Moving together: Interpersonal Coordination and Individual Identification in Dyadic Dance

Thesis submitted in partial fulfillment
of the requirements for the degree of

Masters of Science in
Computer Science and Engineering
by Research

by

Prince Varshney
2020121012
prince.varshney@research.iiit.ac.in

International Institute of Information Technology
Hyderabad - 500 032, INDIA
July 2024

International Institute of Information Technology
Hyderabad, India

CERTIFICATE

It is certified that the work contained in this thesis, titled "Moving together: Interpersonal Coordination and Individual Identification in Dyadic Dance" by Prince Varshney, has been carried out under my supervision and is not submitted elsewhere for a degree.

Date

Adviser: Prof. Vinoo Alluri

To family and friends

Acknowledgments

Writing this thesis acknowledgment brings back memories of my initial apprehension about undertaking research in a dual degree program, worried that I wouldn't have the necessary skills. However, the experience has been a rewarding rollercoaster ride, filled with both challenges and accomplishments. I am incredibly grateful to the numerous individuals without whom this journey wasn't possible.

Foremost, I extend my sincere gratitude to my advisor, Dr. Vinoo Alluri. Your exceptional support, continuous inspiration, and motivation have been invaluable. You granted me the autonomy to pursue my research interests while being readily available to address any queries I had. Your insightful feedback and constructive advice have substantially enhanced the quality of my work, rendering my experience under your mentorship exceptionally rewarding.

I must also express my appreciation to Dr. Martin Hartmann, Dr. Emily Clements, and Dr. Petri Toiviainen. Your assistance throughout my research was invaluable, and your expertise has been instrumental in navigating the intricacies of my research project. I want to give a special thanks to Martin. Thanks for being extremely patient with me, always giving quick replies whenever I needed help dealing with motion capture data, and always pointing me to suitable literature. Your guidance has been as important as that of an advisor.

I would like to express my appreciation to the members of the music cognition group, particularly Pratyaksh and Guru, for your insightful discussions in our weekly meetings. Additionally, I am grateful to my labmates Priya, Vaishnavi, Brijesh, Shlok, and Ipshita (honorary) for creating an enjoyable work environment in the lab.

My acknowledgment would be incomplete without mentioning my friends Hari, Vrinda, VJS, Sreemaee, Ayushman, Shlok, Anjali, Iphsita, and Snehal. Hari, VJS, and Snehal, you were my primary source of support when I needed help, wanted to bounce ideas around, or sought feedback on my work. Vrinda, you made my college life experience unforgettable. I will always cherish our never-ending conversations during our campus walks and food expeditions in Hyderabad. I am glad to have found a best friend during college, Hari. You made every day entertaining with random facts about Spurs, John Green, and, of course, a healthy dose of gossip! Sreemaee, even though I have known you for a very short time, it belies my good memories with you. Thanks for always lending an empathetic ear and for your incredible food recommendations. VJS, thanks for always going the extra mile to help me. Anjali, thanks for always listening to what I have to say with excitement and for enduring my quirks.

At the very core of my gratitude lies my family. I thank Mom, Dad, Rishi Di, and Honey, for your unconditional love and encouragement. Di, thanks for always looking out for what is best for me and sponsoring all the gadgets I own and vacations I have been to. Honey, thanks for always being by my side and making me feel loved.

To all mentioned, your support and belief in my abilities have strengthened and motivated me. I am eternally grateful for your contributions to my journey.

Abstract

We move our bodies as a response to listening to music. These responses could range from simple movements such as head-banging, finger-tapping, and feet-tapping to a complex dance. Dance often occurs in social settings, where individuals entrain their movements to the music being played along with the visual cues derived from others. This phenomenon is termed rhythmic social entrainment. Moving in synchrony with others has been shown to foster social bonding. Dyadic dancing emerges as a first step in investigating rhythmic-social entrainment. This thesis examines the dyadic dance context from two angles: interpersonal coordination and individual identification. We first predict perceived interaction and similarity using kinematic and gestural features. Then, we explore the existence of unique signatures of individuals within dyads amidst interpersonal coordination. As an extension, we also look at the presence of unique movement signatures in a markerless-choreographic setting.

Interpersonal coordination has been studied using two perceptual variables: *Interaction* and *Similarity*. Studies have identified many postural and gestural features that exhibit moderate correlation with perceptual variables. However, several aspects of interpersonal coordination remain underexplored, particularly the role of musical features and energy levels of individuals in dyads. This thesis addresses these gaps by investigating the influence of music's danceability, which is primarily characterized by pulse clarity on interpersonal coordination, revealing a very strong statistically significant correlation between danceability of musical stimulus and mean perceived interaction ratings across all dyads dancing on that stimulus. This finding highlights the facilitative role of danceable music in enabling coupling. Furthermore, this study explores the link between the energy levels of dancers and interpersonal coordination, demonstrating perceived similarity is associated with similar energy levels within a dyad . In addition to this, energetic dyads are also perceived as more interactive, likely due to the impression of enjoyment and engagement they convey and vice versa. As a final step, we take dyad recordings and attempt to classify the perceptual variables into three classes: low, medium, and high. We employ novel features, such as Energy and Covariance Matrix, in addition to the ones from the literature, to train the model. We achieved reasonable accuracies in predicting "Interaction" and "Similarity." We also examined the joints that are important in the classification of these variables. This analysis revealed the significance of hands in predicting interaction relative to other body parts, which is consistent with other modalities, including spoken communication.

Carlson et al. [17] demonstrated the existence of motoric fingerprints by identifying individuals dancing freely to the music stimulus using only the movement features with notably high accuracy.

It is interesting to examine whether an individual has a unique signature even while dancing with a partner in the presence of interpersonal coordination. We achieved noteworthy dancer identification accuracy, signifying the existence of motoric fingerprints in the dyadic contexts. In addition to this, we demonstrated the joint pairs and joints that are key to the classification model. We employed the dyadic model to predict individual dancers based on features extracted from their solo performances. The high identification accuracy achieved indicates a strong consistency of unique movement signatures across both solo and dyadic settings. However, our misclassification analysis identified certain individuals who were not correctly predicted by the dyadic model. This anomaly was explained in terms of empathy dynamics of individuals within dyads.

Studies in the domain of music-induced movements predominantly rely on marker-based methods for movement capture. However, these methods suffer many limitations, with a primary concern being the potential alteration of natural movements due to the presence of markers on subjects. This poses a threat to the ecological validity of such studies. We examined the markerless data of professional dancers executing the same choreography. We investigated the notion of the personal style of a dancer by training a dancer identification model based on movement features. We achieved a dancer identification accuracy at least two times higher than the chance level, signifying the existence of motoric fingerprints in a choreographic-markerless setting.

In summary, this thesis contributes novel insights into interpersonal coordination in dyadic dance. It also shows the presence of motoric fingerprints in dyadic as well choreographic dance contexts, verifying the external validity of Carlson et al. [17]'s methods in dancer identification from movement features and explores the applicability of markerless movement data.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Figure 1.1: Music Induced Movements

Music can evoke emotions, stimulate cognitive processes, increase social bonding, and improve mental and physical health. Among its multifaceted functions, a particularly salient aspect is its ability to induce movements. These music-induced movements, more commonly known as dance, have been central to human experiences in all cultures since time immemorial [65]. After a comprehensive review of dance literature, Christensen et al. [19] suggested six positive effects of dance:

- Dance helps to focus attention on a singular activity. This sustained attention is shown to be beneficial after prolonged sensory overstimulation and stress, perhaps even providing an evolutionary advantage.

- Engaging in or watching dance helps to fulfill our emotional needs.

- Dance can function as a imagery for both the audience and the dancer.
- Dance serves as a medium for storytelling and dissemination of knowledge.
- It enhances self-awareness.
- It fosters social bonding.

He also highlighted some secondary benefits, including aesthetic pleasure, physical and psychological well-being, and advantages in sexual selection. Given these extensive benefits and the undeniable role of music and dance in human life, numerous studies have delved into music-induced movements. We primarily study music-induced movements within the framework of embodied cognition.

1.1 Embodied Cognition

Humans instinctively respond to music through bodily movements like finger tapping, head banging, or air-playing instruments, aligning our movements with the music's rhythm [30, 48, 62, 61]. Specifically, music that maintains a periodic beat at a frequency of 2Hz facilitates motoric entrainment [54]. Evidence also shows that motoric entrainment can occur at various beat levels, with vertical bouncing aligning with each beat and mediolateral swaying occurring at every fourth beat [79, 77]. Lesaffre et al. [50] found that around 95% of participants spontaneously moved in reaction to the music they heard. The intricate connection between sound and movement is deeply rooted in our brains [63, 45]. In an experiment by Bangert [3] involving trained pianists who listened to piano music, fMRI scans revealed activations in brain areas associated with motor control, indicating that listening to music involves more than just auditory processing. Furthermore, research has illustrated that the characteristics of music directly influence the movements it elicits [79, 73, 10].

The relationship between movement and music goes beyond being a response; some argue that movement is integral to parsing and comprehending musical sounds. Basing their argument on music performance studies, sound-tracing studies where listeners depict their impressions of audio stimuli through drawing, and dance movement studies, Godøy et al. [31] offered support for the concept of sound-motion similarity, rooted in the motor theory of perception. This theory suggests that we move our bodies to comprehend the sounds we hear, akin to the movements involved in producing those sounds. This bold assertion aligns with the idea of embodied cognition from psychology.

Embodied cognition challenges the conventional perspective that cognition solely relies on stimuli collected through sensory organs by acknowledging the role of the body in cognitive processes [69, 87]. The embodied cognition framework places sensorimotor processes at the center of understanding human cognition and behavior. According to this framework, actions and perceptions can be affected by one another instead of being linked linearly: perception, computation, and action. The state of our mind affects the way we move our bodies. Research has shown that emotions affect how we walk or dance [66, 38, 55] Conversely, our actions can shape our perceptions. One study illustrated this by finding that individuals who mimicked a smiling expression by holding a pen between their teeth found cartoons funnier compared to those who simulated a frown by holding a pen between their lips [74]. Similarly, another study reported that men perceived themselves as more assertive when they adopted fist-making gestures instead of neutral hand positions [67].

Embodied music cognition, as articulated by Leman [47], stems from the broader concept of embodied cognition. Leman posits a direct experience with music where the listener parses the moving sonic forms in the music through bodily imitation, either internally or externally.

1.2 Dyadic Dancing

Beyond coupling our movements with the audio stimuli, studies have shown that individuals can align their movements with visual stimuli, such as flashing lights [32]. We can also engage in visual coupling with the movements of others, a phenomenon termed social entrainment by Phillips-Silver and Keller [60]. Dyadic dancing becomes the most fundamental avenue to explore rhythmic-social entrainment, where the two individuals in the dyads might entrain their movements to visual stimuli from each other and also to the music stimulus being played.

In this thesis, we explore dyadic dancing from the following angles:

- **Interpersonal coordination**: Predicting interpersonal coordination measured by using perceptual variables using kinematic and gestural features.
- **Individual Identification**: Identifying individuals within dyads using only movement features, signifying the presence of unique signatures despite interpersonal coordination between the individuals.

1.3 Interpersonal coordination

Interpersonal coordination enables us to synchronize seamlessly with others during conversations, typically through turn-taking. The importance of such coordination becomes apparent when disruptions occur. For instance, even minor delays can disrupt the flow of a telephone conversation, leading to interruptions and awkward silences. Under these circumstances, we rely on explicit verbal cues, such as ”Your turn to speak,” to navigate and facilitate smooth communication. Bernieri and Rosenthal [6] define interpersonal coordination as “the degree to which the behaviors in an interaction are nonrandom, patterned, or synchronized in both timing and form.” Interpersonal coordination can be divided into behavior matching and interactional synchrony. Behavior matching involves mirroring actions/postures. Interactional synchrony comprises three elements: rhythm, the co-occurrence of movements, and the seamless integration of actions between individuals.

Some evidence suggests that our ability for interpersonal coordination may be innate. Research involving newborns revealed that infants, only a few days old, could synchronize their movements with human speech, regardless of whether the language was native or foreign to their parents. This synchronization occurred even when the speech was delivered indirectly via tape recordings. A notable discovery was that this synchronization did not manifest when the stimulus consisted of non-speech sounds [21]. Kato et al. [41] later replicated these findings in Japan.

The advantages of interpersonal coordination are well-documented across various studies. Research highlights that during conversations, we move our bodies to synchronize with the rhythm of not only our own speech but also that of the other person. Synchronizing with the speaker's speech rhythm might facilitate smooth communication, allowing for the anticipation of when the speaker will finish without inadvertently interrupting them [24, 25]. Furthermore, Interpersonal coordination, primarily postural and postural shifts congruency, has been linked with enhanced social rapport [23, 44, 43].

1.3.1 Measuring Interpersonal coordination

Various methodologies have been employed in the literature to measure interpersonal coordination. Microanalysis, for example, entails a frame-by-frame examination to identify instances of movement change, referred to as movement boundaries, and then comparing these boundaries between individuals [20]. Spectral analysis is another technique used to assess the periodicity of behaviors [58]. Other metrics include phonetic boundaries, discrepancies in turn-taking, and the degree of behavior matching.

1.3.1.1 Perceptual Ratings

The contemporary approach, the one adopted in this thesis, involves the use of perceptual ratings. This method is predicated on the hypothesis that humans are capable of directly perceiving stimuli in their environment. Baron and Boudreau [4] argue that relational behavior contains stimulus for various social properties. These properties are directly perceptible rather than being inferred, assuming the observer can witness the joint action unfold over time. This methodology was validated by a series of studies by Newtson et al. [58]. Participants in these studies were asked to give interaction ratings while observing a series of coordinated tasks. The tasks could include, for example, two people setting up tents together. These ratings reflected the interactant's coordination while doing the task. Furthermore, the relationship between ratings and coordination deteriorated when observers were engaged in perceptual interference tasks while giving ratings. No such effect was found when a cognitive interference task was given, signifying that the participants were perceiving the Interaction directly and not inferring it.

1.3.2 Interpersonal coordination in dance

Recent research has delved into the study of interpersonal coordination, particularly motoric coupling within dance. This exploration can be approached through two distinct methods: choreographed sequences and naturalistic, free-form movement. Lee et al. [46] noted an enhanced perception of group cohesion and fitness in choreographed performances that are temporally aligned compared to those experimentally misaligned. In a study by Chauvigné et al. [18], professional dancers performed two well-known Greek folk choreographies under four conditions: control, without holding hands, without auditory stimulus, and without visual stimulus. This study aimed to discern the impact of each type of stimulus on group synchrony. It was found that visual coupling has a more significant contribution in the horizontal plane, whereas auditory coupling has a more significant contribution in the vertical axes. Among all three conditions, the absence of haptic feedback was observed to disrupt group synchrony the most and affect all three axes. Concerns regarding the ecological validity of these studies have led to a shift towards investigating naturalistic free-form movements akin to those observed in club or party environments. This approach assesses interpersonal coordination through perceptual ratings, allowing for the identification of relationships between these ratings and kinematic coupling features. Two perceptual variables are employed: "Interaction" and "Similarity". Hartmann et al. [37] studied dyads dancing to pop music and found that Interaction closely relates to torso orientation, measuring the extent to which partners in a dyad face each other, which reflects gaze behavior. They also discovered that temporal-spatial coupling had a stronger correlation with similarity, highlighting the importance of similar movements in space and time for perceived similarity. They also showed that while interaction and similarity ratings correlated strongly, there's a notable pattern in some dyads where high similarity can coexist with low Interaction. This suggests that Interaction may be viewed as a specific instance of similarity where mutual orientation plays a key role. In a subsequent study, Hartmann et al. [36] examined the dyads with high torso orientation to unravel other kinematic and gestural features linked with the perceptual variables. A gestural feature, called postural bounding volume, measuring the similarity of postures of two individuals, was linked with perceived similarity. Exploring the dynamics between sequentiality-simultaneity and interaction-similarity revealed that similarity is associated with slower, simultaneous horizontal movements. In contrast, Interaction correlates with sequential coupling and quicker simultaneous movements. Additionally, the study discovered that mirroring movements foster increased coupling within dyads. Currently, we have many features that exhibit moderate correlations with these perceptual variables. However, further investigation is warranted to deepen our understanding of interpersonal coordination in dyadic dance.

Accordingly, the objectives to investigate interpersonal coordination in dyadic dance are:

- To explore additional features linked with perceived Interaction and Similarity.
- To utilize machine learning for the prediction of these perceptual variables.
- To identify key joints and joint pairs crucial for predicting perceptual variables.

1.4 Individual Identification

The richness of music-induced movements is extensively documented in many studies. Hufschmidt et al. [39] demonstrated that both adults and children can accurately identify gender from dance animations, with physical strength significantly influencing male gender identification by adults but not children. Camurri et al. [11] focused on the emotional expression within dance, instructing dancers to embody the intended emotions of Anger, Fear, Grief, and Joy in a choreography. They found that grief was the most accurately identified emotion, followed by anger, joy, and fear, with all above chance levels. Van Dyck et al. [85] also studied the embodiment of emotion in dance. They induced happy or sad emotions in the participants through a series of imagery sentences and accompanying music of the intended emotion. The participants were then asked to dance to a piece of neutral music. They found that the movements were faster, grander, and more impulsive in a happy state than in a sad-induced state. Luck et al. [52] showcased that music-induced movements reflect the personality type of the dancers and the musical genre that they are dancing to. Carlson et al. [13] found that conscientiousness and extraversion personality trait is linked with how one responds to small tempo changes. Machine learning methods have also been used to predict gender and personality traits [1] and musical genres [17] from free-form dance movements. As music-induced movements can reflect distinct individual characteristics such as gender, personality, and mood, it is reasonable to expect different individuals to move differently to the same music stimulus.

Johansson's [40] seminal study laid the foundation for investigations into the individuality of movement. His findings demonstrated that humans possess the ability to perceive walking from point-light animations featuring key joints. Cutting and Kozlowski's [22] study illustrated that friends could identify each other based on point-light displays of their gait. Troje et al. [81] extended this line of inquiry, revealing that human observers have the ability to learn to distinguish individuals from the point-light animations of their walk. In their investigation, identification performance was measured under various conditions, including displays normalized for size, shape, and walking frequency, as well as rotations of the walker by 90 degrees. Remarkably, the identification performance was three times higher than the chance level. In a subsequent study, Westhoff and Troje [86] used Fourier analysis and removed the first harmonic, which contains the majority of individual information, yet the performance remained above the chance level.

Music-induced movements can also be considered a motoric fingerprint, encompassing information that can be used to identify an individual. Both human observers and machine learning algorithms have demonstrated success in this area. Humans can recognize themselves from their motion-captured dance movements [68, 7]. Carlson et al. [17] employed machine learning methods and achieved a remarkably high accuracy of 94% in identifying individuals from their motion-captured data using only movement features while doing free-form dance movements to the music of eight genres. The presence of unique signatures or motoric fingerprints has not been studied in the dyadic context. In this thesis, we verify whether these movements can retain their status of motoric fingerprints in the dyadic context. As an extension to this work, we also look at individual identification in a markerless-choreographic setting where professional dancers perform identical choreographies.

Accordingly, the objectives to investigate the presence of motoric fingerprints in varied contexts are:

- Identifying individuals in dyadic context and markerless-choreographic settings using only their movement features.

- To identify key joints and joint pairs crucial for the classification.

1.5 Anatomical Planes and Axes

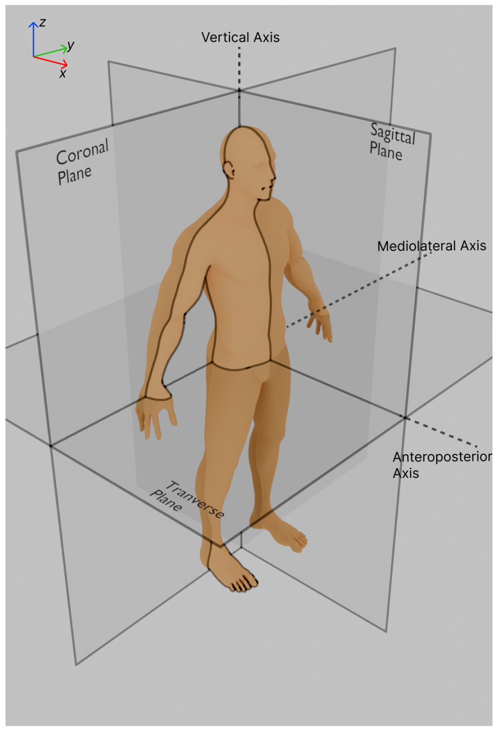


Figure 1.2: Anatomical Planes and Axes

In this thesis, we explore the movements of the human body, necessitating a comprehensive understanding of the anatomical planes and axes, as illustrated in Figure 1.2. These would serve as a reference point to describe directions and transformations applied across any plane. The three anatomical planes are:

- **Sagittal**: This vertical plane divides the body into left and right halves.
- **Transverse**: This horizontal plane divides the body into upper and lower parts.
- **Coronal**: This vertical plane divides the body into anterior (front) and posterior (back) sections.

The three axes are:

- **Anteroposterior** (X-axis): This axes is perpendicular to the coronal plane and runs from the front to the back side of the body.
- **Mediolateral** (Y-axis): This axis is perpendicular to the sagittal plane and runs from one side of the body to another side.

- **Vertical** (Z-axis): This axis is perpendicular to the transverse plane and runs from top to bottom.

1.6 The Scope and Contributions

Dyadic dancing is the topic of interest of this thesis. Our exploration revolves around two key aspects of dyadic dance: Interpersonal Coordination and Individual Identification in the presence of interpersonal coordination. Additionally, we extend our investigation to individual identification within a markerless-choreographic setting. The following are the major contributions of this thesis:

- Additional features, including musical features and energy level of individuals in dyads that are linked with perceived Interaction and similarity.
- Machine learning model to predict perceived Interaction and similarity. We also unveil the importance of hands relative to other body parts in predicting Interaction.
- The presence of motoric fingerprints in the dyadic dance setting.
- The presence of motoric fingerprints in the choreographic dance setting.
- Demonstrating the robustness of Carlson et al. [17]'s methods and findings in individual identification from only movement features.

1.7 Thesis Organization

- Chapter 2 delves into interpersonal coordination in dyadic dancing.
- Chapter 3 explores the presence of motoric fingerprints in dyadic dance settings by predicting individuals using only movement features.
- Chapter 4 explores the presence of motoric fingerprints in markerless-choreographic settings by predicting individuals using only movement features.
- Chapter 5 is about conclusions and future work.

Chapter 2

Interpersonal coordination in Dyadic Dance

Interpersonal coordination is a well-studied topic in social psychology and has recently gained attention in the context of naturalistic movements in dyadic dancing [15, 37, 36]. *Interaction* and *Similarity* are two perceptual variables that are utilized to study interpersonal coordination in a dyadic dance context. While studies have successfully identified numerous features showing moderate correlations with these perceptual variables, many aspects remain yet to be uncovered, necessitating further exploration to enhance our understanding of interpersonal coordination in dyadic dance.

2.1 Objectives and Hypothesis

Despite the significant influence of music on dance, research on interpersonal coordination has primarily focused on movement-related factors, neglecting the potential role of musical features like danceability in facilitating coupling between dancers. Danceable music is characterized by its ability to induce pleasure and the desire to groove, and it typically exhibits high pulse clarity [28], a tempo range of 100-120 BPM [27], a low-frequency range [76], and a moderate level of syncopation [88]. In dyadic dance, individuals allocate their cognitive resources between processing auditory stimuli from the music and visual stimuli from their partner. We propose that danceable music, by virtue of its inherent rhythmic clarity, can alleviate the cognitive demands of rhythmic entrainment, allowing dancers to dedicate more cognitive resources to social entrainment. Therefore, we hypothesize that music's danceability scores positively correlate with perceived interaction.

Dance, as a form of physical activity, demands a significant expenditure of energy. We hypothesize for individuals in dyads to be considered similar, they should be dancing with similar energy levels. In other words, the absolute value of the energy difference between two dancers correlates negatively with similarity. Upon inspection of animations, we hypothesize that dyads with greater levels of energy are perceived as more interactive and vice-versa.

While several features identified in the literature have shown a moderate correlation with perceptual variables [37, 36], no one has tried to predict these variables using machine learning. Incorporating machine learning models to predict these variables with reasonable accuracy could significantly streamline future research in this field, eliminating the need for the time-consuming task of collecting perceptual ratings. Such an advancement could pave the way for studies investigating the relationship between social closeness and motoric coupling or identifying individual characteristics that contribute to higher levels of Interaction within a dyad.

Studies have established the importance of hand gestures in spoken communication [5, 42, 71]. In the context of dyadic dancing, Carlson et al. [15] found that the dyads where individuals move their hands faster are perceived as more interactive and vice-versa. However, the relative role of hands in the entire body coupling has not been studied yet. Understanding this distinction is crucial for evaluating the impact of localized movements, such as hand gestures, compared to more global movements like bouncing. Hartmann et al. [36] calculated coupling features for the whole body and the entire body without hands. However, they couldn't justify the importance of hands in coupling as similar correlations between coupling estimates and perceptual variables were obtained in both cases. We hypothesize that hands play a pivotal role compared to other body parts in predicting Interaction.

In summary, we propose the following hypotheses:

- **H1**: The danceability of the music is positively correlated with mean perceived interaction ratings of the various dyads dancing to that music.
- **H2**: The absolute value of the energy difference between two dancers in the dyad correlates negatively with similarity.
- **H3**: The energy of dyads is positively correlated with interaction.
- **H4**: Hands play a pivotal role compared to other body parts in predicting interaction.

2.2 Methods

2.2.1 Dataset

2.2.1.1 Participants

We utilized Carlson et al. [17]'s dataset in which 73 participants (54 females) aged 19–40 years (M = 25.75, SD = 4.72) were recruited for the motion capture study. The participants were from 24 nationalities and had diverse musical and dance training backgrounds.

2.2.1.2 Apparatus

The motion capture was conducted using a twelve-camera optical system (Qualisys Oqus 5+), tracking 21 reflective body markers in three dimensions of the subject at a frame rate of 120 Hz. Marker locations are represented in Figure 2.1(A).

2.2.1.3 Procedure

The participants were grouped in sets of three or four. Within each group, data was recorded both individually and between every pair in the group. While every participant completed the individual recording, 64 participants (52 in groups of four and 12 in groups of three) completed the dyadic recording. Notably, in some groups of size four, certain markers of a participant were not captured in any of their dyad recordings. These participants were excluded from the analysis, effectively converting the groups to size three. During the recording, participants were instructed to move freely, either individually or in dyads, in response to each musical stimulus, simulating a dance club or party setting.

2.2.1.4 Stimuli Selection

The stimuli were selected based on social tagging data. Social tagging refers to associating free text labels to musical tracks facilitated by music streaming services such as "Last.fm". The selection pipeline ensured that selected stimuli represented the selected genres and were danceable. From the initial pool, 2407 tracks from Last.fm were retained such that they were associated strongly with only one genre label. The tracks were annotated with genre labels using the Short Test of Music Preferences (STOMP-R). The danceability of the musical tracks was ensured using three filtering steps. Firstly, only tracks associated with some dancing-related tags, like "danceable" or 'head banging," were retained. Secondly, tracks with non-zero danceability scores from Echo Nest API were retained. Echo Nest API estimates the danceability of a particular track using its acoustic features. Finally, tracks whose tempos were near the preferred tempo were retained, specifically 120 ± 12 Hz. Four tracks were randomly selected from each genre and manually checked for tempo and style consistency. We got 16 stimuli of length 35 seconds covering eight genres(2 stimuli per genre): Dance, Blues, Country, Metal, Jazz, Reggae, Pop, and Rap. For a detailed account of the selection process, please refer to Carlson et al. [16].

2.2.1.5 Preprocessing

The Motion Capture (MoCap) Toolbox ([9]) was employed for data pre-processing in MATLAB. The movement data for the 21 markers in the three dimensions underwent initial trimming to align with the duration of the musical excerpts. Linear interpolation was applied to address missing data and then resampled to 60Hz. Subsequently, the data was transformed into a set of 20 secondary markers called joints. Figure 2.1(B) illustrates the locations of these 20 joints. Although the majority of joints were situated at the same location as a specific marker, certain joints represented the average location of two or more markers. Please refer to 2.1 for a detailed understanding of the transformation process.

2.2.1.6 Perceptual Experiment

From the 90 dyads, Hartmann et al. [36] selected eight dyads that had maximum average torso orientation across all stimuli for the perceptual rating task to understand other factors linked with perceived

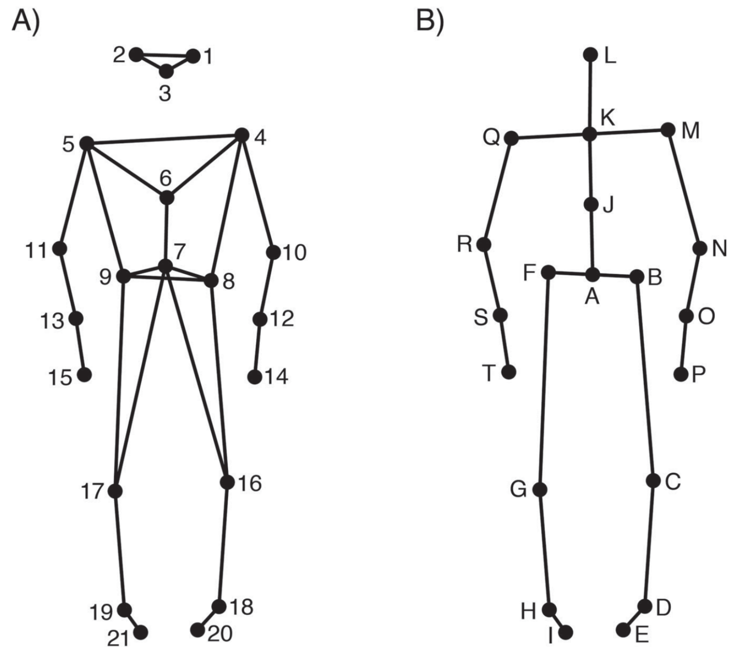


Figure 2.1: A) Marker locations B) Transformed joint locations

Source: Carlson et al. [17]

Interaction and similarity besides torso orientation. Computational details of torso orientation are reported later. The selected dyads are comprised of eleven unique individuals (five individuals were part of two dyads). Animations were created without music, with the markers of the left dancer depicted in green and those of the right dancer in red. Consistency in stick figure size was maintained across all animations. To control for potential biases due to positioning, the left and right dancers were randomly swapped in half of the animations.

To mitigate fatigue associated with rating all animations, the animations were divided into four partitions using a Latin rectangle combinatorial design (see Figure 2.2). This design ensured that each partition comprised of 32 animations had each dyad and each musical stimulus appearing twice.

Participants with diverse dance and musical training backgrounds and belonging to different nationalities were recruited through online platforms for the rating task. Participants were presented with animations in loop mode and asked to indicate their level of agreement or disagreement with two statements using sliders: "These dancers are dancing similarly to each other" and "These dancers are interacting with each other" and thereby rating the given animation on interaction and similarity between 0 and 100. Several measures were implemented to address the potential limitations of collecting ratings outside of a controlled setting to ensure data validity. Outliers were identified and removed, and participants were matched across the four partitions based on age and gender, resulting in 108 raters per partition. Finally, perceived interaction and similarity scores are computed by taking the mean across all raters.

Original	Transformed	Joint
8,9	A	Root
8	B	Right Hip
16	C	Right Knee
18	D	Right Ankle
20	E	Right Toe
9	F	Left Hip
17	G	Left Knee
19	H	Left Ankle
21	I	Left Toe
4,5,8,9	J	Torso
4,5	K	Neck
1,2,3	L	Head
4	M	Right Shoulder
10	N	Right Elbow
12	O	Right Wrist
14	P	Right Fingers
5	Q	Left Shoulder
11	R	Left Elbow
13	S	Left Wrist
15	T	Left Fingers

Table 2.1: Transformation of markers to joints

2.2.2 Features

2.2.2.1 Postural Features

- Torso Orientation [37]: Torso orientation measures the extent to which dancers in a dyad are oriented toward each other. The orientation of a dancer is estimated by a vector perpendicular to the line joining the two shoulders projecting toward the anteroposterior direction. The torso orientation at a particular time frame is calculated as defined in Equation 2.1.

$$\frac{\sum cos(\alpha) + cos(\beta)}{T} \tag{2.1}$$

  Where:

  - α and β are the angles between the projected vector and line connecting the mid-point of two shoulders to the other dancer's torso for each dancer, respectively.
  - T is the number of frames

  The torso orientation measure ranges from -2 (when both dancers are facing away from each other) to $+2$ (when both dancers are facing toward each other). Finally, the average is taken across the time dimension.

- Volumetric Matching [36]: Volumetric matching measures the congruency between the postures of two dancers. In line with previous research [2, 35], we use the convex hull of the position of joints to compute this feature. The convex hull can be considered the smallest convex polygon that bounds all the joints. Volumetric matching is defined as the average of the absolute difference

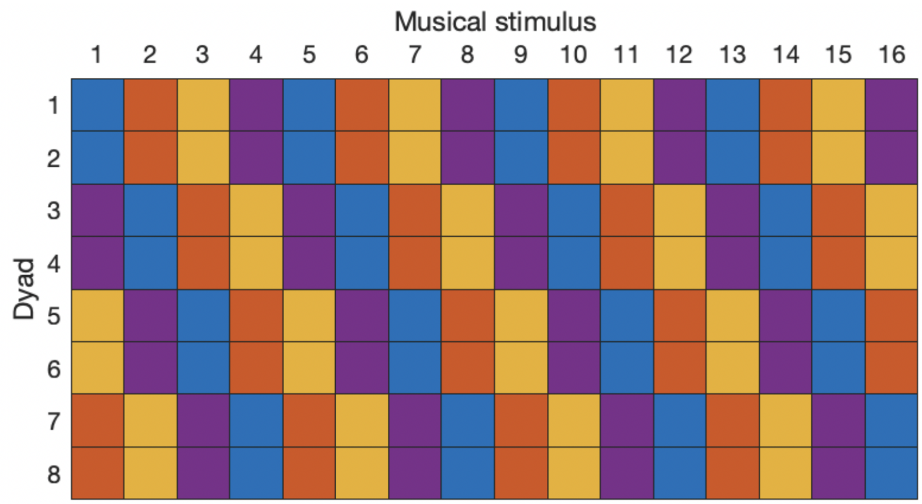


Figure 2.2: Division of 128 animations into 4 partitions where each color depicts a partition.

Source: Hartmann et al. [36]

between the volume of the convex hull of each dancer at each time frame. It ranges from 0 (perfect postural congruency) to $+\infty$ (perfect postural incongruency).

2.2.2.2 Gestural Features

Before extracting the kinematic features, the recordings of each dancer were transformed into a local coordinate system to account for the fact that different individuals may have oriented themselves differently in the recording space. In this coordinate system, the root marker (marker A in Figure 2.1(B) is the origin, and the line joining the hip markers defines the mediolateral axis. Individuals in the dyads tend to mirror each other. Therefore, one of the individuals was randomly mirrored before the gestural features were computed. This was done by following Fujiwara and Daibo [29]'s work that involves the reflection of joints in the sagittal plane and inverting the mediolateral coordinate. Instantaneous velocity was computed through time differentiation and a Butterworth smoothing filter (2nd order; 12 Hz cutoff frequency) for all markers in all three dimensions, as outlined in Burger and Toiviainen [9].

- Imitation [36]: The imitation feature measures sequential coupling between two dancers at different time lags defined by a cross-similarity matrix. Algorithm 1 taken from Hartmann et al. [36] lists step-by-step computation of imitation feature. In order to make the imitation estimate comparable between songs of different tempi, linear interpolation was used to rescale the measure to a beat lag scale. The various symbols used in the algorithm are as follows:
  - A, B : A and B are time series velocity data of both dancers respectively consisting of T number of frames and N number of channels.
  - M : M is the cross similarity matrix

- d : d is the imitation vector.

Algorithm 1: Imitation Algorithm

Input: $\boldsymbol{A} \in \mathbb{R}^{T \times N}, \boldsymbol{B} \in \mathbb{R}^{T \times N}$
Output: $\boldsymbol{d} \in \mathbb{R}^{T}$
$\boldsymbol{M} \in \mathbb{R}^{T \times T} := AB^{\top}$;
//dot products between any two time points
$d_0 = sum(diag(M, 0))$;
for $i \leftarrow 1$ ***to*** $T - 1$ **do**
  $d_i = sum(diag(M, -i)) + sum(diag(M, i))$;
//summing over each diagonal
return d;

- Synchrony [36]: Synchrony feature consists of the temporal average of coupling estimates at different movement frequencies. It is computed using Generalized cross wavelet transform [78], and it involves the conversion of the frequency axis to beat relative scale ranging from 0.5 beat level to 4 beat level to compare the stimulus with different tempi. The step-by-step computation of this feature is listed in the algorithm 2 taken from Hartmann et al. [36]. The various symbols used in the algorithm are as follows:
  - A, B : A and B are time series velocity data of both dancers respectively consisting of T number of frames and N number of channels.
  - U, V : U and V are channel wise wavelet transforms
  - C: Generalized cross wavelet transform
  - d : d is the synchrony vector.

- Covariance Matrix: Covariance Matrix consists of entries having covariance scores between any two marker's velocity time series. Such Covariance-based features have been successfully used in various machine learning classification tasks, including time series classification [26], action recognition [34], pedestrian detection [84], individual identification [17], and the prediction of individual characteristics such as gender and personality [1]. The covariance of velocity between markers of two dancers x_i and x_j across X, Y, and Z dimensions is measured using correntropy—a non-linear covariance measure that is shown in Equation 2.2. With 60 time series for each dancer, we get a 60x60 covariance matrix. Flattening the covariance matrix results in a 3600-length feature vector.

$$K\left(x_i, x_j\right) = e^{\frac{-\left\|x_i - x_j\right\|_2^2}{2\sigma^2 T^2}} \tag{2.2}$$

Where:

Algorithm 2: Synchrony Algorithm

Input: $\boldsymbol{A} \in \mathbb{R}^{T \times N}, \boldsymbol{B} \in \mathbb{R}^{T \times N}$
Output: $\boldsymbol{d} \in \mathbb{R}^{F}$
$\boldsymbol{U} \in \mathbb{C}^{F \times T \times N}, \boldsymbol{V} \in \mathbb{C}^{F \times T \times N}, \boldsymbol{C} \in \mathbb{R}^{F \times T}$;
for $i \leftarrow 1$ ***to*** N **do**
  $U_0 \in \mathbb{C}^{F \times T} := \boldsymbol{cwt}(A[:, i]), V_0 \in \mathbb{C}^{F \times T} := \boldsymbol{cwt}(A[:, i])$;
  $U_0 := \boldsymbol{resample}(U_0), V_0 := \boldsymbol{resample}(V_0)$;
  $U[:, :, i] := U_0, V[:, :, i] := V_0$
//Computing wavelet transform for each channel
for $f \leftarrow 1$ ***to*** F **do**
  for $t \leftarrow 1$ ***to*** T **do**
    $u := \boldsymbol{flatten}(U[f, t, :]), v := \boldsymbol{flatten}(V[f, t, :])$;
    $C[f, t] = \boldsymbol{abs}(\boldsymbol{sqrt}(\boldsymbol{mean}((u \cdot \overline{v})^2)))$;
//Generalized cross wavelet transformation using pair wise approach
$d := \frac{1}{T} \sum_{t=1}^{T} C(:, t)$;
//Computing average for each frequency across all time points
return d;

- $\|x_i - x_j\|_2$ denotes the L2 norm.
- T denotes the number of frames employed to accommodate samples of varying lengths.
- σ governs the steepness of the distribution of the generated features, with higher σ values yielding negatively skewed feature distribution and vice versa. To improve the discriminability of the produced covariance-based features between different subjects, the optimization of σ for each feature separately was achieved through the downhill simplex algorithm, aiming to minimize the absolute value of skewness in the produced features. We utilized the Python library "scipy.optimize" with the "Nelder-Mead" method for this optimization step. Ultimately, these optimized features were normalized to zero mean and one standard deviation.

• Energy: The dancer's energy is defined as the mean square velocity across markers and time frames as shown in Equation 2.3. Consequently, the dyad's energy is the sum of the two dancers' energies.

$$\frac{\sum_i \sum_j \sum_k v_{ijk}^2}{T * 20 * 3} \tag{2.3}$$

Where:

- i denotes a particular marker.
- j denotes a particular dimension.
- k denotes a particular time frame.
- T denotes the number of frames.

2.2.3 Machine Learning Analysis

Joints	Body parts
Root	Root
Hip	Hip (L,R)
Leg	Knee (L,R), Ankle (L,R), Toe (L,R)
Torso	Torso
Neck	Neck
Head	Head
Arm	Shoulder (L,R), Elbow (L,R)
Hand	Wrist (L,R), Fingers (L,R)

Table 2.2: Transformation of joints to body parts

SVM works by identifying the line that optimally segregates data points into distinct classes to correctly classify future data points. This line, known as the Optimal Separating Hyperplane (OSH), is positioned to maximize the margin between the two classes, thereby minimizing the risk of misclassification for new data ([33]; [53]). The data points closest to the hyperplane are referred to as support vectors. It is hard to draw separating lines in real-world data without making some errors. The SVM has a parameter C, which plays a crucial role in balancing the trade-off between training error and margin size. By regulating the penalty for misclassified data points during training, C influences the width of the margin. A smaller C value favors a larger margin, allowing more misclassifications, while a larger C prioritizes minimizing training error, resulting in a narrower margin. Relevant feature identification is important when working with SVM; otherwise, the SVM algorithm might struggle to classify samples accurately. It is essential to note that SVM can work with any number of dimensions and is also not limited to linearly separable classes.

2.2.3.1 Feature Selection

With numerous features, employing feature selection becomes crucial to diminish dimensionality and mitigate the risk of model overfitting. We utilized PCA (principal component analysis) for the feature selection.

2.2.3.2 Classification

We chose to model the problem of predicting perceived interaction and similarity as classification over regression; a key consideration is the nature of the perceptual variables rating themselves. These ratings represent an average across multiple participants, and the specific numerical difference between individual ratings may not be as meaningful as the categorical distinctions in perceived interaction and similarity levels. Consequently, classification offers a more suitable approach, as it allows for the prediction of discrete categories rather than attempting to model precise numerical differences in the ratings. The perceptual variables were categorized based on quantiles: 'low' (< 1/3 quantile), 'medium' (1/3 to 2/3 quantile), and 'high' (> 2/3 quantile).

Linear SVM with an L2 penalty, squared hinged loss, and one vs all strategy was used to predict the dancer from the selected features. A nested cross-validation technique was employed to check the generalizability of the model. The outer cross-validation holds a part of the dataset as the test set. The inner cross-validation helps in hyperparameter tuning, as well as the number of components to be retained in PCA and parameter 'C' of SVM in our case. This approach is superior to the single cross-validation, as it utilizes only a portion of the dataset provided by the outer cross-validation, reducing the risk of overfitting the entire dataset. We also utilized stratified sampling to ensure the same proportion of classes in training and test sets. We take three different feature set groups:

- Covariance Matrix
- Postural and Gestural Features from literature + Energy (Features)
- Combined

We train the model on the covariance matrix feature alone to identify significant joint pairs between dancers for predicting the perceptual variables. This will help us check the relative contribution of various joints in predicting the variables and, therefore, help us verify the hypothesis that hands are relatively more important.

We train each feature set on a two-class classification where we take on "high" and "low" classes as well as three-class classifications.

2.2.3.3 Feature Importance

We used weights from the linear SVM to identify the principal component with the highest weight and then used its loadings' absolute values to measure joint pair importance. In matrix form, the dimensions of the joint pair importance matrix are 60x60. We summed the importance values across the three dimensions, leaving us with a 20x20 matrix. To eliminate the distinction between individual dancers, we summed the values along the diagonal, resulting in a symmetrical matrix. The further transformation involved grouping joints into meaningful body parts and amalgamating left-right distinctions (see table 2.2), ultimately yielding an 8x8 matrix.

2.3 Results

The correlation between the danceability of the musical stimulus (extracted from the Echo nest API) which ranges from 0 to 1 and the mean of the interaction and similarity ratings of all the dyads dancing

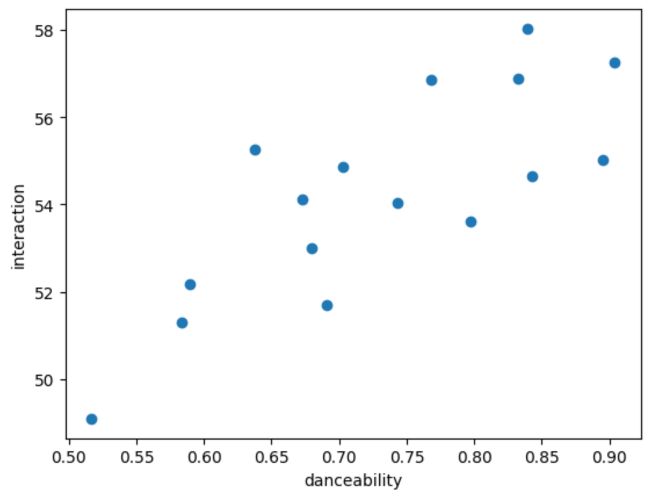


Figure 2.3: Danceability of musical stimulus vs the interaction it elicits

	Interaction	Similarity
Danceability (H1)	0.79^{***}	0.56^{*}
Absolute difference of energy of dyad members (H2)	-0.33^{***}	-0.48^{***}
Energy of dyads (H3)	0.38^{***}	0.24^{***}

Table 2.3: Correlation values for different hypotheses. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

to that stimuli was computed (see Figure 2.3). We also compute the energy of each individual and test the hypothesis H2 and H3. The one-tailed Pearson correlation values for the first three hypotheses are reported in Table 2.3. As there is a strong correlation between interaction and similarity, partial correlations were also computed for the different hypotheses and are reported in Table 2.4.

4-fold mean cross-validation classification accuracies of perceived interaction and similarity are reported in Table 2.5a and Table 2.5b respectively. The feature importance analysis in predicting interaction and similarity is shown in Figure 2.4 and 2.5, respectively.

2.4 Discussion

Interpersonal coordination is a fundamental aspect of human interaction that plays a pervasive role in our daily lives. From navigating busy roads to engaging in conversations, this ability to seamlessly coordinate our actions with others is crucial for effective social interaction and task performance. The study of interpersonal coordination spans diverse contexts, and this chapter specifically focuses on interpersonal coordination in dyadic dance, building upon the work of Hartmann et al. (2023). In his work, perceptual variables such as ”Interaction” and ”Similarity” have been employed to quantify in-

	Interaction	Similarity
Danceability (H1)	0.67^{***}	-0.09
Absolute difference of energy of dyad members (H2)	0.03	-0.37^{***}
Energy of dyads (H3)	0.30^{***}	-0.04

Table 2.4: Partial Correlation values for different hypotheses. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

	Two class	Three class
Covariance	0.72	**0.41**
Features	0.65	0.46
Combined	**0.74**	0.40

(a) Perceived Interaction

	Two class	Three class
Covariance	0.73	**0.51**
Features	0.53	0.34
Combined	**0.77**	0.48

(b) Perceived Similarity

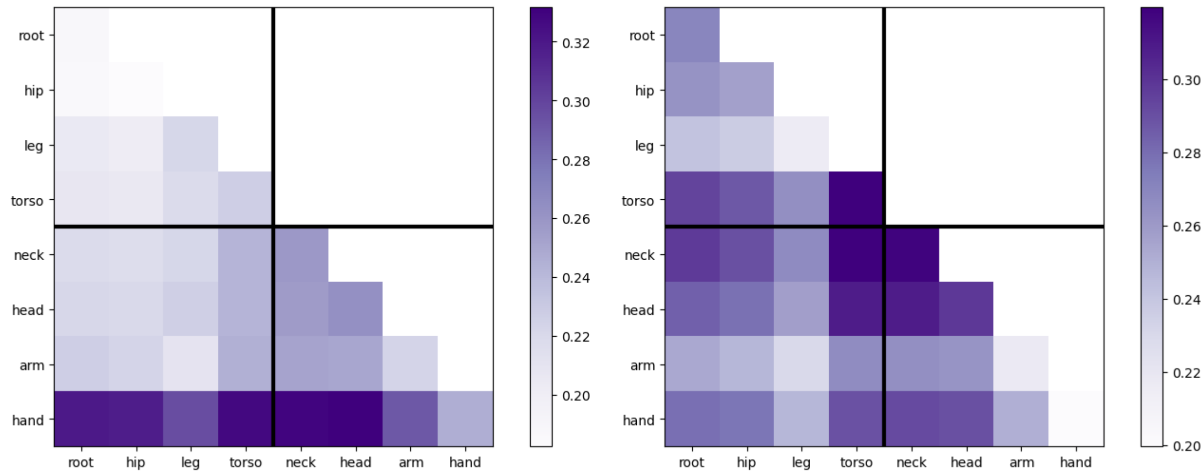
Table 2.5: 4-fold mean cross-validated accuracies

terpersonal coordination in dance, and numerous postural and gestural features exhibiting moderate correlations with these variables have been identified.

Dyad	r
1	0.54^{*}
2	0.60^{*}
3	0.15
4	0.10
5	0.66^{*}
6	0.28
7	0.26
8	0.12

Table 2.6: Correlation between danceability and interaction for each dyad. $^{*}p < 0.05$

Musical features have been overlooked in the study of interpersonal coordination in dyadic dance. We found a strong statistically significant correlation (refer Table 2.3) between danceability scores of musical stimulus mean perceived interaction across all dyads dancing to that music. This correlation remained significant even after controlling for similarity through partial correlation analysis (refer Table 2.4), providing robust support for our first hypothesis. These results indicate that danceable music characterized by high pulse clarity could facilitate coupling between individuals by reducing the effort



(a) Perceived Interaction

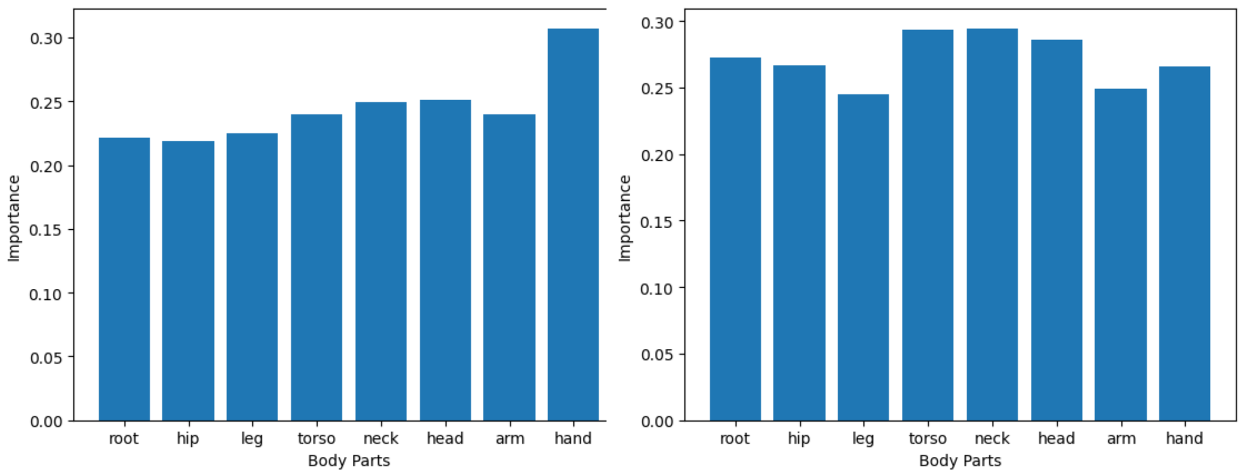
(b) Perceived Similarity

Figure 2.4: Importance of Body-part pairs

required in rhythmic entrainment as the music has more rhythmic clarity and allows more cognitive resources to be allocated to visual coupling, thereby improving interpersonal coordination. As each dyad danced on every musical stimulus, we calculated the correlation between the danceability scores of the music and the corresponding perceived interaction ratings for each dyad. The results, presented in Table 2.6, reveal positive correlations, with several showing statistical significance.

This study introduced a novel feature, termed "Energy," defined as the mean squared velocity across all markers and time frames. The results supported both hypotheses related to this feature. We found statistically significant negative correlations between the absolute value of the difference in energy of both dancers and similarity even after controlling for interaction, thereby validating the second hypothesis. As hypothesized, the energy of the dyads was associated with interaction. This was indicated by statistically significant correlation, even after controlling for similarity. Dyads exhibiting higher energy may be perceived as having more enjoyment and engagement, resulting in higher interaction ratings and vice-versa. The partial correlation between the energy of the dyad and similarity is not significant, as dancers can execute similar movements with varying degrees of energy expenditure.

We also introduced a novel feature called the covariance matrix, where each entry in the matrix represents the non-linear covariance between any two velocity time series of markers of any dimensions of the two dancers in dyads. We employed three different feature sets in our analysis. The first set solely included the covariance matrix to estimate the relative contribution of various joint pairs. The second feature set comprised all postural and gestural features from the literature, along with energy. Finally, the third feature set combined all available features. We conducted both two-class classification (High and Low) and three-class classification tasks. Our best accuracy results were 74% for two-class, and 46% for three-class classification of interaction; 77% for two-class, and 51% for three-class classification of similarity. A notable observation is that our feature sets are more effective at predicting similarity compared to interaction. The relatively lower accuracy in the three-class classification task can be



(a) Perceived Interaction (b) Perceived Similarity

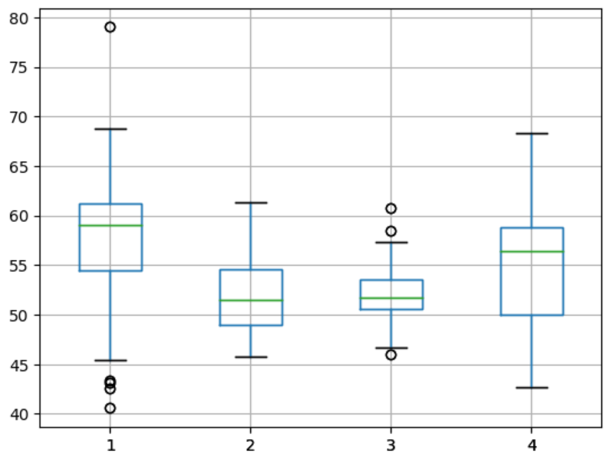
Figure 2.5: Importance of various body parts

justified by considering that final perceptual ratings are averages across multiple raters. Consequently, small numerical differences between two ratings may not accurately reflect the actual difference in the coupling, leading to misclassifications at the boundaries in the three-class scenario.

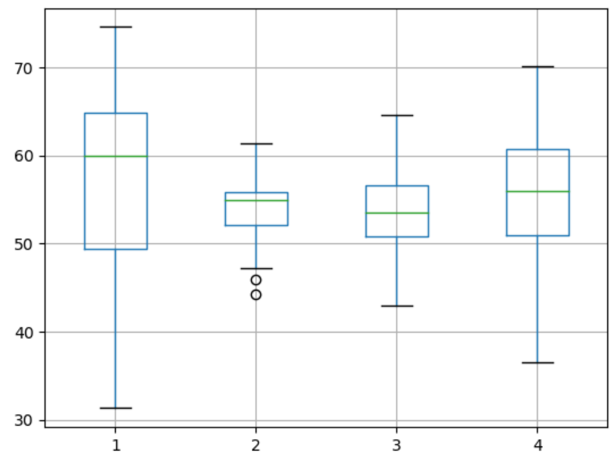
The weights of the two-class SVM model trained on only the covariance matrix were used to select the prime principal component. The loading values of this component were taken as representative of feature importance. As we had features between every pair of joints of any direction, this could be visualized as a 60X60 matrix. Finally, through a series of transformations, we get an 8X8 body-parts pair importance matrix (see Figure 2.4). The importance of each body part is visualized in Figure 2.2, revealing the importance of hands in predicting interaction relative to other body parts, showing support for the hypothesis 4.

To mitigate rater fatigue, the animation ratings were collected in four partitions. Each individual was matched for age and gender in all partitions. Each partition contained two animations per dyad and two animations per musical stimulus. Therefore, it is reasonable to expect similar rating distributions across partitions (see Figures 2.6a and 2.6b). However, Kruskal-Wallis ANOVA revealed statistically significant differences between partitions for perceived interaction. Post-hoc Dunn's test further identified significant differences between Partitions 1 and 2, as well as Partitions 1 and 3. A closer examination of the animations within each partition showed that while the ordering of animations by interaction values made perceptual sense within each group, the overall order across partitions did not. This indicated that the values from different partitions were not directly comparable. Kruskal-Wallis ANOVA showed no group differences in perceived similarity. This suggests that perceived similarity is a more stable and well-defined notion compared to perceived interaction.

To address the problem of different distribution of interaction ratings in different partitions, a normalization solution was implemented, aiming to adjust the ratings within each partition such that the mean ratings would reflect the same level of coupling across all groups. This measure was based on



(a) Perceived Interaction



(b) Perceived Similarity

Figure 2.6: Distribution of perceptual variables across various partitions.

	Original	Normalized	Good Partitions	Good Partitions Normalized
Covariance	0.72	**0.78**	0.70	0.74
Features	0.65	0.70	0.61	0.75
Combined	**0.74**	**0.78**	**0.82**	**0.79**

(a) Two class classification

	Original	Normalized	Good Partitions	Good Partitions Normalized
Covariance	**0.41**	**0.47**	**0.42**	0.48
Features	**0.46**	0.41	0.41	0.39
Combined	0.40	**0.47**	0.41	**0.53**

(b) Three class classification

Table 2.7: 4-fold mean cross-validated accuracies in predicting Interaction under different conditions

the assumption that the average rating in each partition represents the same coupling. The prediction results after normalization are presented in Table 2.7 and show improved accuracy in both two-class and three-class classifications. Poor internal consistency for perceived interaction was observed in partitions 2 and 3 (see Table 2.8). Consequently, only the data from partitions 1 and 4, exhibiting good internal consistency, were utilized for further analysis, with the results also reported in Table 2.7.

In our current study, we observed a strong and statistically significant correlation between the danceability of musical stimuli and the mean perceived interaction ratings for those stimuli. However, dyad-level analysis indicated that this correlation does not hold consistently across all individuals. This discrepancy highlights an anomaly that warrants further investigation in future research. One significant limitation identified in the utilized dataset is the comparability of ratings across different partitions. The normalization solution that is suggested in this chapter is based solely on the assumption that average ratings in each partition reflect equivalent levels of interaction, which may not be accurate. To address this, future studies linking danceability and interaction in dyads should consider having all ratings for a

	I	II	III	IV
Similarity	.75	.68	.72	.74
Interaction	.74	.66	.63	.79

Table 2.8: Internal consistency of perceptual variables partition-wise

particular dyad conducted by the same evaluator. To mitigate fatigue involved in rating multiple dyads, dyads could be allocated to different partitions. This approach would also facilitate a more detailed analysis of the effects of time on these interaction ratings, providing deeper insights into how temporal factors influence perceived dynamics within dyads.

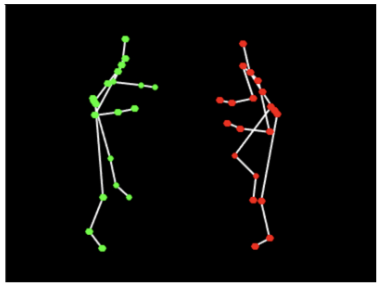
Chapter 3

Individual Identification in Dyadic Dance

Music-induced movements embody rich information about gender [39], mood, emotion, personality [11, 52, 85, 13, 14], and culture [80]. Therefore, it is reasonable to expect different individuals to move differently to the same music stimulus. Historically, the individuality of movements was first examined through gait analysis, which has been shown to contain individual identifying information [22, 81, 86]. Sevdalis and Keller [68] demonstrated that the uniqueness of movements extends beyond mere walking. In their research, participants were asked to engage in dancing, clapping, and walking. Subsequently, they were required to determine whether motion capture recordings of these three activities were of themselves or others. The self-identification accuracy surpassed random chance for all activities, with dancing exhibiting the highest accuracy. Bläsing and Sauzet [7] studied action recognition in addition to self-recognition in dance through a two-part experiment. In the first phase, participants were either blindfolded and instructed to create movements, learned movements that others created while being blindfolded, or simply observed the movements of others. In the subsequent phase, participants were tasked with categorizing recordings into four distinct categories: unfamiliar, only observed, observed and learned, and self-created. Findings indicated that self-recognition was influenced by action recognition, as participants more accurately identified themselves from the movements that they created despite having no visual memory of the creation process as compared to those they had learned from others. Carlson et al. [17] employed machine learning methods and achieved a remarkably high accuracy of 94% in identifying individuals from their motion-captured data using only movement features while doing free-form dance movements to the music of eight genres. We have evidence to conclude that music-induced movements can serve as motoric fingerprints, encompassing information that can be used to identify an individual [17, 68, 7]. Dance is a social activity and, therefore, often occurs in groups. In particular, we look at the presence of motoric fingerprints in the dyadic dance context.

3.1 Objectives and Hypothesis

Interpersonal coordination between individuals of dyads leads to mimicking and mirroring of movements either simultaneously or after some delay [36]. Carlson et al. [14] showed that individuals in



Dyads	27
Dancers	54
Musical Stimuli	16 (8 musical genres)

Figure 3.1: A portion of the dyadic dataset for the dancer identification study.

dyads tend to move their hands more than when dancing alone for the same music stimulus, and this difference is statistically significant. They also showed individuals with higher empathy attune their movements according to their partner. The individuality of movements in the context of dyadic settings has not been studied yet. We hypothesize that the unique signature of individual movements persists when individuals dance with a partner.

3.2 Methods

3.2.1 Dataset

We use the entire dyadic part of the dataset that was described in Chapter 2. An individual can dance in two or three dyads depending on whether they were part of groups of sizes three or four, respectively. In order to ensure balanced classes, we selected two dyads with the highest torso orientation from groups of size four and one dyad from groups of size three, ensuring each participant was present in only one dyad. Following these filtering steps, we arrived at a total of 864 recordings comprising 27 dyads (54 participants) dancing to 16 musical stimuli for the final classification analysis. The recordings were preprocessed using Motion Capture (MoCap) Toolbox ([9]) in Matlab as outlined in Chapter 2, and instantaneous velocity was computed for each joint for further feature engineering.

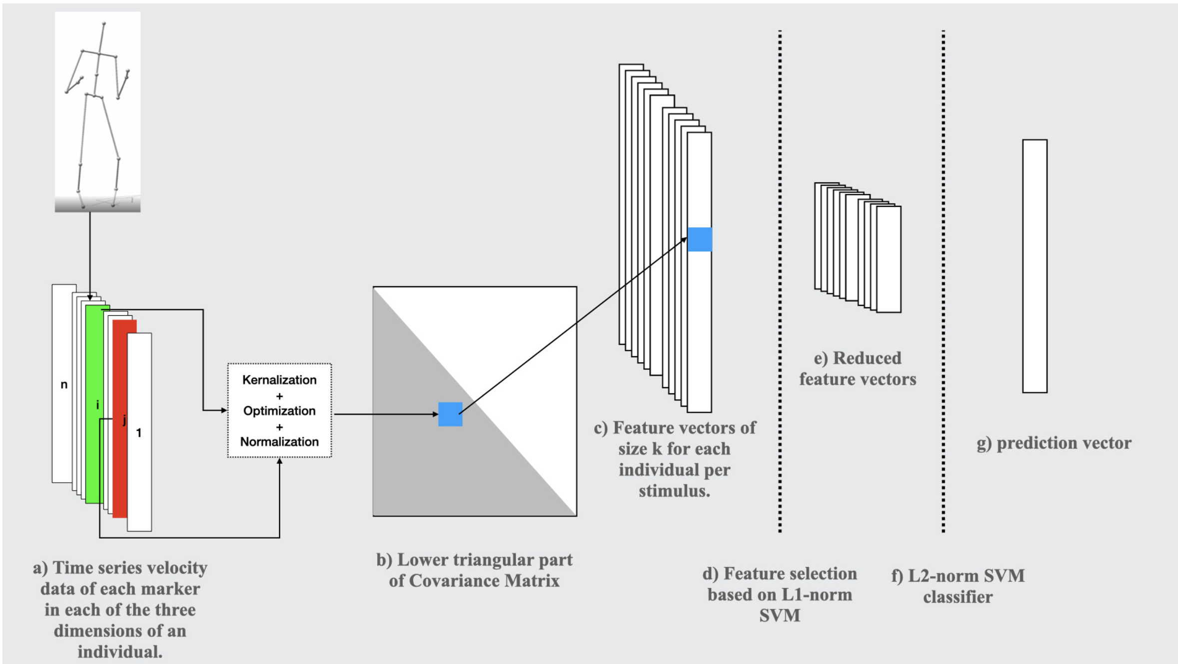


Figure 3.2: Machine Learning pipeline

3.2.2 Machine Learning Analysis

A support vector machine (SVM) algorithm was used to identify the dancer. We rigorously adhered to the machine-learning pipeline outlined in Carlson et al. [17], which is illustrated in Figure 3.2.

3.2.2.1 Feature Extraction

Based on the findings by Troje et al. [81] that show how markers move in relation to each other (as opposed to their spatial relationships) plays some role in the human perceptual identification of walkers, the covariance measure between the velocity of any two marker time series of any dimension was used for feature extraction. The process to compute the covariance matrix is the same as in Chapter 2. The resulting covariance matrix is symmetric, and the diagonals have zero values. Consequently, only the lower triangular portion is flattened to generate a feature vector. This vector has a length of 1770, given 20 markers in three dimensions.

3.2.2.2 Feature Selection

With numerous features, employing feature selection becomes crucial to diminish dimensionality and mitigate the risk of model overfitting. We utilized Linear SVM with an L1 penalty, squared hinged loss, and one vs all strategy for feature selection. ([89]). This approach results in many feature weights learned by the model having almost zero values, facilitating the filtering of irrelevant features. The L1

Table 3.1: Dancer Identification accuracy when that musical genre was held as test-set (Dyadic)

Musical Genre	Accuracy
Reggae	98.15
Pop	99.07
Metal	87.96
Jazz	95.37
Dance	100.00
Country	96.30
Blues	99.07
Rap	98.11

norm SVM is better suited at feature selection than the L2 norm, and it also avoids overfitting ([59]). Employing the one-vs-all strategy for training the classifier yields feature weights for each class. The L1 norm of these weights is computed across each class and feature, serving as a measure of the importance of that particular feature. The classifier is trained on features iteratively introduced in decreasing order of importance. The number of features to be retained becomes a hyperparameter.

3.2.2.3 Classification

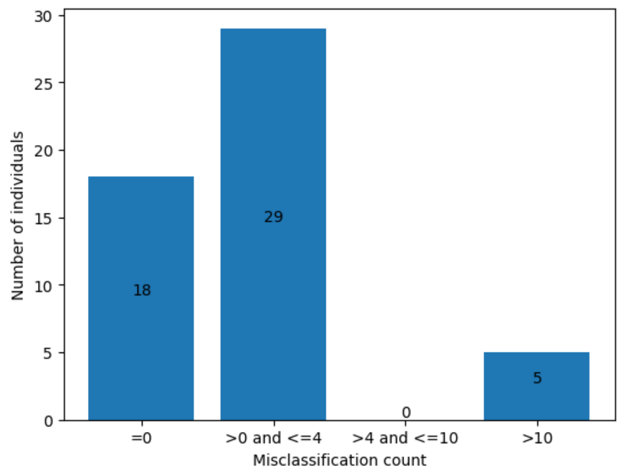
Linear SVM with an L2 penalty, squared hinged loss, and one vs all strategy was used to predict the dancer from the selected features. A nested cross-validation technique was employed to check the generalizability of the model. It is important to note that optimal hyperparameters were computed using only the training fold in the outer cross-validation. The dyadic dataset used the leave-one-genre-out technique for the outer cross-validation to ensure the generalizability of dancer identification in new genres.

3.3 Results

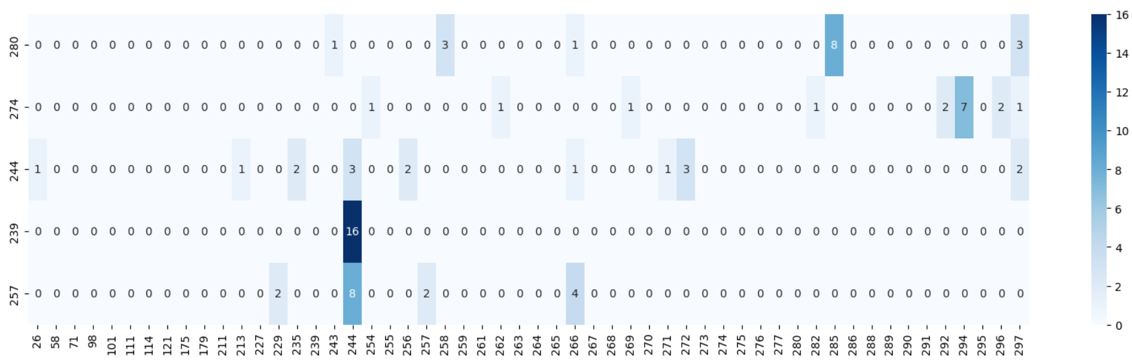
We attained an exceptionally high mean cross-validation accuracy of 96.75% with a standard deviation of 0.04. Accuracy, when each musical genre was held as a test set, is depicted in Table 3.1.

3.4 Discussion

Inspired by the findings of Carlson et al. [17], who demonstrated high accuracy in predicting individuals solely based on their movement features using machine learning, our study delves into the exploration of whether music-induced movements retain their status as motoric fingerprints within the novel dyadic dance context.



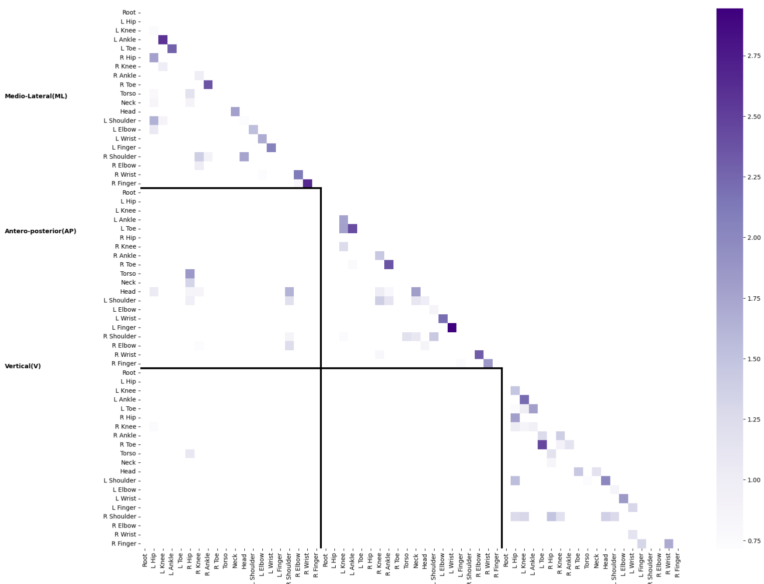
(a) Number of individuals vs Misclassification count.



(b) The confusion matrix for the five individuals with > 10 misclassification in (a), namely 257, 239, 244, 274, 280, is depicted. The color bar shows the number of misclassifications (darker being more misclassifications)

Figure 3.3: Misclassification analysis when utilizing features from individual dance settings to predict dancers using the dyadic model.

We accomplished a mean cross-validation accuracy of 96.75% in the dyadic dataset. To further unravel the factors contributing to surpassing the accuracy Carlson et al. [17] achieved, we extended our analysis to include training on the individual portion of the dataset. Employing a nested cross-validation approach, we achieved a notable mean cross-validation accuracy of 97.04%, surpassing Carlson et al. [17] results. This improvement can be attributed to the use of nested cross-validation rather than single cross-validation. These results underscore the notion that our movements retain unique characteristics even when dancing with a partner. We observed lower accuracy in metal and jazz genres compared to pop, rap, and others, which is consistent with Carlson et al. [17]. Metal and jazz genres are associated with stereotypical moves, which may cause individuals to move more similarly than in other genres. Headbanging is common in the Metal subculture [72, 8, 75], and moves like Charleston and swing are common in the Jazz subculture [49, 56]. The influence of distinctive movement stereotypes within these subcultures may contribute to a higher degree of movement similarity and potentially impact identification accuracy.



(a) Importance of Joint pairs

Marker	Importance
R Hip(ML)	18.25
R Shoulder(ML)	16.93
Head(AP)	16.35
L Shoulder(AP)	15.33
L Hip(ML)	14.44
L Knee(ML)	13.50
R Elbow(AP)	13.47
R Shoulder(V)	12.78
R Knee(ML)	12.41
R Ankle(AP)	11.96
L Ankle(AP)	11.94
L Shoulder(ML)	11.69
L Knee(AP)	11.44
L Toe(V)	11.42
Head(V)	11.40

(b) Importance of joints

Figure 3.4: Feature Importance analysis

Feature importance analysis revealed a pattern consistent with Carlson et al. [17], illustrated in Figure 3.4. Joint pairs aligned in the same direction, such as both Antero-Posterior (AP), hold greater importance compared to pairs in different directions, like one in AP and the other vertical. Proximity also plays a role, with pairs occurring in nearby locations, such as fingers and wrists, being more important. Further, individual joint importance was determined by summing the importance values of pairs involving that joint. The analysis underscores the significance of limb joints, including shoulders, ankles, wrists, and knees, along with hips, in predicting individuals.

Utilizing data encompassing both individual and dyadic dance performances for the same individuals, we employed the model trained on dyadic performances to predict individuals based on features extracted from their individual performances. Two individuals from the dyadic dataset were excluded due to labeling issues in their individual recordings. Remarkably, we achieved an accuracy of 86.44% using the important features derived from the dyadic model for the prediction task. This high level of dancer identification accuracy suggests that unique movement signatures are consistently maintained across solo and dyadic dancing scenarios. Our subsequent misclassification analysis revealed intriguing patterns (refer to Figure 3.3a). Notably, certain individuals were accurately predicted by the dyad model without any errors, while others remained unpredicted by the dyad model altogether.

We further analyzed five individuals (257, 239, 244, 274, 280) who were poorly predicted by the dyadic model. We found that the two dyads make up four out of the five individuals, and the one remaining individual's partner was accurately predicted. This suggests that the prediction challenges may be related to specific dynamics within these dyadic interactions. We focused on three dyads: A(274-280), B(244-239), and C(257,268). Upon reviewing the performances of these individuals in both solo and dyadic settings, Dyad A was excluded from further analysis due to significant recording errors, even though the data markers were fully recorded. For Dyads B and C, a common pattern emerged: one individual displayed minimal movement, such as swaying left to right with limited hand and leg movements and occasional walking, while the other performed complex movements involving both hands and legs in a solo setting. When dancing together, they exhibited strongly coupled sways with minimalistic hand and leg movements. In both dyads, the individuals who adjusted their movements from complex in solo settings to more limited in dyadic settings had higher empathy levels, as measured by EQ scores, compared to their partners. This finding supports previous research suggesting that dyads comprising one individual with a higher EQ score and the other one with lower EQ tend to be more interactive compared to dyads where both individuals have high or low EQ scores [15]. This pattern indicates that individuals with higher empathy may be more attuned to their partners with lower empathy, adapting their movements to facilitate a more harmonious interaction, which is in line with the findings of Carlson et al. [14]. Interestingly, the individual exhibiting complex movements in solo settings was not recognized by the dyadic model. This discrepancy likely arises because the interaction within the dyad forces these individuals to adopt restrained movements that do not capture their unique signature. In Dyad B, the individual with minimalistic movements was also not recognized, being misidentified as their partner instead, as illustrated by the confusion matrix 3.3.

Analyzing the animations of the dyads where individuals were accurately predicted by the dyadic model revealed that either there is very little interaction between the individuals or the interaction is there, but the movements are free enough to capture their unique signature.

We selected dyads based on high torso orientation scores to ensure that the individuals in the dyads were interacting with each other. High torso orientation scores only ensure that the individuals in the dyad are facing each other. The future work could involve selecting dyads based on high perceptual ratings of interaction, strengthening the validity of the obtained results.

In this chapter, we examined the presence of motoric fingerprints in dyadic settings and found that these unique movement patterns are generally consistent for an individual across both solo and dyadic setting, with few exceptions. As an extension of this work, the next chapter delves into the challenge of individual identification in a constrained choreographic setting, where dancers perform the same routine.

Chapter 4

Indiviudal Identification in Markerless-choreographic setting

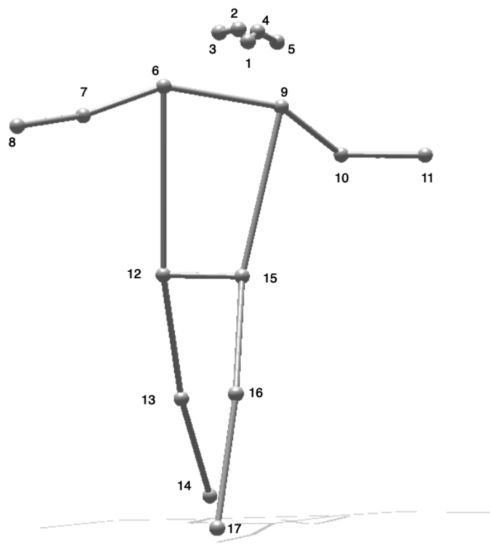
In the previous chapter, we have shown the presence of motoric fingerprints in the dyadic dance context. In this chapter, we attempt the problem of individual identification in a constrained choreographic setting, where professional dancers are dancing to the same routine in markerless motion capture.

4.1 Objectives and Hypothesis

Dancer identification has been studied in the context of naturalistic free-form movements [17, 68, 7]. These movements could be influenced by gender [39], mood, emotion, personality [11, 52, 85, 13, 14], and culture [80]. This leads to an intriguing question: *To what extent can we identify individuals based on their movements within a more constrained setting, where each subject performs the same routine?* Unlike free-form environments, constrained settings offer limited scope for variations between subjects. The distinctions, if they exist, are subtle and may manifest, for instance, in aspects like *flowy versus jerky* renditions of the same choreography [1]. To investigate choreographic settings, it becomes imperative to involve professional dancers with substantial experience, ensuring strict adherence to the prescribed choreography. This work tries to verify the notion of the personal style of a dancer.

These dancer identification studies have employed a marker-based motion capture system for recording dance movements. However, there are several challenges associated with using marker-based systems for capturing movements. Marker-based systems necessitate extensive preparation time for subjects, limiting their practicality. They cannot be used in environments where their placement could impede the studied activity, such as sports. Furthermore, the placement of markers can alter the naturalness of subjects' movements. Hence, it is important to capture music-induced movements in a markerless setting. Advancements in computer vision, particularly leveraging modern deep learning methods, have significantly enhanced the efficacy of human pose estimation in a markerless setting. Human pose estimation and tracking, a computer vision task, includes detecting, associating, and tracking semantic key points such as "right shoulders" and "left knees" from images and videos. OpenPose stands out as

[1] Please note that in the two videos video1 and video2, despite doing the same routine, each dancer infuses a personal touch to it



(a) 17 key-points locations

Keypoint No.	Name
1	Nose
2	Left Eye
3	Left Ear
4	Right Eye
5	Right Ear
6	Left Shoulder
7	Left Elbow
8	Left Wrist
9	Right Shoulder
10	Right Elbow
11	Right Wrist
12	Left Hip
13	Left Knee
14	Left Ankle
15	Right Hip
16	Right Knee
17	Right Ankle

(b) 17 key-points names

Figure 4.1: COCO Format

a noteworthy library capable of 2D/3D pose estimation ([12]). Despite the challenges in human pose estimation, including occlusions due to viewing angles, several research studies have substantiated the accuracy of these systems in tracking the key points. Nakano et al. [57] conducted a study involving participants engaging in activities like walking, countermovement jumping, and ball throwing, utilizing both marker-based and Openpose-based markerless motion capture systems to record the activity. The differences were quantitatively analyzed using mean absolute errors, revealing that 80% of the errors were less than 30mm. Notably, recent studies have underscored the success of markerless systems, offering a promising alternative in overcoming the limitations associated with marker-based approaches. Hence, we should be able to validate Carlson et al. [17]'s findings in a markerless setting. We hypothesize that despite dancing to the same choreography, there is a unique signature of a dancer that is identifiable in a markerless setting.

4.2 Methods

4.2.1 Dataset

We employed the AIST++ dataset from Li et al. [51] for the choreographic-markless setting. AIST++ is a large-scale 3D dance motion dataset generated from the AIST dance database ([83]). AIST is just a collection of dance videos in 9 camera angles without any 3D information. AIST++ provides 17

			Dance Genres
Basic Dance (10x3x10x4) 	Dancers	3	ballet jazz
			street jazz
	Choreographies	10	krump
			house
	Impressions	4	LA-Style hip hop
Advanced Dance (10x3x7) 	Dancers	3	middle hip hop
			waack
			lock
	Choreographies	7	pop
			break

Figure 4.2: Summary of AIST++ dataset

COCO-format [64] human joint locations in 2D and 3D for each frame, along with camera and SMPL pose parameters. 17 COCO-format joint locations are depicted in the Figure 4.1.

The subjects of the study were 40 professional dancers (15 females) with more than five years of experience specializing in a particular dance genre. It is a rich dataset covering ten dance genres: Ballet Jazz, Street Jazz, Krump, House, LA-style Hip-hop, Middle Hip-hop, Waack, Lock, Pop, and Break.

The dataset, summarized in Figure 4.2, comprises 1,408 dance sequences; basic dance constitutes 85%, and advanced dance includes the remaining 15%. Within the basic dance category, each genre includes ten choreographies performed by three dancers in four impressions: intense, loose, hard, and soft. In the advanced dance category, dancers were asked to choreograph their own moves. Some dancers shared their choreographies with others. For each genre, there were seven choreographies performed by three dancers. We will be using the basic dance category for our analysis.

4.2.2 Preprocessing

The recordings were initially in the form of (T,17,3), with T representing the number of frames. It was then flattened to (T,54) to facilitate further processing in the Motion Capture (MoCap) Toolbox [9]. Gaps were filled linearly, and an additional step of smoothing was carried out using a Butterworth smoothing filter (2nd order; 12 Hz cutoff frequency) for all markers in all three dimensions. Finally, instantaneous velocity was computed for each joint for further feature engineering.

4.2.3 Machine Learning Analysis

We utilized the machine-learning pipeline from the previous chapter. In this dataset, we have 17 key points as opposed to 20 from the dyadic dataset. Therefore, the resulting feature vector generated from flattening the lower triangular half of the covariance matrix has a length of 1275.

4.2.3.1 Classification

We aim to capture the subtle differences between dancers following the same routine. Hence, it is crucial to train the model for each dance genre while also ensuring that no choreography overlaps between the training and test datasets. We have 120 data points for each genre covering 3 participants. Employing a Stratified k-fold with five folds for the outer cross-validation ensures that classes are evenly distributed in both training and test sets. We also trained a dance genre classifier using the same pipeline and employed a Stratified K-fold with three folds. We also ensured that different participants were present in both the training and test sets for the dance genre classifier. Consequently, dancer identification becomes a two-step process: initially predicting the dance genre from the movements and subsequently using the model specific to that genre to predict the individual dancer. Additionally, a dancer identification model was trained on the entire dataset for comparison purposes.

4.3 Results

We attained a mean cross-validation accuracy of 47.6% with a standard deviation of 0.05 for dancer identification across the entire dataset. The dance genre classifier demonstrated a mean cross-validation accuracy of 88.25% with a standard deviation of 0.02. The mean cross-validation dancer identification accuracy for each dance genre is presented in Table 4.1.

Dance Genres	Mean	Std
Break	63.89	0.07
House	91.67	0.08
B Jazz	79.07	0.12
S Jazz	65.51	0.14
Krump	67.72	0.11
L Hip-hop	81.67	0.09
Lock	85.00	0.09
M Hip-hop	86.67	0.07
Pop	70.00	0.12
Waack	78.90	0.14

Table 4.1: Cross-validation dancer identification accuracy for each dance genre (AIST++)

4.4 Discussion

This study delves into the notion of the personal style of a dancer by trying to identify the dancer in a choreographic setting, where each dancer performs the same routine using only the movement features. It is a more constrained setting than the free-form setting. In an effort to overcome the limitations associated with marker-based systems, we opted for markerless data acquisition in this setting. AIST++ dataset [51] was utilized, where there are ten dance genres. Each dance genre consists of 3 individuals performing ten choreographies.

In our training over the entire AIST++ dataset, we attained an accuracy of 47.6%, surpassing the chance level of 3.33%. However, given the model's dual challenge, that is, discriminating between individuals across diverse genres with distinct movements and discerning subtle differences within each genre, the achieved accuracy is reasonable.

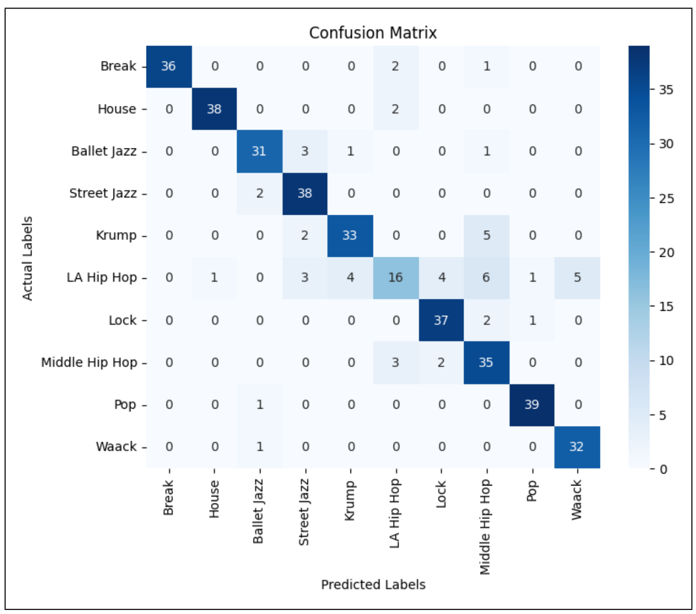


Figure 4.3: Confusion matrix of dance genre prediction model.

Recognizing these complexities, we adopted a two-step approach to dancer identification. First, we performed genre classification using the same machine learning pipeline, achieving an impressive 88.25% accuracy (against a 10% chance level). We also performed a misclassification analysis using a confusion matrix (see Figure 4.3) to examine which dance genres were often confused with each other. It is important to note that we don't have 40 data points for some of the genres because certain videos were placed on the ignore list by the dataset's author due to inaccurate keypoint captures. The confusion matrix shows that most genres were accurately predicted, with some confusion between Ballet Jazz and Street Jazz, as well as LA style hip hop and Middle Hip Hop. This is expected since these pairs are variants of the same dance styles and share common elements.

In the second step, we use the model specific to that genre to predict the individual. The dancer identification accuracy ranged from 63.89% to 91.67% across different dance genres, all surpassing the chance level of 33.33%. These outcomes underscore the existence of a distinctive personal style for each dancer, further corroborating the efficacy and robustness of Carlson's methods in dancer identification.

Feature importance analysis conducted across all dance genres within the AIST++ dataset reveals a consistent pattern with Carlson et al. [17] and dyadic setting, providing a consistent narrative across studies. Please refer to the Appendix A for the detailed visualization of important key points/joints pairs and key points/joints.

In our current study, the sample was limited to just three dancers for each dance genre performing the same routine. To broaden the scope and enhance the generalizability of our findings, future research could replicate these studies with a larger cohort. One promising approach is to leverage short video platforms like TikTok, Instagram Reels, and YouTube Shorts, where millions of users engage with trending dance challenges. This would provide a rich dataset of varied dance styles and routines followed by a diverse group of individuals. Recent advancements in technology have also shown a growing interest in generating smooth dance animations that synchronize well with music, as evidenced by research utilizing the AIST++ dataset [51, 70, 82]. Moving forward, an exciting direction would be to generate dance animations that not only replicate a given choreography but also incorporate the unique style of a particular dancer.

Chapter 5

Expert-based Qualitative Analysis

In this chapter, we take a qualitative approach to gain deeper intuition about the perceived interaction between dyads and the personal style of a dancer from a movement perspective. To achieve this, a dance expert was interviewed and shown animations from both datasets.

Regarding the concept of perceived interaction, the expert explained that an individual's technical strength lies in the lower body, while creativity and the tendency to interact with others are expressed through the upper body. Our feature importance analysis in predicting interaction supports this assertion, as the importance scores for the upper body parts, namely the head, neck, and hand, were higher than those for the lower body parts, namely the root, hip, and leg. He also added that not only mirrored movements or similar movements done simultaneously or after a lag lead to enhanced perceived interaction, but the different movements done at the same time also add to the notion of perceived interaction. For example, if one individual sway left and right and the other individual takes a spin at the same time, then this movement pattern would be considered interactive.

Regarding the dancer's personal style in the choreographic setting, the expert noted that a choreography can be broken down into postures and the transitions between them. A dancer's technical strength and personal style are actually captured in these transitions—the way they move from one posture to another. Future studies could delve into pinpointing the personal style of a dancer within the time domain and thereby verifying the notion that it's the transition that captures the personal style of the dancer as opposed to the postures.

Chapter 6

Conclusions

This thesis contributes to the dyadic dance literature in various aspects. The first half of the thesis examines dyadic dance from the perspective of interpersonal coordination. Interpersonal coordination plays a pervasive roles in our life. It has been studied in dyadic dance context using two perceptual dimensions: Interaction and Similarity. However, the role of musical features and energy levels of individuals in dyads in interpersonal coordination has not received attention in the literature. Our findings indicate a strong and statistically significant correlation between the danceability of a music stimulus and the average perceived interaction it elicits across all dyads. Furthermore, we found that similarity is associated with synchronized energy levels among individuals, while interaction correlates with the overall energy of the dyads. Finally, we introduced a new feature called the covariance matrix and trained a machine-learning model to predict interaction and similarity. The model achieved accuracies above the chance level, with our feature sets proving to be more effective at predicting similarity than interaction. The feature importance analysis revealed the importance of hands in predicting interaction relative to other body parts. We also discovered that interaction ratings across all partitions were not directly comparable. To address this issue, we implemented a normalization solution that enhanced the accuracy of our models.

The implications of these findings are significant: danceable music not only gets us moving but also facilitates interaction among individuals, making it particularly effective for engaging people in party settings. Furthermore, for live group performances or dance animations to captivate an audience, it is crucial that all involved individuals maintain high and similar energy levels.

The analysis linking danceability and interaction for each dyad revealed that the correlation was not statistically significant for some dyads. Future studies should investigate the reasons behind these anomalies while also ensuring that the anomalies are not simply due to group differences between partitions. To address this, it is crucial for future research to involve having all ratings for a specific dyad conducted by the same evaluator. Additionally, future studies could explore how interpersonal coordination evolves over time for each dyad.

The second half of the thesis focused on individual identification within dyadic dance settings. Previous research has shown that each individual possesses a unique movement signature or a motoric fingerprint when dancing freely. Our study extends these findings to dyadic interactions, demonstrating that these unique signatures persist even when dancing with partners. This is evident by high dancer identification in dyadic settings. We also used the dyadic model to predict individual dancers based on features extracted from their solo performances. We achieved high identification accuracy that underscores the consistency of movement signatures across both solo and dyadic contexts. However, our misclassification analysis revealed that specific individuals were not accurately predicted by the dyadic model. Further investigation involved analyzing these individuals' animations in both solo and dyadic settings. We observed a common pattern: in these dyads, an individual with high empathy was paired with a less empathetic partner who barely danced. This pairing forced the empathetic dancer to adopt constrained movements to synchronize with their partner, which did not reflect their unique movement signature from solo performances. Future research could explore selecting dyads based on interaction ratings rather than torso orientation to prevent samples where there is no interaction.

As an extension of the thesis, we examined markerless data from professional dancers, each following the same choreography. We addressed the problem of dancer identification within this context and discovered that the accuracy was at least twice as high as the chance level, varying with the dance genre. This analysis substantiates the concept of a "personal style" of a dancer and carries significant implications for dance generation technologies. These findings suggest that dance generation solutions can progress beyond merely producing animations that synchronize well with music. Now, they can also incorporate a dancer's personal style, enhancing the authenticity and appeal of the generated animations. Additionally, given that this study was conducted in a markerless setting, the approach holds the potential for further validation on a broader scale. Platforms like TikTok, Instagram Reels, and YouTube Shorts, where millions of users participate in dance trend challenges, provide an ideal environment for testing our findings across diverse and extensive samples.

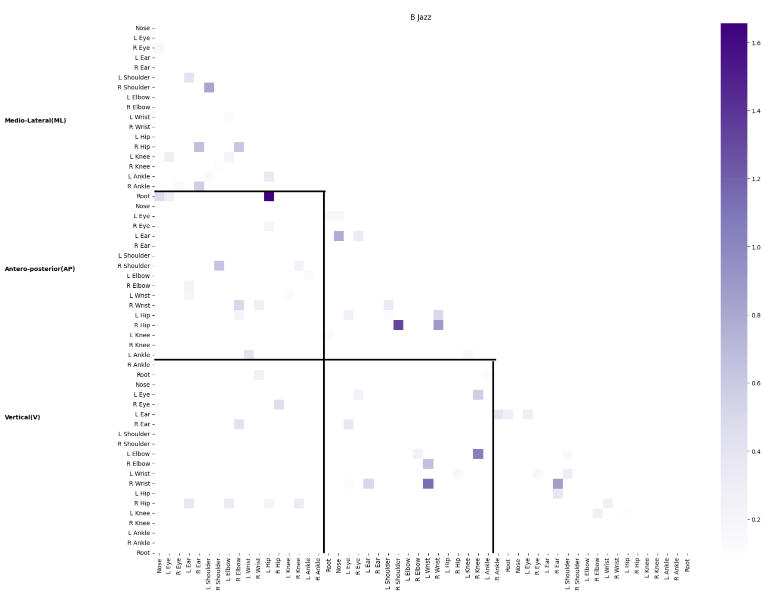
To conclude, this thesis examined dyadic dancing from two angles: Interpersonal coordination and Individual Identification. As an extension, we also looked into the problem of individual identification in the markerless-choreographic setting. We showed how danceability and energy level of individuals are linked with interpersonal coordination and these findings have significant real-world implications. Further, we proposed machine learning models to predict Interaction and Similarity. We also showed the presence of motoric fingerprints in both dyadic and markerless-choreographic settings. Our research also uncovered that how high levels of empathy could sometimes restrain an individual's movements in a dyad, potentially changing their unique signatures from solo setting. The successful adaptation of markerless data in this thesis paves the way for leveraging the vast array of video content available on platforms like TikTok, Instagram Reels, and YouTube Shorts to further enhance and expand dance research.

Appendix A

Supplementary Material

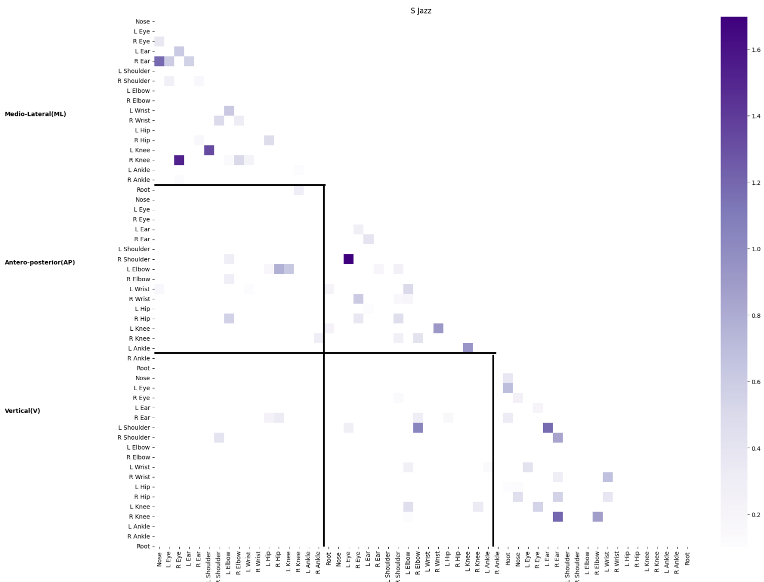
A.1 Feature Importance Analysis of AIST++ Dataset

For each dance genre, we show the importance of key-point pairs and keypoints in the predicition of the dancer for that particular genre.



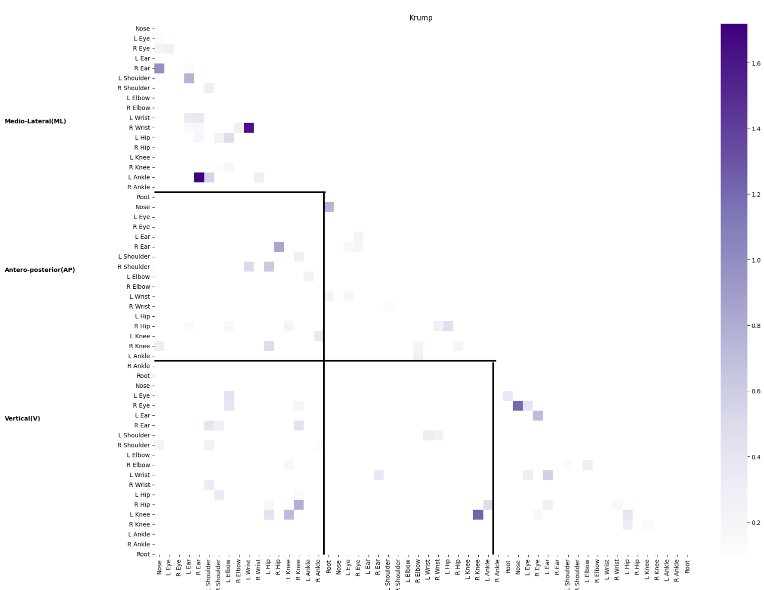
Marker	Importance
Nose(AP)	2.73
R Hip(V)	2.72
L Hip(AP)	2.50
L Hip(ML)	2.39
L Knee(AP)	2.39
R Wrist(AP)	2.22
L Elbow(AP)	2.20
R Shoulder(V)	2.00
R Elbow(ML)	1.84
R Hip(ML)	1.74
L Ankle(AP)	1.64
R Shoulder(ML)	1.64
R Ear(AP)	1.64
R Knee(V)	1.55
L Wrist(V)	1.44

Figure A.1: Ballet Jazz



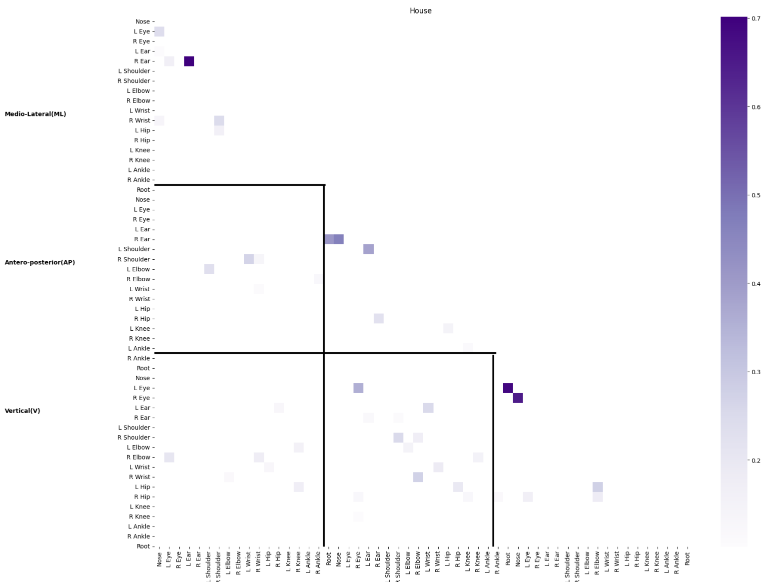
Marker	Importance
R Shoulder(V)	4.23
R Elbow(AP)	3.53
L Elbow(AP)	3.31
R Knee(ML)	2.91
R Ear(ML)	2.72
R Eye(ML)	2.65
L Elbow(V)	2.58
R Ankle(V)	2.18
L Wrist(AP)	2.09
R Knee(AP)	2.09
R Eye(AP)	1.97
L Knee(ML)	1.97
L Elbow(ML)	1.95
L Hip(AP)	1.88
L Hip(V)	1.85

Figure A.2: Street Jazz



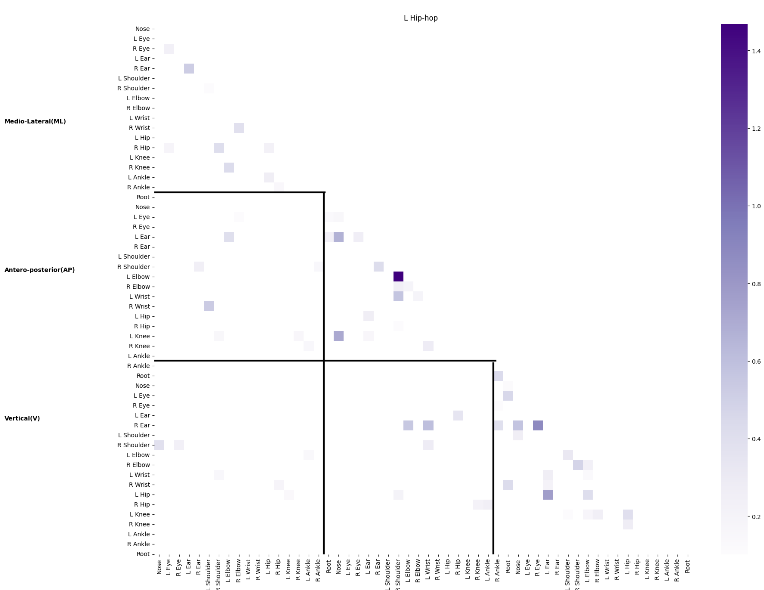
Marker	Importance
R Ear(ML)	3.53
R Ear(V)	3.15
L Ankle(V)	3.13
L Ankle(ML)	3.03
L Wrist(ML)	2.88
L Hip(ML)	2.62
L Shoulder(ML)	2.57
R Wrist(ML)	2.57
L Ankle(AP)	2.55
R Knee(ML)	2.30
R Knee(V)	1.94
Nose(ML)	1.83
L Elbow(ML)	1.78
L Ear(ML)	1.67
L Shoulder(V)	1.66

Figure A.3: Krump



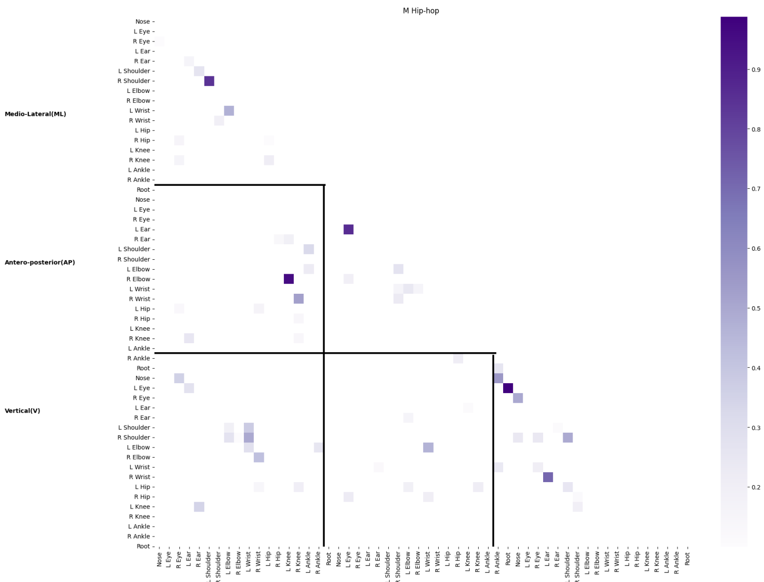
Marker	Importance
R Wrist(V)	1.27
L Ear(V)	1.26
L Shoulder(AP)	1.14
R Ear(ML)	0.95
R Wrist(ML)	0.93
L Ear(ML)	0.87
L Elbow(AP)	0.85
L Eye(V)	0.77
R Knee(V)	0.76
L Ear(AP)	0.74
L Knee(V)	0.72
R Eye(V)	0.69
L Wrist(AP)	0.69
L Eye(ML)	0.66
R Ear(V)	0.66

Figure A.4: House



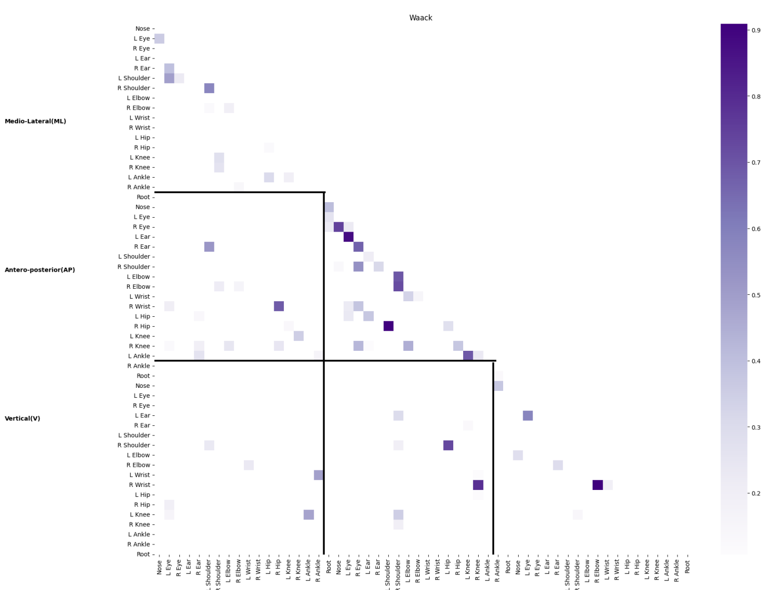
Marker	Importance
L Elbow(AP)	3.42
R Shoulder(V)	3.03
L Knee(V)	2.24
R Elbow(AP)	2.22
R Ear(AP)	1.99
R Wrist(AP)	1.92
L Eye(AP)	1.69
L Shoulder(V)	1.64
L Eye(V)	1.48
L Wrist(V)	1.41
R Elbow(V)	1.39
R Knee(AP)	1.36
R Hip(ML)	1.25
R Ear(V)	1.15
Nose(V)	1.11

Figure A.5: LA-style hip hop



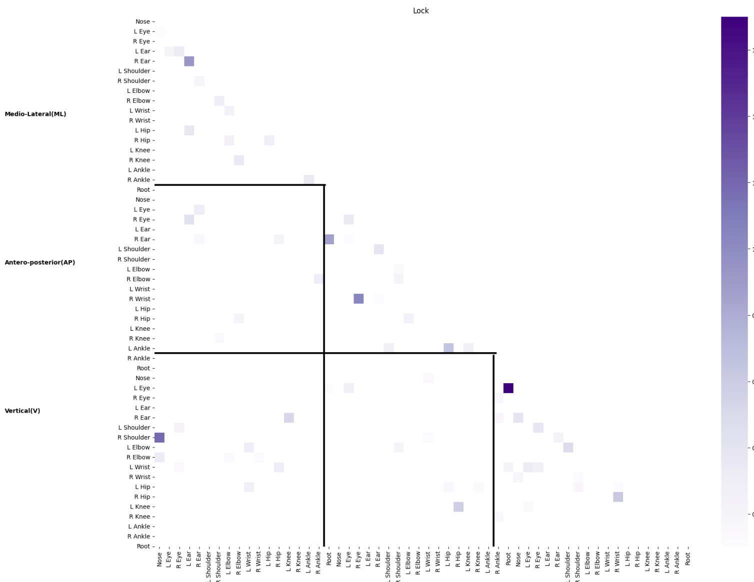
Marker	Importance
R Elbow(V)	2.20
R Eye(V)	1.79
L Wrist(ML)	1.65
L Wrist(AP)	1.59
L Elbow(V)	1.46
R Knee(ML)	1.46
L Eye(V)	1.32
Nose(V)	1.32
L Ear(V)	1.30
R Eye(AP)	1.28
R Wrist(AP)	1.28
L Knee(ML)	1.21
L Shoulder(ML)	1.16
R Shoulder(ML)	1.15
L Wrist(V)	1.12

Figure A.6: Middle Hip hop



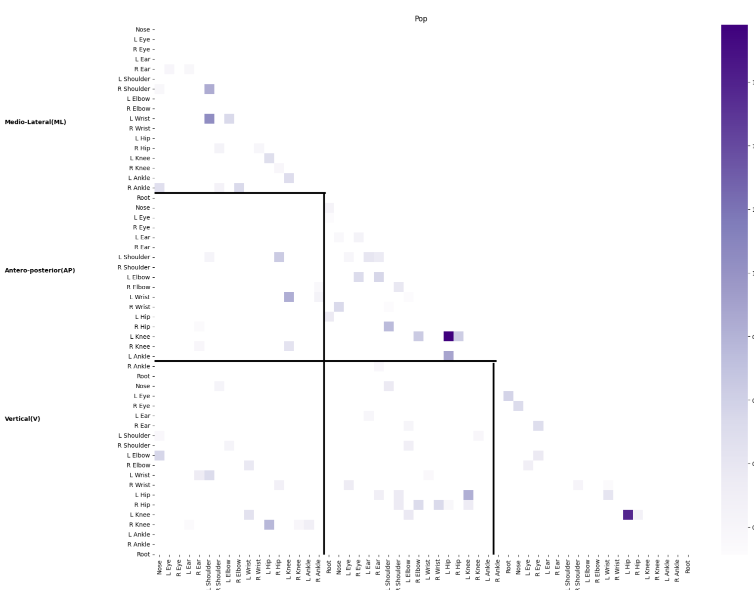
Marker	Importance
L Elbow(AP)	3.66
L Ankle(AP)	3.50
L Ear(AP)	3.20
L Shoulder(ML)	2.17
R Hip(V)	2.02
L Eye(ML)	1.92
R Eye(AP)	1.92
L Knee(AP)	1.85
R Hip(AP)	1.85
R Ear(AP)	1.60
L Shoulder(AP)	1.60
R Elbow(AP)	1.58
L Hip(AP)	1.56
R Ankle(AP)	1.47
R Wrist(V)	1.43

Figure A.7: Waack



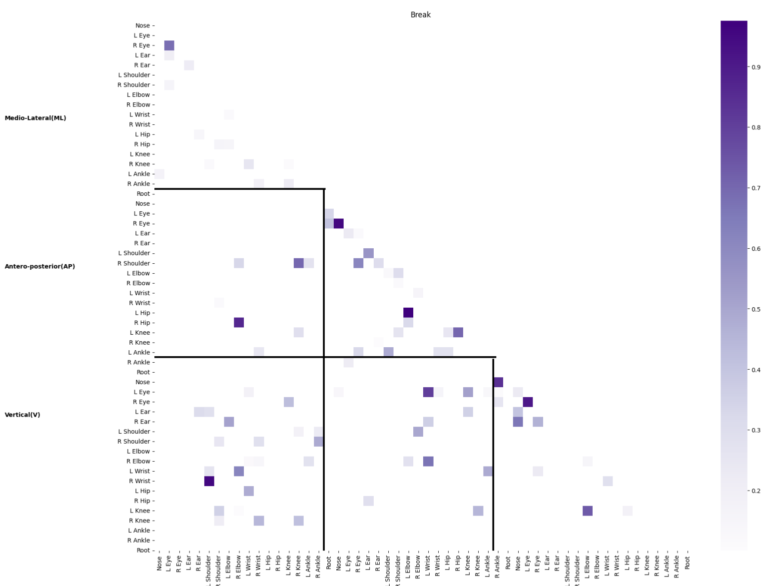
Marker	Importance
L Ear(V)	2.53
L Ear(ML)	2.25
L Eye(V)	1.91
L Shoulder(AP)	1.84
R Elbow(V)	1.84
L Ear(AP)	1.81
R Ear(ML)	1.70
Nose(ML)	1.64
R Shoulder(V)	1.32
R Ankle(AP)	1.29
L Hip(V)	1.26
L Hip(AP)	1.20
L Elbow(V)	1.17
R Eye(AP)	1.08
Nose(AP)	1.08

Figure A.8: Lock



Marker	Importance
R Knee(AP)	4.24
L Knee(V)	3.46
R Hip(AP)	3.30
R Shoulder(AP)	3.04
L Ankle(V)	2.82
L Shoulder(ML)	2.65
L Wrist(ML)	2.44
L Knee(ML)	2.24
R Knee(V)	2.17
R Elbow(AP)	2.14
R Shoulder(ML)	1.74
R Ankle(ML)	1.72
L Wrist(AP)	1.70
L Hip(V)	1.65

Figure A.9: Pop



Marker	Importance
L Ear(V)	3.35
L Elbow(AP)	2.90
R Knee(AP)	2.56
L Ear(AP)	2.52
R Ear(V)	2.48
R Ankle(AP)	2.24
R Eye(V)	2.14
R Knee(ML)	2.10
L Hip(V)	2.00
R Elbow(AP)	2.00
R Shoulder(V)	2.00
R Wrist(AP)	1.99
R Elbow(ML)	1.94
L Knee(AP)	1.92
L Ankle(V)	1.82

Figure A.10: Break

Related Publications

- **Prince Varshney**, Martin Hartmann, Emily Carlson, Vinoo Alluri. "Predicting Perceived Interaction in Dancing Dyads: A Machine Learning Approach." In 10th Annual Conference of Cognitive Science (ACCS). 2023.
- **Prince Varshney**, Martin Hartmann, Emily Carlson, Petri Toiviainen, Vinoo Alluri."Exploring Individuality in Dance: Unveiling Unique Signatures of Dancers in Choreographic and Dyadic Dance Settings." In 46th Cognitive Science Society Annual Conference (COGSCI). 2024.

Bibliography

[1] Y. Agrawal, E. Carlson, P. Toiviainen, and V. Alluri. Decoding Individual differences and musical preference via music-induced movement. *Scientific Reports*, 12(1):2672, Feb. 2022. Number: 1 Publisher: Nature Publishing Group.

[2] I. Ajili, Z. Ramezanpanah, M. Mallem, and J.-Y. Didier. Expressive motions recognition and analysis with learning and statistical methods. *Multimedia Tools and Applications*, 78(12):16575–16600, June 2019.

[3] M. Bangert, T. Peschel, G. Schlaug, M. Rotte, D. Drescher, H. Hinrichs, H.-J. Heinze, and E. Altenmüller. Shared networks for auditory and motor processing in professional pianists: Evidence from fmri conjunction. *NeuroImage*, 30:917–26, 05 2006.

[4] R. M. Baron and L. A. Boudreau. An ecological perspective on integrating personality and social psychology. *Journal of Personality and Social Psychology*, 53(6):1222–1228, 1987. Place: US Publisher: American Psychological Association.

[5] P. Bernardis and M. Gentilucci. Speech and gesture share the same communication system. *Neuropsychologia*, 44:178–90, 02 2006.

[6] F. J. Bernieri and R. Rosenthal. Interpersonal coordination: Behavior matching and interactional synchrony. In *Fundamentals of nonverbal behavior*, Studies in emotion & social interaction, pages 401–432. Cambridge University Press, New York, NY, US, 1991.

[7] B. E. Bläsing and O. Sauzet. Frontiers | My Action, My Self: Recognition of Self-Created but Visually Unfamiliar Dance-Like Actions From Point-Light Displays, 2018.

[8] B. Bryson. ”anything but heavy metal”: Symbolic exclusion and musical dislikes. *American Sociological Review*, 61(5):884–899, 1996.

[9] B. Burger and P. Toiviainen. Mocap toolbox - a matlab toolbox for computational analysis of movement data. 08 2013.

[10] B. Burger and P. Toiviainen. Embodiment in electronic dance music: Effects of musical content and structure on body movement. *Musicae Scientiae*, 24:186–205, 01 2020.

[11] A. Camurri, I. Lagerlöf, and G. Volpe. Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies*, 59(1-2):213–225, 2003. Place: Netherlands Publisher: Elsevier Science.

[12] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, May 2019. arXiv:1812.08008 [cs].

[13] E. Carlson, B. Burger, J. London, M. R. Thompson, and P. Toiviainen. Conscientiousness and Extraversion relate to responsiveness to tempo in dance. *Human Movement Science*, 49:315–325, Oct. 2016.

[14] E. Carlson, B. Burger, and P. Toiviainen. Dance Like Someone is Watching: A Social Relations Model Study of Music-Induced Movement - Emily Carlson, Birgitta Burger, Petri Toiviainen, 2018, 2018.

[15] E. Carlson, B. Burger, and P. Toiviainen. Empathy, Entrainment, and Perceived Interaction in Complex Dyadic Dance Movement. *Music Perception*, 36(4):390–405, Apr. 2019. Publisher: University of California Press.

[16] E. Carlson, P. Saari, B. Burger, and P. Toiviainen. Personality and musical preference using social-tagging in excerpt-selection. *Psychomusicology: Music, Mind, and Brain*, 27(3):203–212, 2017. Place: US Publisher: Educational Publishing Foundation.

[17] E. Carlson, P. Saari, B. Burger, and P. Toiviainen. Dance to your own drum: Identification of musical genre and individual dancer from motion capture using machine learning. *Journal of New Music Research*, 49:1–16, 01 2020.

[18] L. A. S. Chauvigné, A. Walton, M. J. Richardson, and S. Brown. Multi-person and multisensory synchronization during group dancing. *Human Movement Science*, 63:199–208, Feb. 2019.

[19] J. Christensen, C. Cela-Conde, and A. Gomila. Not all about sex: neural and biobehavioral functions of human dance: Neural and biobehavioral functions of human dance. *Annals of the New York Academy of Sciences*, 1400:8–32, 07 2017.

[20] W. S. Condon and W. D. Ogston. SOUND FILM ANALYSIS OF NORMAL AND PATHOLOGICAL BEHAVIOR PATTERNS. *The Journal of Nervous and Mental Disease*, 143(4):338, Oct. 1966.

[21] W. S. Condon and L. W. Sander. Synchrony demonstrated between movements of the neonate and adult speech. *Child Development*, 45(2):456–462, 1974. Place: United Kingdom Publisher: Blackwell Publishing.

[22] J. E. Cutting and L. T. Kozlowski. Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9(5):353–356, May 1977.

[23] J. M. Dabbs Jr. Similarity of gestures and interpersonal influence. *Proceedings of the Annual Convention of the American Psychological Association*, 4(Pt. 1):337–338, 1969. Place: US Publisher: American Psychological Association.

[24] A. T. Dittmann and L. G. Llewellyn. Relationship between vocalizations and head nods as listener responses. *Journal of Personality and Social Psychology*, 9(1):79–84, 1968. Place: US Publisher: American Psychological Association.

[25] A. T. Dittmann and L. G. Llewellyn. Body movement and speech rhythm in social conversation. *Journal of Personality and Social Psychology*, 11(2):98–106, 1969. Place: US Publisher: American Psychological Association.

[26] H. Ergezer and K. Leblebicioglu. Time series classification with feature covariance matrices. *Knowledge and Information Systems*, 55, 06 2018.

[27] T. Etani, A. Marui, S. Kawase, and P. E. Keller. Optimal Tempo for Groove: Its Relation to Directions of Body Movement and Japanese nori. *Frontiers in Psychology*, 9:462, 2018.

[28] W. T. Fitch. Dance, Music, Meter and Groove: A Forgotten Partnership. *Frontiers in Human Neuroscience*, 10, Mar. 2016. Publisher: Frontiers.

[29] K. Fujiwara and I. Daibo. Empathic accuracy and interpersonal coordination: behavior matching can enhance accuracy but interactional synchrony may not. *The Journal of Social Psychology*, 162(1):71–88, Jan. 2022.

[30] R. I. Godøy, E. Haga, and A. R. Jensenius. Playing "air instruments": mimicry of sound-producing gestures by novices and experts. *International Gesture Workshop*, pages 256–267, 2005.

[31] R. I. Godøy, M. Song, K. Nymoen, M. R. Haugen, and A. R. Jensenius. Exploring Sound-Motion Similarity in Musical Experience. *Journal of New Music Research*, 45(3):210–222, July 2016. Publisher: Routledge _eprint: https://doi.org/10.1080/09298215.2016.1184689.

[32] J. A. Grahn. See what I hear? Beat perception in auditory and visual rhythms. *Experimental Brain Research*, 220(1):51–61, July 2012.

[33] G. Guo, S. Li, and K. Chan. Face recognition by support vector machines. *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, pages 196–201, 2000.

[34] K. Guo, P. Ishwar, and J. Konrad. Action recognition in video by covariance matching of silhouette tunnels. pages 299–306, 10 2009.

[35] K. Hachimura, K. Takashina, and M. Yoshimura. Analysis and evaluation of dancing movement based on lma. In *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, pages 294–299, 2005.

[36] M. Hartmann, E. Carlson, A. Mavrolampados, B. Burger, and P. Toiviainen. Postural and Gestural Synchronization, Sequential Imitation, and Mirroring Predict Perceived Coupling of Dancing Dyads. *Cognitive Science*, 47(4):e13281, 2023. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cogs.13281.

[37] M. Hartmann, A. Mavrolampados, E. Allingham, E. Carlson, B. Burger, and P. Toiviainen. Kinematics of perceived dyadic coordination in dance. *Scientific Reports*, 9(1):15594, Oct. 2019. Number: 1 Publisher: Nature Publishing Group.

[38] H. Hicheur, H. Kadone, J. Grèzes, and A. Berthoz. *The Combined Role of Motion-Related Cues and Upper Body Posture for the Expression of Emotions during Human Walking*, pages 71–85. 01 2013.

[39] C. Hufschmidt, B. Weege, S. Röder, K. Pisanski, N. Neave, and B. Fink. Physical strength and gender identification from dance movements. *Personality and Individual Differences*, 76:13–17, Apr. 2015.

[40] G. Johansson. Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(2):201–211, June 1973.

[41] T. Kato, E. Takahashi, K. Sawada, N. Kobayashi, T. Watanabe, and T. Ishii. A computer analysis of infant movements synchronized with adult speech. *Pediatric Research*, 17(8):625–628, Aug. 1983.

[42] R. M. Krauss, Y. Chen, and P. Chawla. Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In *Advances in experimental social psychology, Vol. 28*, pages 389–450. Academic Press, San Diego, CA, US, 1996.

[43] M. LaFrance. Posture mirroring and rapport. *Interaction rhythms: Periodicity in communicative behavior*, 279298, 1982.

[44] M. Lafrance and M. Broadbent. Group Rapport: Posture Sharing as a Nonverbal Indicator. *Group & Organization Studies*, 1(3):328–333, Sept. 1976.

[45] A. Lahav, E. Saltzman, and G. Schlaug. Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27:308–14, 02 2007.

[46] H. Lee, J. Launay, and L. Stewart. Signals through music and dance: Perceived social bonds and formidability on collective movement. *Acta Psychologica*, 208:103093, July 2020.

[47] M. Leman. Embodied Music Cognition and Mediation Technology, 2008.

[48] M. Leman and R. Godøy. Why study musical gestures? pages 3–11, 01 2010.

[49] J. C. Lena and R. A. Peterson. Classification as culture: Types and trajectories of music genres. *American Sociological Review*, 73(5):697–718, 2008. Place: US Publisher: Sage Publications.

[50] M. Lesaffre, L. D. Voogdt, M. Leman, B. D. Baets, H. D. Meyer, and J.-P. Martens. How potential users of music search and retrieval systems describe the semantic quality of music. *Journal of the American Society for Information Science and Technology*, 59(5):695–707, 2008. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.20731.

[51] R. Li, S. Yan g, D. A. Ross, and A. Kanazawa. AI Choreographer: Music Conditioned 3D Dance Generation With AIST++. pages 13401–13412, 2021.

[52] G. Luck, S. Saarikallio, B. Burger, M. R. Thompson, and P. Toiviainen. Effects of the Big Five and musical genre on music-induced movement. *Journal of Research in Personality*, 44(6):714–720, Dec. 2010.

[53] A. Mammone, M. Turchi, and N. Cristianini. Support vector machines. *Wiley Interdisciplinary Reviews: Computational Statistics*, 1:283 – 289, 11 2009.

[54] B. Merker. Groove or swing as distributed rhythmic consonance: introducing the groove matrix. *Frontiers in Human Neuroscience*, 8, 2014.

[55] J. Michalak, N. Troje, J. Fischer, P. Vollmar, T. Heidenreich, and D. Schulte. Embodiment of sadness and depression—gait patterns associated with dysphoric mood. *Psychosomatic medicine*, 71:580–7, 06 2009.

[56] T. Monaghan. Why Study the Lindy Hop? - University of Guelph, 2001.

[57] N. Nakano, T. Sakura, K. Ueda, L. Omura, A. Kimura, Y. Iino, S. Fukashiro, and S. Yoshioka. Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. *Frontiers in Sports and Active Living*, 2:50, 2020.

[58] D. Newtson, J. Hairfield, J. Bloomingdale, and S. Cutino. The structure of action and interaction. *Social Cognition*, 5(3):191–237, 1987. Place: US Publisher: Guilford Publications.

[59] A. Ng. Feature selection, l 1 vs. l 2 regularization, and rotational invariance. *Proceedings of the Twenty-First International Conference on Machine Learning*, 09 2004.

[60] J. Phillips-Silver and P. Keller. Searching for Roots of Entrainment and Joint Action in Early Musical Interactions. *Frontiers in Human Neuroscience*, 6, 2012.

[61] B. Repp and Y.-H. Su. Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic bulletin & review*, 20, 02 2013.

[62] B. H. Repp. Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, 12(6):969–992, Dec. 2005.

[63] G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi. Premotor cortex and the recognition of motor actions. *Brain Research. Cognitive Brain Research*, 3(2):131–141, Mar. 1996.

[64] M. R. Ronchi and P. Perona. Benchmarking and error diagnosis in multi-instance pose estimation. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 369–378, 2017.

[65] A. P. Royce. *The Anthropology of Dance*. 01 2002.

[66] M. Sawada, K. Suda, and M. Ishii. Expression of emotions in dance: Relation between arm movement characteristics and emotion. *Perceptual and motor skills*, 97:697–708, 01 2004.

[67] T. Schubert and S. Koole. The embodied self: Making a fist enhances men's power-related self-conceptions. *Journal of Experimental Social Psychology*, 45:828–834, 07 2009.

[68] V. Sevdalis and P. E. Keller. Self-recognition in the Perception of Actions Performed in Synchrony with Music. *Annals of the New York Academy of Sciences*, 1169(1):499–502, 2009. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1749-6632.2009.04773.x.

[69] L. Shapiro. The Embodied Cognition Research Programme - Shapiro - 2007 - Philosophy Compass - Wiley Online Library, 2007.

[70] L. Siyao, W. Yu, T. Gu, C. Lin, Q. Wang, C. Qian, C. C. Loy, and Z. Liu. Bailando: 3D Dance Generation by Actor-Critic GPT With Choreographic Memory. pages 11050–11059, 2022.

[71] J. I. Skipper, S. Goldin-Meadow, H. C. Nusbaum, and S. L. Small. Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language*, 101(3):260–277, June 2007.

[72] D. Snell and D. Hodgetts. Heavy metal, identity and the social negotiation of a community of practice. *Journal of Community & Applied Social Psychology*, 17, 11 2007.

[73] R. Solberg and A. Jensenius. Pleasurable and intersubjectively embodied experiences of electronic dance music. *Empirical Musicology Review*, 11:301, 04 2017.

[74] F. Strack, L. Martin, and S. Stepper. Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, 54:768–777, 01 1988.

[75] W. Straw. Characterizing Rock Music Cultures: The Case of Heavy Metal. *Canadian University Music Review / Revue de musique des universités canadiennes*, (5):104–122, 1984. Publisher: Canadian University Music Society / Société de musique des universités canadiennes.

[76] J. Stupacher, M. J. Hove, and P. Janata. Audio Features Underlying Perceived Groove and Sensorimotor Synchronization in Music. *Music Perception*, 33(5):571–589, June 2016. Publisher: University of California Press.

[77] P. Toiviainen and E. Carlson. Embodied Meter Revisited: Entrainment, Musical Content, and Genre in Music-Induced Movement. *Music Perception*, 39(3):249–267, Feb. 2022.

[78] P. Toiviainen and M. Hartmann. Analyzing multidimensional movement interaction with generalized cross-wavelet transform. *Human Movement Science*, 81:102894, Feb. 2022.

[79] P. Toiviainen, G. Luck, and T. Marc. Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception: An Interdisciplinary Journal*, 28, 09 2010.

[80] H. Tommi and T. Marc R. Learning and Synchronising Dance Movements in South African Songs – Cross-cultural Motion-capture Study. *Dance Research*, 29(supplement):305–328, Nov. 2011. Publisher: Edinburgh University Press.

[81] N. F. Troje, C. Westhoff, and M. Lavrov. Person identification from biological motion: effects of structural and kinematic cues. *Perception & Psychophysics*, 67(4):667–675, May 2005.

[82] J. Tseng, R. Castellon, and K. Liu. EDGE: Editable Dance Generation From Music. pages 448–458, 2023.

[83] S. Tsuchida, S. Fukayama, M. Hamasaki, and M. Goto. AIST DANCE VIDEO DATABASE: MULTI-GENRE, MULTI-DANCER, AND MULTI-CAMERA DATABASE FOR DANCE INFORMATION PROCESSING. 2019.

[84] O. Tuzel, F. Porikli, and P. Meer. Pedestrian detection via classification on riemannian manifolds. *IEEE transactions on pattern analysis and machine intelligence*, 30:1713–27, 11 2008.

[85] E. Van Dyck, P.-J. Maes, J. Hargreaves, M. Lesaffre, and M. Leman. Expressing induced emotions through free dance movement. *Journal of Nonverbal Behavior*, 37(3):175–190, 2013. Place: Germany Publisher: Springer.

[86] C. Westhoff and N. F. Troje. Kinematic cues for person identification from biological motion. *Perception & Psychophysics*, 69(2):241–253, Feb. 2007.

[87] A. Wilson and S. Golonka. Embodied Cognition is Not What you Think it is. *Frontiers in Psychology*, 4, 2013.

[88] M. A. G. Witek, E. F. Clarke, M. Wallentin, M. L. Kringelbach, and P. Vuust. Syncopation, Body-Movement and Pleasure in Groove Music. *PLOS ONE*, 9(4):e94446, Apr. 2014. Publisher: Public Library of Science.

[89] J. Zhu, S. Rosset, T. Hastie, and R. Tibshirani. Support vector machines. 01 2003.