained in a defined area located in the right Russell's quadrant. Please note that we averaged the data for better visualization.

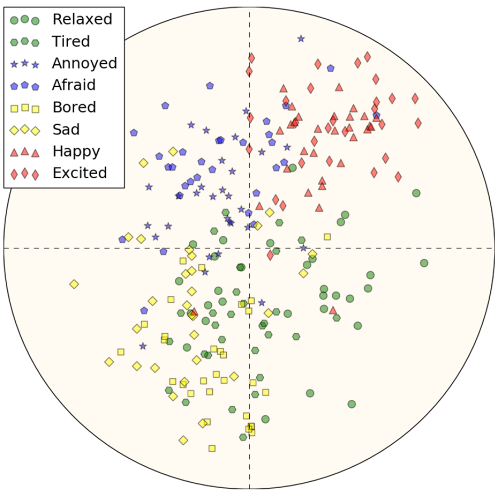


Figure 5.10 – Results of the neural network function on the test data after training phase. Each color correspond to one emotion from all performers. Each point correspond to a LMA measurements done within 35frame window.

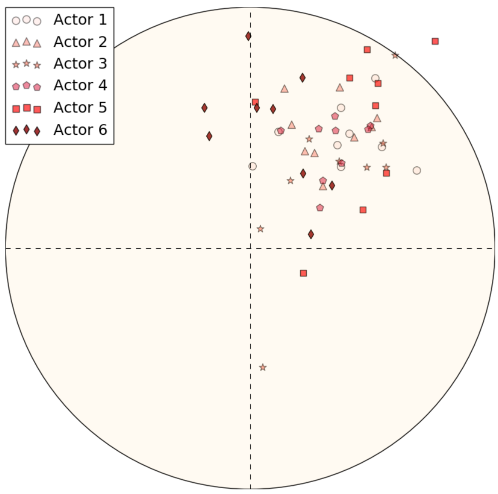


Figure 5.11 – Output of the neural network for only the Happy emotion. The distribution of each actor is shown. Each point correspond to a LMA measurements done within 35frame window

Looking at the clouds individually and separating the people, as for the emotion Happy on the figure 5.11, it's evident that the distribution is almost the same and shows a good acting homogeneity among the actors. Additionally, the covered area is still placed on the correct quadrant of Russell's model. To improve the visualization of the results, we draw the mean values and the standard deviation of the clouds and overlay it to the RCM, as shown in figure 5.12. One can observe the tendency to the center. This is due to the averaging effect of the neural network. It can be reduced by increasing the quality of the data. Another point is that the clouds seem to be aligned along a diagonal. The difference between activation and deactivation is easier to produce as an actor and detects via motion capture than pleasant and unpleasant aspects. The experiment showed us that this dimension might be more carried by facial expression. However, it can be observed that the values are quite close to their reference area. They are indeed in their respective parts of the angular space, except for tired that shows a tendency to be seen as more unpleasant than it should be. This can be explained by the difficulty to express emotions that have more deactivation: Since the actor doesn't move much, it has to be very subtle motion.

Continuous recognition To test the performance and relevance of the trained neural network for continuous emotion analysis, we registered a sequence of movements that embed four emotions coming from 4 different quadrants of Russell's model (Sad, Afraid, Happy, Calm). The actor was asked to perform 10s of each emotion in the previously mentioned order and switch from one to another in a continuous way. We recorded her/his motion and extracted the LMA components for the whole sequence. It represents around 60 LMA vector values. Then we used the trained neural network to visualize the corresponding output.

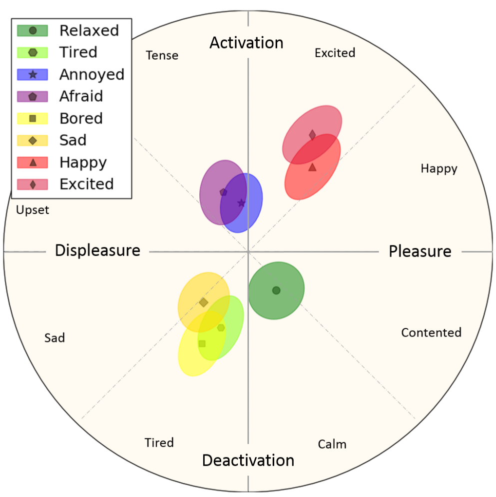


Figure 5.12 – Mean value and standard deviation for each emotion, from all performers. For such recognition system, it is difficult to present a percentage of success because of its continuity aspect: we can however see a qualitative feedback as the order along the two axis respect the Russell's area references.

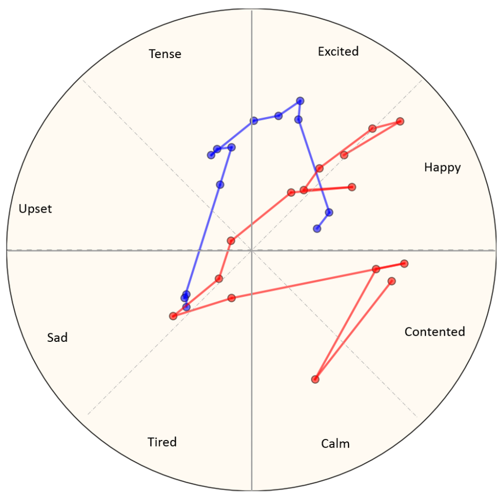


Figure 5.13 – Output of the neural network for two sequences (blue, red) of successive emotion. The actors express a series of specific emotions (Calm, Sad, Afraid, and Happy in this order) during a 40s recording (10s each). We can see the trajectory of the expressed emotion over time. The arrows indicate the flow of time.

The results are presented in figure 5.13, where the trajectories of the detected emotion are shown over time (the arrows mean the direction of time). We can visualize the emotion transition as a trajectory on the diagram. Although some parts of the trajectory can vary a lot, by averaging, there is an apparent continuous variation along with the quadrants. This continuous detection of the emotion's variation allows us, therefore, to know that the performers changed their emotional state. The presence of noise obliged us to average still in a strong manner, but we can imagine having more fine transition detection with better processing. Our method validates the continuous body emotion recognition by a neural network for theater performances.

Conclusion In this work, we have proposed a framework that uses LMA components and machine learning to map an input motion to the Russel Circumplex Model. Results show that we can recognize multiple emotions and place them onto Russell's 2D diagram with sufficient accuracy that respects the initial classification. Thus, we show that LMA and neural networks are suitable for continuous analysis of emotion and estimation in terms of Intensity and Valence. Analysis of the sequence of emotions shows a progressive change in the diagram, which is an illustration of the continuous shift in emotions over time. We show that we can perform continuous analysis of emotions in terms of trajectories of Russell's diagram. This can be seen as the study of emotion's dynamic. Beyond the ability to see the emotional trajectory of an agent, it is possible to use this framework in the context of animation on a motion synthesizing perspective to achieve more natural emotion behavior. To improve the reliability of the recognition, the first idea is to improve the quality of the input data by reducing the inherent problem of depth cameras. Future work would be to enhance the capabilities of our capturing system by using a multi-device architecture to deal with larger angles and occlusions, such as Kitsikides et al. [70] work. Making a statistic analysis of the impact of the noise on the quality of the data would also help. Besides, as previously mentioned, the pleasure dimension is difficult to recognize due to the multimodality of the context of theater that makes the performers using other channels such as facial expression. Additionally, since our system seems better at picking up the activation dimension, facial expressions will add much information toward discriminating of the pleasure dimension. Therefore, we will integrate facial expressions and combine them with the existing body analysis to get a better emotion recognition that leads to complete multimodal emotional behavior analysis. LMA can be used as a complimentary system for our partner dance performance evaluation.

5.4.2.2 Motivation and chosen structure

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 assess the learning skills of our platform, we learn motion characteristics based on the LMA principles [75], drawing from the framework described in Aristidou [11]. This framework was strategically designed to capture the diversity of stylistic and geometric characteristics of a set of dancing motions [4], and has been used to analyze and compare folkloric dances [10]. 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 (ϕ_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.

Table 5.3 – 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 [11]

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

Thereafter, and similarly to Aristidou et al. [9], 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 5.3.

5.5 Results

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.

5.5.1 Musical motion features 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, before normalisation to a range 0-1. 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.

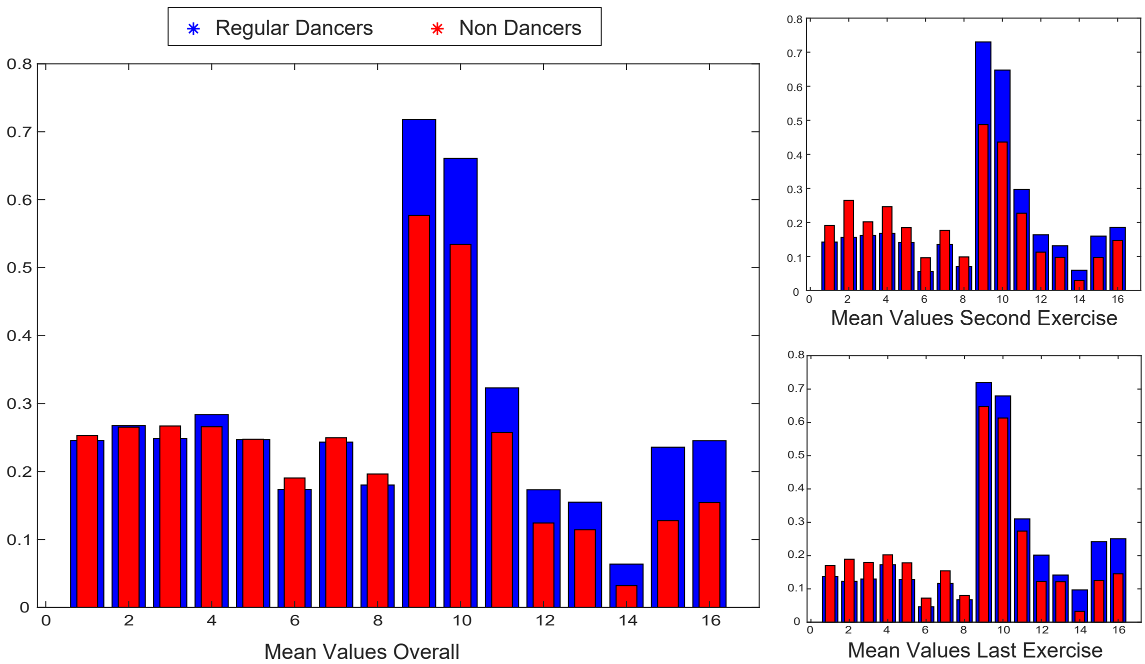


Figure 5.14 – The mean values over all exercises for the MMF-derived feature measurements μ_j is shown on the left. On the right top is the mean values of the second exercise, and at the bottom is the mean values of the last exercise.

Figure 5.14 left shows the mean values of the MMF measurements μ_j of the performers for the two classes for all exercises, while on figure 5.14 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.

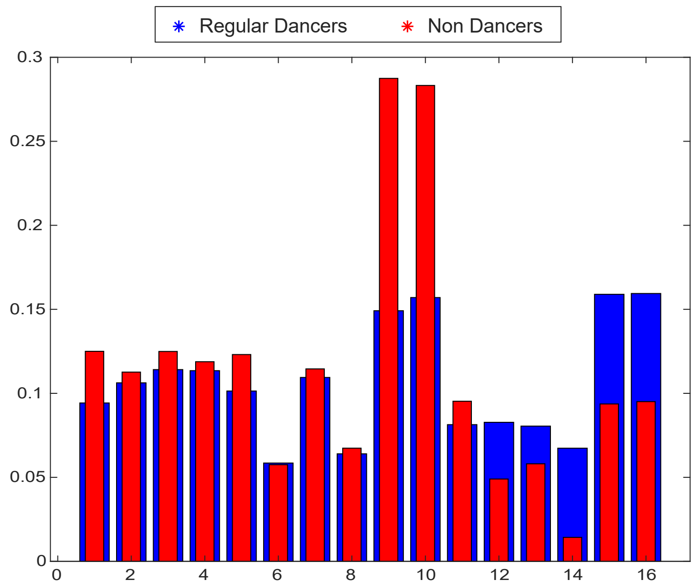


Figure 5.15 – The standard deviation of the MMF-derived feature measurements μ_j for the *regular dancers* (blue) and *non dancers* (red).

Another remarkable notice, as shown in figure 5.15, 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.

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) [78]. We used t-SNE for dimensionality reduction, rather than the Multi-Dimensional Scaling (MDS) [130], since it is particularly well suited for the visualization of high-dimensional datasets such as ours. Figure 5.16 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 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.

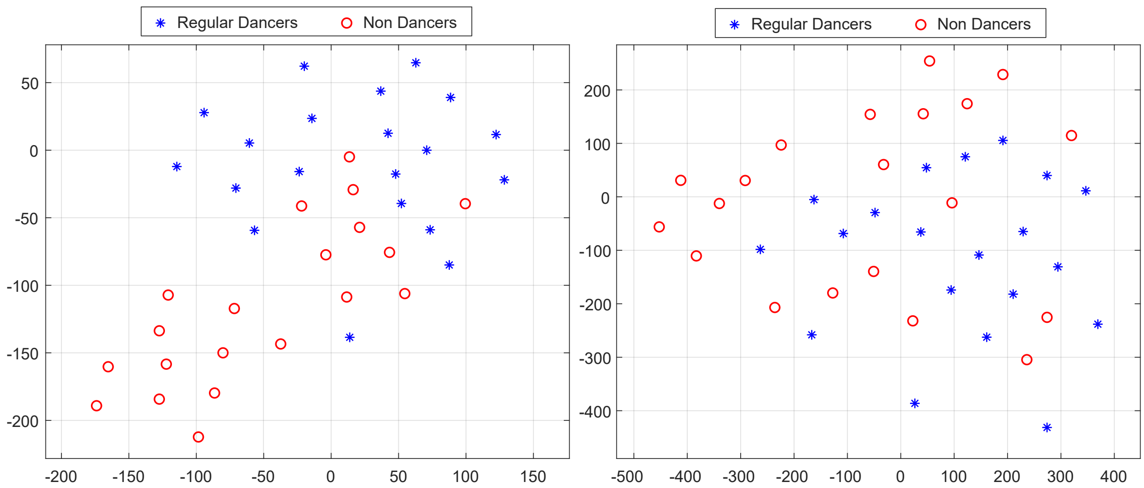


Figure 5.16 – t-SNE dimension reduction of the MMFs for the second exercise (left) and the last exercise (right).

5.5.2 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 5.17 shows the mean values of the LMA-derived feature measurements ϕ_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 (std) of the LMA feature measurements for the *regular dancers* is larger than those of the *non dancers* (refer to Figure 5.18). 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

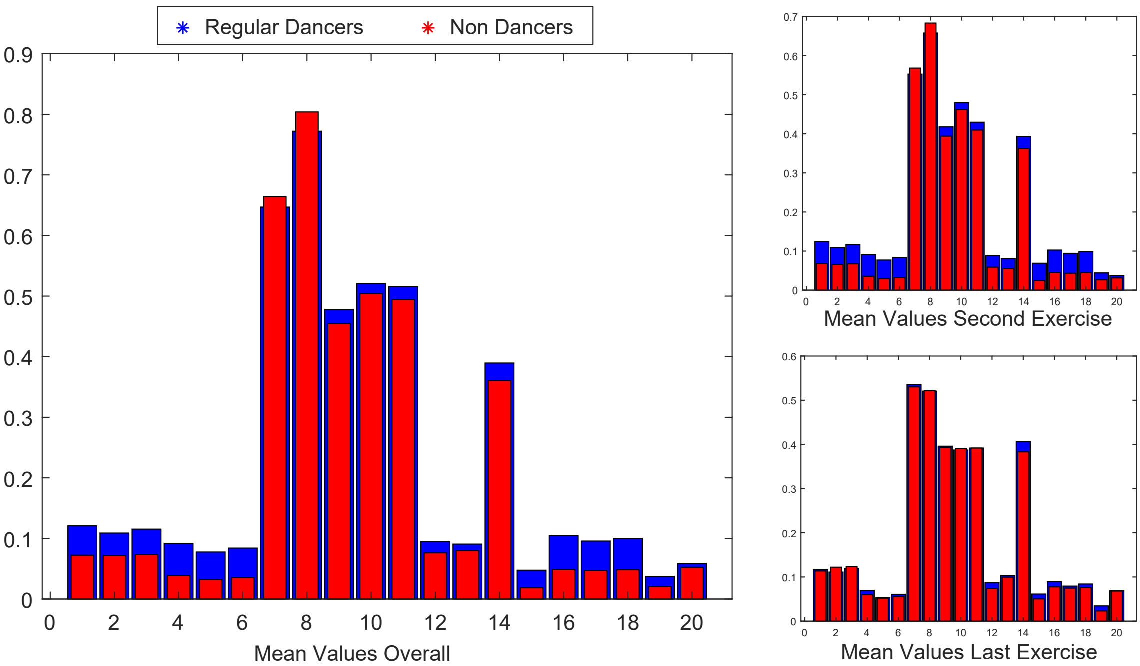


Figure 5.17 – The mean values for the LMA-derived feature measurements ϕ_j. The mean values of the performers for all exercises are shown on the left. On the right top is the mean values of the second exercise, and at the bottom is the mean values of the last exercise.

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 Figure 5.18, 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 5.5.1, we visualize the differences between the two classes, using t-SNE. Figure 5.19 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 are mixed (it is more difficult to be separated). Figure 5.19 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.* [11], we represented

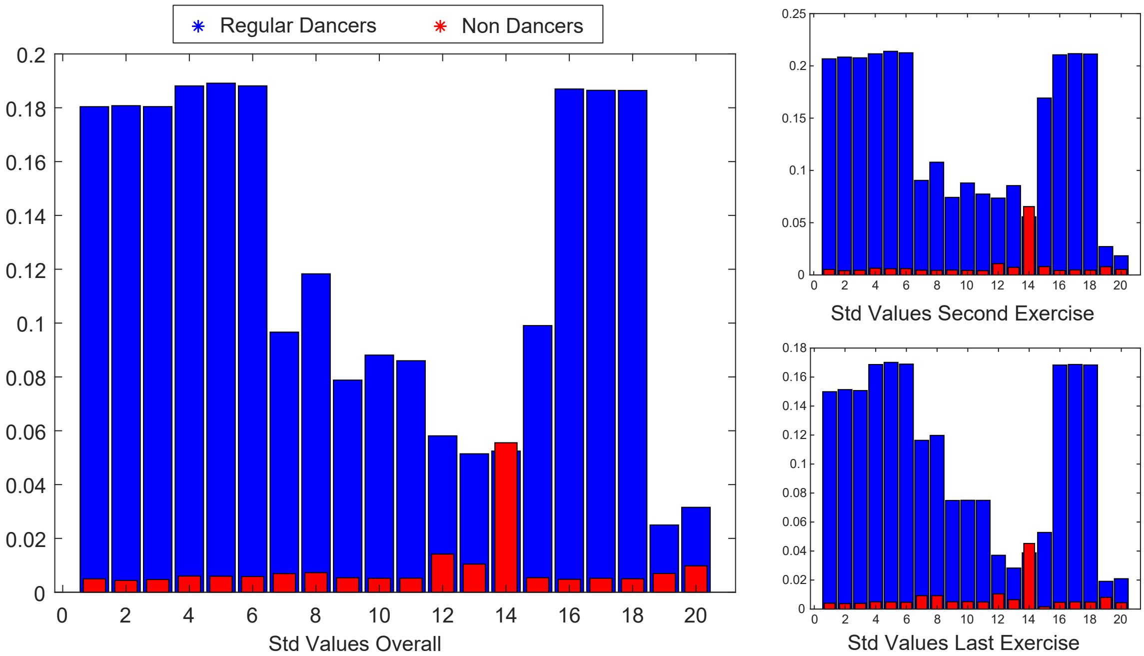


Figure 5.18 – The average standard deviation of the LMA features for the *regular dancers* (blue) and *non dancers* (red) over all exercises (left image). On the right, top image, is the standard deviation of the LMA features for the second exercise, and at the bottom, for the last exercise.

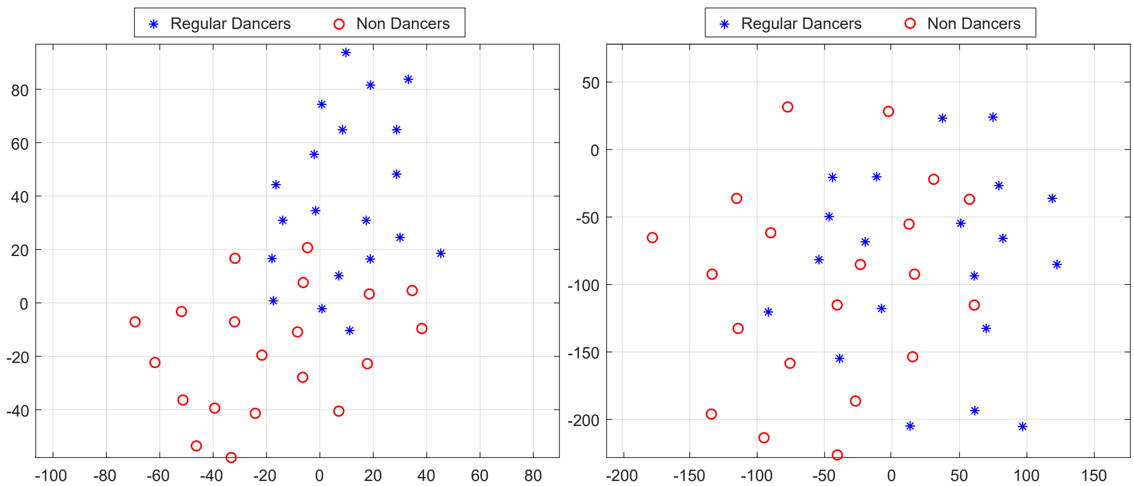


Figure 5.19 – The 2D embedding of the two classes using the mean LMA-derived arrays for the second (left) and final exercises (right). It can be observed that, for the early exercises of the experiment, the two classes can be separated based on the performers learning skills, but as we move forward to later exercises, the dancing skills of the *non dancers* are converging to those of the *regular dancers*.

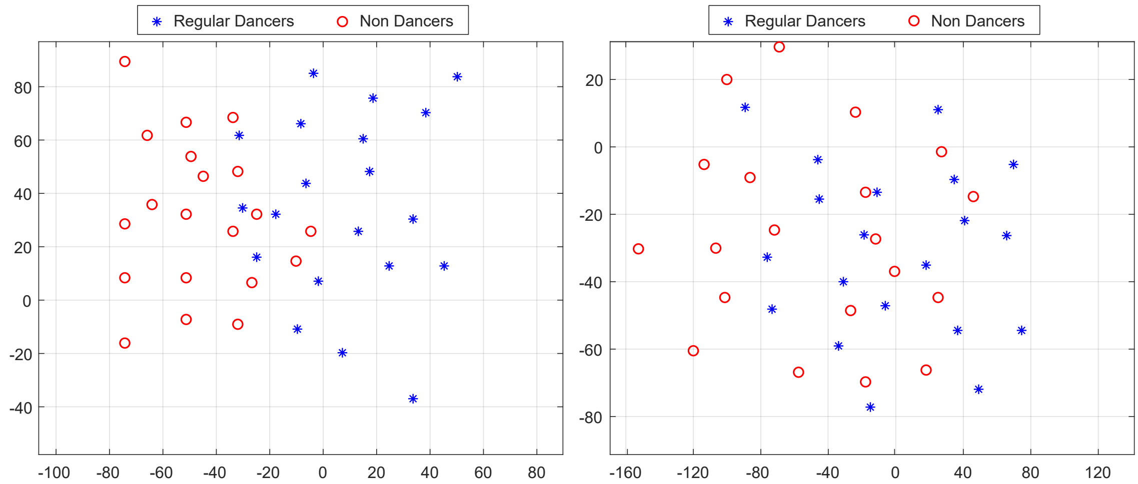


Figure 5.20 – The 2D embedding using t-SNE, when the two classes are represented by the distribution of their LMA-derived arrays [11]. Again, it can be observed that the two classes can be partly separated in the early exercises (left), but as we move forward to the latest exercises (right), the two classes are mixed together. This indicates that our dance VR application helps the participants to improve their dancing stylistic behavior.

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 [130], 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 is LMA-derived arrays; stylistically similar performances have a resemblance distribution, while stylistically dissimilar performances have a different distribution. The distance between these LMA-derived arrays was computed using the Earth Mover's Distance (EMD) metric [119]; note that, EMD performs better than the Euclidean distance, or the Pearson Correlation Coefficient that was originally used in [4]. 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 Figure 5.20. 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.

5.6 Conclusion

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 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.

Chapter 6

Conclusion

"You have to love dancing to stick to it. It gives you nothing back, no manuscripts to store away, no paintings to show on walls and maybe hang in museums, no poems to be printed and sold, nothing but that single fleeting moment when you feel alive." Merce Cunningham, dancer.

6.1 Contributions

Facing the challenges for learning and understanding partner dance, we proposed a solution under the form of an interactive learning system that includes a virtual partner with which you can hold hands and guide it to transitioned between different dance pattern on the music, a learning method with defined objectives of guiding your partner and follow the music, and a motion analysis based on customized motion feature that are able to evaluate a dance performance in terms of dance skills. The overall system has been tested and validated through a user study and statistical analysis, and take his base from a solid expert field study and thorough literature review.

This work contributes to the domain of motion analysis by proposing a new approach that takes music-synchronized motion and two partner interaction into account. It can be used for motion data indexing and classification. There is also a contribution for partner dance understanding, as we defined the main dance skills to be learned in salsa, from interview of experts. We also validate the design of a virtual reality salsa simulation. This can be used for developing new pedagogical solution for teaching partner dance and adds to the dance description. This work also contributes to learning domain as we developed and validate a learning method.

Our system can be utilized as a base to develop a partner dance online learning platform, that can be easily connected to the network to send information and add new content or upgrade the lessons and exercises.

6.2 Limitations

Our system presents some limitations. First, we took only three dance skills for the dance evaluation. Indeed the three other dance skills can be difficult to define properly. Secondly, the learning level of people that participated in our experiment is varying a lot as some of them were more less used to move their body in such dance situation. For example, non-dancer girls were more interested to participate in the experiment than non-dancer men. Then, the variation of dance pattern sequence during the recording add some noise for the upcoming processing. Finally, the total duration of the training was very short and for a better learning evaluation, a longer period of time over maybe few weeks with frequent training session is required.

6.3 Further research

Future work includes the study of the three other dance skills (Fluidity, Sharing and Musicality) and their integration into an extended set of MMF. Testing the system with another partner dance can be a good idea to see if a generalization is possible or just if the research would apply on another case study. An extensive learning study with control groups and over a long period of time can enhance the learning evaluation. The virtual partner have its motion a bit sharp, like a robot. Another work can be to improve the naturalness of the virtual partner motion when dancing and when in contact with the user. Beyond the two basic dance pattern "mambo" and "cucaracha", other new ones can be added and especially turns, which means to extend the decision gesture management to another level.

Appendix A

Acronyms

MMF Music-related Motion Features

HCI Human Computer Interaction

HRI Human Robotic Interaction

IMU Inertial Measurement Unit

LMA Laban Motion Analysis

BPM Beats Per Minute

AI Artificial Intelligence

SVM Support Vector Machine

FPS Frames Per Second

IK Inverse Kinematics

KNN k-nearest neighbours

RF Random Forest

VR Virtual Reality

VP Virtual Partner

Appendix B

Hardware

This section lists an overview of all the hardware mentioned in more detail.

(i) Workstation, Intel I7 4770 CPU at 3.40GHz, 24Go RAM, NVidia GeForce 970 GTX

(ii) Workstation, Intel I7 6780 CPU at 3.40GHz, 24Go RAM, NVidia GeForce 1080Ti

(iii) Vicon Motion capture system

(iv) MX13 and MX13+ Vicon camera IR at 120Hz

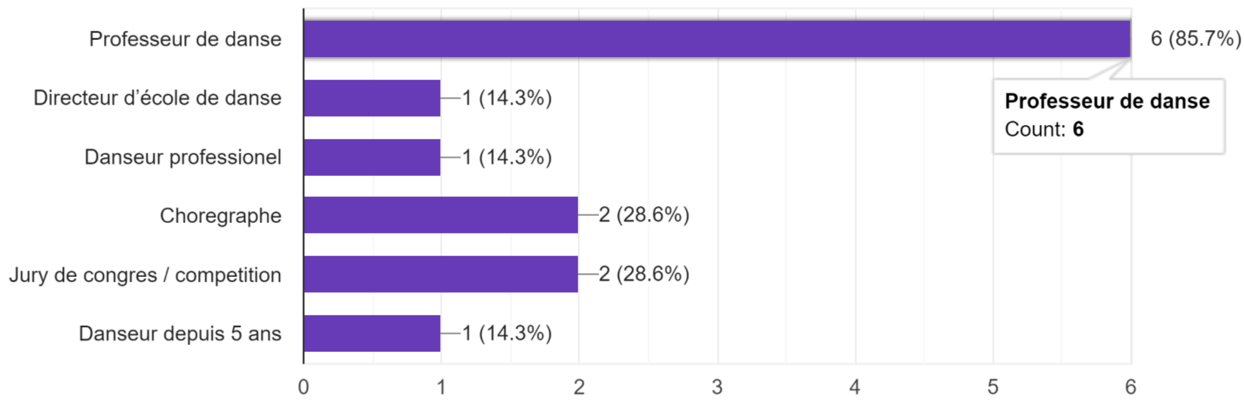
(v) Kinect for windows V2

Appendix C

Supplementary material

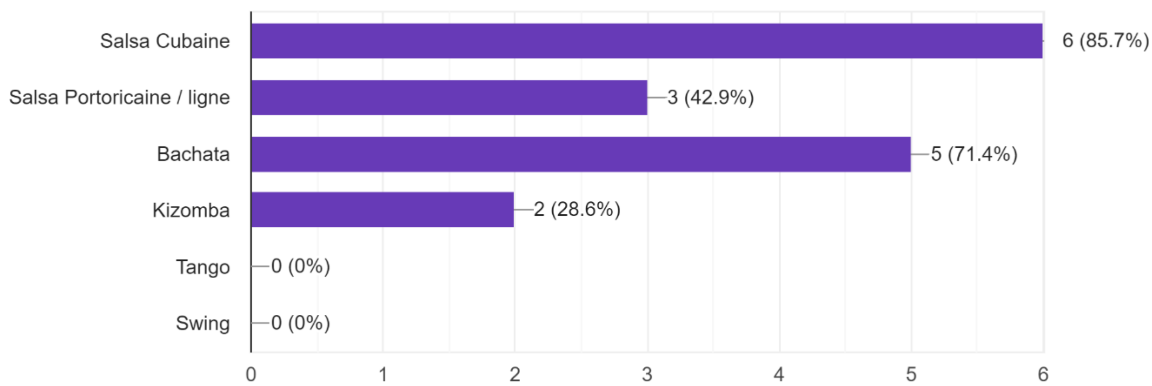
Quel est votre profession / statut par rapport a la danse ?

7 responses



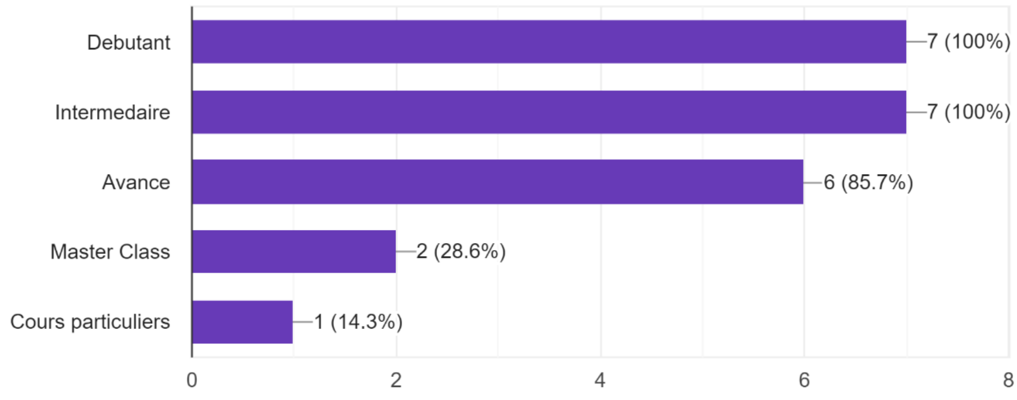
Quels types de danse pratiquez vous ?

7 responses



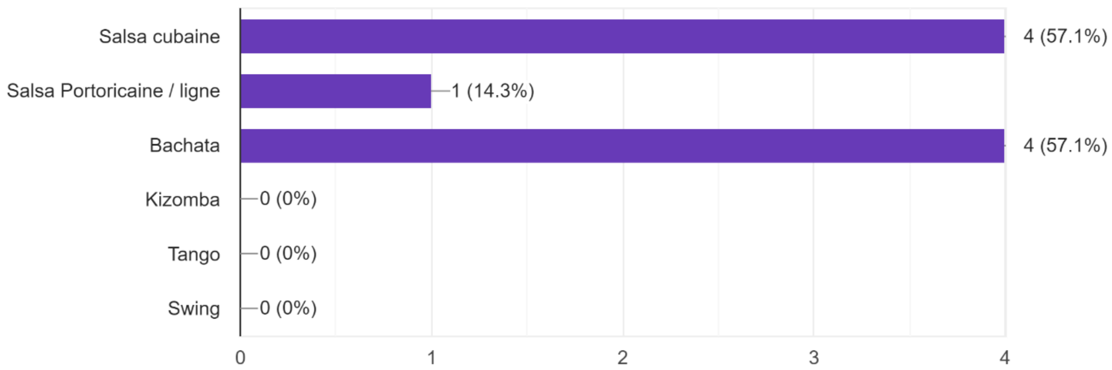
Quels niveaux enseignez vous ?

7 responses



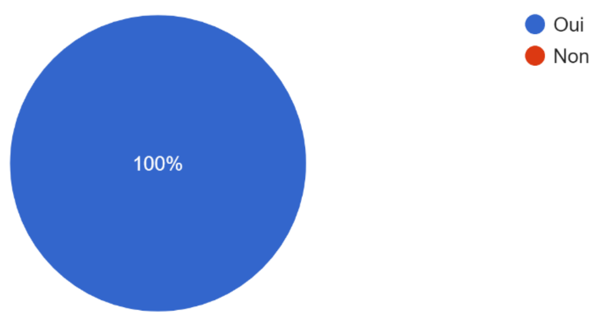
Quel type de danse enseignez vous ?

7 responses



Danser sur le rhythme

7 responses



En quoi ce critère est il important pour le bon développement de l'élève ?

5 responses

1. La connection entre les deux danseur est rythmé et les informations donné par l'homme doivent être sur le rythme de la musique pour être en adéquation avec les pas de la femme. si les danseur ne sont pas sur le rythme ils vont avoir beaucoup de mal à creer une bonne connexion, et donc un bon guidage. 2. Etre sur le rythme signifie être sur la musique pouvoir donc danser avec la musique. Et cela est un des fondement de la danse. 3. Danser sur le rythme signifie aussi garder une régularité dans les pas, et du coup va aider l'élève a fluidifier ses mouvements.

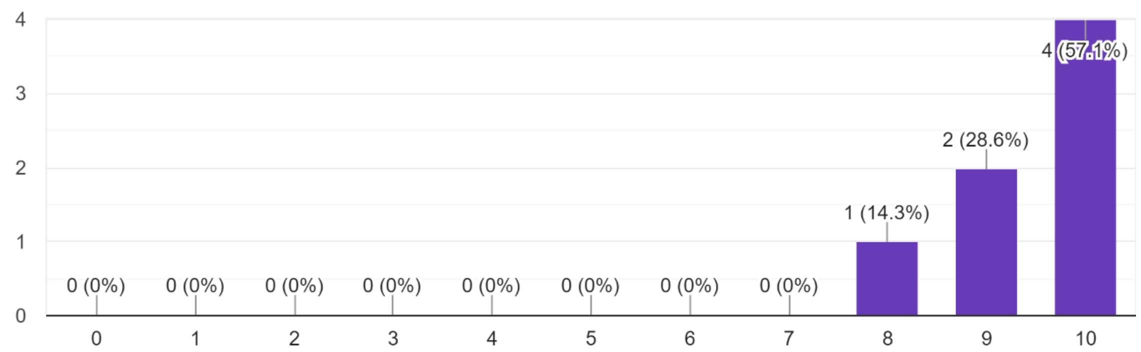
C'est l'un des critères les plus importants et pourtant le plus difficile à transmettre. Vu que nous parlons de danse de couple, le rythme constitue le "langage commun" (de base) qui va permettre de faire le lien virtuel entre les 2 danseurs.

Sans respect du rythme, le(a) partenaire de danse en subit les conséquences, de du manque de fluidité dans les mouvements jusqu'à la bousculade

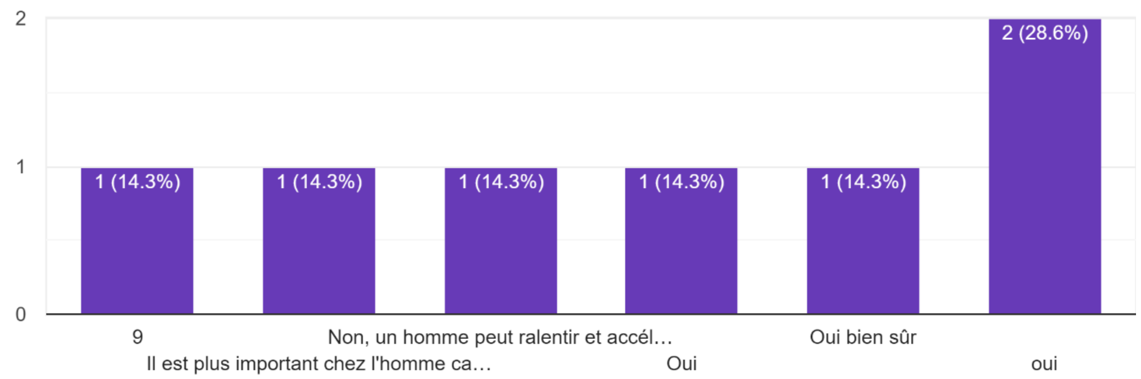
car il est le liant entre le danseur et la musique

C'est une des bases fondamentales, apprendre à écouter la musique est important pour profiter à fond d'une danse et partager un maximum avec notre partenaire

Veuillez annoter l'importance de ce critère sur une échelle de 0 a 10 pour votre évaluation globale.
7 responses

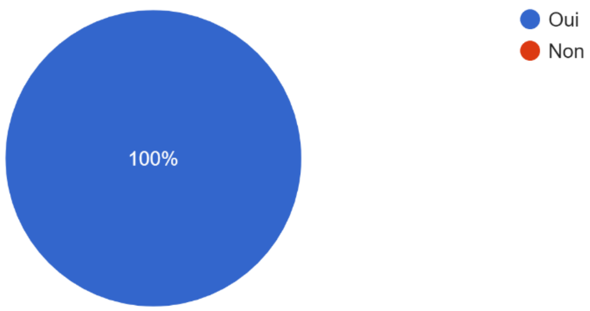


Ce critère as t il la même importance pour les femmes (danse follow) et les hommes (danse lead) ?
7 responses



Guidage / Suivis

7 responses



En quoi ce critère est il important pour le bon développement de l'élève ?

6 responses

1. Pour la connexion entre les deux partenaires. pour que deux personne puisse danser ensemble, il faut un "leader" et un "follower", donc chacun doit être très précis et claire dans son role, sinon le couple aura de la peine a effectuer certaine figures, certain déplacement. 2. Lorsque l'on apprend a danser en cours, on est avec des personne qui on l'habitude des figures que l'on effectue, lorsque l'on va en soirée, on se confronte a de nouvelles personne qui ne connaissent pas les même figures ou n'ont pas les mêmes habitudes. Le seul moyen de pouvoir bien danser de manière fluide avec quelqu'un que l'on ne connait pas, et qui a pas le même background, c'est d'avoir un guidage/suivis clair et précis. 3. ce n'est que lorsque le guidage/suivis, est précis et claire, et qu'il ne soulève pas de doute, que les danseur peuvent profiter pleinement de leur danse et penser au styling, s'amuser. Le mauvais guidage/suivis, est un frein a ca car on doit tout le temps réfléchir ou accentuer le guidage/suivis pour palier aux imprécisions.

Le guidage a une importance clé dans la danse de couple. C'est également important de Savoie se laisser guider et de comprendre les signaux de l'homme.

Si le rythme est le lien virtuel entre les danseurs, le guidage / following est lui le lien tactile. Les bonnes indications de guidage au bon moment va permettre à un danseur va permettre d'effectuer les mouvements / figures. Pour les danseuses, la faculté de following est parfois sous-estimée mais elle est toute aussi importante que le guidage car une danseuse qui ne l'a pas peut faire qu'une danse "ne fonctionne pas", quelque soit le niveau du danseur.

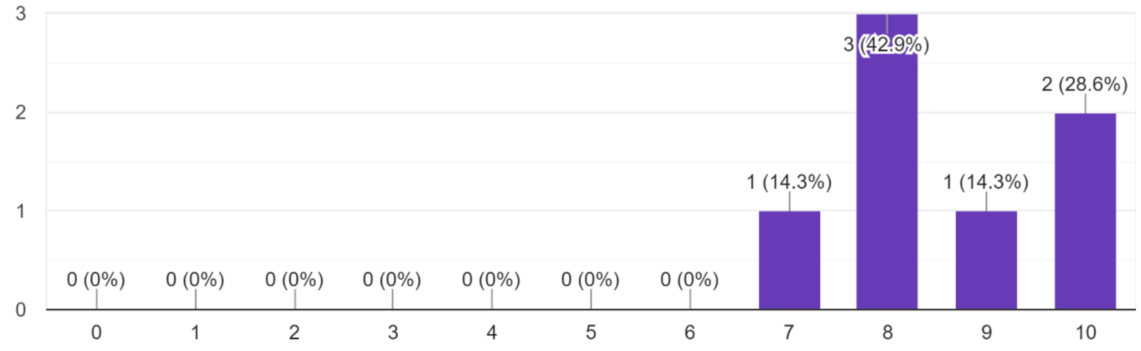
L'homme doit prendre l'habitude de prévoir ses prochains mouvements afin de guider au mieux la danseuse. La femme a besoin de laisser son équilibre être gérer par le guidage du danseur afin de suivre / opérer / reconnaître les mouvements.

car il est le liant entre le leader et le follower

C'est le plus important c'est ce qui crée le rapport d'échange entre les deux parties

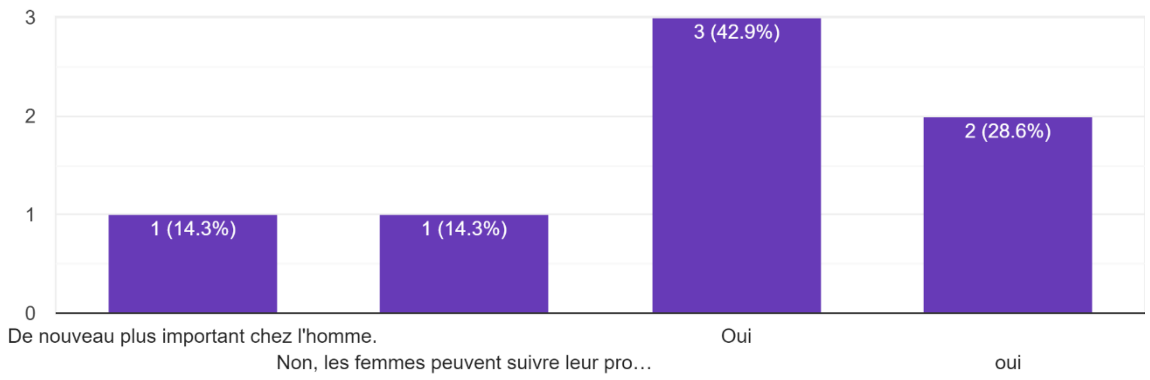
Veuillez annoter l'importance de ce critère sur une échelle de 0 a 10 pour votre evaluation globale.

7 responses



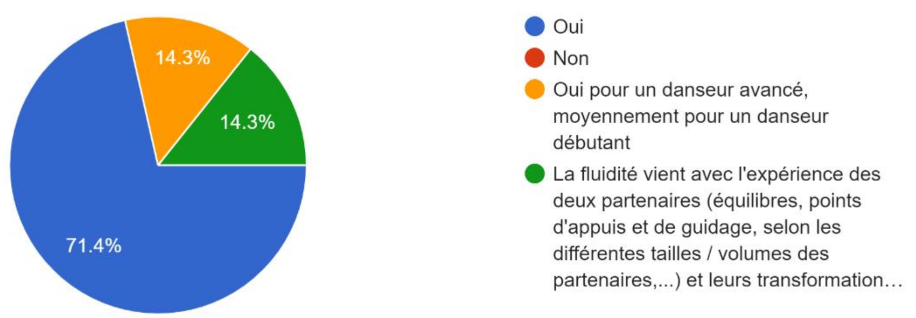
Ce critère as t il la même importance pour les femmes (danse follow) et les hommes (danse lead) ?

7 responses



Fluidite

7 responses

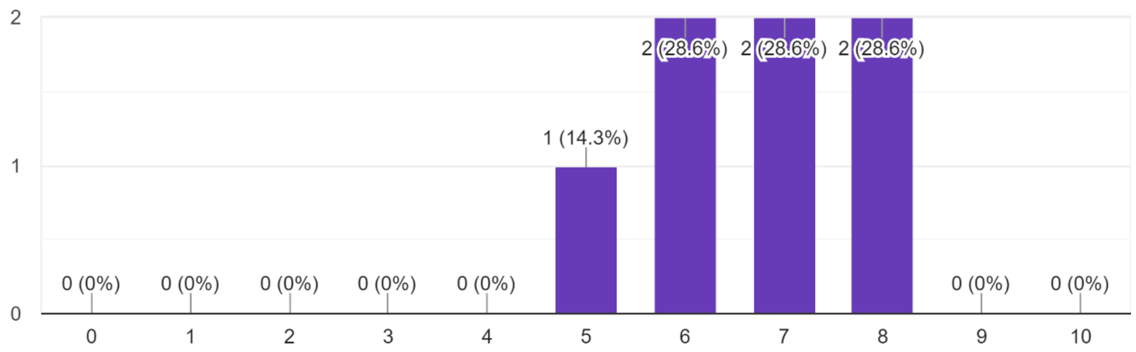


En quoi ce critère est il important pour le bon développement de l'élève ? 5 responses

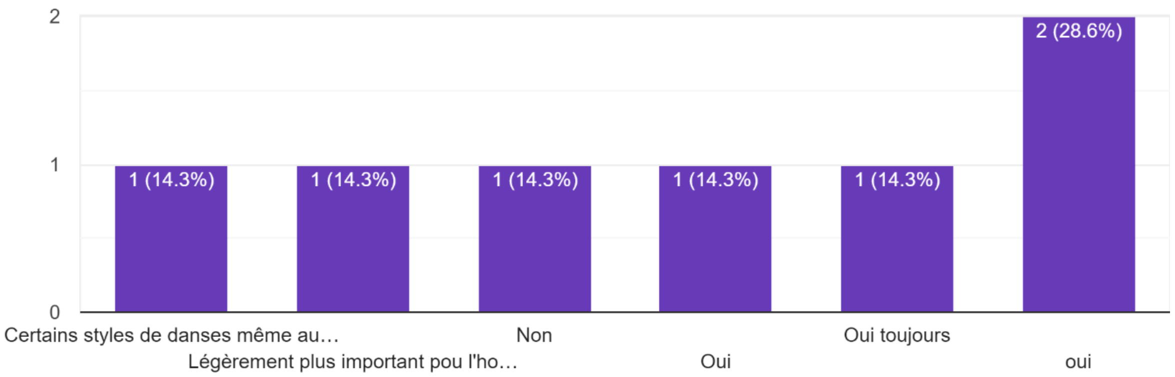
Cela fait parti pour moi d'un bon guidage/suivit, discuter dans la question précédente. mais il est vrais que l'on peu avoir un guidage très précis et juste, sans pour autant qu'il soit fluide.

Être fluide dans les mouvements et dans le guidage/suivis, permet aux deux danseurs de se sentir très confortable dans leur danse et donc de pouvoir être beaucoup plus libre dans le styling, les variations, mieux profiter de la danse. C'est aussi un critère très important our le "visuel" de la danse, un couple ou une personne seul sera plus beau a regarder danser s'il est fluide.
C'est un critère qui rend la danse plus agréable au ressenti et visuellement mais ne constitue pas un critère indispensable selon moi
La fluidité chez le danseur apparaît avec l'expérience du danseur et de la danseuses par un guidage plus souple et une interprétation plus libre. Le risque est que les danseurs(es) n'apprennent jamais à devenir fluide en dehors des séquences de mouvements que les professeurs leurs ont appris.
oui mais pas au début
Pour ne pas que la danse soit brutale et désagréable

Veuillez annoter l'importance de ce critère sur une échelle de 0 a 10 pour votre évaluation globale.
7 responses



Ce critère as t il la même importance pour les femmes (danse follow) et les hommes (danse lead) ?
7 responses



Style

7 responses

En quoi ce critère est il important pour le bon développement de l’élève ? 5 responses

C'est un critère particulier, car il es a la fois très important dans le développement de la danse, du rythme, des variations, et pour l'évolution de l'aisance au mouvement. Mais en même temps il n'est pas fondamentale dans la danse en couple, car on peu avoir un très bon guidage/suivis, un très bonne technique, et ne jamais faire de styling. Je dirai donc que c'est un critère très important pour ceux qui désire travailler leur danse au sens plus large, leur gestuelle, aisance au mouvement etc. Mais qu'il est secondaire pour quelqu'un qui veut danser en sociale, et apprendre a avoir un bonne connexion, une bonne connaissance de figures etc.

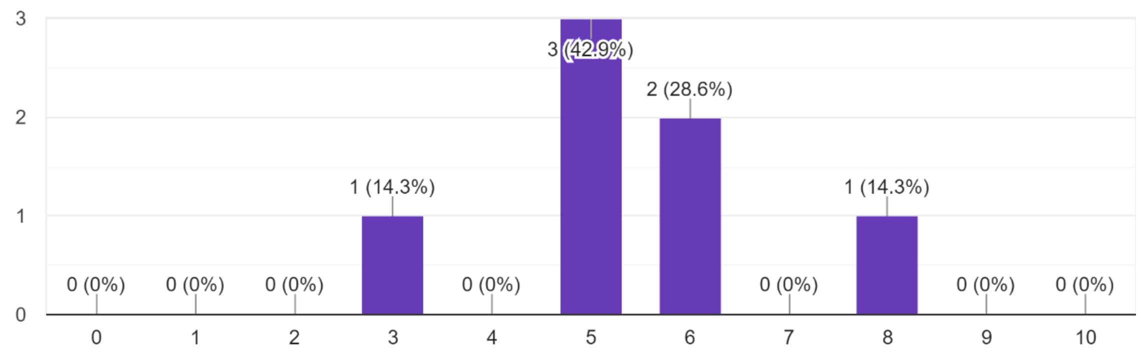
Là il s'agit d'un critère dont l’importance grandi proportionnellement au niveau du danseur

Le style ne peut et ne doit apparaître qu'une fois les bases débutantes, et voir même intermédiaire, sont acquises. Sinon les danseurs risques de prendre des mauvaises habitudes juste pour tenter d'être perçus comme les experts (jusqu'au risque de blesser les autres et soi-même)

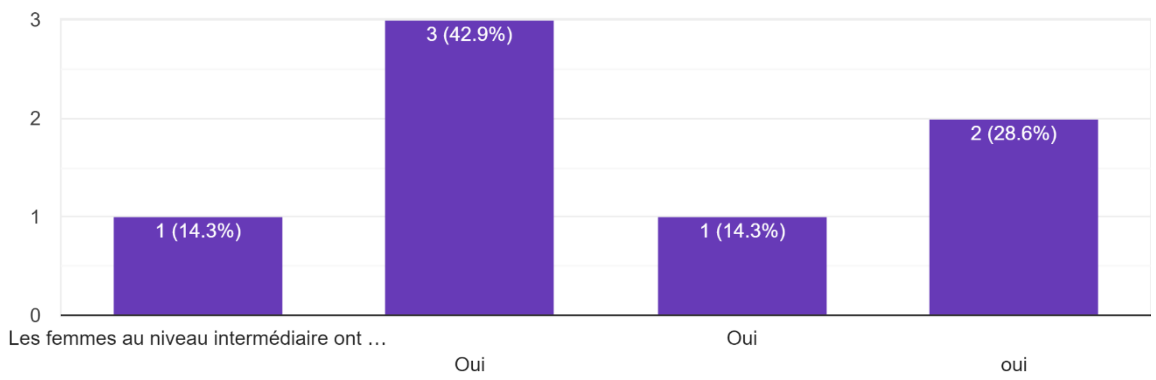
pour qu'il développe sa propre identité

Il est important pour le développement personnel mais il ne faut pas en faire trop ni pour l’homme ni pour la femme... en social nous ne faisons pas le show

Veuillez annoter l'importance de ce critère sur une échelle de 0 a 10 pour votre évaluation globale.
7 responses

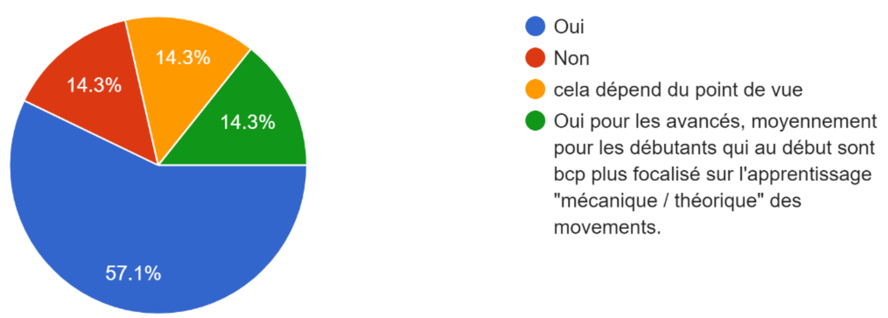


Ce critère as t il la même importance pour les femmes (danse follow) et les hommes (danse lead) ?
7 responses



Intention / Partage

7 responses



En quoi ce critère est il important pour le bon développement de l'élève ? 5 responses

Pour compléter ma première réponse : "Cela dépend du point de vue", voilà deux point de vue différent Si l'on regarde la dance plus techniquement, et simplement le "mouvement" à effectuer dans une figures, entre le guidage, les déplacements et les tours, c'est simplement un mouvement a comprendre et a effectuer, en rythme. donc cela n'implique pas de notion de partage, ou d'intention. d'un point de vue donc purement technique, le partage et l'intention de sont pas important pour apprendre a danser et se perfectionner, car on peu très bien progresser dans le guidage, dans le nombres de figures connues, dans la technique de tour etc, sans ne jamais y prêter attention. D'un point de vue un peu plus réaliste et personnel, le partage est l'intention est le fondement de la danse de couple, quel qu'elle soit. Pour créer une connexion, un quelque chose de spéciale entre les deux danseur durant la danse, il faut avoir envie de partager. Pour donner de l'importance a ces mouvement, au guidage/suivis, et au jeux qui s'installe entre les deux danseurs, il faut de l'intention. C'est donc pour moi peut-être le critère le plus important de tous. Mais compte tenu des deux point de vue expliquer, afin de rester objectif je noterai 5/10, avec un avis personnel à 10/10

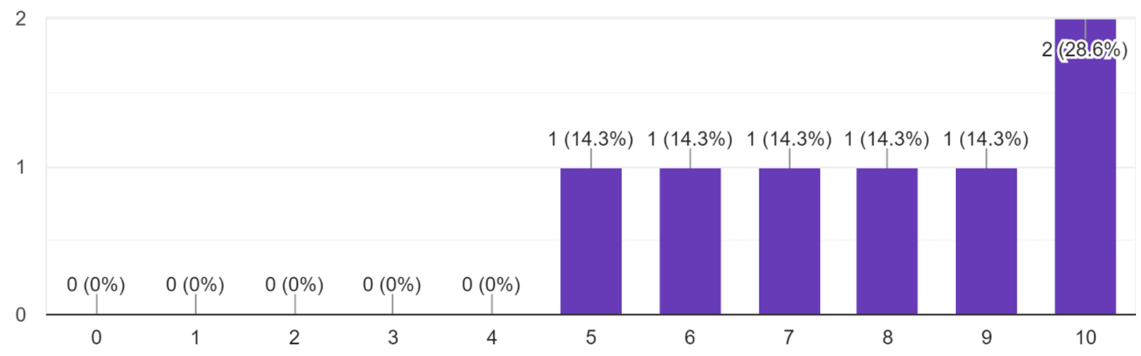
S'agissant de danse de couples c'est effectivement un critère très important qui arrive souvent plus tard dans l'apprentissage, lorsque les danseurs ont déjà un niveau qui leurs permets un certains lâché prise.

Il y a deux sortes d'évaluation : celle du point de vue externe au couple qui évalue ce que le couple partage aux spectateurs, et celle du point de vue interne au couple qui évalue ce que le(a) partenaire à sa/son partenaire. Les regards, la direction du corps, les sourires, ... sont dirigés vers les spectateurs en cas d'une forme de compétition (du "m'as tu vu" jusqu'au concours) ou vers sa/son partenaire hors compétition. Le cas des ruedas et/autres moments de groupes exigent un mix des deux (les spectateurs étant ici les autres membres du groupe, en plus de la foule si compétition).

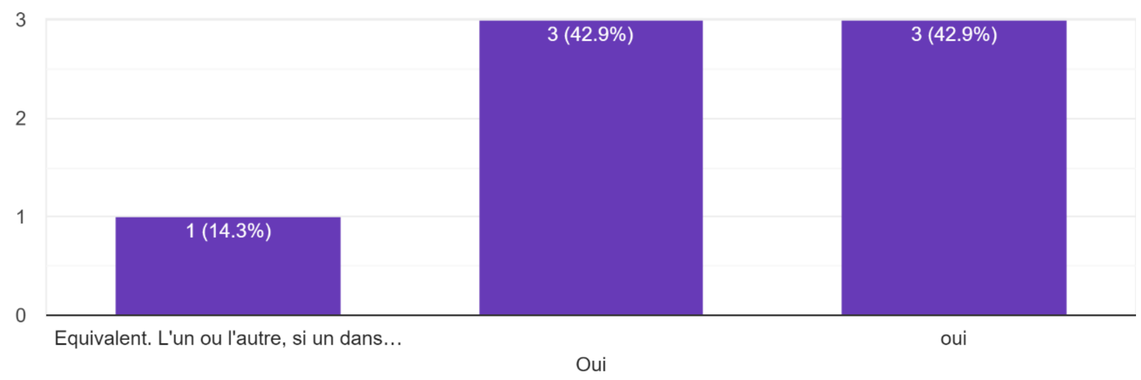
oui car qui dit couple dit partage

Fondamental pour que ça se passe bien, il faut partager tout en restant respectueux

Veuillez annoter l'importance de ce critère sur une échelle de 0 a 10 pour votre évaluation globale.
7 responses

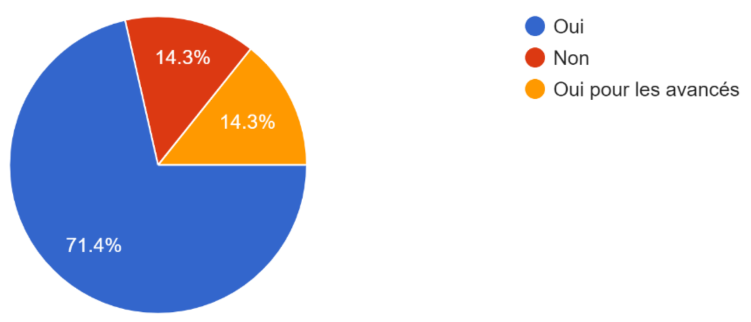


Ce critère as t il la même importance pour les femmes (danse follow) et les hommes (danse lead) ?
7 responses



Musicality

7 responses



En quoi ce critère est il important pour le bon développement de l'élève ? 5 responses

On danse sur la musique, elle nous donne tout, le rythme, l'intensité, les changement de dynamique, l'évolution, les changement de rythmes... Si l'on veut être en adéquation avec la musique, et en rythme, il faut apprendre à l'écouter et à la suivre. Chaque mouvement que l'on peu faire peu être soit inapproprié par rapport a la musique ou simplement sans fondement dans la musique, soit totalement sublimé par la musique s'il est placé au moment opportun et dans la bonne dynamique. La musicalité ouvre une nouvelle porte pour le styling et la connexion avec la partenaire. Après bien sur ca n'est pas primordiale, car on peu très bien danser simplement sur le rythme sans prêter plus attentions a la musique et ca n'est pas "faux" mais la musique apporte quelque chose de beaucoup plus puissant a la danse et a la connexion que l'on peu avoir avec la partenaire.

Arrivé a un certain niveau c'est un critère essentiel au perfectionnement, et la prise de plaisir

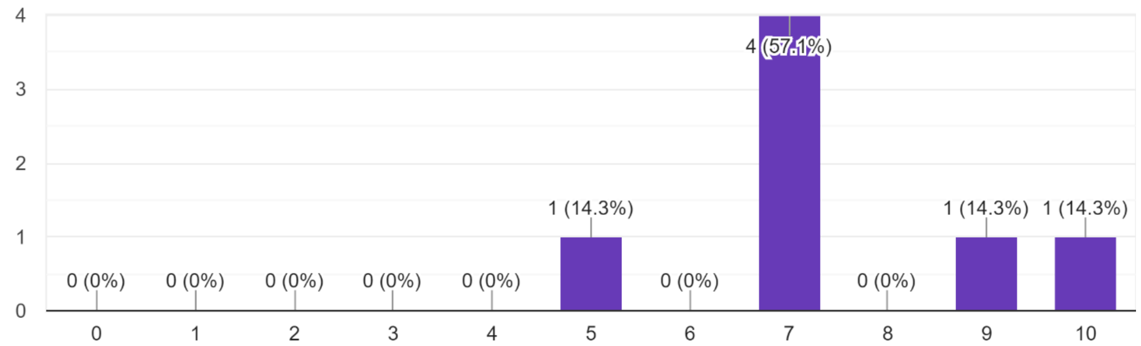
Celui-ci ne devient important qu'après le niveau intermédiaire atteint, en accord avec le développement du style personnel. Il permettra de danser différemment selon chaque musique et donc de permettre de faire profiter au mieux son/sa partenaire surtout en cas de fans de la chanson (puisque certaines nous touchent plus que d'autres, au point de se retrouver à danser avec quelqu'un qui la chante).

pareil que la première question mais au début c'est dur

Au même titre que le rythme

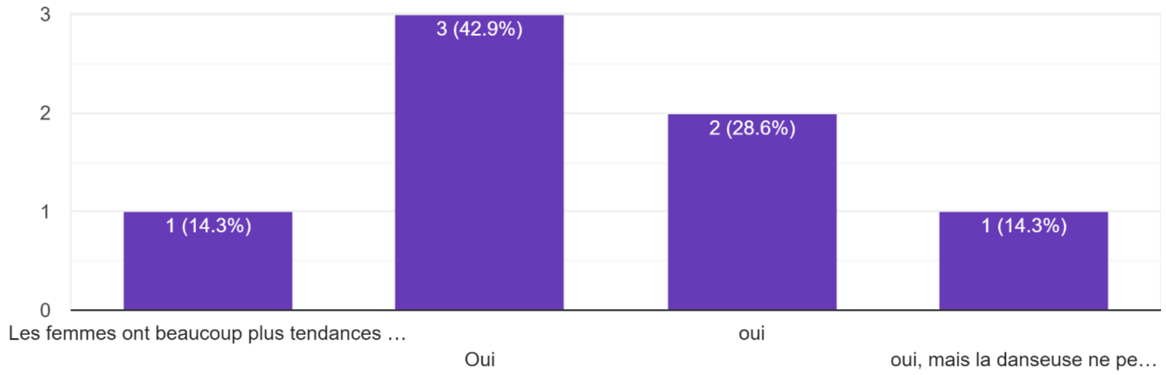
Veuillez annoter l'importance de ce critère sur une échelle de 0 a 10 pour votre évaluation globale.

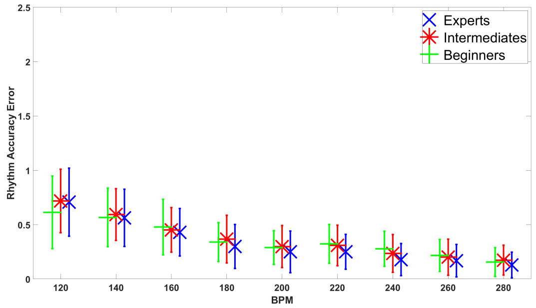
7 responses



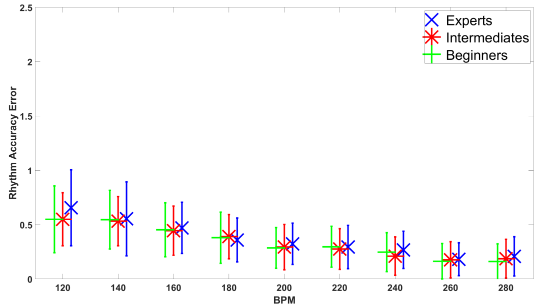
Ce critère as t il la même importance pour les femmes (danse follow) et les hommes (danse lead) ?

7 responses

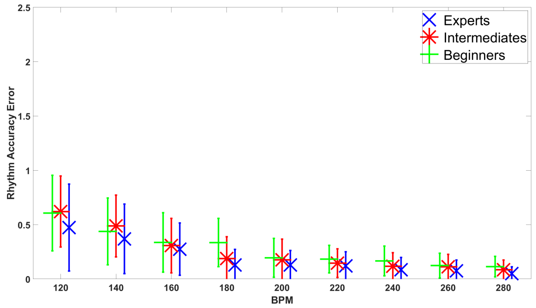




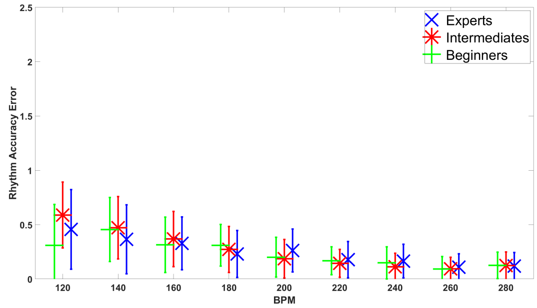
(a) Leader left foot velocity peak time difference with the beat 1 for the basic sequence (μ_1)



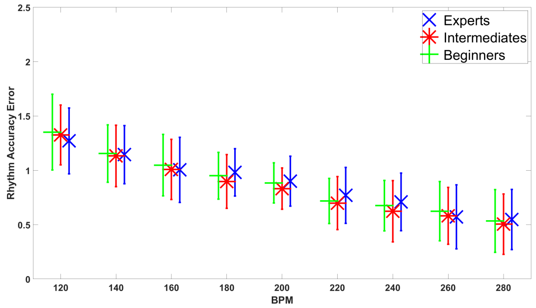
(b) Leader left foot velocity peak time difference with the beat 1 for the improvisation sequence (μ_1)



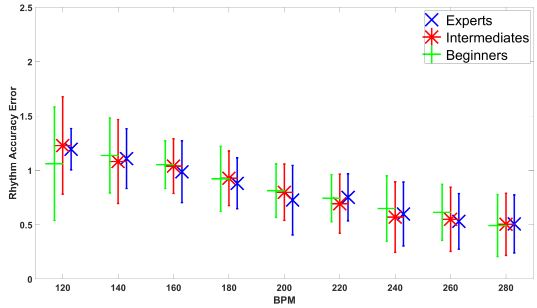
(c) Leader left foot velocity peak time difference with the beat 3 for the basic sequence (μ_2)



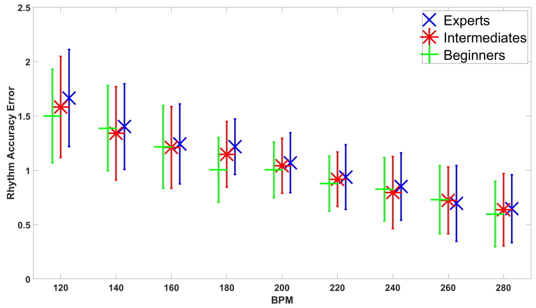
(d) Leader left foot velocity peak time difference with the beat 3 for the improvisation sequence (μ_2)



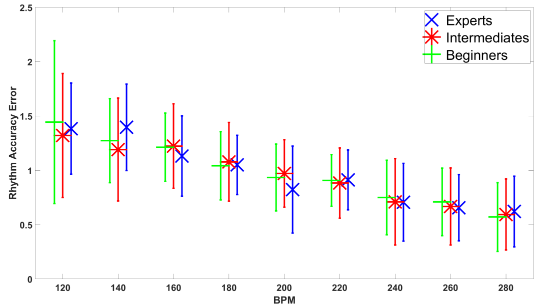
(e) Leader right foot velocity peak time difference with the beat 5 for the basic sequence (μ_3)



(f) Leader right foot velocity peak time difference with the beat 5 for the improvisation sequence (μ_3)

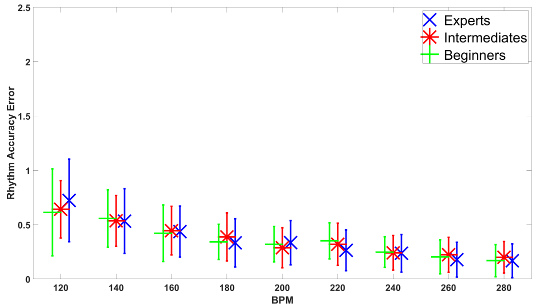


(g) Leader right foot velocity time difference with the beat 7 for the basic sequence (μ_4)

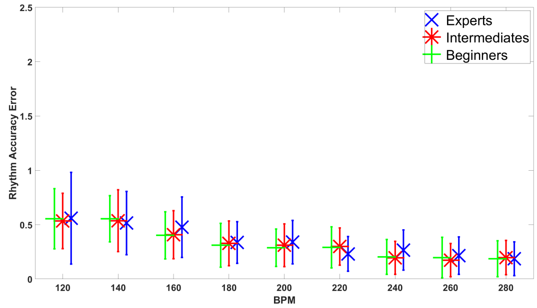


(h) Leader right foot velocity time difference with the beat 7 for the improvisation sequence (μ_4)

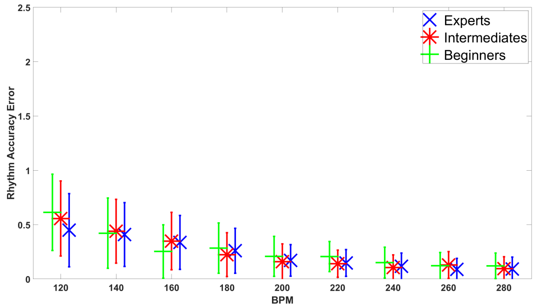
Figure C.1 – Average values and standard deviation per category of the time difference between musical beat and detected leader foot velocity peak μ_{1-4}



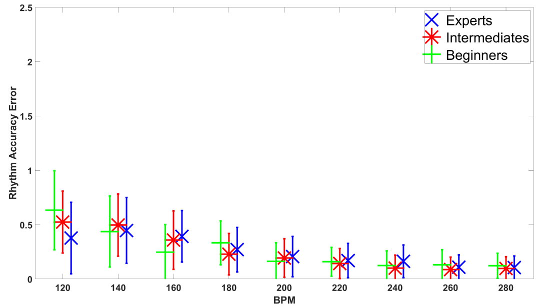
(a) Follower right foot velocity peak time difference with the beat 1 for the basic sequence(μ_5)



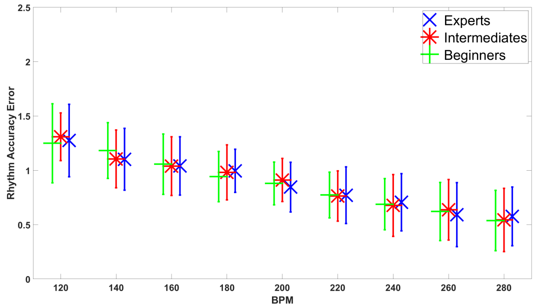
(b) Follower right foot velocity peak time difference with the beat 1 for the improvisation sequence (μ_5)



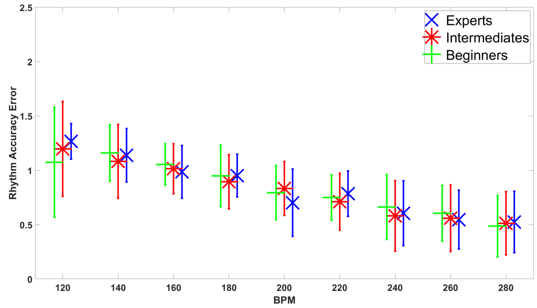
(c) Follower right foot velocity peak time difference with the beat 3 for the basic sequence (μ_6)



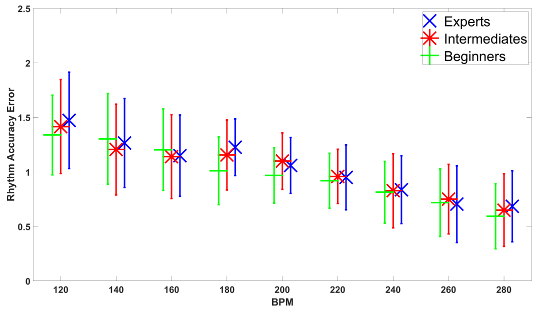
(d) Follower right foot velocity peak time difference with the beat 3 for the improvisation sequence (μ_6)



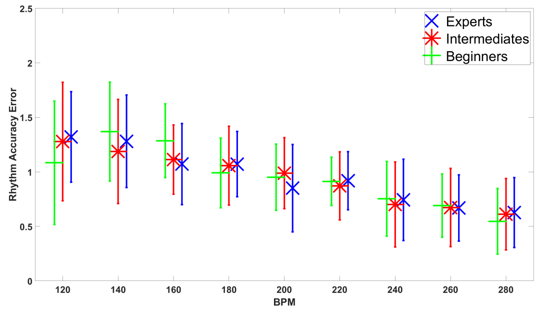
(e) Follower left foot velocity peak time difference with the beat 5 for the basic sequence(μ_7)



(f) Follower left foot velocity peak time difference with the beat 5 for the improvisation sequence (μ_7)

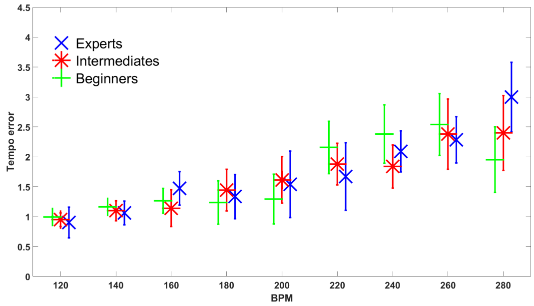


(g) Follower left foot velocity peak time difference with the beat 7 for the basic sequence (μ_8)

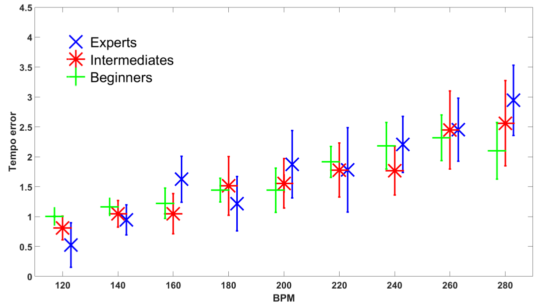


(h) Follower left foot velocity peak time difference with the beat 7 for the improvisation sequence (μ_8)

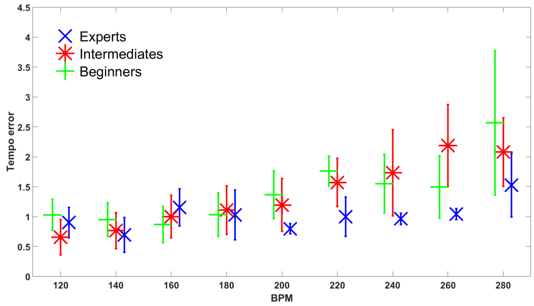
Figure C.2 – Average values and standard deviation per category of the time difference between musical beat and detected follower foot velocity peak μ_{5-8}



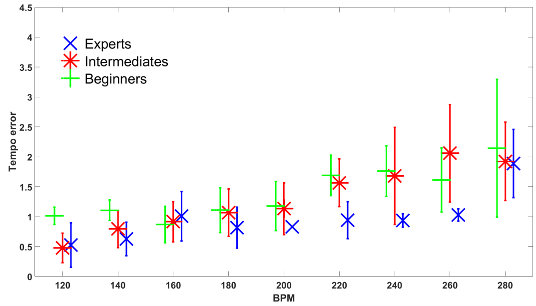
(a) Leader main foot velocity spectrum frequency for the basic sequence (μ_9)



(b) Leader main foot velocity spectrum frequency for the improvisation sequence (μ_9)

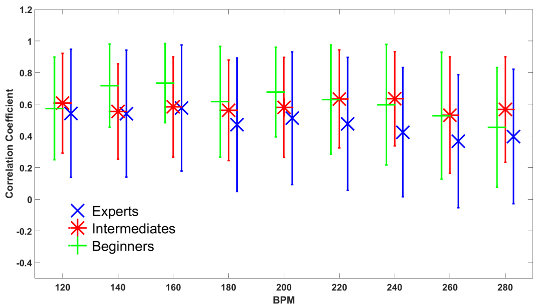


(c) Follower main foot velocity spectrum frequency for the basic sequence (μ_{10})

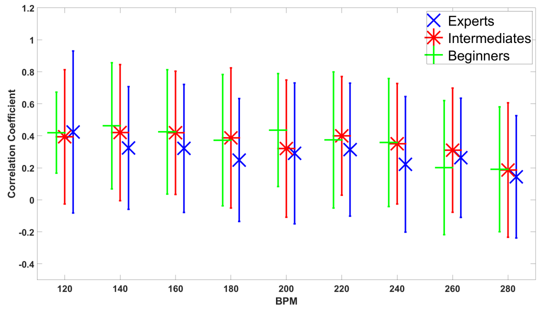


(d) Follower main foot velocity spectrum frequency for the improvisation sequence (μ_{10})

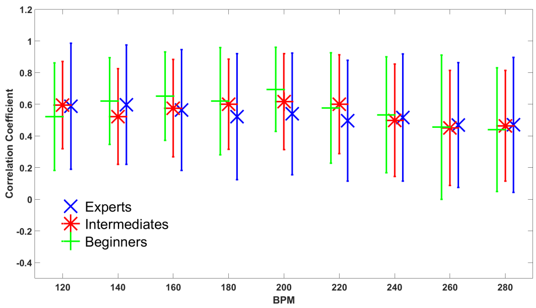
Figure C.3 – Average values and standard deviation over each tempo of the main foot velocity spectrum frequency, as the average tempo, μ_{9-10}, averaged per category.



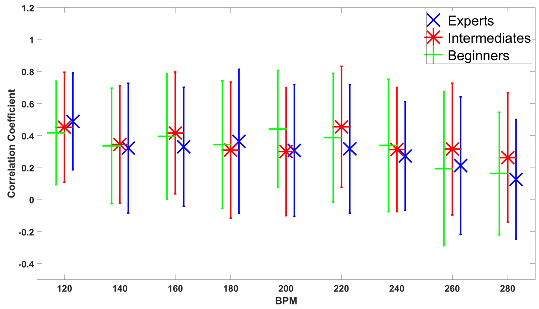
(a) Linear correlation coefficient for the beat 1 to 3 of the basic sequence (μ_{11})



(b) Linear correlation coefficient for the beat 1 to 3 of the improvisation sequence (μ_{11})



(c) Linear correlation coefficient for the beat 5 to 7 of the basic steps sequence (μ_{12})



(d) Linear correlation coefficient for the beat 5 to 7 of the improvisation sequence (μ_{12})

Figure C.4 – Average values and standard deviation per category of the linear correlation coefficient between foot velocity of leader and follower μ_{11-12}

Table C.1 – Motion analysis (1/3)

	Ref	Frame-work	Motion features - measurements	Experiment	Motion qualities - High level space	Validation technique	Results
1	Aristidou [4, 9]	LMA (Effort, Shape, Space and Body)	27 motion features in 86 sub measurements	Contemporary dance	Bidimensional space of Energy and Activation (emotion)	Comparison with manual annotation - random forest	87% recognition
2	Alexiadis [3]	Huawei 3DLife/EMC	Joint position, joint velocities, 3D flow error	13 amateurs dancers compare to two professional for Salsa dance.	Signal similarity (Choreography, Musical timing, Body balance)	Algorithm ranking compare to expert ranking	Strong correlation in ranking
3	Alborno [1]	Multilayer DANCE project [34]	Velocity peaks of upper limbs	Analyst annotated performances of 2 professional dancer expressing motion qualities. Total time is 5 minutes and segments are 5s long in average.	Intra-personnal synchronisation (fluidity, impulsivity and rigidity)	Comparison with experts ranking	Ranking of movement synchronisation: Impulsive > Rigid > Fluid
4	Niewiadom-ski [94]	Multilayer DANCE project [34]	Multimodal (IMU, EMG, Camera and microphone)	Pop, Contemporary, classic	Lightness and Fragility	Naive Baiyes classifier	77% recognition
5	Niewiadom-ski [93]	Multilayer DANCE project [34]	Biomechanical efficiency, Static posture, Intrapersonnal coordination (16 measurements)	Karate's Kata performance annotated by expert	Movement quality score	Pearson correlation coefficient between experts and algorithm	Up to 84% correlation
6	Piana [109]	Multilayer DANCE project [34]	Postures, Kinematic and Geometry	Emotion expressed by children with autism	Six basic emotions	SVM classifier	61%
7	Piana [108]	Multilayer DANCE project [34]	Mean Jerk value of Shoulders, Elbow and Hands from dancer and from simulated model	Two professional dancers provided 42 full body movements while asked to have fluid or non fluid motion. Each sequence annotated by 41 user.	Fluidity computed with a reference model of Spring-mass that does the same movements with perfect fluidity	Comparison with perceptual study	high correlation between the algorithm and the user review

Table C.2 – Motion analysis (2/3)

	Ref	Frame-work	Motion features - measurements	Experiment	Motion qualities - High level space	Validation technique	Results
8	Dong [38]		Hilbert Huang transform	Waltz and Salsa	Dance pattern		Outperform previous approach
9	Faridee [45]		Inertial sensors over the body	Lasya traditional indian dance (1 experts and 4 learners)	Dance microsteps	Deep neural network	Outperform Random Forest by 7%
10	Fourati [50]		114 Body reduced to 11 body cues	Emotion variation within daily actions (Walk, sitting down, etc.)	8 basic emotions	Random forest classifier	67% to 97% recognition
11	Kim		Angles from skeletal data reduced by PCA	400 motion of K-pop dance	Dance pattern, dance style	Extreme machine learning	94.5% recognition
12	Kitsikidis [69]		kinect Skeletal data, Waist horizontal motion	Geek Tsamiko dance, 3 experts and 9 students	Dance pattern	Exemplar based hidden Markov model compared to manual annotation	5% max error
13	Masurelle [80]	Huawei 3DLife/EMC	Footstep impact detection plus skeletal joint position	Salsa dancer doing 6 different dance pattern	Dance pattern	Hidden Markov Model	74% F-measure
14	Kar-avarsamis [63]	Huawei 3DLife/EMC	Footstep impact detection plus skeletal joint position	Salsa dancer doing 6 and 20 different dance pattern	Dance pattern	Random Forest after PCA	94% F-measure for 6 dance pattern, 45% for 20 dance pattern

Table C.3 – Motion analysis (3/3)

	Ref	Frame-work	Motion features - measurements	Experiment	Motion qualities - High level space	Validation technique	Results
15	Chantas [24]		Motion capture data: joints in space	30 experts and 30 beginners doing 20 Salsa dance pattern and Greek Tsamiko	Performance (Execution accuracy and Synchronisation)	Multi Entity Baeysian network - Comparing experts to beginners	For Salsa: 13% difference in execution accuracy and 5% difference for synchronisation
16	Morioka [89]		Power, Speed, Stability, Spread and Regularity	Correlation between music and motion that carry the emotions	7 Emotion (happy, lonely, sharp, etc.)	Correlation between manual annotation and algorithm	90% in average for all emotion
17	Ofli [95]		Chroma feature of music and body motion capture	Turkish dance Kasik	Dance pattern	Hidden Markov Model	Qualitative production of motion to suit the dance
18	Ozcimder [98]		Motion primitives from data	20 couple dance pattern of Salsa annotated by 15 judges	Artistic merit (phrase complexity, distance covered)	Correlation between judges and the two metrics	81% correlation
19	Samanta [124]		Posture based on optical flow from video	Indian classical dance - database of youtube video	Dance pattern	SVM with intersection kernel	86.67% recognition
20	Shiratori [136]		Center of mass and hands speed, music onset detection	Traditional Japanese dance, pauses helps the segmentation	Dance pattern		Qualitative results show good match
21	Saha [120]		Distance and angles of joints	Traditional Indian dance Odissi performed by 7 dancers	Posture	Mapping onto a 4 sided polygon then introduced into a SVM	92.7% recognition rate with a 80/20% classification strategy

Table C.4 – Interactive Dance Learning System (1/3)

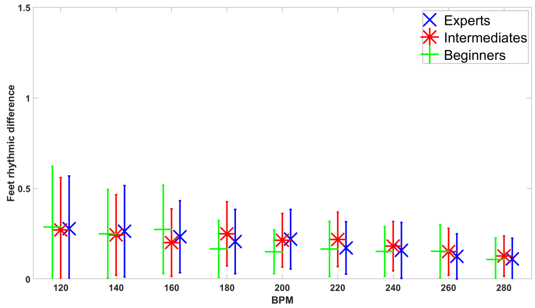
	Ref	Dance	Experiment	Motion features	Evaluation	Feedback	Result
1	Kyan [74], Sun [139]	Ballet	Two trainees repeating three dance movements	LMA components: Body Shape and Effort, and Rhythm	Comparison with a trainer performance, Motion accuracy, Musical Timing	Visual overlay correct motion, side view, numeric score, slow motion playback, score as graph, use of a metronome	4-10% improvements
2	Raheb [113]	Contem-porary, Improvisa-tion	Wide space with various parameters of the visualisation controllable by the dancer. Three professional ballet and contemporary dancer and 12 practitioners.	Kinematics of the body joints	Questionnaire	Different avatar visualization, scene, visual effect, and virtual objects.	System, perceived as stimulating and attracting. Different avatar lead to different movements and inner emotion.
3	Molina [86]	Ballet	Delay mirror present during a five day intensive ballet course for adult, with 8 participants.	Video capture	n/a	Image of the dancer while practicing delayed of 15s, allowing self visualization	Positive appreciation from teacher and dancer: act as objective feedback, is ephemeral.
4	Alexiadis [3]	Salsa dance	13 amateurs dancers compare to two professional for Salsa dance doing six choreography.	Signal similarity (Choreography, Musical timing, Body balance)	Algorithm ranking compare to expert ranking visualizing the performance	Avatars side by side, Score	Strong correlation in ranking
5	Aristidou [12, 10]	Cypriot Folk	User reproduce the motion showed by a 3D avatar. One expert and a few beginners.	LMA component: Body, Shape, Effort, Space: 70 related measurements on 35 frame sliding window	Correlation between LMA component between expert and beginners for each time window.	LMA components score.	The high correlation shows validates the learning system
6	Kitsikidis [70]	Greek Folk (Tsamiko)	1 Expert, 2 experienced and eight beginner dancers during two recording sessions of 4min. Sequences manually annotated.	Metrics: Knee and Ankle distance. Pattern recognition with HCRF.	Comparison with expert - Two-level fuzzy inference system. One dimensional performance score.	3D environment with visualization	The score matched the ranking of experienced and beginner category.

Table C.5 – Interactive Dance Learning System (2/3)

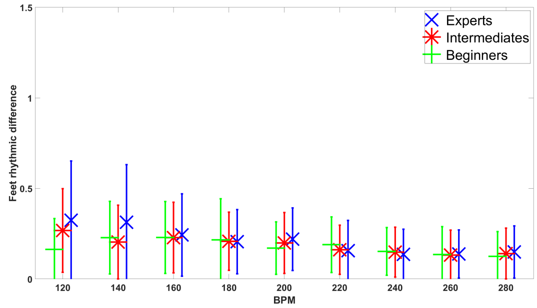
	Ref	Dance	Experiment	Motion features	Evaluation	Feedback	Result
7	Thalmann [76]	Traditional dance (UK, Bulgaria, Greece)	Reproduce the movements shown by an avatar	n/a	Experts from different countries, students filling a questionnaire	Visualizing the motion-captured recording with online interface	Questionnaire show good results in term of motivation to use the system - 77%.
8	Camurri [20]	Any	The user performs with the sensors.	Dynamic Symmetry from two inertial sensors on wrists.	Absolute value of the motion feature	Musical rewarding when good value	Project demo at a Workshop
9	Chan, Tang	Hip-hop, A-go-go	4 learners try to reproduce motion they see on-screen for 15 min compare to 4 learners with videos only.	Motion accuracy, Musical Timing	Motion matching to expert motion with euclidean distance between joint angle sets.	Immediate feedback (3D representation of user), score report and slow-motion	Ctrl: 37.08 => 37.92, Exp: 40.58 => 51.41 Questionnaire: Easier to learn with the system
10	Santos [128, 127]	Forro	Using an app on smartphone, 10 students completed a 3-week course with additional training provided by the app, 18 times in total.	Rhythm consistency, Rhythm BPM, Practice and Body motion	Comparision with a reference values	Summary, narrative, Social Comparison, and Visualisation	Narrative and Summary feedback are highly appreciated. high correlation between app usage and performance.
11	Drobny [40]	Slow Waltz	4 experienced couples, 4 beginner couples dancing 5 songs. One control group.	Step detection with force sensors, rhythm consistency, standard deviation of steps timing.	Estimation of the rhythm consistency	Linear increase of the rhythm music track if errors are detected	Improvement of the skills of the experiment group. Questionnaire prove high acceptance.
12	Yang [155]	Latin dance and Hip-hop	Each user have to learn first a series of dance patterns and then in a second phase assemble them into choreography. 52 Students participated.	3D position in time	Self similarity matrix from teacher data	n/a	Questionnaire shows better satisfaction with the system

Table C.6 – Interactive Dance Learning System (3/3)

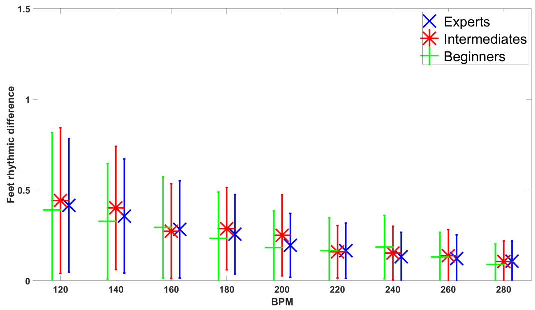
	Ref	Dance	Experiment	Motion features	Evaluation	Feedback	Result
13	Ros [117, 118, 116]	Creative dance	11 children in hospital participated during 11 days. A first phase where concept motion is demonstrated by a robot. A second phase with robot asking to show the concept.	Analysis of the response from the children	Wether if the children are able to show the required motion or if he needs another demonstration	The robot is reacting to the children	Children show more engagement with the robot.
14	Granados [101]	Waltz	12 Volunteers novice in dance. After tutorial, they practice 30min with the robot.	Force applied to the robot	Referenced to expert dancers	Progressive teaching: adaptation to the level of the user.	Questionnaire shows that progressive teaching module increase Comfort, Peace of Mind and Performance.
15	Nakamura [91]	Traditional Japan dance Tsug-arabushi	10 beginners are testing two experiments: one moving screen for translational information and vibration wristband for timing information. A control group performed with video-only.	Walked distance, timing error	Comparison with control group	Wrist vibration, moving screen in space.	50% error reduction and variance decrease for walked distance, 20% better timing.
16	Tsuchida [148, 149]	Group dance	9 dancers with experience doing three choreography of 8 seconds with moving screen, moving curtain, and a control group with only video and only robot.	Dual dancer motion: Approaching, parallel translation and intersection	Comparison with control group	Moving screen with video	Movable curtain presents the best results as it does not carry collision fear



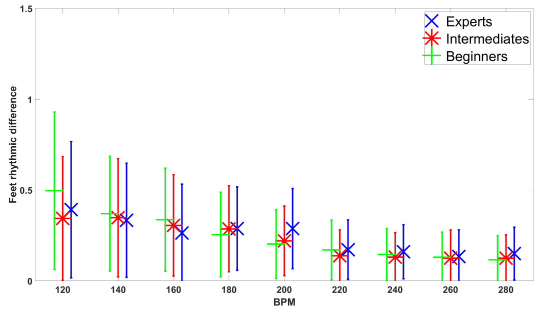
(a) Foot velocity peak time difference between partner beat 1 for the basic sequence (μ_{13})



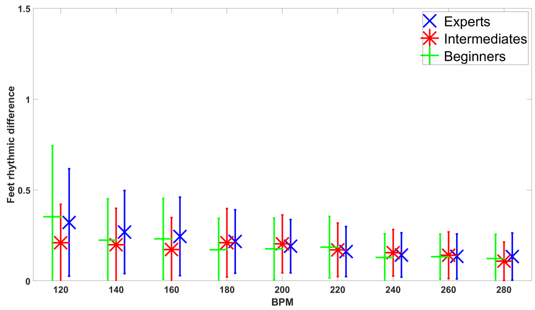
(b) Foot velocity peak time difference between partner beat 1 for the improvisation sequence (μ_{13})



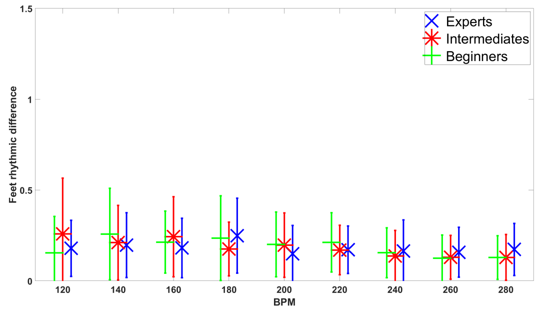
(c) Foot velocity peak time difference between partner beat 3 for the basic steps sequence (μ_{14})



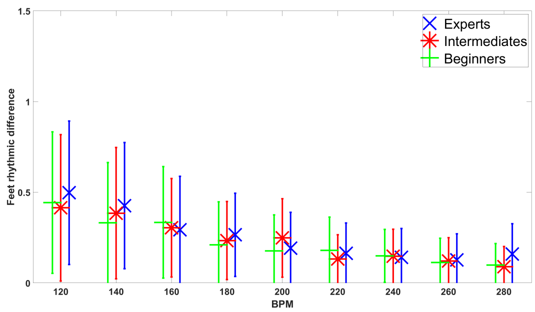
(d) Foot velocity peak time difference between partner beat 3 for the improvisation sequence (μ_{14})



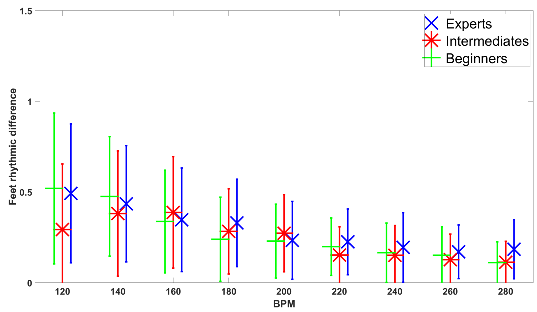
(e) Foot velocity peak time difference between partner beat 5 for the basic steps sequence (μ_{15})



(f) Foot velocity peak time difference between partner beat 5 for the improvisation sequence (μ_{15})

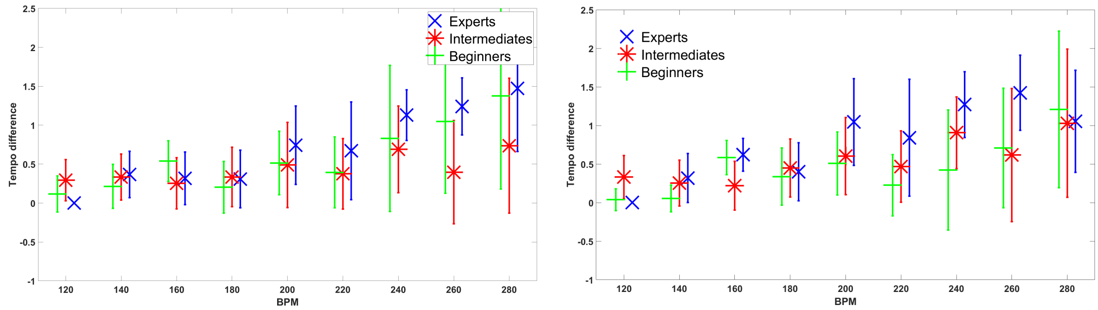


(g) Foot velocity peak time difference between partner beat 7 for the basic steps sequence (μ_{16})



(h) Foot velocity peak time difference between partner beat 7 for the improvisation sequence (μ_{16})

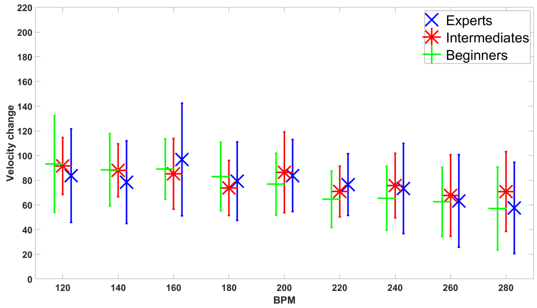
Figure C.5 – Average values and standard deviation per category of foot velocity peak time difference between partners μ_{13-16}



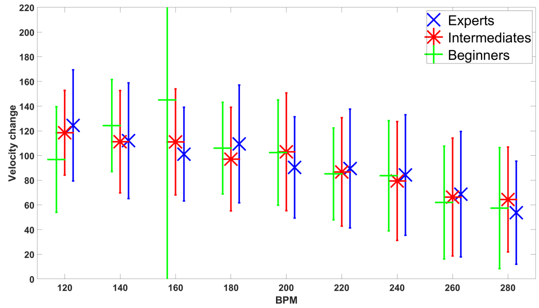
(a) Partners frequency difference between the main peak of the foot velocity spectrum for the basic steps sequence (μ_{17})

(b) Partners frequency difference between the main peak of the foot velocity spectrum for the improvisation sequence (μ_{17})

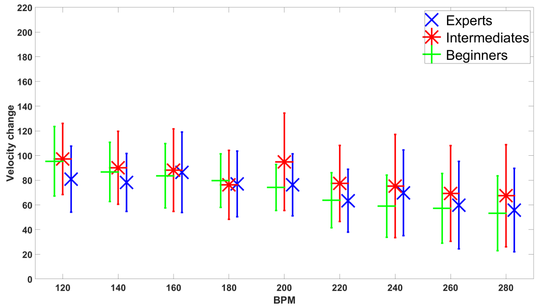
Figure C.6 – Average values and standard deviation per category of foot velocity main spectrum frequency difference between the main peaks of leader and follower μ_{17}



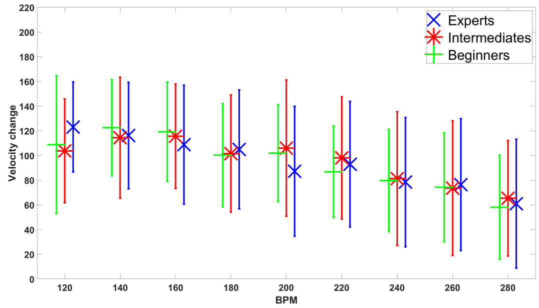
(a) Leader net left foot velocity change for the basic sequence (μ_{18})



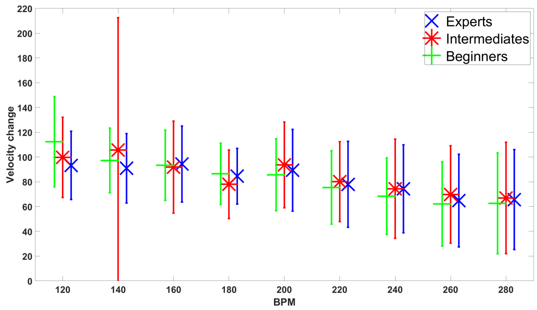
(b) Leader net left foot velocity change for the improvisation sequence (μ_{18})



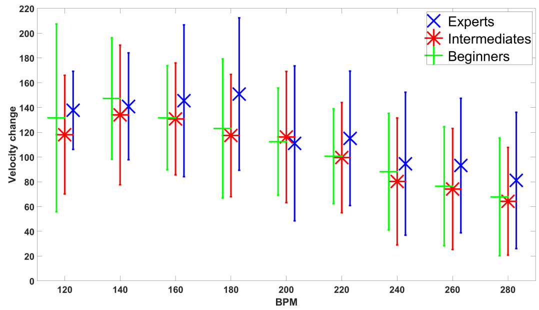
(c) Leader net right foot velocity change for the basic sequence (μ_{19})



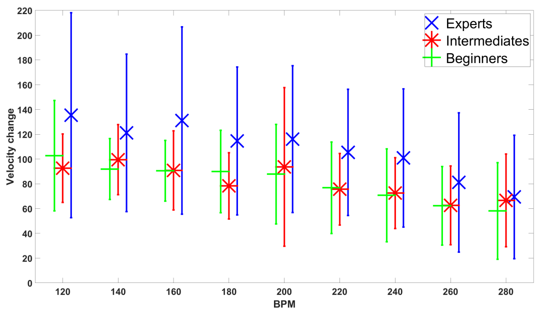
(d) Leader net right foot velocity change for the improvisation sequence (μ_{19})



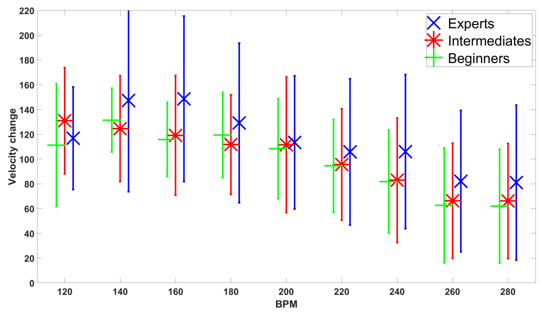
(e) Follower net right foot velocity change for the basic sequence (μ_{20})



(f) Follower net right foot velocity change for the improvisation sequence (μ_{20})

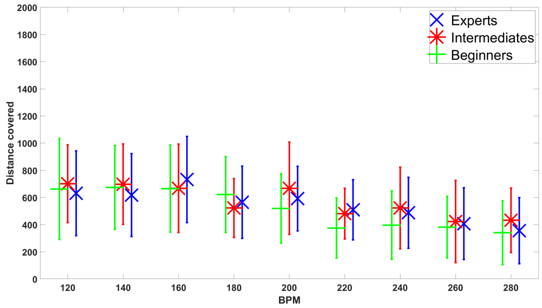


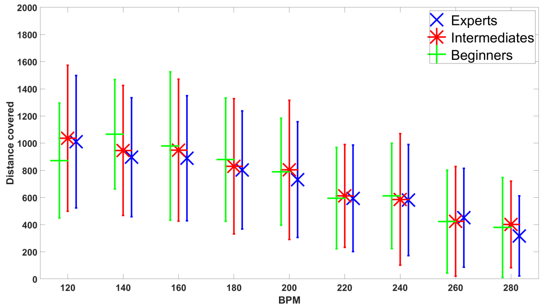
(g) Follower net left foot velocity change for the basic sequence (μ_{21})



(h) Follower net left foot velocity change for the improvisation sequence (μ_{21})

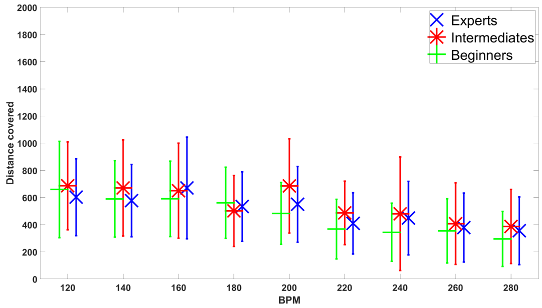
Figure C.7 – Average values and standard deviation per category of the feet net velocity change μ_{18-21}

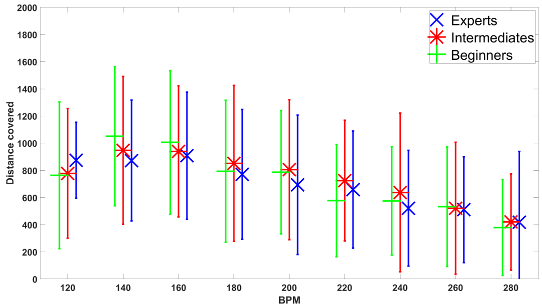




(a) Leader area covered for the left foot the basic sequence (μ_{22})

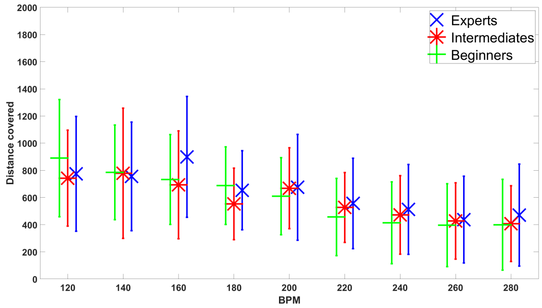
(b) Leader area covered for the left foot for the improvisation sequence (μ_{22})

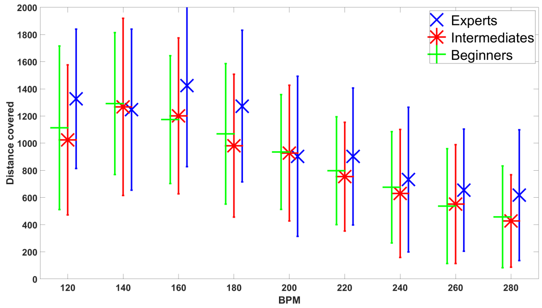




(c) leader area covered for the right foot for the basic sequence (μ_{23})

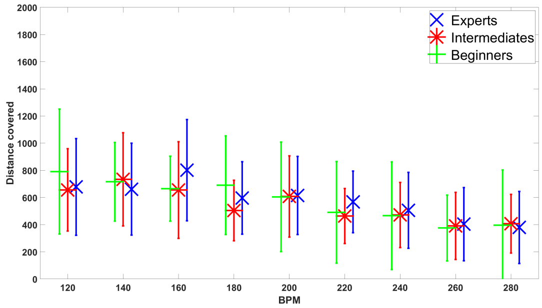
(d) Leader area covered for the right foot for the improvisation sequence (μ_{23})

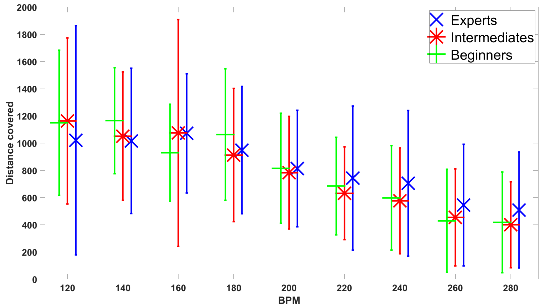




(e) Follower area covered for the left foot for the basic sequence (μ_{24})

(f) Follower area covered for the left foot for the improvisation sequence (μ_{24})

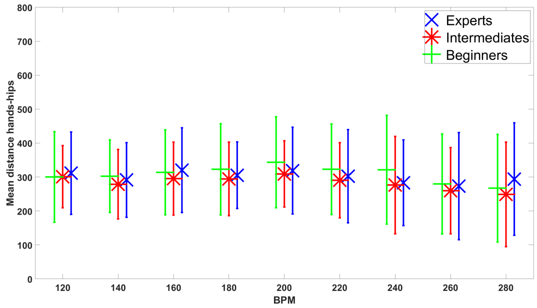




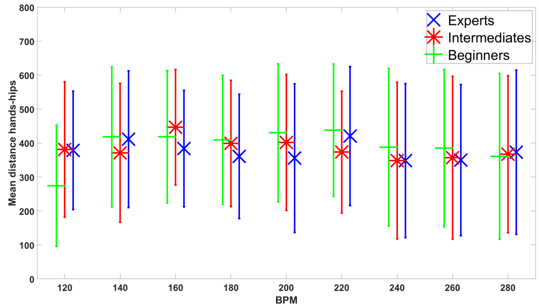
(g) Follower area covered for the right foot for the basic sequence (μ_{25})

(h) Follower area covered for the right foot for the improvisation sequence (μ_{25})

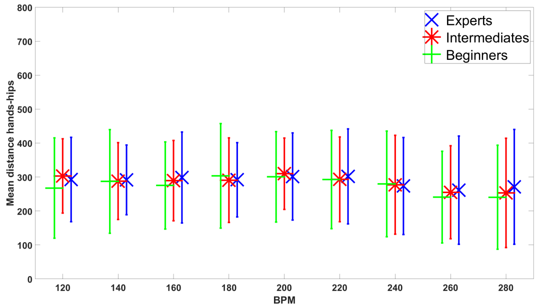
Figure C.8 – Average values and standard deviation per category of the area covered by left or right foot μ_{22-25}



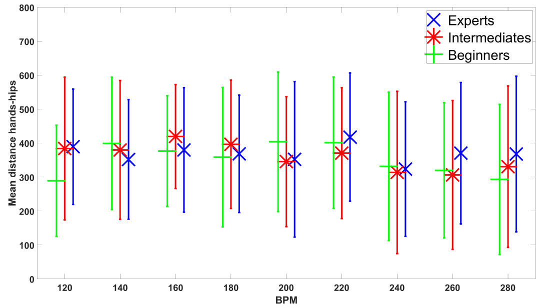
(a) Leader mean distance hands-hips (beat 1 to 3) for the basic sequence (μ_{26})



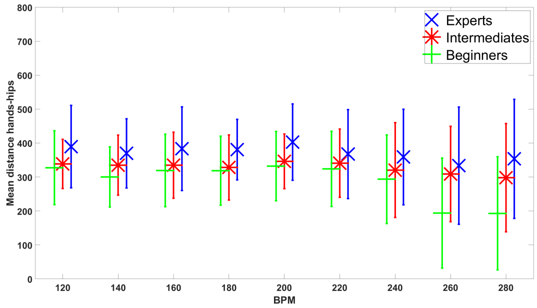
(b) Leader mean distance hands-hips (beat 1 to 3) for the improvisation sequence(μ_{26})



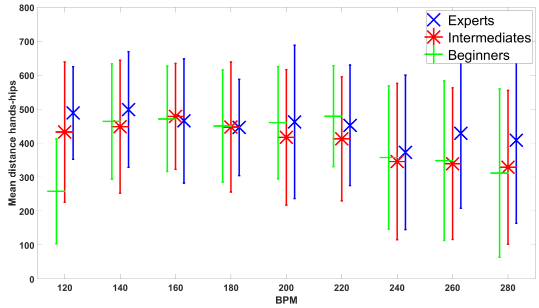
(c) Leader mean distance hands-hips (beat 5 to 7) for the basic sequence(μ_{27})



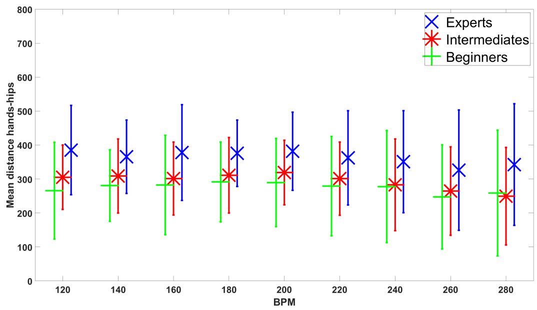
(d) Leader mean distance hands-hips (beat 5 to 7) for the improvisation sequence(μ_{27})



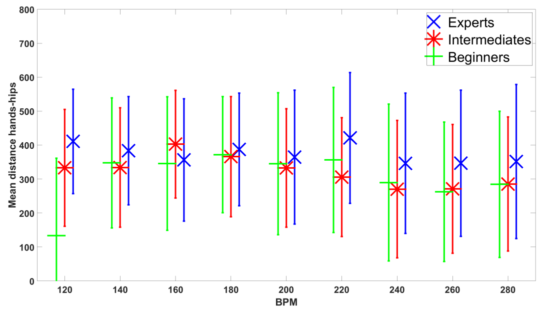
(e) follower mean distance hands-hips (beat 1 to 3) for the basic sequence (μ_{28})



(f) Follower mean distance hands-hips (beat 1 to 3) for the improvisation sequence (μ_{28})



(g) Follower mean distance hands-hips (beat 5 to 7) for the basic sequence (μ_{29})



(h) Follower mean distance hands-hips (beat 5 to 7) for the improvisation sequence (μ_{29})

Figure C.9 – Average values and standard deviation per category of the mean distance hands to hips μ_{26-29}

Appendix D

Preliminary research performed for other European Research Projects

Several projects supported the work presented in this thesis. Each of these projects brought their domain, limitations, and requirements to the research topic. It broadened the subject, made it more varied and diverse as a whole level of complexity in the overall design of the framework was added, which made it challenging and intriguing. Each section presents a description of the project and includes a section with the contributions correlated with the work presented in this thesis. Although the thematic can be different, the tools, techniques, and knowledge developed in these projects brought their contribution to the Ph.D. subject.

D.1 Preliminary work on interactive application: Cultural heritage

This Marie-Curie European project ITN-DCH is about digitizing cultural heritage content for preservation, documentation, and archiving. A main part of the contribution concerns the case study of the church of Asinou located in Cyprus. Beyond the church monument itself, digitizing the priest and liturgy that occurs in this church as part of intangible heritage brought many challenges.

D.1.1 Avatar modelling and 3D digitization

The priest responsible for the service in the church shows us his clothes and the different parts of the religious liturgy. A virtual avatar was made using body reconstruction techniques and face generation (FaceGen software). The virtual character was rigged to allow playing Avatar and Cloth simulation. In order to create an immersive 3D avatar, we added a layer of clothes which we simulate using the physics simulation module available in our 3D Engine (Unity 3D). The face was made using photos acquired on-site and applied in a predefined standard face mesh using FaceGen software techniques, visualized in figure D.1 and D.2. We merged the face on a pre-made modified body that we updated with the priest garment and made a supplementary layer of rough geometry for clothes simulation.

In a second step, a reconstruction of the whole body of the priest, including his clothes, was realized. A 3D reconstruction of the Church of Asinou with its surroundings was also performed using a collection of photos taken on-site and photogrammetry techniques. Measurements of the different parts of the church were also done to help keep the same proportion. (Distance between the holy altar and wooden arcade, etc.). An example of the resulting 3D avatar and church is visualized in figure D.3.

Figure D.1 – Photo of the current Priest of Asinou Church, Cyprus

Figure D.2 – 3D reconstruction the priest face from photos. The hair are not modeled and therefore appear as flat texture.

Figure D.3 – Virtual priest in the reconstructed church environment

Since it was possible to record the movements on the site of the church, it was done at the University of Nicosia, partner of the project. It was, therefore, difficult for the priest to recreate the liturgical movements with accuracy and naturalness as we could imagine what he could fall in the technological environment of the motion capture room. We decided to focus on the specific movement of the priest, corresponding to the critical part of the liturgy â€" the first blessing using incense around the altar and the benediction by giving water drop to the people. Motion capture was done using PhaseSpaceX technology. Due to the technology limitation, we could only record in a relatively small space. Some movement implying the priest to move in the entire space of the church were not taken into account. After acquiring the raw motion capture data, we need to process it to smooth the motion and add the animation structure, also known as skeleton or RIG. The Rig should be the same as the one used to make the virtual character, and so we rigged the animation using the one defined in Mixamo's software. As we had the motion and the body of the priest, we integrated both of them to the virtual environment in our project's augmented environment, shown in figure D.4 and D.5.

Figure D.4 – Virtual priest in the inside of Asinou via android device

Figure D.5 – Virtual priest outside the Asinou church, from android device

D.1.2 Interactive Application

The final part of our contribution to the ITN-DCH EU project is an application that recreates a liturgy from the Asinou church in Cyprus to serve as documentation for people. The inside and outside of the church have been digitized and reconstructed, as well as the priest and his movements in the church's space. We use the metadata information we collect from other fellows to ensure the veracity of our simulation. We integrated the priest inside the church, at someplace corresponding to the different steps of the liturgy. The interior 3D Model was acquired using photogrammetry. In the next step, it will be needed to do processing to make a low-resolution textured mesh out of the data. A calibration also will be necessary to adapt the liturgy motion of the priest to the virtually recreated environment.

In a first screen, we used a specific police font for the screen title as it corresponds to a historical application. There is a short waiting time that allows loading of all needed assets, textures, 3d models etc. The title showed is: *"Inside the Unesco's 16th century church: Meeting with Kyriakos, Priest of Asinou"*.

After a few seconds, a first video transition shows the exterior of the church, and the camera view is moved inside where the priest is standing. In a second screen illustrated in the figure D.6, to give a better impression of freedom, we proposed a menu on which we can choose between two actions: the first one is the recreation of a part of the liturgy. The second choice leads to an interactive mode where the user can ask questions to the priest himself.

On this next screen show in figure D.7, we can see the priest making movements with the holy Evangelion book. These movements have been recorded from the real priest. At the bottom, we can read an explanation of the context. The lights are real global illumination from the various candles located in the room.

This last screen is shown in figure D.8 Here, we can type and send a question to the priest that will respond with an answer, giving the impression of a communication.

D.1.3 Conclusion

Our involvement in this project brings us important knowledge about:

- Designing multiple screen interactive application
- 3D modeling and motion capture of a priest

Figure D.6 – Choice menu

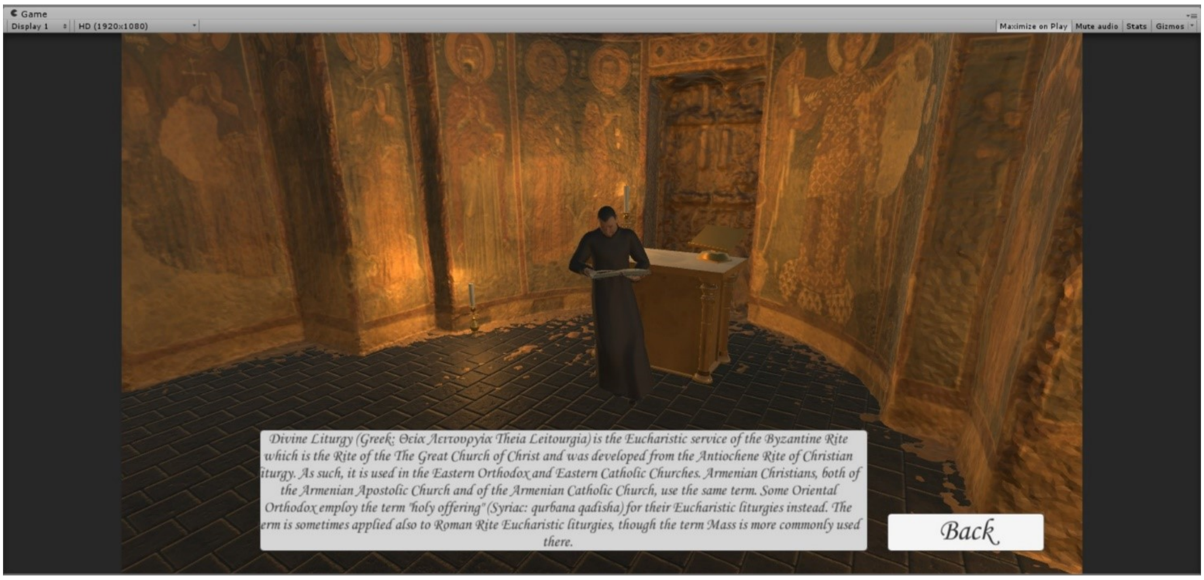


Figure D.7 – Liturgy explanation and demo

- The integration of cultural content and the challenges related

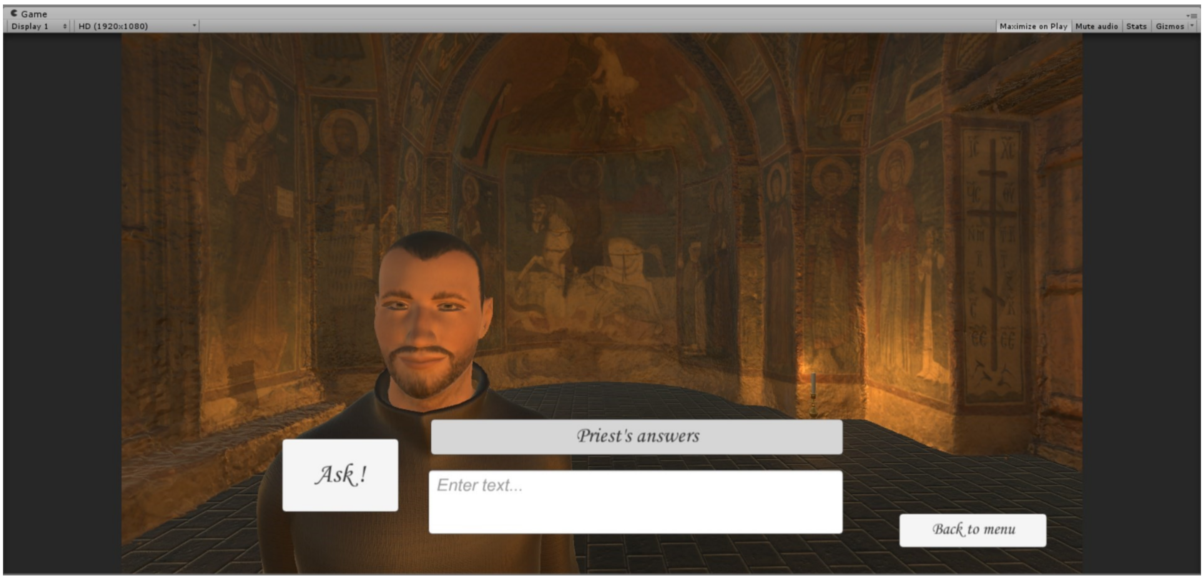


Figure D.8 – Interactive talk with the priest

D.2 Preliminary work on interactive application: Dance game for Elderly

This EU-supported AAL project aims at providing a solution for the elderly (65+ years old) that enhances their health (balance and strength) through the use of a game. Together with the consortium, an application was developed in which users can follow lessons of Tai-Chi and Latin dances. The author was used as a reference for the dance motion capture.

D.2.1 Design and Data capture

The first phase of the project was about finding the right exercise for the elderly that is also fun to do and repeat trough the time, while also enhancing health. A set of 6 styles of dances (Disco, Salsa, Bachata, Jive, Waltz, and DiscoFox) with three levels of difficulties for each style has been chosen and recorded. In Total, 18 dance files have been produced. An example of the motion capture and results on the in-game avatar is available on the figure D.9.

D.2.2 The Exergame

The application consists of a device that can be plugged into a TV and four sensors placed on wrists and ankles using Velcro. The user can then point to the TV to choose between the different exercises and accomplish their weekly advised routine, visible in figure D.10. The whole pipeline of this project is visualized in figure D.11.

D.2.3 Conclusion

Our involvement in this project brings us important knowledge about:

- Designing suitable dance exercises of Latin dances
- The design consideration required to make an application easy to use and to satisfy

- The efficiency of human-machine interface for interactive dance systems

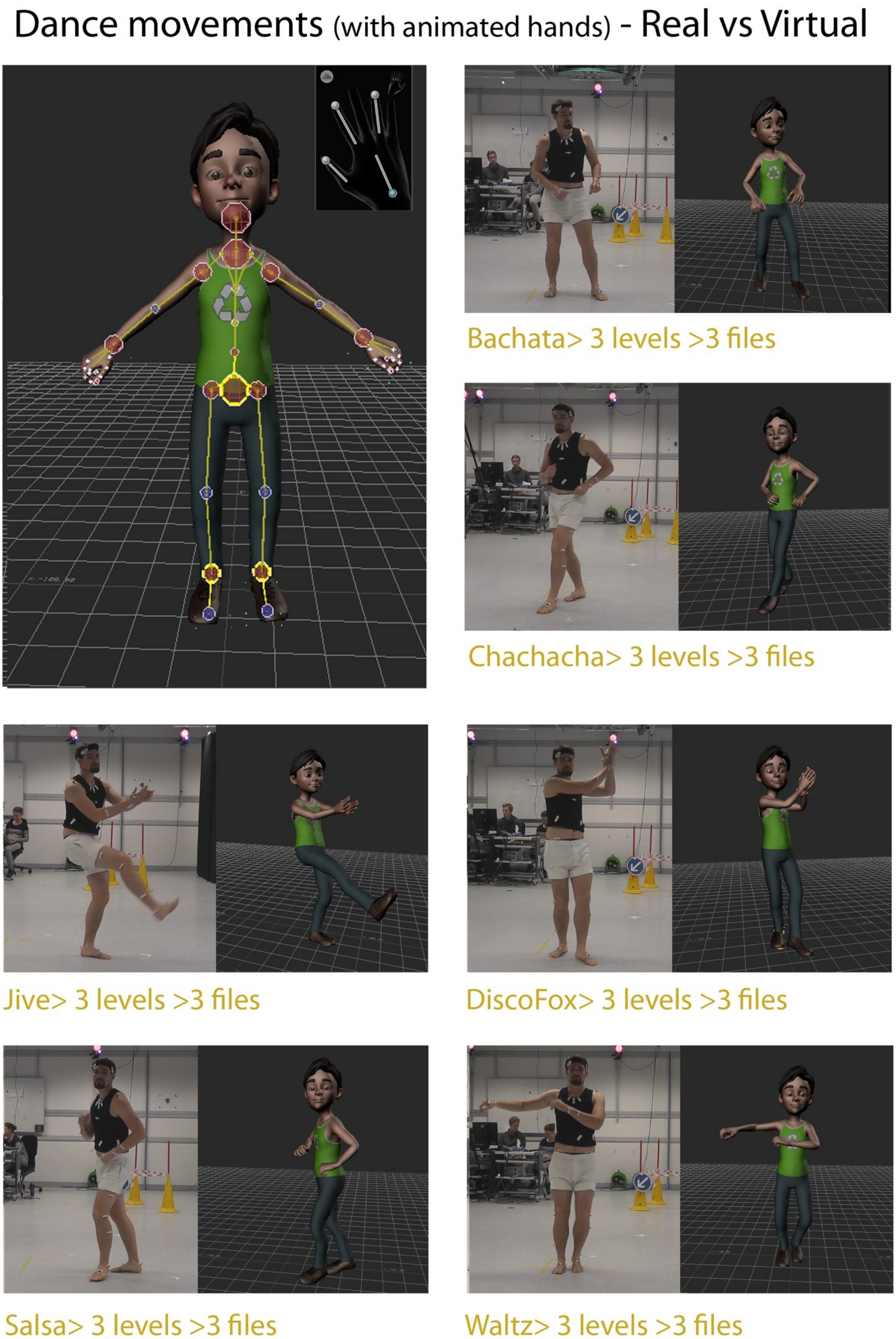


Figure D.9 – Retargeting the chosen dance motions onto the game Avatar. The user will have to subsequently follow its movements.

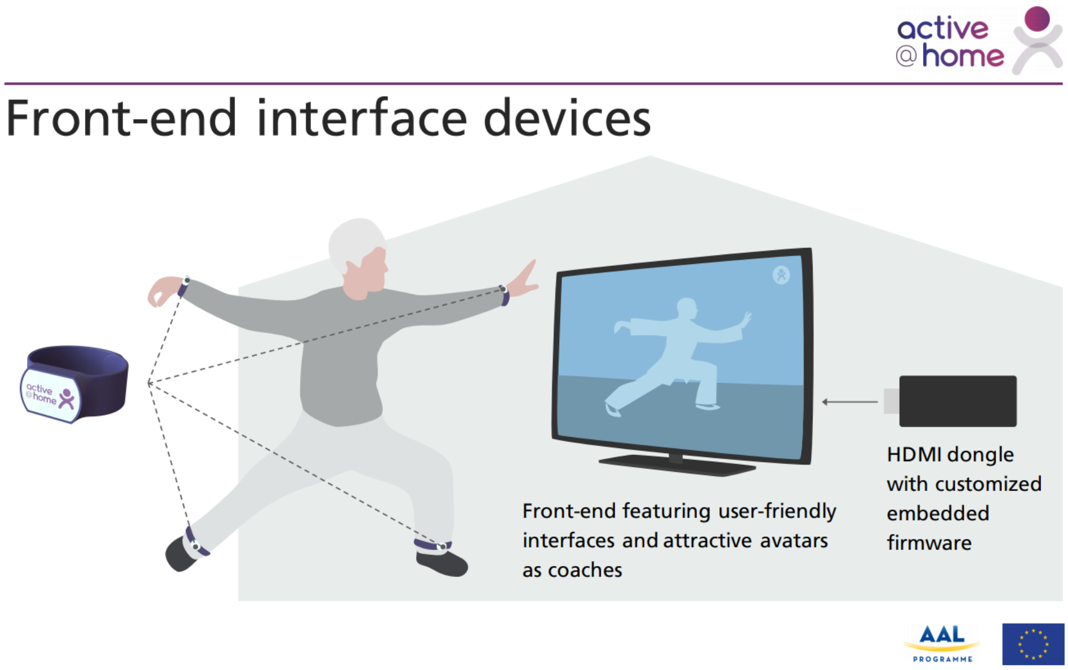


Figure D.10 – Principle of the exergame, outcome of the project. An elderly person plays weekly at home and improve its health.

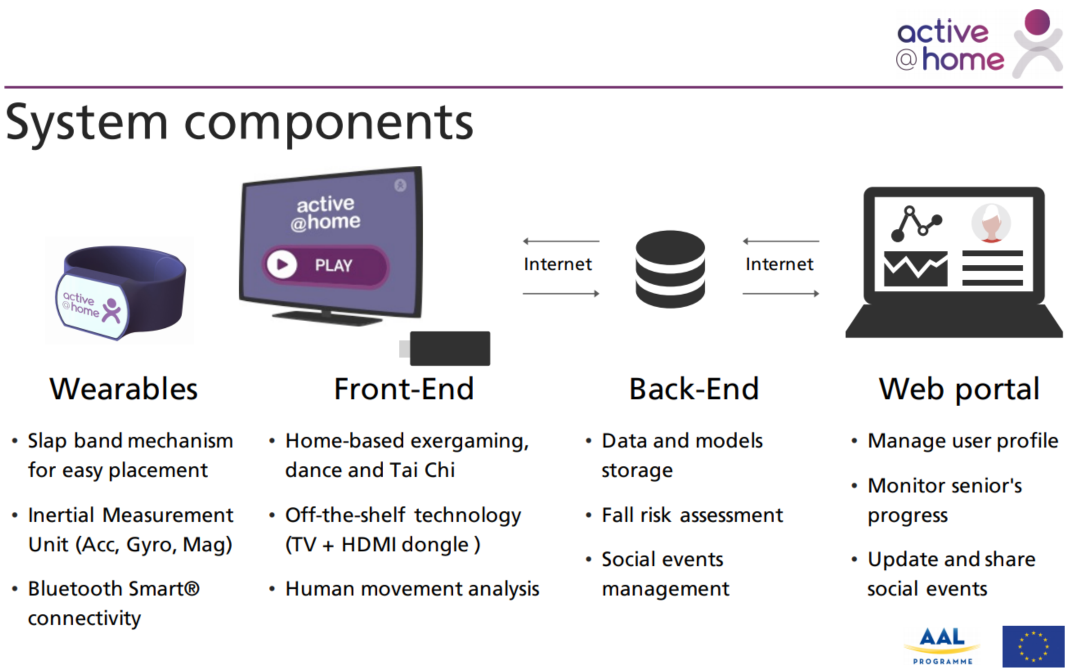


Figure D.11 – Detail of the pipeline of the project outcome.

Bibliography

[1] P. Alborno, S. Piana, and A. Camurri. "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. New York, New York, USA: ACM Press, 2016, pp. 136–143.

[2] D. S. Alexiadis et al. "Evaluating a dancer's performance using kinect-based skeleton tracking". In: *Proceedings of the 19th ACM international conference on Multimedia - MM '11*. 2011, p. 659.

[3] D. S. Alexiadis et al. "Real- Time Classification of Evoked Emotions using Facial Feature Tracking and Physiological Responses". In: *International Journal of Human-Computer Studies* 66.5 (2007), pp. 303–317.

[4] A. Aristidou, P. Charalambous, and Y. Chrysanthou. "Emotion Analysis and Classification: Understanding the Performers' Emotions Using the LMA Entities". In: *Computer Graphics Forum* 34.6 (2015), pp. 262–276.

[5] A. Aristidou and Y. Chrysanthou. "Feature Extraction for Human Motion Indexing of Acted Dance Performances". In: *GRAPP 2014 - International Conference on Computer Graphics Theory and Applications* (2013), pp. 277–287.

[6] A. Aristidou and Y. Chrysanthou. "Motion indexing of different emotional states using LMA components". In: *SIGGRAPH Asia 2013 Technical Briefs on - SA '13* (2013), pp. 1–4.

[7] A. Aristidou, E. Stavrakis, and Y. Chrysanthou. "LMA-Based Motion Retrieval for Folk Dance Cultural Heritage". In: *Euromed* (2014), pp. 207–216.

[8] A. Aristidou et al. "Deep motifs and motion signatures". In: *ACM Transactions on Graphics* 37.6 (2018), pp. 1–13.

[9] A. Aristidou et al. "Emotion control of unstructured dance movements". In: 2017, pp. 1–10.

[10] A. Aristidou et al. "Folk Dance Evaluation Using Laban Movement Analysis". In: *Journal on Computing and Cultural Heritage* 8.4 (2015), pp. 1–19.

[11] A. Aristidou et al. "Style-based motion analysis for dance composition". In: *Visual Computer* 34.12 (2018), pp. 1725–1737.

[12] a Aristidou, E Stavrakis, and Y Chrysanthou. "Motion Analysis for Folk Dance Evaluation". In: *EUROGRAPHICS* (2014).

[13] J. Baillieul and K. Ozcimder. “The control theory of motion-based communication: Problems in teaching robots to dance”. en. In: *2012 American Control Conference (ACC)*. Montreal, QC: IEEE, June 2012, pp. 4319–4326.

[14] A. Bastanfard, H. Takahashi, and M. Nakajima. “Toward E-Appearance of Human Face and Hair by Age, Expression and Rejuvenation”. In: *Proceedings of the 2004 International Conference on Cyberworlds*. CW â€™04. USA: IEEE Computer Society, 2004, 306â€“311.

[15] R. Bellini, Y. Kleiman, and D. Cohen-Or. “Dance to the Beat: Enhancing Dancing Performance in Video”. In: ().

[16] U. Bernardet et al. “Assessing the reliability of the Laban Movement Analysis system”. en. In: *PLOS ONE* 14.6 (June 2019). Ed. by E. S. Cross, e0218179.

[17] H. Brenton, A. Kleinsmith, and M. Gillies. “Embodied Design of Dance Visualisations”. In: *Proceedings of the International Workshop on Movement and Computing* (2014), pp. 124–129.

[18] C. Brown and G. Paine. “Interactive Tango Milonga: designing internal experience”. en. In: *Proceedings of the 2nd International Workshop on Movement and Computing - MOCO ’15*. Vancouver, British Columbia, Canada: ACM Press, 2015, pp. 17–20.

[19] G. Buondonno et al. “A model predictive control approach for the Partner Ballroom Dance Robot”. en. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. Seattle, WA, USA: IEEE, May 2015, pp. 774–780.

[20] A. Camurri et al. “A system to support the learning of movement qualities in dance”. In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp ’16*. New York, New York, USA: ACM Press, 2016, pp. 973–976.

[21] A. Camurri et al. “The Dancer in the Eye”. In: *Proceedings of the 3rd International Symposium on Movement and Computing - MOCO ’16*. New York, New York, USA: ACM Press, 2016, pp. 1–7.

[22] A. Camurri et al. “WhoLoDancE: Towards a Methodology for Selecting Motion Capture Data Across Different Dance Learning Practice”. In: *Proceedings of the 3rd International Symposium on Movement and Computing* (2016), 43:1–43:2.

[23] J. C. Chan et al. “A virtual reality dance training system using motion capture technology”. In: *IEEE Transactions on Learning Technologies* 4.2 (2011), pp. 187–195.

[24] G. Chantas et al. “A Probabilistic, Ontological Framework for Safeguarding the Intangible Cultural Heritage”. en. In: *Journal on Computing and Cultural Heritage* 11.3 (Aug. 2018), pp. 1–29.

[25] W. Choensawat, M. Nakamura, and K. Hachimura. “GenLaban: A tool for generating Labanotation from motion capture data”. en. In: *Multimedia Tools and Applications* 74.23 (Dec. 2015), pp. 10823–10846.

[26] R. E. Cisneros et al. “Virtual Reality and Choreographic Practice: The Potential for New Creative Methods”. en. In: *Body, Space & Technology* 18.1 (Mar. 2019), pp. 1–32.

[27] R. E. Cisneros et al. “WhoLoDancE: digital tools and the dance learning environment”. en. In: *Research in Dance Education* 20.1 (Jan. 2019), pp. 54–72.

[28] Congres1. “www.salsadancecongresses.com”. In: ().

[29] Congres2. *salsa.wikia.com/wiki/Salsa_congresses*.

[30] A. Curtis et al. “Dance dance Pleo: developing a low-cost learning robotic dance therapy aid”. en. In: *Proceedings of the 10th International Conference on Interaction Design and Children - IDC ’11*. Ann Arbor, Michigan: ACM Press, 2011, pp. 149–152.

[31] S. Cuykendall. “Untying the Knot of Dance Movement Expertise: An Enactive Approach”. In: 2016.

[32] S. Cuykendall, E. Soutar-Rau, and T. Schiphorst. “POEME: A Poetry Engine Powered by Your Movement”. In: *Proceedings of the TEI ’16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction* (2016), pp. 635–640.

[33] S. Cuykendall et al. “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* (2016), pp. 234–238.

[34] *Dance*. URL: `http://dance.dibris.unige.it/`.

[35] M. E. P. Davies and M. D. Plumbley. “Context-Dependent Beat Tracking of Musical Audio”. en. In: *IEEE Transactions on Audio, Speech and Language Processing* 15.3 (Mar. 2007), pp. 1009–1020.

[36] A. Davis and M. Agrawala. “Visual rhythm and beat”. en. In: *ACM Transactions on Graphics* 37.4 (July 2018), pp. 1–11.

[37] L. Deng et al. “Real-time mocap dance recognition for an interactive dancing game”. en. In: *Computer Animation and Virtual Worlds* 22.2-3 (Apr. 2011), pp. 229–237.

[38] R. Dong, D. Cai, and N. Asai. “Nonlinear dance motion analysis and motion editing using Hilbert-Huang transform”. en. In: *Proceedings of the Computer Graphics International Conference on - CGI ’17*. Yokohama, Japan: ACM Press, 2017, pp. 1–6.

[39] Y. Dong et al. “SalsaAsst: Beat Counting System Empowered by Mobile Devices to Assist Salsa Dancers”. In: *2017 IEEE 14th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*. Orlando, FL: IEEE, Oct. 2017, pp. 81–89.

[40] D. Drobny, M. Weiss, and J. Borchers. “Saltate!: a sensor-based system to support dance beginners”. en. In: *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems - CHI EA ’09*. Boston, MA, USA: ACM Press, 2009, p. 3943.

[41] V. M. Dyaberi et al. “Phrase structure detection in dance”. en. In: *Proceedings of the 12th annual ACM international conference on Multimedia - MULTIMEDIA ’04*. New York, NY, USA: ACM Press, 2004, p. 332.

[42] K. El Raheb and Y. Ioannidis. “Dance in the world of data and objects”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 7990 LNCS. 2013, pp. 192–204.

[43] K. El Raheb and Y. Ioannidis. “From Dance Notation to Conceptual Models”. In: *Proceedings of the 2014 International Workshop on Movement and Computing - MOCO ’14*. New York, New York, USA: ACM Press, 2014, pp. 25–30.

[44] K. El Raheb et al. “BalOnSe: Temporal aspects of dance movement and its ontological representation”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 10250 LNCS. Springer, Cham, 2017, pp. 49–64.

[45] A. Z. M. Faridee et al. “HappyFeet: Recognizing and Assessing Dance on the Floor”. In: *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications - HotMobile ’18*. New York, New York, USA: ACM Press, 2018, pp. 49–54.

[46] S. Ferguson, E. Schubert, and C. Stevens. “Dynamic dance warping: Using dynamic time warping to compare dance movement performed under different conditions”. In: *Proceedings of the 2014 \ldots* (2014), pp. 94–99.

[47] K. Forbes and E. Fiume. “An Efficient Search Algorithm for Motion Data Using Weighted PCA”. In: *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. SCA ’05. Los Angeles, California, 2005, pp. 67–76.

[48] N. Fourati and C. Pelachaud. “Collection and characterization of emotional body behaviors”. In: *Proceedings of the 2014 International Workshop on Movement and Computing - MOCO ’14* (2014), pp. 49–54.

[49] N. Fourati and C. Pelachaud. “Multi-level classification of emotional body expression”. In: *IEEE* (2015).

[50] N. Fourati and C. Pelachaud. “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)* (2015), pp. 267–273.

[51] N. Fourati and C. Pelachaud. “Toward new expressive movement characterizations”. In: *Proc. Motion in Games* (2012), pp. 5–8.

[52] S. Gentry and E. Feron. “Musicality experiments in lead and follow dance”. en. In: *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583)*. Vol. 1. The Hague, Netherlands: IEEE, 2004, pp. 984–988.

[53] M. Gowing et al. “Enhanced visualisation of dance performance from automatically synchronised multimodal recordings”. en. In: *Proceedings of the 19th ACM international conference on Multimedia - MM ’11*. Scottsdale, Arizona, USA: ACM Press, 2011, p. 667.

[54] C. F. Griggio, M. Romero, and G. Leiva. “Towards an Interactive Dance Visualization for Inspiring Coordination Between Dancers”. In: *Extended Abstracts of the ACM CHI’15 Conference on Human Factors in Computing Systems*. Vol. 2. New York, New York, USA: ACM Press, 2015, pp. 1513–1518.

[55] M. Hajarian et al. “A personalized gamification method for increasing user engagement in social networks”. In: *Social Network Analysis and Mining* 9 (2019), pp. 1–14.

[56] D. Herbison-Evans. “Dancing Backwards in High Heels and Other Problems in Animation”. en. In: *2006 7th International Conference on Information Technology Based Higher Education and Training*. Ultimo, Australia: IEEE, July 2006, pp. 927–930.

[57] L. Heydrich et al. “Visual capture and the experience of having two bodies â€“ Evidence from two different virtual reality techniques”. In: *Frontiers in Psychology* 4 (2013).

[58] J. D. Hicks et al. “Bayesian-Game Modeling of C2 Decision Making in Submarine Battle-Space Situation Awareness”. In: (2004), p. 60.

[59] E. S. L. Ho et al. “Interactive partner control in close interactions for real-time applications”. In: *ACM Transactions on Multimedia Computing, Communications, and Applications* 9.3 (2013), pp. 1–19.

[60] J. Holldampf, A. Peer, and M. Buss. “Synthesis of an interactive haptic dancing partner”. en. In: *19th International Symposium in Robot and Human Interactive Communication*. Viareggio, Italy: IEEE, Sept. 2010, pp. 527–532.

[61] J. Hoysniemi. “International survey on the Dance Dance Revolution game”. In: *Computers in Entertainment* 4.2 (2006), p. 8.

[62] IDF. “Olympic certification”. In: ().

[63] S. Karavarsamis et al. “Classifying Salsa dance steps from skeletal poses”. In: *2016 14th International Workshop on Content-Based Multimedia Indexing (CBMI)*. IEEE, 2016, pp. 1–6.

[64] D. Kim, D. H. Kim, and K. C. Kwak. “Classification of K-pop dance movements based on skeleton information obtained by a kinect sensor”. In: *Sensors (Switzerland)* 17.6 (2017), p. 1261.

[65] D.-H. Kim and K.-C. Kwak. “A motion analysis and classification based on PCA and ELM classifier”. In: *2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)*. IEEE, 2017, pp. 1–4.

[66] K. H. Kim, S. W. Bang, and S. R. Kim. “Emotion recognition system using short term monitoring of physiological signals”. In: *Medical Biological Engineering and computing* 42.Journal Article (2004), pp. 419–427.

[67] T.-h. Kim, S. I. Park, and S. Y. Shin. “Rhythmic-motion synthesis based on motion-beat analysis”. In: *ACM Transactions on Graphics* 22.3 (2003), p. 392.

[68] W. H. Kim et al. “LMA based emotional motion representation using RGB-D camera”. In: *HRI '13: Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction* (2013), pp. 163–164.

[69] A Kitsikidis, N. V. Boulgouris, and K Dimitropoulos. “Unsupervised Dance Motion Patterns Classification from Fused Skeletal Data using Exemplar-based HMMs Unsupervised Dance Motion Patterns Classification from Fused Skeletal Data using Exemplar-based HMMs”. In: 4.2 (2015), pp. 209–220.

[70] A. Kitsikidis et al. “Dance Analysis using Multiple Kinect Sensors”. In: *VISAPP2014, Lisbon, Portugal* (2014), pp. 789–795.

[71] A. Kitsikidis et al. “Multi-sensor Technology and Fuzzy Logic for Dancer?s Motion Analysis and Performance Evaluation within a 3D Virtual Environment”. In: *Universal Access in Human-Computer Interaction. Design and De- velopment Methods for Universal Access*. 2014, pp. 379–390.

[72] K. Kosuge. “Dance Partner Robot: An engineering approach to human-robot interaction”. en. In: *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. Osaka, Japan: IEEE, Mar. 2010, pp. 201–201.

[73] K. Kosuge. “Human-robot interaction - What we learned from robot helpers and dance partner robots”. en. In: *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*. Munich: IEEE, Aug. 2008, pp. v–v.

[74] M. Kyan et al. “An Approach to Ballet Dance Training through MS Kinect and Visualization in a CAVE Virtual Reality Environment”. In: *ACM Transactions on Intelligent Systems and Technology* 6.2 (Mar. 2015), pp. 1–37.

[75] R. Laban and L. Ullmann. *The Mastery of Movement*. 4th ed. Dance Books Ltd, 2011.

[76] J. Lee et al. “Interactive Control of Avatars Animated with Human Motion Data”. en. In: (2002), p. 10.

[77] Z. Liu et al. “A new control algorithm for a passive type dance partner robot”. en. In: *2010 IEEE Conference on Robotics, Automation and Mechatronics*. , Singapore: IEEE, June 2010, pp. 486–491.

[78] L. van der Maaten and G. Hinton. “Visualizing Data using t-SNE”. In: *Journal of Machine Learning Research* 9 (2008), pp. 2579–2605.

[79] N. Magnenat-Thalmann, D. Protopsaltou, and E. Kavakli. “Learning How to Dance Using a Web 3D Platform”. In: *Advances in Web Based Learning â€“ ICWL 2007*. Ed. by H. Leung et al. Vol. 4823. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 1–12.

[80] A. Masurelle, S. Essid, and G. RIchard. “MULTIMODAL CLASSIFICATION OF DANCEMOVEMENTS USING BODY JOINT TRAJECTORIES AND STEP SOUNDS”. In: *Image Analysis for Multimedia Interactive Services (WIAMIS), 2013 14th International Workshop on*. IEEE, 2013, pp. 3–6.

[81] E. A. Matthews and G. B. Matthews. “Procedural generation of Cuban dance motion”. In: *Proceedings of CGAMES’2011 USA - 16th International Conference on Computer Games: AI, Animation, Mobile, Interactive Multimedia, Educational and Serious Games*. IEEE, 2011, pp. 293–297.

[82] J. McMains. *Spinning Mambo into Salsa: Caribbean Dance in Global Commerce*. 2015.

[83] D. Merom et al. “Can social dancing prevent falls in older adults? a protocol of the Dance, Aging, Cognition, Economics (DAnCE) fall prevention randomised controlled trial”. In: *BMC Public Health* 13.1 (2013), p. 477.

[84] D. Merom et al. “Cognitive Benefits of Social Dancing and Walking in Old Age: The Dancing Mind Randomized Controlled Trial.” In: *Frontiers in aging neuroscience* 8 (2016), p. 26.

[85] D. Merom et al. “Social Dancing and Incidence of Falls in Older Adults: A Cluster Randomised Controlled Trial”. In: *PLOS Medicine* 13.8 (2016). Ed. by C. Brayne, e1002112.

[86] L. Molina-Tanco, C. GarcÃa-BerdonÃ©s, and A. Reyes-Lecuona. “The Delay Mirror: a Technological Innovation Specific to the Dance Studio”. en. In: *Proceedings of the 4th International Conference on Movement Computing - MOCO ’17*. London, United Kingdom: ACM Press, 2017, pp. 1–6.

[87] H. Mori, S. u. Ohta, and J. Hoshino. “Automatic dance generation from music annotation”. en. In: *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology - ACE ’04*. Singapore: ACM Press, 2004, pp. 352–353.

[88] M. Mori, K. MacDorman, and N. Kageki. “The Uncanny Valley [From the Field]”. In: *IEEE Robotics & Automation Magazine* 19.2 (June 2012), pp. 98–100.

[89] H. Morioka, M. Nakatani, and S. Nishida. “Proposal of an algorithm to synthesize music suitable for dance”. en. In: *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology - ACE '04*. Singapore: ACM Press, 2004, pp. 296–301.

[90] C. Mousas. “Performance-Driven Dance Motion Control of a Virtual Partner Character”. In: *25th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2018 - Proceedings* (2018), pp. 57–64.

[91] A. Nakamura et al. “Multimodal presentation method for a dance training system”. en. In: *CHI '05 extended abstracts on Human factors in computing systems - CHI '05*. Portland, OR, USA: ACM Press, 2005, p. 1685.

[92] D. Nakayama, K. Kosuge, and Y. Hirata. “Human-adaptive step estimation method for a dance partner robot”. en. In: *2009 IEEE International Conference on Automation and Logistics*. Shenyang, China: IEEE, Aug. 2009, pp. 191–196.

[93] R. Niewiadomski et al. “Analysis of Movement Quality in Full-Body Physical Activities”. en. In: *ACM Transactions on Interactive Intelligent Systems* 9.1 (Feb. 2019), pp. 1–20.

[94] R. Niewiadomski et al. “Low-intrusive recognition of expressive movement qualities”. en. In: *Proceedings of the 19th ACM International Conference on Multimodal Interaction - ICMI 2017*. Glasgow, UK: ACM Press, 2017, pp. 230–237.

[95] F Ofli et al. “MULTI-MODAL ANALYSIS OF DANCE PERFORMANCES FOR MUSIC-DRIVEN CHOREOGRAPHY SYNTHESIS”. In: (2010).

[96] N. Okada. “Expressive Dance Motion Generation”. In: *ACM SIGGRAPH 2013 Posters on - SIGGRAPH '13* (2013), p. 4503.

[97] K. Ozcimder. “Communication through motion in dance with topological constraints”. en. In: *2014 American Control Conference*. Portland, OR, USA: IEEE, June 2014, pp. 178–183.

[98] K. Ozcimder et al. “Investigating group behavior in dance: An evolutionary dynamics approach”. In: *Proceedings of the American Control Conference*. Vol. 2016-July. IEEE, 2016, pp. 6465–6470.

[99] K. Ozcimder et al. “Perceiving Artistic Expression: A Formal Exploration of Performance Art Salsa”. en. In: *IEEE Access* 6 (2018), pp. 61867–61875.

[100] D. F. Paez Granados and K. Kosuge. “Design of a Male-type Dance Partner Robot for leading a physical Human-Robot Interaction”. In: *2015 IEEE International Conference on Mechatronics and Automation, ICMA 2015*. IEEE, 2015, pp. 1234–1240.

[101] D. F. Paez Granados et al. “Dance Teaching by a Robot: Combining Cognitive and Physical Humanâ€“Robot Interaction for Supporting the Skill Learning Process”. en. In: *IEEE Robotics and Automation Letters* 2.3 (July 2017), pp. 1452–1459.

[102] D. F. Paez Granados et al. “Guiding Human Motions in Physical Human â€“ Robot Interaction through COM Motion Control of a Dance Teaching Robot”. In: *IEEE International Conference on Humanoid Robots (Humanoids)*. IEEE, Nov. 2016, pp. 279–285.

[103] P. Palacio and D. Bisig. “Piano&Dancer: Interaction Between a Dancer and an Acoustic Instrument”. en. In: *Proceedings of the 4th International Conference on Movement Computing - MOCO '17*. London, United Kingdom: ACM Press, 2017, pp. 1–8.

[104] P. Patil et al. “Clinical Human Gait Classification: Extreme Learning Machine Approach”. In: *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)* (2019), pp. 1–6.

[105] J. Peeters and A. Trotto. “Designing Expressions of Movement Qualities”. en. In: *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS ’18*. Hong Kong, China: ACM Press, 2018, pp. 679–690.

[106] J. Peeters, A. Trotto, and S. Kuenen. “MoCap Tango: Traces of Complexity”. en. In: *Proceedings of the TEI ’16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction - TEI ’16*. Eindhoven, Netherlands: ACM Press, 2016, pp. 545–550.

[107] S. Piana. “Movement Fluidity Analysis Based on Performance and Perception”. In: *CHI Extended Abstracts on Human Factors in Computing Systems* (2016), pp. 1629–1636.

[108] S. Piana et al. “Adaptive Body Gesture Representation for Automatic Emotion Recognition”. In: *ACM Transactions on Interactive Intelligent Systems* 6.1 (2016), pp. 1–31.

[109] S. Piana et al. “Real-time Automatic Emotion Recognition from Body Gestures”. In: *Computer repository* (2014).

[110] S. Piana et al. “Towards a Multimodal Repository of Expressive Movement Qualities in Dance”. In: *Proceedings of the 3rd International Symposium on Movement and Computing - MOCO ’16*. New York, New York, USA: ACM Press, 2016, pp. 1–8.

[111] R. S. U. Powers. *Brief Histories of Social Dance*.

[112] K. E. Raheb, V. Katifori, and A. Rc. “HCI challenges in Dance Education”. In: 3.9 (2016), pp. 6–10.

[113] K. E. Raheb et al. “Choreomorphy: a whole-body interaction experience for dance improvisation and visual experimentation”. In: *Proceedings of the 2018 International Conference on Advanced Visual Interfaces - AVI ’18*. the 2018 International Conference. Castiglione della Pescaia, Grosseto, Italy: ACM Press, 2018, pp. 1–9.

[114] K. E. Raheb et al. “Dance Interactive Learning Systems: A Study on Interaction Workflow and Teaching Approaches”. en. In: (2019), p. 35.

[115] A. Rizzo et al. “WhoLoDancE: Whole-body Interaction Learning for Dance Education”. In: 2018, p. 11.

[116] R. Ros and Y. Demiris. “Creative dance: An approach for social interaction between robots and children”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 8212 LNCS. Springer International Publishing, 2013, pp. 40–51.

[117] R. Ros et al. “Adapting Robot Behavior to User ’ s Capabilities : a Dance Instruction Study âˆ—”. In: *Proceedings of the 6th international conference on Human-robot interaction - HRI ’11* (2011), pp. 235–236.

[118] R. Ros et al. “Behavioral accommodation towards a dance robot tutor”. In: *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction - HRI ’14* (2014), pp. 278–279.

[119] Y. Rubner, C. Tomasi, and L. J. Guibas. “The Earth Mover’s Distance As a Metric for Image Retrieval”. In: *Int. J. Comput. Vision* 40.2 (Nov. 2000), pp. 99–121.

[120] S. Saha et al. “Human skeleton matching for e-learning of dance using a probabilistic neural network”. In: *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2016, pp. 1754–1761.

[121] S. Saha et al. “Identification of Odissi dance video using Kinect sensor”. en. In: *2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. Mysore: IEEE, Aug. 2013, pp. 1837–1842.

[122] Y. Sakai et al. “Collision avoidance based on estimated step of other dance couples for male-type dance partner robot”. In: *IEEE International Conference on Intelligent Robots and Systems*. IEEE, Oct. 2007, pp. 3264–3269.

[123] M. Sakashita et al. “Materialization of motions: tangible representation of dance movements for learning and archiving”. en. In: *ACM SIGGRAPH 2017 Studio on - SIGGRAPH ’17*. Los Angeles, California: ACM Press, 2017, pp. 1–2.

[124] S. Samanta, P. Purkait, and B. Chanda. “Indian Classical Dance classification by learning dance pose bases”. In: *Proceedings of IEEE Workshop on Applications of Computer Vision*. IEEE, Jan. 2012, pp. 265–270.

[125] A. D.P. d. Santos et al. “You Are Off The Beat!: Is Accelerometer Data Enough for Measuring Dance Rhythm?” en. In: *Proceedings of the 5th International Conference on Movement and Computing - MOCO ’18*. Genoa, Italy: ACM Press, 2018, pp. 1–8.

[126] A. D. P. dos Santos, L. Loke, and R. Martinez-Maldonado. “Exploring video annotation as a tool to support dance teaching”. en. In: *Proceedings of the 30th Australian Conference on Computer-Human Interaction - OzCHI ’18*. Melbourne, Australia: ACM Press, 2018, pp. 448–452.

[127] A. Dias Pereira dos Santos, K. Yacef, and R. Martinez-Maldonado. “Forro Trainer: Automated Feedback for Partner Dance Learning”. en. In: *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization - UMAP ’17*. Bratislava, Slovakia: ACM Press, 2017, pp. 103–104.

[128] A. dos Santos, K. Yacef, and R. Martinez-Maldonado. “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*. New York, New York, USA: ACM Press, 2017, pp. 183–191.

[129] T. Schiphorst, S. Mah, and J. Crawford. “STILL DANCING: interacting inside the dance”. en. In: *Conference companion on Human factors in computing systems - CHI ’94*. Boston, Massachusetts, United States: ACM Press, 1994, pp. 61–62.

[130] G. A. F. Seber. *Multivariate Observations*. John Wiley & Sons, Inc., 2008.

[131] V. B. Semwal, K. Mondal, and G. C. Nandi. “Robust and accurate feature selection for humanoid push recovery and classification: deep learning approach”. In: *Neural Computing and Applications* 28 (2015), pp. 565–574.

[132] S. Senecal, N. Nijdam, and N. Thalmann. “Classification of Salsa Dance Level using Music and Interaction based Motion Features”. In: *GRAPP 2019 - International Conference on Computer Graphics Theory and Applications*. 2019, pp. 100–109.

[133] A. Serino et al. “Peripersonal Space: An Index of Multisensory Bodyâ€“Environment Interactions in Real, Virtual, and Mixed Realities”. In: *Frontiers in ICT* 4 (Jan. 22, 2018), p. 31.

[134] N. Shikanai and K. Hachimura. “A fundamental study on synchronous dance movements between two dancers in face-to-face”. en. In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. Paris, France: IEEE, Sept. 2012, pp. 1039–1044.

[135] T. Shiratori, A. Nakazawa, and K. Ikeuchi. “Rhythmic motion analysis using motion capture and musical information”. In: *Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, MFI2003.* (2003), pp. 89–94.

[136] T. Shiratori, A. Nakazawa, and K. Ikeuchi. “Dancing-to-music character animation”. In: *Computer Graphics Forum* 25.3 (2006), pp. 449–458.

[137] T. Shiratori, A. Nakazawa, and K. Ikeuchi. “Synthesizing dance performance using musical and motion features”. In: *Proceedings - IEEE International Conference on Robotics and Automation*. Vol. 2006. IEEE, 2006, pp. 3654–3659.

[138] H. P. H. Shum et al. “Interaction patches for multi-character animation”. en. In: *ACM Transactions on Graphics* 27.5 (Dec. 2008), p. 1.

[139] G. Sun et al. “A Virtual Reality Dance Self-learning Framework using Laban Movement Analysis”. In: *Journal of Engineering Science and Technology Review* 10.5 (2017), pp. 25–32.

[140] Takahiro Takeda, Yasuhisa Hirata, and Kazuhiro Kosuge. “Dance partner robot cooperative motion generation with adjustable length of dance step stride based on physical interaction”. en. In: *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*. San Diego, CA, USA: IEEE, Oct. 2007, pp. 3258–3263.

[141] T. Takeda, Y. Hirata, and K. Kosuge. “Dance Step Estimation Method Based on HMM for Dance Partner Robot”. en. In: *IEEE Transactions on Industrial Electronics* 54.2 (Apr. 2007), pp. 699–706.

[142] T. Takeda et al. “HMM-based Error Detection of Dance Step Selection for Dance Partner Robot -MS DanceR-”. en. In: *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Beijing, China: IEEE, Oct. 2006, pp. 5631–5636.

[143] J. K. T. Tang, J. C. P. Chan, and H. Leung. “Interactive dancing game with real-time recognition of continuous dance moves from 3D human motion capture”. en. In: *Proceedings of the 5th International Confernece on Ubiquitous Information Management and Communication - ICUIMC '11*. Seoul, Korea: ACM Press, 2011, p. 1.

[144] T. Tang, J. Jia, and H. Mao. “Dance with Melody: An LSTM-autoencoder Approach to Music-oriented Dance Synthesis”. en. In: *2018 ACM Multimedia Conference on Multimedia Conference - MM '18*. Seoul, Republic of Korea: ACM Press, 2018, pp. 1598–1606.

[145] L. Theodorou, P. G. T. Healey, and F. Smeraldi. "Exploring Audience Behaviour During Contemporary Dance Performances". en. In: *Proceedings of the 3rd International Symposium on Movement and Computing - MOCO '16*. Thessaloniki, GA, Greece: ACM Press, 2016, pp. 1–7.

[146] Y. Tongpaeng et al. "Thai dance knowledge archive framework based on Labanotation represented in 3D animation". In: *2nd Joint International Conference on Digital Arts, Media and Technology 2017: Digital Economy for Sustainable Growth, ICDAMT 2017*. IEEE, 2017, pp. 66–70.

[147] G. Tsampounaris et al. "Exploring Visualizations in Real-time Motion Capture for Dance Education". In: *Proceedings of the 20th Pan-Hellenic Conference on Informatics - PCI '16*. New York, New York, USA: ACM Press, 2016, pp. 1–6.

[148] S. Tsuchida, T. Terada, and M. Tsukamoto. "A system for practicing formations in dance performance supported by self-propelled screen". en. In: *Proceedings of the 4th Augmented Human International Conference on - AH '13*. Stuttgart, Germany: ACM Press, 2013, pp. 178–185.

[149] S. Tsuchida, T. Terada, and M. Tsukamoto. "A system for practicing formations in dance performance using a two-axis movable electric curtain track". en. In: *Proceedings of the 5th Augmented Human International Conference on - AH '14*. Kobe, Japan: ACM Press, 2014, pp. 1–2.

[150] C. M. Univ. *CMU, 2003. MoCap Database.*

[151] H. Wang and K. Kosuge. "An inverted pendulum model for reproducing human's body dynamics in waltz and its applications in a dance partner robot". en. In: *2010 IEEE/SICE International Symposium on System Integration*. Sendai, Japan: IEEE, Dec. 2010, pp. 182–187.

[152] H. Wang and K. Kosuge. "Towards an understanding of dancers' coupled body dynamics for waltz". In: *IEEE International Conference on Intelligent Robots and Systems*. IEEE, Sept. 2011, pp. 2008–2013.

[153] *Wholodance*. URL: `http://www.wholodance.eu/`.

[154] J. Won et al. "Generating and ranking diverse multi-character interactions". In: *ACM Transactions on Graphics* 33.6 (2014), pp. 1–12.

[155] Y. Yang et al. "Generating a two-phase lesson for guiding beginners to learn basic dance movements". In: *Computers and Education* (2013).

[156] Zhao Liu et al. "Motion control of a passive robot for dancing with a male dancer". en. In: *2010 8th World Congress on Intelligent Control and Automation*. Jinan, China: IEEE, July 2010, pp. 6590–6595.