Title
Exploring Individuality in Dance: Unveiling Unique Signatures of Dancers in Choreographic and Dyadic Dance Settings

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Varshney, Prince
Hartmann, Martin
Carlson, Emily
et al.

Publication Date
2024

Peer reviewed

Exploring Individuality in Dance: Unveiling Unique Signatures of Dancers in Choreographic and Dyadic Dance Settings

Prince Varshney (prince.varshney@research.iiit.ac.in)
Vinoo Alluri (vinoo.alluri@iiit.ac.in)
Cognitive Science Lab, IIIT Hyderabad, India

Martin Hartmann (martin.hartmann@jyu.fi)
Emily Carlson (emjocarl@jyu.fi)
Petri Toiviainen (ptoiviai@jyu.fi)
University of Jyväskylä, Jyväskylä, Finland

Abstract

How we move and interact with our surroundings can reveal a lot about us as an individual. This study delves into the interplay of music, movement, and individual identity within the framework of embodied cognition. Drawing inspiration from Carlson et al. (2020)'s work, which showcased remarkably high accuracy in identifying individuals based on free-form dance movements in a marker-based setting, our investigation extends their work into two novel contexts: markerless-choreographic and marker-based-dyadic dance settings. In the choreographic setting, professional dancers perform identical routines in a markerless setting. In the dyadic setting, individuals danced with a partner. We found that the dancer identification accuracy was at least two times better than the chance level in the choreographic setting and notably high accuracy in the dyadic setting. These results showcase the robustness of Carlson et al. (2020)'s method in generalizing to new settings and the presence of motoric fingerprints in choreographic as well as dyadic settings.

Introduction

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. In a study by Lesaffre et al. (2008), approximately 95% of participants reported engaging in spontaneous movements in response to the music they heard. 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. (2016) 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.(Shapiro, 2007; Wilson and Golonka, 2013). Embodied music cognition, as articulated by Leman (2008), 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.

Music-induced movements embody rich information about gender (Hufschmidt et al., 2015), mood, emotion, personality (Camurri et al., 2003; Luck et al., 2010; Van Dyck et al., 2013; Carlson et al., 2016; Carlson et al., 2018), and culture (Tommi and Marc R., 2011). Machine learning methods have also been used to predict gender and personality traits from free-form dance movements (Agrawal et al., 2022). Therefore, it is reasonable to expect different individuals to move differently to the same music stimulus. Johansson (1973)'s 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 (1977)'s study illustrated that friends could identify each other based on point-light displays of their gait. Troje et al. (2005) 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 (2007) 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(Sevdalis and Keller, 2009; Bläsing and Sauzet, 2018). Carlson et al. (2020) 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.

All the above-mentioned studies have used naturalistic free-form movements as a medium for dancer identification. These movements could be influenced by an individual's anatomy, diverse dancing training, and other factors outlined previously. This leads to an intriguing question: *To what extent can we identify individuals based on their movements*

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

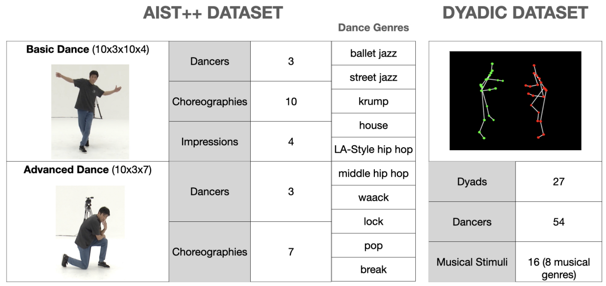

Figure 1: Summary of both datasets

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. In addition to this, these 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 a noteworthy library capable of 2D/3D pose estimation (Cao et al., 2019). 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. (2020) 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. (2020)'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.

Dance is a social activity and, therefore, often occurs in groups. In particular, we look at dyadic dance, which is a good example of rhythmic social entrainment. In contrast to solo dancing, where movements are primarily influenced by the music, in dyadic settings, dance movements are not just influenced by the auditory cues but also the visual information derived from the partner. We have evidence that dyadic dancing is different from dancing alone. Carlson et al. (2018) found that individuals in 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 that the same individual can move differently when dancing with different partners, and this happens particularly in individuals with high self-reported empathy scores. Interpersonal coordination has also been studied in dyadic dancing with two perceptual variables: interaction and similarity. Hartmann et al. (2019) showed that high interaction is linked with how closely the dyads are horizontally oriented toward each other, referred to as torso orientation. In the subsequent study, Hartmann et al. (2023) studied mirroring, sequential, and simultaneous coupling in the dyadic context. We can conclude from the above-mentioned studies that participants tend to synchronize their movements with their partners and sometimes also change their movement patterns depending on the 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.

In summary, this work endeavors to assess the external validity, a critical yet often overlooked aspect, of Carlson's study, particularly within markerless-choreographic and marker-dyadic settings. By extending the investigation into these contexts, the aim is to evaluate the robustness, broader applicability, and generalizability of Carlson's findings.

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

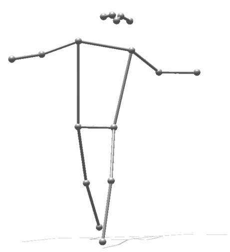

Figure 2: 17 key-point's locations in COCO-format

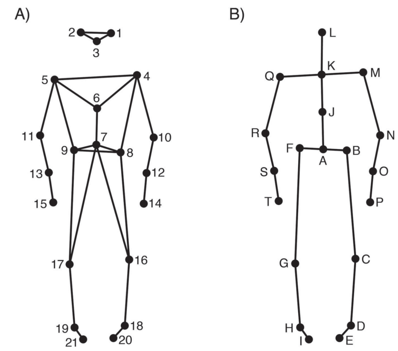

Figure 3: A) Marker locations B) Transformed joint locations

Methods

Datasets

AIST++ We employed the AIST++ dataset from Li et al. (2021) for the choreographic-markless setting. AIST++ is a large-scale 3D dance motion dataset generated from the AIST dance database (Tsuchida et al., 2019). AIST is just a collection of dance videos in 9 camera angles without any 3D information. AIST++ provides 17 COCO-format (Ronchi and Perona, 2017) 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 2.

The subjects of the study were 40 professional dancers (15 females) with more than five years of experience. 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. Each dancer specializes in a particular dance genre. Notably, the choreography undertaken by dancers in one genre differs from that of dancers performing in other genres.

The dataset 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.

Dyadic Carlson et al. (2020)'s dataset was utilized for the dyadic setting. 73 participants (54 females) aged 19–40 years (M = 25.75, SD = 4.72) were recruited for the motion capture study. The participants, hailing from 24 different nationalities and possessing diverse musical and dance training backgrounds.

The motion capture was conducted using a twelve-camera optical system (Qualisys Oqus 5+), tracking 21 reflective body markers in three-dimension of the subject at a frame rate of 120 Hz. Marker locations are represented in Figure 3(A).

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. 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. Please refer to Hartmann et al. (2019) for torso orientation computation. 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.

During the recording, participants were instructed to move freely, either individually or in dyads, in response to musical stimuli, simulating a dance club or party setting. The stimuli covered 8 genres: Dance, Blues, Country, Metal, Jazz, Reggae, Pop, and Rap. There were 2 stimuli per genre. These stimuli represents the selected genres accurately and were selected based on social tagging data. For a detailed account of the selection process, please refer to Carlson et al., 2017.

Pre-processing

The Motion Capture (MoCap) Toolbox (Burger and Toiviainen, 2013) was employed for data pre-processing in MATLAB. In the case of the dyadic dataset, 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 3(B) illustrates the locations of these 20 joints. For a detailed understanding of the transformation process, please refer to Carlson et al. (2020).

In the AIST++ dataset, each recording was 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 MoCap Toolbox. Gaps were filled linearly, and an additional step of smoothing was carried out using a Butterworth

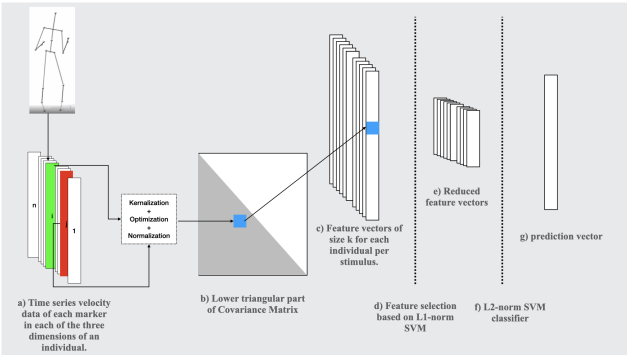

Figure 4: Machine Learning pipeline

smoothing filter (2nd order; 12 Hz cutoff frequency) for all markers in all three dimensions.

For both datasets, 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 (2013). Subsequently, for the dyadic data, 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 3(B) is the origin, and the line joining the hip markers defines the mediolateral axis.

Machine learning analysis

Support Vector Machine (SVM) algorithm was used for the dancer identification task. SVM works by identifying the line that optimally segregates data points into distinct classes such that future data points are classified correctly. 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 (G. Guo et al., 2000; Mammone et al., 2009). 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. We rigorously adhered to the machine-learning pipeline outlined in Carlson et al. (2020), which is illustrated in the Figure 4.

Feature Extraction Based on the findings by Troje et al. (2005) 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. Such Covariance-based features have been used in various applications, including time series classification (Ergezer and Leblebicioglu, 2018), action recognition (K. Guo et al., 2009), pedestrian detection (Tuzel et al., 2008), and the prediction of individual characteristics such as gender and personality (Agrawal et al., 2022). Carlson et al. (2020) found that using correntropy or Radial Basis Function(RBF) kernel as a covariance measure yielded much higher accuracy than linear covariance. Hence, only correntropy is used in this study. Correntropy between any two-time series x_i and x_j was computed as (Liu et al., 2007):

$$K(x_i, x_j) = e^{\frac{-\left\|x_i - x_j\right\|_2^2}{2\sigma^2 T^2}} \tag{1}$$

$\left\|x_i - x_j\right\|_2$ denotes the L2 norm; T denotes the number of frames employed to accommodate samples of varying lengths. The parameter sigma governs the steepness of the distribution of the generated features, with higher sigma 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 sigma 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.

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. In the dyads dataset, this vector has a length of 1770, given 20 markers in three dimensions. For the AIST++ dataset, which features 17 markers in three dimensions, the resulting feature vector has a length of 1275.

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 for feature selection (Zhu et al., 2003). This approach results in many feature weights learned by the model having almost zero values, facilitating the filtering of irrelevant features. The L1 norm SVM is better suited at feature selection than the L2 norm and it also avoids overfitting (Ng, 2004). 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 the entire dataset on features iteratively introduced in decreasing order of importance. The number of features to be retained becomes a hyperparameter. 100 features were selected for the dyadic dataset, and 200 features were selected for AIST++ dataset.

Classification Linear SVM with an L2 penalty 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, parameter 'C' of SVM in our case. This approach is superior to the single cross-validation used by Carlson et al. (2020), as it utilizes only a portion of the dataset provided by the outer cross-validation, reducing the risk of overfitting the entire dataset. 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. For the AIST++ dataset, 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.

Results

AIST++

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 1.

Dyadic

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 test set is depicted in Table 2.

Table 1: Cross-validation dancer identification accuracy for each dance genre (AIST++)

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 2: 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

Discussion

Inspired by the findings of Carlson et al. (2020), 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 novel contexts. We initially ventured into a choreographic setting (AIST++ dataset)—a more restrictive environment compared to free-form dance—anticipating greater challenges in distinguishing individuals when everyone adheres to the same routine. In an effort to overcome the limitations associated with marker-based systems, we opted for markerless data acquisition in this setting. Subsequently, we extended our investigation to the dyadic setting, introducing an additional layer of complexity over dancing individually as individuals now synchronize their movements with a partner.

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. Recognizing these complexities, we adopted a two-step approach to dancer identification. First, we performed genre classification using the same machine learning pipeline, achieving

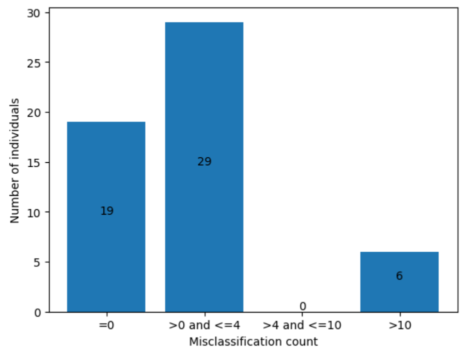

Figure 5: Number of individuals vs Misclassification count when utilizing features from individual dance settings to predict dancers using the dyadic model

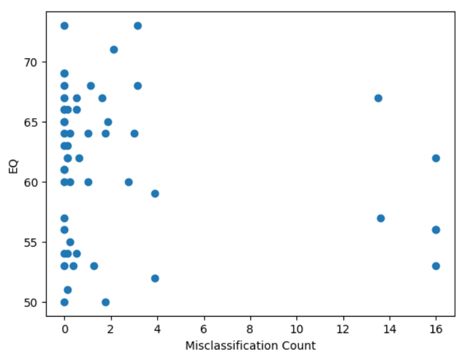

Figure 6: Empathy vs Misclassification count

an impressive 88.25% accuracy (against a 10% chance level). Then, 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.

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. (2020) 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. (2020) 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. (2020). 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 (Snell and Hodgetts, 2007; Bryson, 1996; Straw, 1984), and moves like Charleston and swing are common in the Jazz subculture (Lena and Peterson, 2008; Monaghan, 2001). The influence of distinctive movement stereotypes within these subcultures may contribute to a higher degree of movement similarity and potentially impact identification accuracy.

Feature importance analysis conducted on dyadic settings and across all dance genres within the AIST++ dataset reveals a consistent pattern. 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. This aligns with the findings of Carlson et al. (2020), providing a consistent narrative across studies. Please refer to the appendix for the detailed visualization of important key-points/joints pairs and key-points/joints.

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. Remarkably, we achieved an accuracy of 83.23% using the important features derived from the dyadic model for the prediction task. Our subsequent misclassification analysis revealed intriguing patterns (refer to Figure 5). Notably, certain individuals were accurately predicted by the dyad model without any errors, while others remained unpredicted by the dyad model altogether. This suggests a variance in how individuals express their movements between solo and partnered settings. While initial suspicions pointed towards empathy as a potential factor, plotting Empathizing Quotient (EQ) scores against misclassification count (see Figure 6) did not reveal a discernible pattern. This finding needs further warrant and investigation. Future research should explore how the interaction between personality traits, the empathy levels of individuals in dyads, and the music genre may collectively contribute to these observations.

References

Agrawal, Y., Carlson, E., Toiviainen, P., & Alluri, V. (2022). Decoding Individual differences and musical preference via music-induced movement [Number: 1 Publisher: Nature Publishing Group]. *Scientific Reports*, *12*(1), 2672. https://doi.org/10.1038/s41598-022-06466-3

Bläsing, B. E., & Sauzet, O. (2018). Frontiers — My Action, My Self: Recognition of Self-Created but Visually Unfamiliar Dance-Like Actions From Point-Light Displays. Retrieved January 31, 2024, from https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01909/full

Bryson, B. (1996). "Anything But Heavy Metal": Symbolic Exclusion and Musical Dislikes. *American Sociological Review*, *61*(5), 884–899. Retrieved January 31, 2024, from http://www.jstor.org/stable/2096459

Burger, B., & Toiviainen, P. (2013). MoCap Toolbox - A Matlab toolbox for computational analysis of movement data. *Proceedings of the 10th Sound and Music Computing Conference, Stockholm, Sweden*.

Camurri, A., Lagerlöf, I., & Volpe, G. (2003). Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques [Place: Netherlands Publisher: Elsevier Science]. *International Journal of Human-Computer Studies*, *59*(1-2), 213–225. https://doi.org/10.1016/S1071-5819(03)00050-8

Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2019, May). OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields [arXiv:1812.08008 [cs]]. https://doi.org/10.48550/arXiv.1812.08008

Carlson, E., Burger, B., London, J., Thompson, M. R., & Toiviainen, P. (2016). Conscientiousness and Extraversion relate to responsiveness to tempo in dance. *Human Movement Science*, *49*, 315–325. https://doi.org/10.1016/j.humov.2016.08.006

Carlson, E., Burger, B., & Toiviainen, P. (2018). Dance Like Someone is Watching: A Social Relations Model Study of Music-Induced Movement - Emily Carlson, Birgitta Burger, Petri Toiviainen, 2018. Retrieved January 31, 2024, from https://journals.sagepub.com/doi/10.1177/2059204318807846

Carlson, E., Saari, P., Burger, B., & Toiviainen, P. (2017). Personality and musical preference using social-tagging in excerpt-selection [Place: US Publisher: Educational Publishing Foundation]. *Psychomusicology: Music, Mind, and Brain*, *27*(3), 203–212. https://doi.org/10.1037/pmu0000183

Carlson, E., Saari, P., Burger, B., & Toiviainen, P. (2020). 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. https://doi.org/10.1080/09298215.2020.1711778

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, *9*(5), 353–356. https://doi.org/10.3758/BF03337021

Ergezer, H., & Leblebicioglu, K. (2018). Time series classification with feature covariance matrices. *Knowledge and Information Systems*, *55*. https://doi.org/10.1007/s10115-017-1098-1

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

Guo, G., Li, S., & Chan, K. (2000). Face recognition by support vector machines. *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, 196–201. https://api.semanticscholar.org/CorpusID:14555654

Guo, K., Ishwar, P., & Konrad, J. (2009). Action Recognition in Video by Covariance Matching of Silhouette Tunnels, 299–306. https://doi.org/10.1109/SIBGRAPI.2009.29

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

Hartmann, M., Mavrolampados, A., Allingham, E., Carlson, E., Burger, B., & Toiviainen, P. (2019). Kinematics of perceived dyadic coordination in dance [Number: 1 Publisher: Nature Publishing Group]. *Scientific Reports*, *9*(1), 15594. https://doi.org/10.1038/s41598-019-52097-6

Hufschmidt, C., Weege, B., Röder, S., Pisanski, K., Neave, N., & Fink, B. (2015). Physical strength and gender identification from dance movements. *Personality and Individual Differences*, *76*, 13–17. https://doi.org/10.1016/j.paid.2014.11.045

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*(2), 201–211. https://doi.org/10.3758/BF03212378

Leman, M. (2008). Embodied Music Cognition and Mediation Technology. Retrieved January 31, 2024, from https://mitpress.mit.edu/9780262122931/embodied-music-cognition-and-mediation-technology/

Lena, J. C., & Peterson, R. A. (2008). Classification as culture: Types and trajectories of music genres [Place: US Publisher: Sage Publications]. *American Sociological Review*, *73*(5), 697–718. https://doi.org/10.1177/000312240807300501

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

Li, R., Yan g, S., Ross, D. A., & Kanazawa, A. (2021). AI Choreographer: Music Conditioned 3D Dance Generation With AIST++, 13401–13412. Retrieved August 1, 2023, from https://openaccess.thecvf.com/content/ICCV2021/html/Li_AI_Choreographer_Music_Conditioned_3D_Dance_Generation_With_AIST_ICCV_2021_paper.html

Luck, G., Saarikallio, S., Burger, B., Thompson, M. R., & Toiviainen, P. (2010). Effects of the Big Five and musical genre on music-induced movement. *Journal of Research in Personality*, *44*(6), 714–720. https://doi.org/10.1016/j.jrp.2010.10.001

Mammone, A., Turchi, M., & Cristianini, N. (2009). Support vector machines. *Wiley Interdisciplinary Reviews: Computational Statistics*, *1*, 283–289. https://doi.org/10.1002/wics.49

Monaghan, T. (2001). Why Study the Lindy Hop? - University of Guelph. Retrieved January 31, 2024, from https://ocul-gue.primo.exlibrisgroup.com

Nakano, N., Sakura, T., Ueda, K., Omura, L., Kimura, A., Iino, Y., Fukashiro, S., & Yoshioka, S. (2020). Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. *Frontiers in Sports and Active Living*, *2*, 50. https://doi.org/10.3389/fspor.2020.00050

Ng, A. (2004). Feature selection, L 1 vs. L 2 regularization, and rotational invariance. *Proceedings of the Twenty-First International Conference on Machine Learning*. https://doi.org/10.1145/1015330.1015435

Ronchi, M. R., & Perona, P. (2017). Benchmarking and Error Diagnosis in Multi-instance Pose Estimation. *2017 IEEE International Conference on Computer Vision (ICCV)*, 369–378. https://doi.org/10.1109/ICCV.2017.48

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

Shapiro, L. (2007). The Embodied Cognition Research Programme - Shapiro - 2007 - Philosophy Compass - Wiley Online Library. Retrieved January 31, 2024, from https://compass.onlinelibrary.wiley.com/doi/10.1111/j.1747-9991.2007.00064.x

Snell, D., & Hodgetts, D. (2007). Heavy Metal, identity and the social negotiation of a community of practice. *Journal of Community & Applied Social Psychology*, *17*. https://doi.org/10.1002/casp.943

Straw, W. (1984). Characterizing Rock Music Cultures: The Case of Heavy Metal [Publisher: Canadian University Music Society / Société de musique des universités canadiennes]. *Canadian University Music Review / Revue de musique des universités canadiennes*, (5), 104–122. https://doi.org/10.7202/1013933ar

Tommi, H., & Marc R., T. (2011). Learning and Synchronising Dance Movements in South African Songs – Cross-cultural Motion-capture Study [Publisher: Edinburgh University Press]. *Dance Research*, *29*(supplement), 305–328. https://doi.org/10.3366/drs.2011.0022

Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: effects of structural and kinematic cues. *Perception & Psychophysics*, *67*(4), 667–675. https://doi.org/10.3758/bf03193523

Tsuchida, S., Fukayama, S., Hamasaki, M., & Goto, M. (2019). AIST DANCE VIDEO DATABASE: MULTI-GENRE, MULTI-DANCER, AND MULTI-CAMERA DATABASE FOR DANCE INFORMATION PROCESSING.

Tuzel, O., Porikli, F., & Meer, P. (2008). Pedestrian Detection via Classification on Riemannian Manifolds. *IEEE transactions on pattern analysis and machine intelligence*, *30*, 1713–27. https://doi.org/10.1109/TPAMI.2008.75

Van Dyck, E., Maes, P.-J., Hargreaves, J., Lesaffre, M., & Leman, M. (2013). Expressing induced emotions through free dance movement [Place: Germany Publisher: Springer]. *Journal of Nonverbal Behavior*, *37*(3), 175–190. https://doi.org/10.1007/s10919-013-0153-1

Westhoff, C., & Troje, N. F. (2007). Kinematic cues for person identification from biological motion. *Perception & Psychophysics*, *69*(2), 241–253. https://doi.org/10.3758/bf03193746

Wilson, A., & Golonka, S. (2013). Embodied Cognition is Not What you Think it is. *Frontiers in Psychology*, *4*. Retrieved January 31, 2024, from https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00058

Zhu, J., Rosset, S., Hastie, T., & Tibshirani, R. (2003). Support Vector Machines. https://doi.org/10.1007/978-1-4614-7138-7_9