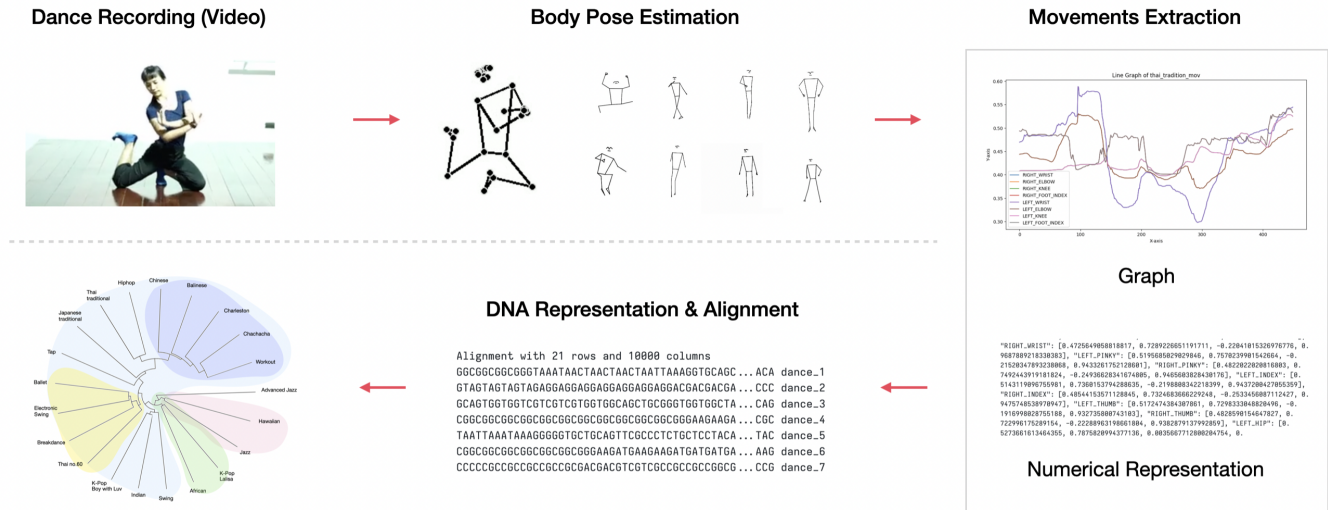# Phylogenetic Tree of Dance

Pat Pataranutaporn
MIT Media Lab
Cambridge, MA
patpat@media.mit.edu

Chayapatr Archiwaranguprok
Creatorgarten
Bangkok, Thailand
pub@creatorsgarten.org

Pichet Klunchun
Pichet Klunchun Dance Company
Bangkok, Thailand
pichetklunchun@gmail.com

Figure 1: A pipeline for the Phylogenetic Tree Construction of Dance. Movement from each dance recording is extracted and covered to intermediate representation as "motion DNA," which is then aligned. The distance matrix is calculated to construct a phylogenetic tree construction using UPGMA.

ABSTRACT

As motion capture technology enables the digitization and precise quantification of dance movement vocabularies across cultures, we propose extending phylogenetic analysis to visually map evolutionary relationships between various cultural dance lineages based on their digitized kinetic "genomic" data. This computational phylogenetics approach derived from bioinformatics can unveil intriguing genealogical patterns linking global cultural dance traditions, revealing surprising connections of ancestry and inheritance through the ages. Beyond analyzing superficial visual similarities, classifying intrinsic embodied knowledge encoded within the dynamics of traditional movement styles offers windows into deep structures reflecting shared cultural histories. Techniques adapted from genetic sequence alignment and phylogenetic tree construction promise new transcultural perspectives that may challenge, support, or complicate our assumptions regarding development, migration, convergence, divergence and other forces shaping the intricate kinships binding humanity's diverse dance heritage. By computationally excavating these lively artifacts of intangible heritage, this research aims to cultivate appreciation for the flowing interconnectedness of cultural evolution and spotlight possibilities for respectful computational co-creation that propels endangered traditions into the future.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Systems and tools for interaction design**; **Collaborative and social computing systems and tools**; • **Applied computing** → **Performing arts**; **Media arts**; **Fine arts**.

KEYWORDS

Phylogenetic Tree, Dance, Taxonomy, Motion Capture

1 INTRODUCTION

From ritual origins echoing prehistory to the stage of modern pop culture, dance represents a vibrant form of intangible cultural heritage, with traditional movement vocabularies encoding deep layers of embodied knowledge reflecting millennia of development. Recent innovations in motion capture technology have enabled the digitization and precise quantification of the intricacies underlying diverse global dance lineages. Motion capture (mocap) has shown tremendous versatile applications across diverse domains, from sports [1, 16], medical science [10, 12], to performing arts. In dance choreography [9, 14, 18, 19, 19, 21], mocap enables quantitative recording and examination of dance movements to study technique styles and variability. Tracking technologies, including optical markers, depth sensors, and computer vision, process the limb positions of dancers to obtain precise positional information.

However, suitable computational procedures are necessitated to derive meaningful observations from such complex multi-dimensional motion data.

One possibility for utilizing mocap in motion computing is regarding the classification of dances. Inspired by phylogenetic analysis in evolutionary biology, which constructs tree diagrams reflecting the molecular trajectory of living species One noteworthy observation that can take advantage of mocap in motion computing is regarding the classification of dances. Inspired by phylogenetic analysis in evolutionary biology, which constructs tree diagrams reflecting the molecular trajectory of living species evolution [23]. We propose adapting computational phylogenetics to build a "dance phylogenetic tree" taxonomizing choreographic styles. This presents dance lineages based on quantifiable feature vectors extracted from mocap tracking choreographic motifs.

Computational phylogenetics can reveal surprising developmental patterns and forces like migration, adaptation, convergence and divergence that shaped the flourishing diversity of dance. Appreciating these cultural evolutionary kinships fosters respect and possibilities for digitally revitalizing endangered traditions. Overall this technique promises fresh transcultural insight into the flowing interconnectedness of global dance inheritance.

Overall, this paper introduces a pipeline to generate phylogenetic trees of dances over history. The aim is not to reconstruct an accurate evolutionary tree based solely on documented evidence. Instead, it provides a medium to potentially uncover previously overlooked connections across dances that transcend the origin of genres. Challenges still need to be addressed regarding feature conceptualization and scalability.

2 RELATED WORKS

2.1 Dance Classification

Works on dance classification can be broadly organized into two main types: the context of the dances and the movement.

Works on contextual classification commonly use discrete factors to derive the taxonomy; Bisig [3] presents the taxonomy of generative dance based on six topics revolving around the context of dances, i.e., Domain, Contribution, Manifestation, Autonomy, Representation, and Process. Peng et al. [17] categorized robotic dance in social robotics into four categories: cooperative human-robot dance, imitation of human dance motions, synchronization for music, and creation of robotic choreography.

Meanwhile, the movement classification of dance typically uses some feature-extracting function to derive the classification from the formatted representative data (for example, vectors from the mocap dataset or symbolic notation) and their metadata [4, 6]. Notably, A. Aristidou et al. [1] semantically classified dance data by creating a bag-of-motif representation of dances and using quartet-based analysis to organize dance data into a categorization tree while using the metadata to set the parent-child relationship.

In this paper, we focus on the dance's movement features to transcend the original context and observe the similarities in underlying patterns.

2.2 Phylogenetic Tree

Understanding the evolutionary relationships between species is essential for many areas of biological research. A representation of the evolution such as phylogenetic tree is crucial to make sense of major transitions during the process, like the development of new body structures or metabolic pathways [13, 15]. Although the desire to reconstruct the evolutionary relationships between all living things traces back to early attempts at classification by Aristotle, Linnaeus and others, this endeavor began in earnest in the 19th century with the advent of Darwinian theory. While early phylogenetic trees relied heavily on morphological traits, biomolecules like nucleic acids and proteins have since become a far richer source of information for tree-building [7]. Since DNA sequencing allowed molecular data to be used for phylogenetic analyses, our understanding of the tree of life has been radically transformed [13].

Outside of biology, phylogenetic tree construction has been adapted to represent the evolution and lineages of various human-generated systems, finding applications in diverse domains beyond modeling organic life forms.

For example, work by Shu et al. [20] has focused on identifying problems and introducing refinements for building language phylogenies in linguistics, indicating the existing usage of phylogenetic approaches in this field. In the area of motion computing, Chen et al. [5] employed a quartet-based analysis technique presented by Huang et al. [11] to produce a phylogenetic tree for the relational organization of poses in dance dataset.

Phylogenetic tree construction algorithms vary in their principles and methodologies. In this paper, we utilized the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) as our algorithm of choice. Hierarchical clustering assigns branches to join pairs of sequences or entities based on their pairwise distances, creating a tree structure. The simplicity of this approach provides a solid foundation for our study exploration [22].

3 METHODOLOGY

To construct the dance phylogeny, we develop a pipeline to extract the movement data from recordings using markerless motion capture, convert the data vectors into an intermediate "genotypes" format, and utilize phylogenetic tree construction algorithms to formulate their evolutionary relationships. Details of each process in the pipeline are described below.

3.1 Dataset

The demonstrative dataset is created from 21 dance videos from online sources, with each video representing a different dance genre. While this dataset provides an initial sampling of some popular dance styles, it only contains a single exemplar video for each genre. For more comprehensive analysis, future iterations of this dataset should aim to incorporate multiple example videos per genre. Expanding to include a wider diversity of cultural dance traditions would also make this a more inclusive and representative dataset. Building up the dataset to contain several examples of each style, especially those from underrepresented regions, should be a priority as this research continues to progress.
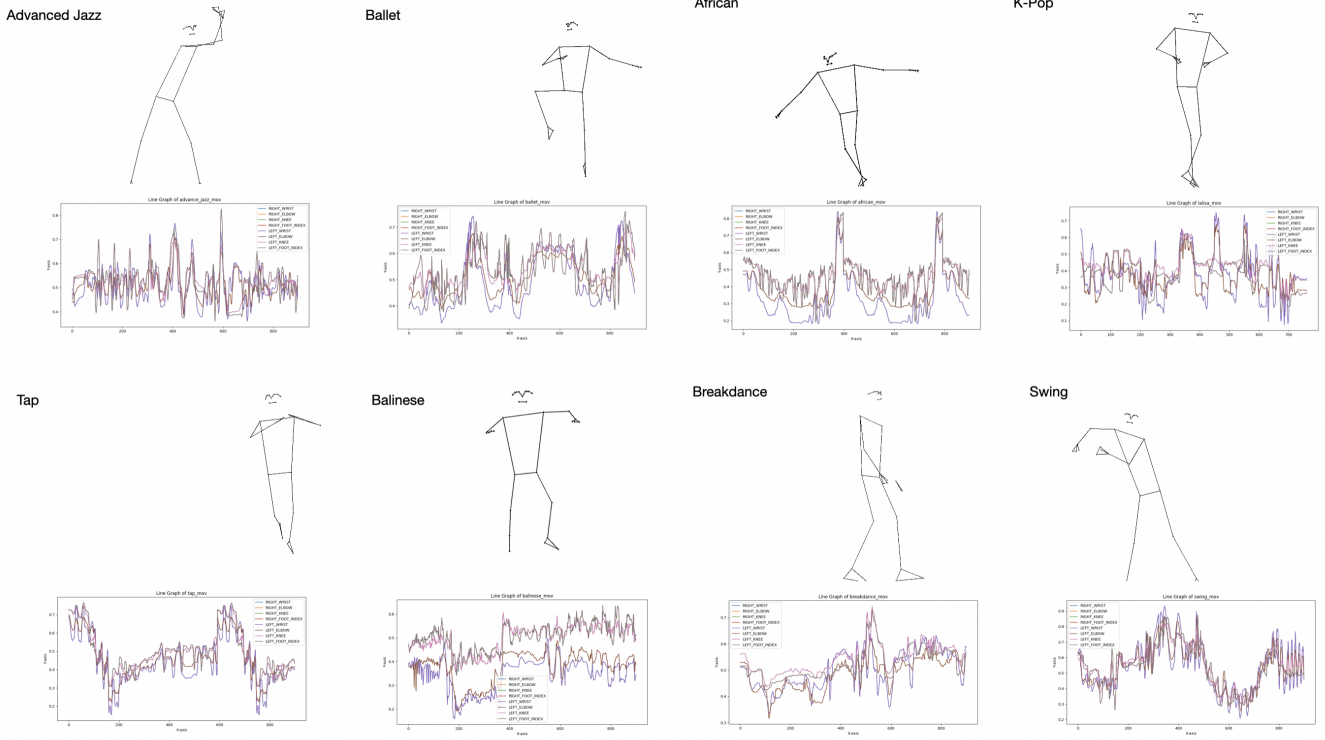
Figure 2: Utilizing the MediaPipe Pose framework, we utilize open-source machine learning methods for 3D pose estimation. Cartesian coordinates for eight key joints (shoulders, elbows, wrists, hips, knees, and ankles) are extracted from the annotation.

3.2 Feature Extraction and Vector Encoding

A preprocessing step for each video within our dataset involves accelerating the footage to enhance the temporal dynamics. Then, we employ the MediaPipe Pose [2, 8], an open-source ML for 3D pose estimation, to annotate 33 landmark points on the dancer's body on each frame of the dance videos.

Coordinates in Cartesian space for eight critical joints, including shoulders, elbows, wrists, hips, knees, and ankles, are extracted from the entire annotation. Subsequently, the axis coordinates of each point are concatenated into a single vector. We encoded the vector into a DNA sequence. It takes in a list of decimals and converts each digit into its equivalent binary representation. Then, it iterates through the binary string, extracting each 2-bit chunk, and maps them to the nucleotide bases A, T, C, or G, depending on the 2-bit combination. This process results in a unique "motion DNA" representation for each dance.

3.3 Phylogenetic Reconstruction

The phylogenetic tree construction in this paper mainly employs Biopython cite10.1093/bioinformatics/btp163 for biological computation. The encoded "motion DNA" representations are aligned, and then the divergence between dances is calculated from the alignment to generate a distance matrix. The result serves as base data for phylogenetic tree construction using UPGMA [22].

The resulting phylogenetic tree diagrams the pseudo-evolutionary relationships between the 21 dance genres sampled, with branch lengths representing estimated degrees of accumulated motion divergence between styles. This serves as an alternative visualization medium to reveal and compare similarities, differences, and connections in choreographic vocabularies across cultures.

4 RESULT

The phylogenetic tree reveals intriguing initial insights into the relationships and shared characteristics across diverse dance genres. We observe some notable clusters, like the proximity of Hawaiian and Electro Swing dances, suggesting potential commonalities in choreographic styles. The close branching of Tap and Breakdance implies a shared evolutionary path, perhaps tied to similar movement patterns or historical connections.

Additionally, we see some dances with related cultural origins grouping together, like Traditional Thai and Japanese dances - likely reflecting their common roots. Upon closer inspection, we may also find specific shared attributes there, like the relatively slow pacing. Conversely, the grouping of genres with minimal cultural connection like Ballet, No. 60 (contemporary Thai dance), and Electro Swing points to underlying similarities possibly tied to features like pronounced arm movements.

However, we acknowledge that this phylogenetic tree is a preliminary visualization and may not fully capture all the nuances of historical connections and contextual influences between these
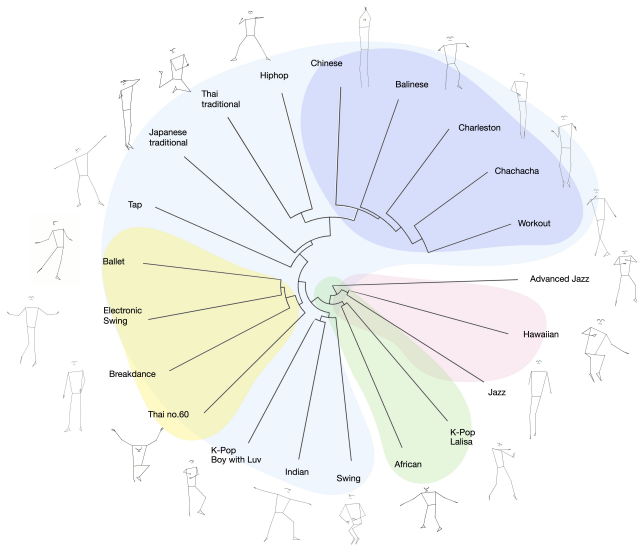
Figure 3: Phylogenetic Tree of Dance constructed from 21 dance samples.

dance forms. As a next step, we will discuss ways to refine the analysis approach.

5 DISCUSSION

While a motion-derived phylogenetic tree of dance is an innovative concept, constructing an accurate one would require overcoming significant data, standardization, and algorithmic challenges. Computational analysis can uncover intriguing high-level patterns for further investigation. But the intricate details shaping these creative forms can only be fully understood through qualitative cultural anthropology, oral histories, archival practices that center communities, and preservation of collective dance knowledge often marginalized from mainstream documentation. Thus, This computational phylogeny should supplement rather than supplant profound anthropological and ethnomusicological scholarship on dance history and cross-cultural exchange gathered through countless academic studies as well as the preservation of indigenous knowledge. Dance evolution is complex, arising from many intertwined forces beyond just motions – such as music, costumes, rituals, social context, and more. Reducing these culturally integral dance styles down to biomechanical features alone risks overlooking Origins and nuances essential to understanding their origins and global dissemination.

At the same time, this computational analysis has unveiled some preliminary transcultural groupings that motivate deeper investigation and introspection. The unexpected clustering of genres with minimal cultural overlap suggests potential universal patterns arising from the innate human drive to move expressively. Furthermore, the quantitative abstraction of motion signatures begins to break down assumptions around rigid stylistic boundaries. What we categorize as distinct genres today have always involved intercultural exchange and inspiration even when not directly visible.

These initial phylogeny results prompt profound reflection on the interconnectedness of global heritage. The quantitative patterns hint at universal themes binding these creative manifestations despite geographic separation. Qualitative scholarship must expand on this potential insight through nuanced exploration of tacit transfers of rhythm, form, and meaning throughout history. Computational analysis provides an opening to restore lost context by inspiring anthropologists, activists and descendant communities to trace and honor these invisible threads tying the destiny of cultures across centuries and continents. We hope that this phylogeny framework would motivate richer dialogue and research from both computational and social science perspectives.

5.1 Limitations & Future Works

the computational phylogeny provides an incomplete picture without incorporating ethnological scholarship on the human stories behind dance evolution. In future investigations, our aim is to enhance our study by refining the intermediate representation to a more inclusive format and including meta-data such as cultural context and geographical data. This will deepen the analysis of dance evolution. Furthermore, we plan to explore advanced phylogenetic tree construction algorithms, potentially transitioning from sequence-based to topological-based approaches, as suggested by [the source]. [5].

In addition, expanding the dataset to include a more diverse collection of the choreographic dataset should enable more comparison between genres and comparison within a specific genre, providing nuanced insights into the patterns, similarities, and differences within and across genres. Expanding the diversity and volume of digitized dance data will empower more finely-grained comparative analysis between genres, lineages, styles, and geographies.

Furthermore, we aim to enhance interactive visualizations that allow dancers, choreographers, anthropologists, and other experts to fluidly navigate the unfolding "tree of life" linking dance vocabularies. Overall, the next phase will catalyze a profound appreciation for dance as the art form that perhaps most elementally binds civilizations across boundaries through deeply ingrained cultural memories coded into the personalized poetry of moving bodies.

6 CONCLUSION

In conclusion, the phylogenetic tree serves as a powerful medium for uncovering dance similarities through a unique lens. The observed patterns prompt further exploration into the cultural, historical, and stylistic influences shaping each dance form's evolution. The tree's capacity to capture shared roots and common choreographic elements across diverse genres offers a nuanced perspective on the intricate tapestry of dance evolution, paving the way for deeper analyses within the realm of movement and expression.

REFERENCES

[1] Andreas Aristidou, Ariel Shamir, and Yiorgos Chrysanthou. 2019. Digital Dance Ethnography: Organizing Large Dance Collections. *J. Comput. Cult. Herit.* 12, 4, Article 29 (nov 2019), 27 pages. https://doi.org/10.1145/3344383

[2] Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, and Matthias Grundmann. 2020. BlazePose: On-device Real-time Body Pose tracking. arXiv:2006.10204 [cs.CV]

[3] Daniel Bisig. 2022. Generative Dance - a Taxonomy and Survey. In *Proceedings of the 8th International Conference on Movement and Computing* (Chicago, IL, USA)

(MOCO '22). Association for Computing Machinery, New York, NY, USA, Article 3, 10 pages. https://doi.org/10.1145/3537972.3537978

[4] Seyed Hossein Chavoshi, Bernard De Baets, Tijs Neutens, Guy De Tré, and Nico Van de Weghe. 2015. Exploring Dance Movement Data Using Sequence Alignment Methods. *PLOS ONE* 10, 7 (Jul 2015), e0132452. https://doi.org/10.1371/journal.pone.0132452

[5] Songle Chen, Zhengxing Sun, and Yan Zhang. 2015. Scalable Organization of Collections of Motion Capture Data via Quantitative and Qualitative Analysis *(ICMR '15)*. Association for Computing Machinery, New York, NY, USA, 411–418. https://doi.org/10.1145/2671188.2749384

[6] Katerina El Raheb and Yannis Ioannidis. 2012. A Labanotation Based Ontology for Representing Dance Movement. In *Gesture and Sign Language in Human-Computer Interaction and Embodied Communication*, Eleni Efthimiou, Georgios Kouroupetroglou, and Stavroula-Evita Fotinea (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 106–117.

[7] Walter M Fitch and Emanuel Margoliash. 1967. Construction of phylogenetic trees: a method based on mutation distances as estimated from cytochrome c sequences is of general applicability. *Science* 155, 3760 (1967), 279–284.

[8] Google. 2024. MediaPipe Pose. https://github.com/google/mediapipe/blob/master/docs/solutions/pose.md

[9] Mohd Firdaus Mohd Herrow and Nur Zaidi Azraai. 2021. Digital Preservation of Intangible Cultural heritage of Joget Dance Movement Using Motion Capture Technology. *Int. J. Herit. Art Multimed* 4 (2021), 1–13.

[10] Kiyoshi Hoshino, Sota Sugimura, Motomasa Tomida, Naoki Igo, Isao Kawano, and Masahiko Sumitani. 2016. Hand Motion Capture for Medical Usage *(ICBBE '16)*. Association for Computing Machinery, New York, NY, USA, 40–45. https://doi.org/10.1145/3022702.3022712

[11] Shi-Sheng Huang, Ariel Shamir, Chao-Hui Shen, Hao Zhang, Alla Sheffer, Shi-Min Hu, and Daniel Cohen-Or. 2013. Qualitative Organization of Collections of Shapes via Quartet Analysis. *ACM Trans. Graph.* 32, 4, Article 71 (jul 2013), 10 pages. https://doi.org/10.1145/2461912.2461954

[12] Atsushi Yamada Iwori Takeda and Hiroshi Onodera. 2021. Artificial Intelligence-Assisted motion capture for medical applications: a comparative study between markerless and passive marker motion capture. *Computer Methods in Biomechanics and Biomedical Engineering* 24, 8 (2021), 864–873. https://doi.org/10.1080/10255842.2020.1856372

arXiv:https://doi.org/10.1080/10255842.2020.1856372 PMID: 33290107.

[13] Paschalia Kapli, Ziheng Yang, and Maximilian J Telford. 2020. Phylogenetic tree building in the genomic age. *Nature Reviews Genetics* 21, 7 (2020), 428–444.

[14] Iris Kico, Nikos Grammalidis, Yiannis Christidis, and Fotis Liarokapis. 2018. Digitization and visualization of folk dances in cultural heritage: a review. *Inventions* 3, 4 (2018), 72.

[15] Arne O Mooers and Stephen B Heard. 1997. Inferring evolutionary process from phylogenetic tree shape. *The quarterly review of Biology* 72, 1 (1997), 31–54.

[16] Basilio Pueo Ortega and José M Jiménez Olmedo. 2017. Application of motion capture technology for sport performance analysis. *Retos: nuevas tendencias en educación física, deporte y recreación* 32 (2017), 241–247.

[17] Hua Peng, Changle Zhou, Huosheng Hu, Fei Chao, and Jing Li. 2015. Robotic Dance in Social Robotics—A Taxonomy. *IEEE Transactions on Human-Machine Systems* 45, 3 (2015), 281–293. https://doi.org/10.1109/THMS.2015.2393558

[18] Ioannis Rallis, Athanasios Voulodimos, Nikolaos Bakalos, Eftychios Protopapadakis, Nikolaos Doulamis, and Anastasios Doulamis. 2020. Machine learning for intangible cultural heritage: a review of techniques on dance analysis. *Visual Computing for Cultural Heritage* (2020), 103–119.

[19] MR Reshma, B Kannan, VP Jagathy Raj, and S Shailesh. 2023. Cultural heritage preservation through dance digitization: A review. *Digital Applications in Archaeology and Cultural Heritage* (2023), e00257.

[20] Kevin Shu, Sharjeel Aziz, Vy-Luan Huynh, David Warrick, and Matilde Marcolli. 2018. *Syntactic Phylogenetic Trees*. Springer International Publishing, Cham, 417–441. https://doi.org/10.1007/978-3-319-64813-2_14

[21] Maria Skublewska-Paszkowska, Pawel Powroznik, Jakub Smolka, Marek Milosz, Edyta Lukasik, Dilbar Mukhamedova, and Elzbieta Milosz. 2021. Methodology of 3D scanning of intangible cultural heritage—The example of Lazgi dance. *Applied Sciences* 11, 23 (2021), 11568.

[22] Peter HA Sneath and Robert R Sokal. 1973. Unweighted pair group method with arithmetic mean. *Numerical Taxonomy* (1973), 230–234.

[23] Tujin Zhu and Yunpeng Cai. 2021. Applying Neural Network to Reconstruction of Phylogenetic Tree. In *Proceedings of the 2021 13th International Conference on Machine Learning and Computing* (Shenzhen, China) *(ICMLC '21)*. Association for Computing Machinery, New York, NY, USA, 146–152. https://doi.org/10.1145/3457682.3457704