Digital Dance Ethnography: Organizing Large Dance Collections

ANDREAS ARISTIDOU*, University of Cyprus, Cyprus and RISE Research Center, Cyprus
ARIEL SHAMIR, The Interdisciplinary Center Herzliya, Israel
YIORGOS CHRYSANTHOU, University of Cyprus, Cyprus and RISE Research Center, Cyprus

Fig. 1. Our method can efficiently organize large dance motion collections based on their degree of similarity. Inferred metadata information are then applied to semantically organize data into chronological, or geographical order, unveiling potential similarities in terms of the evolution of dance in time and at different countries. In this example, we illustrate similar Cypriot dances that were popular at different time periods.

Folk dances often reflect the socio-cultural influences prevailing in different periods and nations; each dance produces a meaning, a story with the help of music, costumes and dance moves. However, dances have no borders; they have been transmitted from generation to generation, along different countries, mainly due to movements of people carrying and disseminating their civilization. Studying the contextual correlation of dances along neighboring countries, unveils the evolution of this unique intangible heritage in time, and helps in understanding potential cultural similarities. In this work we present a method for contextually motion analysis that organizes dance data semantically, to form the first digital dance ethnography. Firstly, we break dance motion sequences into some narrow temporal overlapping feature descriptors, named *motion* and *style words*, and then cluster them in a high-dimensional features space to define *motifs*. The distribution of those motion and style motifs creates *motion* and *style signatures*, in the content of a bag-of-motifs representation, that implies for a succinct but descriptive portrayal of motions sequences. Signatures are time-scale and temporal-order invariant, capable of exploiting the contextual correlation between dances, and distinguishing fine-grained difference between semantically similar motions. We then use *quartet*-based analysis to organize dance data into a *categorization tree*, while inferred information from dance metadata descriptions are then used to set parent-child relationships. We illustrate a number of different organization trees, and portray the evolution of dances over time. The efficiency of our method is also demonstrated in retrieving contextually similar dances from a database.

CCS Concepts: • **Computing methodologies** → **Animation**; • **Applied computing** → *Ethnography*; *Digital libraries and archives*; • **Information systems** → Information retrieval;

*This is the corresponding author

Authors' addresses: Andreas Aristidou, University of Cyprus, 75, Kallipoleos, Nicosia, Cyprus, 1678, RISE Research Center, Nicosia, Cyprus, a.aristidou@ieee.org; Ariel Shamir, The Interdisciplinary Center Herzliya, Kanfei Nesharim, Herzliya, Israel, 4610101, arik@idc.ac.il; Yiorgos Chrysanthou, University of Cyprus, 75, Kallipoleos, Nicosia, Cyprus, 1678, RISE Research Center, Nicosia, Cyprus, yiorgos@cs.ucy.ac.cy.

XXXX-XXXX/2019/1-ART1
https://doi.org/10.1145/3344383

Additional Key Words and Phrases: Dance, Intangible Cultural Heritage, Motion and Style Signatures, Categorization Tree.

ACM Reference Format:
Andreas Aristidou, Ariel Shamir, and Yiorgos Chrysanthou. 2019. Digital Dance Ethnography: Organizing Large Dance Collections. *ACM J. Comput. Cult. Herit.* 1, 1, Article 1 (January 2019), 26 pages. https://doi.org/10.1145/3344383

1 INTRODUCTION

Most countries in the Mediterranean, among others, are rich in history and cultural heritage. Their tradition has been transmitted-to and influenced-by various neighboring civilizations. Over the last few decades, and due to the recent technological advances, many nations have attempted to record, e-document, preserve, protect and disseminate tangible cultural heritage. However, apart from tangible artifacts, cultural heritage also encompasses a range of important assets that includes collective knowledge of communities, skills, practices, expressions, art, fashion and representations that do not have a tangible form. Intangible Cultural Heritage (ICH), as stated in [98], is a mainspring of humanity's cultural diversity and its maintenance is a guarantee for continuing creativity. Over the years, certain elements of ICH have been lost forever, while others are at high risk of disappearing, mainly due to globalization, wars, financial crisis, and the movement of people, that cause the diminishment of unique cultural assets.

This work focuses on documenting, analyzing and visualizing aspects of dance heritage. Folk dance is one of the most important aspects of ICH. The indigenous dance traditions of most nations are still alive, and continue to influence dance in modern years. Thus, appreciating and understanding dance and other structured movement systems is important in the larger scheme of cultural forms. Dance forms respect no borders: they can be transmitted through different civilizations and cultures, creating different variations that reflect the differences of each country's sociopolitical specificities. Many folk dance creations have been modified over time through the process of collective recreation, and/or changes in the way of life over the years. Studying the evolution of dances and their relation among neighboring countries is of high importance to the cultural heritage community. Our work provides the algorithmic means for contextual dance motion analysis, unveiling similarities of dance heritage in different neighboring countries and at different time periods, paving the way for computing the first digital dance ethnography.

Dance ethnography refers to the textual presentation of data, including ethnographic descriptions on indigenous perspectives. Dance has been proven to be an indicator of sociocultural circumstances with political and religious influences, often identifying points of conflict and driving transitions. In order to be understood as dance, movements must be grammatical; the grammar of a movement idiom involves structure, style, and meaning. One must learn to recognize the movements that make up the system, how they can be stylistically modified, and what is their syntax (rules about how they can be put together to form motifs, phrases, larger forms, and whole pieces) [52]. In this work, motion sequences are divided into motion and style words that are clustered together to form motifs. The motions are then characterized by the frequency and distribution of those motifs. Representing motion in such a succinct and descriptive form allows to automatically unveil the social and cultural correlation between neighboring countries in terms of their context.

In recent years, the attention of a number of scientists and media artists has been drawn to utilizing existing techniques, and developing new ones for the preservation and propagation of the dance heritage, as well as its distribution through media technologies [7, 19, 42, 44, 57] . More specifically, motion acquisition has been enabled using motion capture systems (e.g., optical, inertial or RGB-Depth motion capture systems). These systems have proven to be an effective technology and a convenient tool for capturing, and digitizing complex human dynamic movements, advancing the ability to digitally store, curate, present and re-use intangible creations, such as folk dancing [59, 92]. With the emergence of motion capture data and the large availability of motion datasets, such as the Dance Motion Capture Database (DMCD) [29], content-based motion analysis techniques have become essential for organizing dance motion collections. Such techniques should allow effective searching in the datasets and unveil contextual similarities and differences. However, motion data does not contain labels, annotation or semantics to assist organization. In addition, the large diversity of motions, and their complexity, makes automatic motion indexing and clustering challenging, especially for highly dynamic, heterogeneous, and stylized motions, such as dancing. Most motion clustering techniques rely on motion skeletal [13, 56, 62] or relational [54, 75, 82] information, and fail to assess some important aspects of human action, such as synchronization and scaling. In fact, each performer's improvisation, experience, and talent may result in different variations of the same dance, while the same dance can vary in the temporal order, duration, and time, even if they are performed by the same dancer.

In this work, we introduce the algorithmic framework for contextual analysis, organization, and comparison of dances. Our context-based motion organization approach, which exploits the geometric and stylistic relation between motions, automatically group similar dance performances and can be used to form the digital dance ethnography, to unveil potential similarities between dances from neighboring countries, to study the evolution of dance, and many other applications. As a first step, we have enriched the DMCD, in collaboration with dance schools and cultural workshops, with a large number of high quality folk dances originated from the wider region of the eastern Mediterranean, the Balkans and Pontus. We describe how dance motion can be acquired and documented holistically so as to enable the extraction of semantic, cultural, and contextual correlations. We have defined a descriptive representation for motions based on *signatures* [5], which rely on the distribution of some narrow temporal overlapping feature descriptors, named *motion* and *style words*. Signatures are time-scale and temporal-order invariant, capable of exploiting the contextual correlation between dances. Our dance collection is then organized into a *categorization tree* via *quartet*-based analysis [49]. In close collaboration with dance experts, we identified which metadata are useful for archiving, curating, presenting and re-using dance motion data. We designed a holistic metadata scheme to drive further studies of dances from an anthropology, and ethnology perspective. Moreover, inferred information from those metadata descriptions is used to establish semantic links between dances. The semantic links create parent-child relationships, in a hierarchical mode, to establish chronological and geographical correlations in our dance collection, paving the way for creating the first digital dance ethnography. It is important to note that we only focus on the geometric and stylistic characteristics of movements, while the hierarchy is based on semantic links derived from metadata information. We visualize the evolution of dances over time and at different countries/cultures. We demonstrate the efficiency of our method on a number of folk, modern and other dances (taken from three different online dance libraries), organizing motion collections and retrieving similar dances from a database. Our method can find applications for motion indexing in virtual dance museums, and can be used for studying and analyzing the cultural and social similarities/differences of the dance heritage in neighboring countries.

2 RELATED WORK

The related work in divided into methods for (a) motion analysis, (b) motion organization and summarization, and (c) a review on computing methods with specific applications in dancing and intangible cultural heritage.

Motion Analysis: Keyword queries [25], or annotation [3] are cost-effective methods for motion retrieval and context-based data organization. However, these methods require manual labelling, and cannot apprehend the complexities and particularities of motion data. Another way to retrieve similar motion is by matching poses [62, 66, 81] or other geometric distances [13, 20, 24, 95]. Indeed, these methods provide an intuitive means of query specification, but cannot capture the temporal evolution, and dynamics of human motion, while they have difficulties in handling heterogeneous and complex motions. Small temporal windows may extract the spatiotemporal features of motion, but since similar motion segments can vary in duration and speed, we cannot simply compare fixed-length time windows. One way to deal with some aspects of temporal evolution in motion is to include both geometric and dynamic features [17], to employ uniform scaling, a global stretching or shrinking of the time series [56], or Dynamic Time Warping (DTW) to temporally synchronize motion sequences [74, 79]. However, due to the high dimensionality of human motion, these methods are very expensive in terms of computational cost. To accelerate motion retrieval in large scale databases, Kovar and Gleicher [61] use match webs as an index structure to find numerically similar motions, Chai and Hodgins [17] build a graph to allow fast nearest neighbor search, while others decompose motions into body parts and use a hierarchical motion representation (R-trees and kd-trees) [28, 56, 63, 68]. Another way to deal with time inconsistency in motion is to employ relational [75, 82] or qualitative features [4, 54, 87]. Indeed, these features extract the dynamic and spatiotemporal information of motion, but not the numerical similarity between poses.

Several methods deal with the high complexity and dynamic feature of motion by defining a succinct representation for motion sequences, either by applying principal component analysis (PCA) [38], by presenting motion as Boolean values at selected keyframes [74], or by training deep autoencoders [100]. These methods, however, can only handle short-time sequences, they cannot deal with complex and dynamic motion sequences, motions with temporal variation in duration and speed, or motions with extreme reordering, a common characteristic of dances (e.g., in folk dancing, dancers may perform similar pirouettes, but at different times and in arbitrary order). On the other hand, Kapsouras and Nikolaidis [55], and later Fotiadou *et al.* [39], use a Bag-of-Words (BoW) model to define motion codebooks for human action recognition or to recognize Greek folk dances among other actions, respectively. However, they treat motion as a set of individual poses and not as a sequence, resulting in losing the semantic information and the temporal evolution of motion. In this work, we learn a high-dimensional universal feature space of short-time motions sequences using a deep network, in a similar way to Aristidou *et al.* [5], and define motion and style signatures which are time-scale and temporal-order invariant, offering a succinct and descriptive representation of motion sequences.

Motion Organization: One of the best ways to organize a set of elements is clustering analysis [37]. To achieve a relying clustering, it is important to utilize an efficient distance metric that can quantitatively measure the similarity among the elements of the collection, i.e., [40]. Ordination or neighborhood construction is another important analysis for organizing data collection that puts similar elements near to each other, and dissimilar farther apart [46], while it is the basis for data visualization, overview and exploration. In this direction, Bernard *et al.* [15] developed MotionExplorer to cluster and display motions as a hierarchical tree structure. The authors used the self-organizing map (SOM) on joint position features to train a topology preserving the grid of poses. More recently, they introduced a visual-interactive approach for labeling human motion data that represents motion as sequences of motion classes [14]. Müller *et al.* [73] propose to express motion as an explicit matrix, while Chen *et al.* [21] used a number of low level pose features to perform data abstraction, and then applied topological constrains to generate multiple layers of data aggregations. In contrast, we use multiple distance measures, and similarly to Huang *et al.* [49], we organize our dance motion collection using a set of quartets that set topological constrains and contribute to the design of a categorization tree; inferred semantic information is also used to portray dances in a chronological order.

Dance and Intangible Cultural Heritage: A number of methods have been proposed for motion comparison, indexing, retrieval and summarization that explicitly deal with dancing e.g., for modern [4, 32], and folk dance [8, 30, 80], resulting in numerous dance-oriented applications, e.g., dance synthesis, teaching, games, annotation etc. [101]. For instance, Shiratori *et al.* [89] synthesize dance movements to match given music, Fukayama *et al.* [41] use machine learning to generate music-driven dance movements, while Aristidou *et al.* [9] build, in the context of Motion Graphs [62], an LMA-derived motion analysis framework that eliminates potentially problematic transitions and synthesizes style-coherent dance motions. The LMA framework has also been used for dance motion evaluation [50], as well as emotion recognition and stylization in dancing [11]. Other methods focus on automatically evaluating dance performances compared to pre-captured samples, aiming to provide visual feedback to the performer in a 3D virtual environment, e.g., in martial arts [51], dance [1, 2, 18, 27, 36, 94, 96], ballet [64], and folk dancing [7, 22, 60, 65, 67]. Other works introduce different whole-body interaction interfaces for exploring various visualizations in dance learning and gamification [35, 70, 97]. Iris *et al.* [57] have recently presented a detailed review on methods for digitization and visualization of folk dances in cultural heritage.

3 DANCE MOTION CAPTURE DATABASE

In this section, we describe how dance motion can be acquired and documented holistically using emerging technologies, so as to enable safeguarding of intangible heritage creations. In close collaboration with dance experts, we identify and present the metadata information that is useful for archiving, curating, presenting, analyzing, and re-using dance motion data, aiming to expose semantic, cultural, and contextual correlations among dances from neighboring countries. Moreover, we describe the framework for dance motion acquisition, and the way these data and metadata information have been used to enrich DMCD.

3.1 Metadata for Dance ICH Data

Metadata is the data that provides information about other data; in other words, metadata is the documentation that describes data. There are different types of metadata, and can be distinguished into five main categories: *descriptive*, *structural*, *administrative*, *reference*, and *statistical* metadata, each one aiming to describe different resource for purposes [48]. Descriptive metadata describes a resource for purposes such as discovery and identification (e.g., title, abstract, author, keywords); structural metadata is about containers of data (e.g., types, versions, relationships, other characteristics of digital materials); administrative metadata provides information to help manage a resource (e.g, date of creation, file type, authentication, access and administration rights); reference metadata describes the contents and quality of statistical data; and statistical metadata describes practices that collect, process, or produce statistical data.

The best way to systematically and structurally organize data is to use and define metadata schemas as a logical plan that shows the relationships between metadata elements [90]. Metadata main purpose is to assist users to locate information, discover resources, and allow further studies with regard to the content, history, structure etc. of the data, and are vital for electronic resource organization and the digital preservation of resources and information. Information professionals, creators and users of digital content, agree that the most complete metadata information provided enables accessibility, interoperability and preservability of cultural heritage creations, and contributes to archiving, disseminating, studying, and reusing of cultural heritage information objects [12].

Over the years, numerous ontological and metadata representations have been proposed to allow further investigation, studies, and research in dance heritage, e.g., [44, 53, 58, 92, 102]. An extended research about metadata in digital folklore collections, and the main metadata standards to represent cultural heritage collections is given by Lourdi *et al.* [69]. The digitization and documentation of folk dancing consists a large variety of different and complex data, including descriptive data, with regard to the description of the dance being performed, the history, notations, but most importantly, the multimedia recordings. In this direction, many researchers focused on ways to document and encode different multimedia elements of the dance creations, e.g. the MusciXML [43] to describe and encode musical motifs in XML-based, the LabanXML [76], MovementXML [47], and more recently, the DanceXML [31, 32], to represent, again in XML-based, the movements using Labanotation [45]. To provide online access to digital dance creations, the OWL ontology was used to assign annotations representing movement sequences of a dance-recording video [34], or the Multimedia Web Ontology Language (MOWL) [71], to correlate heritage resources and multimedia data. A comprehensive report on metadata and schemas specially designed for Intangible Cultural Heritage creations is given in [42], where all metadata schemas, and standards for 3D virtual representations are extensively discussed.

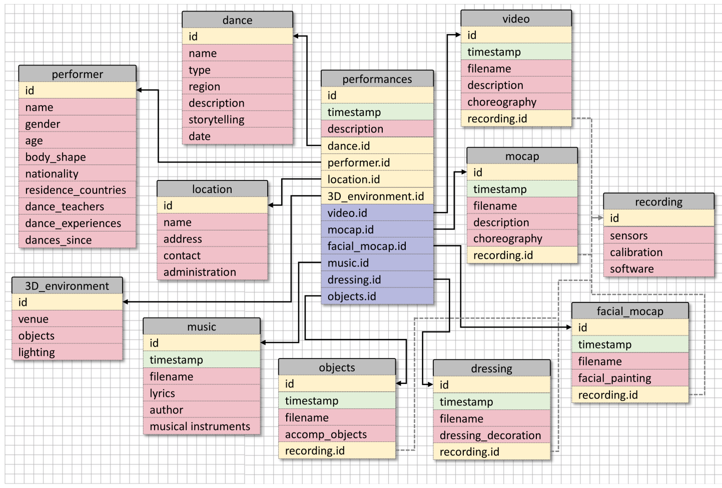

Fig. 2. Our folk dance database schema.

Before constructing the metadata schema, it is important to accurately define the data to be collected. Folk dancing is a rich and diverse intangible element that requires numerous and complex data to be fully defined, including descriptive data, with regard to the description of the dance being performed, the dancer experience and background, and the multimedia recordings. In this direction, we have consulted folk dance professionals to assist in defining the data and metadata information needed to holistically describe, document, and archive folk dancing, but most importantly, to provide the basis for creating the digital dance ethnography, and thus forming the time and space evolution of dances. For each dance creation, it is recommended to collect information about the dance, such as textual description, the story-telling, the country/region of origin, the date that the dance first became known, the type of the dance (e.g., solo or group), and its history. The multimedia collection includes the motion capture data, both body and facial, the video recordings, and music (including the lyrics, composer, singer, etc.) of individual performances. It is also important to collect data with regard to the technology used for capturing, such as the motion capture system, and the kind of sensors used, the calibration parameters, the recording software and the data from recording. In addition to these recordings, metadata information about the dancers appearing in performances (e.g., name, gender, age, height, weight, nationality, countries of residence, dance teachers, years of dance experience, dance background and years of experience in different genres), the locations where these dances are performed, the traditional costumes, accompanied objects (e.g., swords, sticks), the 3D environment of the digital representation, the lighting etc. All these metadata information allows to study each of the performance parameters holistically. They also contribute in understanding the experience and background of the dancer, and are particularly useful for further cultural, anthropological, and ethnological studies. Other metadata information useful for processing, analysis, and management of dance data can be extracted using fusion analysis. A set of low or medium-level multimedia features can be applied to exploit information across different modalities, e.g., stylistic properties based on the Laban theories on kinesiology and choreography (the Laban Movement Analysis system) [4, 23, 54], Labanotation [33, 84], or music, rhythm analysis [88, 89], while the context and content are integrated to transform data into a level of interpretation that is understandable by humans. In this work, we employed a relational database schema to structure the information within the archive. Figure 2 shows our ideally database schema that structures the dance data and metadata information. Note that, this is an ongoing project, and our dance collection will be enriched with more data and metadata information as

Fig. 3. An animated virtual character that performs the *Antikristos* dance and wears the traditional Cypriot uniform, at the Hadjigeorgakis Kornesios Mansion in the medieval town of Nicosia, Cyprus.

they become available; currently, some information is missing (e.g., facial expressions) due to the lack of software and hardware facilities in the lab.

3.2 Dance Data Acquisition

Over the last few months, we devoted a considerable effort to enrich the existing Dance Motion Capture Database [29] by capturing and digitizing folk dances from the wider area of the Easter Mediterranean, the Balkans and Pontus. DMCD currently provide access to different types of content (e.g. text, audio, images, video, 3D graphics, motion capture data) from various intangible cultural heritage resources, mainly for research, studies, and education purposes. The data has been captured using an eight cameras PhaseSpace Impulse X2 motion capture system [78], which allows for high-frequency (up to 960Hz) optical tracking of the dance performers using modulated LEDs. Such system is able to acquire 3D motion data, and maintain the correct human proportions and the naturalness of the action, however our system ability is only for a single character over time. In order to ensure that these important and valuable intangible cultural creations are sufficiently well documented, recorded and archived, we captured male and female performers who are experienced dancers and active members of cultural organizations and dance schools. These quality and culturally important datasets have been uploaded to DMCD, whereas this publicly accessible online dance repository currently stores more than 180 dance performances (of which more than 30 are folk dances), being one of the most complete digital dance libraries in the world. Note that, some of these folk dance creations have been harvested into EUROPEANA[1], the European Digital Library of Cultural Heritage. Figure 3 shows snapshots of an animated virtual character who performs a Cypriot folk dance from the database. To facilitate the long term maintenance of the database, and since folk dance digitization is an evolving project, the archive has been designed in a way that is scalable and can easily accumulate new data as they become available.

For the purpose of this work, we have created a database $\mathcal{D}$ of 72 dances; data has been taken from the DMCD [29], the CMU [25] and NUS [77] motion capture databases. Our datasets comprise folk dances with different origin, such as from *Cyprus*, *Greece*, *Serbia*, *Egypt*, and *Spain*. It also includes other dances, e.g., *Latin* dances, *Waltz*, *Capoeira*, *Ballet*, *Contemporary* dances, etc. All dances, with their metadata, used in this work are listed in the Appendix, Table 2. The motion capture data are in BVH (Biovision Hierarchical Data) format, they

[1]Europeana, EU digital platform for cultural heritage: https://www.europeana.eu/

were originally sampled between 120 to 480 frames per second, but then downsampled to 24 frames per second (to reduce the computational time) without much loss of the temporal information (see Forbes and Fiume [38]). It is important to note that our data have been retargeted to a single BVH skeleton with standard body proportions, to enable uniform processing of all the acquired motion capture data. Also, more than one capture is in general recommended, for each dance. This is because each performer has its own dance accent that is highly related to hir/her experiences, background, body type, age, emotion etc. No standardization or normalization of dances should be taken into account since these dances correspond to the idiosyncracy and experience of the performer.

4 MOTION ANALYSIS

In order to organize a categorization tree for dances, we need a succinct but descriptive representation of motion sequences. Such a representation enables similarity comparison of complex, heterogeneous, and highly dynamic movements that are time-scale and temporal-order variable, such as the dance movements. The representation should also apprehend the geometric, dynamic, and stylistic aspects of motion. In this paper we tackle these challenges by utilizing *motion* and *style signatures*; the core idea is that motion sequences can be broken down to smaller movements, and can then be characterized by the distribution of such movements. Thus, in a similar manner to Aristidou *et al.* [5], we first extract a set of overlapping *motion* and *style words* from the whole motion sequence, and then distill those words to a set of motifs, which are descriptive and frequent words. Movements are therefore represented by their motion and style signatures which are defined by the frequency and distribution of their motion and style words, respectively.

4.1 Geometric Motion Analysis

To construct motion signatures, we first divide motion sequences into motion words. A motion word is defined as a narrow temporal window of all joint rotations around a given frame that represent the local evolution of pose [6]; each joint defines three rotation values that are in the range of [0; 360] degrees. Motion words divide a motion sequence into smaller, overlapping, feature descriptors defining a local spatiotemporal descriptor. Motion words, in our experiments, are defined using the $m = 16$ most informative joints with their relative joint angles. Similarly to [5], we use 16 frames window to define motion words, that reflects to 0.66 seconds, with a skip of 4 frames. This length proved to be long enough to cover simple movements, but short enough to promote similarities.

To define the vocabulary of our universal motion words feature-space, we gather motion words from our dance database $\mathcal{D}$. Motion words are embedded into a feature space using a deep neural network. The embedding places semantically similar motion words close together, and semantically different words far apart. Instead of computing the distance, which is computational and time expensive, we learn the embedding using a triplet-loss network [85]. We train the network using positive examples, which are either motion words that appear temporarily close in the training data, or words that match using dynamic time warping, and negative examples, which are random motion words that are either temporally or postural dissimilar. The network maps all motions words of the dataset into the d-dimensional universal feature space $\mathbb{R}^d$. Note that our network creates a 736×1 embedding by integrating the Inception model [93]. We use K-means clustering algorithm and group the motion words into K (empirically, $K = 100$) mutually exclusive clusters[2]. Each cluster, is represented by a *motif* motion word which is the centroid of the cluster. Thus, motion signatures are defined as a bag-of-motifs, which models the distribution of its motion-motifs. In other words, given a motion sequence, we first extract all its motion words, map them to the universal feature space, assign each word to its representative motif, count the number of words in each of the K clusters, and divide by the total number of words in the motion sequence. This creates a

[2]Note that, for large dance databases, with highly dynamic motion sequences, we believe that a larger number of clusters might be required ($K > 200$)

comparable signature for every motion sequence regardless of its length. More details about the definition of the embedding space, the parameters, structure and training of the triplet-loss network can be found in [5]. Two motion sequences are considered similar if their signatures have similar characteristics, which means that they have a similar distribution of motion motifs.

4.2 Stylistic Motion Analysis

Style plays an important role in dance motion evaluation [4]. Thus, apart from the use of motion words to define motion signatures, as presented in [5], in this work we also introduce *style words* and *style signatures*, aiming to appreciate the stylistic variations of movements. The stylistic representation enhances mostly a set of dynamic features, with regard to the body, effort, shape and space components of the Laban Movement Analysis (LMA) system, and allows finding stylistic similarities within motion sequences. Style words are one dimensional arrays that encode a number of LMA-derived features (114 measurements ϕ_i, please refer to Table 1), from selected key joints, within a short temporal-window around a given frame. Similarly to motion words, style words are local spatiotemporal descriptors that divide motion into smaller, overlapping components. Style words are computed in right anchored windows of 16 frames, with stride 4 frames (to be similar to our motion word definition), using a number of high-level motion features based on the LMA principles (see Aristidou *et al.* [11]). The distance between two style words is computed using the Earth Mover's Distance (EMD) metric [83]. We position style words into d-dimensional space using a robust Multi-Dimensional Scaling (MDS) [16] that is optimal to outliers (we tried different dimensionalities, $\mathbb{R}^d$, and the quality of embedding seems to be equally well for $d \geq 10$), and cluster the words in this space using K-means (again $K = 100$). Similarly to motion words, we define the centroid of each cluster as *style motif*, and define *style signatures* as the normalized histogram of the frequency of style words in all K clusters. Style words help us apprehend differences in effort and space (e.g., dynamics of motion), even if movements appear similar in pose.

4.3 Computing the Relation between Dance Sequences

Our motion and style analysis is based on a Bag-of-Motifs motion representation that is invariant to the temporal-ordering of motion and style motifs. Being oblivious to the motif's order is important in dancing since it allows comparing motions with similar frequencies of motifs regardless of their exact temporal location. Our motion and style signatures allow comparing sequences of different speed and duration, enriching diversity in comparisons and analysis. A common problem although in BoW implementation is the naive frequency counting. Highly frequent motion and style words dominate the dataset, but may not be as informative or descriptive as some less common motion words. To deal with this problem, we rescale the frequency of the motion and style motifs by the frequency of the total motifs in the corpus, for each one separately. We search for the statistically significant motion and style parts which occur regularly, and makes them unique. Those elements are both repeating, i.e., occur often in a motion group, and spatiotemporal discriminative, i.e., occur much more often in a specific motion group than other motion groups. The reoccurrence of such motion words across a motion sequence induces strong and meaningful affinities. Thus, we re-weight motion signatures using tf-idf (term frequency – inverse document frequency [26]). The importance of a motif is proportional to its frequency in the clip and inversely proportional to its frequency in the corpus. Figure 4 shows the motion and style signatures for two families of three dances.

Once we have the signatures (both motion and style), we define the similarity between dance motion sequences as the distance between those signatures. We use the Earth Mover's Distance (EMD) [83] to compare two signatures as they represent distributions. We have three metrics to compute the relation between dances, the geometric that is computed using the motion signatures, the stylistic that is computed using the style signatures, and the combination of the two aforementioned relationships. We allow users to manually tune the weight of

Table 1. The LMA-derived feature measurements used for defining style words and style signatures, as listed in [10].

	Features		Measurements			
	f^i	Description	f^i_{max}	f^i_{min}	f^i_{σ}	f^i_{μ}
Body	f^1	Left foot-hip distance	ϕ_1	ϕ_2	ϕ_3	ϕ_4
		Right foot-hip distance	ϕ_5	ϕ_6	ϕ_7	ϕ_8
	f^2	Left hand-shoulder distance	ϕ_9	ϕ_{10}	ϕ_{11}	ϕ_{12}
		Right hand-shoulder distance	ϕ_{13}	ϕ_{14}	ϕ_{15}	ϕ_{16}
	f^3	Hands distance	ϕ_{17}	ϕ_{18}	ϕ_{19}	ϕ_{20}
	f^4	Left hand-head distance	ϕ_{21}	ϕ_{22}	ϕ_{23}	ϕ_{24}
		Right hand-head distance	ϕ_{25}	ϕ_{26}	ϕ_{27}	ϕ_{28}
	f^5	Hip-ground distance	ϕ_{29}	ϕ_{30}	ϕ_{31}	ϕ_{32}
	f^6	Hip-ground minus feet-hip	ϕ_{33}	ϕ_{34}	ϕ_{35}	ϕ_{36}
	f^7	Centroid-ground distance	ϕ_{37}	ϕ_{38}	ϕ_{39}	ϕ_{40}
	f^8	Centroid-pelvis distance	ϕ_{41}	ϕ_{42}	ϕ_{43}	ϕ_{44}
	f^9	Gait size	ϕ_{45}	ϕ_{46}	ϕ_{47}	ϕ_{48}
Effort	f^{10}	Head orientation	ϕ_{49}	ϕ_{50}	ϕ_{51}	ϕ_{52}
	f^{11}	Deceleration peaks				ϕ_{53}
	f^{12}	Pelvis velocity	ϕ_{54}		ϕ_{55}	ϕ_{56}
	f^{13}	Left-hand velocity	ϕ_{57}		ϕ_{58}	ϕ_{59}
		Right-hand velocity	ϕ_{60}		ϕ_{61}	ϕ_{62}
	f^{14}	Left foot velocity	ϕ_{63}		ϕ_{64}	ϕ_{65}
		Right foot velocity	ϕ_{66}		ϕ_{67}	ϕ_{68}
	f^{15}	Pelvis acceleration	ϕ_{69}		ϕ_{70}	
	f^{16}	Left-hand acceleration	ϕ_{71}		ϕ_{72}	
		Right-hand acceleration	ϕ_{73}		ϕ_{74}	
	f^{17}	Left foot acceleration	ϕ_{75}		ϕ_{76}	
		Right foot acceleration	ϕ_{77}		ϕ_{78}	
	f^{18}	Jerk	ϕ_{79}		ϕ_{80}	
Shape	f^{19}	Volume (5 joints)	ϕ_{81}	ϕ_{82}	ϕ_{83}	ϕ_{84}
	f^{20}	Volume (upper body)	ϕ_{85}	ϕ_{86}	ϕ_{87}	ϕ_{88}
	f^{21}	Volume (lower body)	ϕ_{89}	ϕ_{90}	ϕ_{91}	ϕ_{92}
	f^{22}	Volume (left side)	ϕ_{93}	ϕ_{94}	ϕ_{95}	ϕ_{96}
	f^{23}	Volume (right side)	ϕ_{97}	ϕ_{98}	ϕ_{99}	ϕ_{100}
	f^{24}	Torso height	ϕ_{101}	ϕ_{102}	ϕ_{103}	ϕ_{104}
	f^{25}	Hands level				$\phi_{105-107}$
Space	f^{26}	Total distance				ϕ_{108}
	f^{27}	Total area				ϕ_{109}
	f^{28}	Total volume				ϕ_{110}
	f^{29}	Total volume	ϕ_{111}	ϕ_{112}	ϕ_{113}	ϕ_{114}

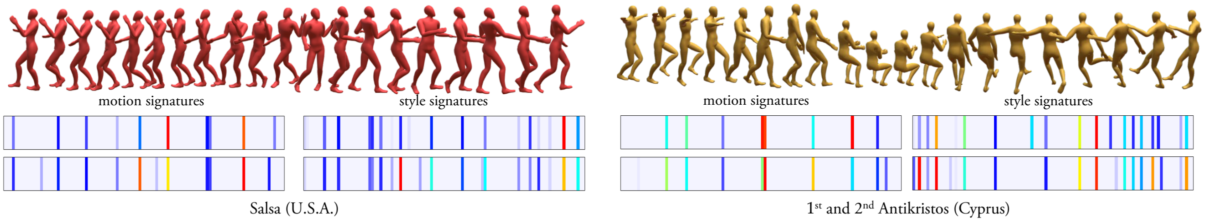

Fig. 4. Motion and style signatures for two families of dances; the first family consists two *salsa* dances (on the left), and the second family two Cypriot folk dances, the *1st* and *2nd Antikristos* (on the right). The frequency of motion and style motifs is illustrated by the colors (hot colors for high frequency, cold colors for low frequency). As can be seen, dances that belong to similar families have similar distribution of motifs in their motion and style signatures.

influence of the similarity descriptors, in motion comparison, so as to tilt the sensitivity toward posture or stylistic correlation. Recall that signatures are independent of the length of the motion sequence and can be applied to sequences from a few seconds to several minutes. This allows to understand that two sequences belong to the same semantic group even if they differ in length, and without requiring temporal alignment or exact matching.

It is important to note that important metadata information, such as the accompanied music, the story-telling, etc., is not taken into consideration in the distance metric in this version of the work. As stated in section 7, our future directions will see the enrichment of our semantic measurement metric with a larger number of metadata information, taking into consideration audio features (e.g., the music rhythm), semiological information (e.g., the story-telling), as well as the social, cultural, economical, and religion aspects of the dances.

5 CATEGORIZATION TREE

Given a collection of folk dances $\mathcal{D}$, our target is to organize and visualize them based on their similarity. One way to deal with is to apply dimensionality reduction [46] on their signatures or to directly cluster the collection using a distance metric [37]. In this paper, and similarly to Huang *et al.* [49], we first organize dances into a categorization tree that assesses dances on the basis of their contextual similarity, and later set parental-children relations among dances based on their ancestry to portray the evolution of dances. The categorization tree is defined by utilizing two different measurements, the geometric and/or stylistic relations between dances, while the parent-child relationships are implied by imposing semantic information.

The main idea behind structuring a categorization tree is to denote any sets of four dances in the collection as *quadruplets*, and then perform a series of tests on quadruplets to find a subset of *quartets*. A quartet is a sub-tree that consists of four leaves, and expresses a quantitative relation among a set of nodes that can be used as a constraint for the construction of the categorization tree. More specifically, it defines the topological constraints between two pairs of dances so that the distance between the elements of each pair is small, but the distance to the elements of the other pair is large. In other words, a quartert separates two pairs of dances. A set of quartets are then used to build one global tree, where all the dances in the collection $\mathcal{D}$ reside at the leaves of the tree, and the number of edges between two leaves reflects their degree of separation within the given collection. In a similar way, Chen *et al.* [21] used phylogenetic tree built from quartets to summarize a motion sequence into poses; the distances between poses were calculated using some low-level and high-level relative geometry features. In contrast, in our work we do not look only at pose but use dances as whole sequence of motions, whereas the distances among dances are calculated using time-scale and temporal-order measurements that take into consideration both the geometric and stylistic features of the dance as a sequence.

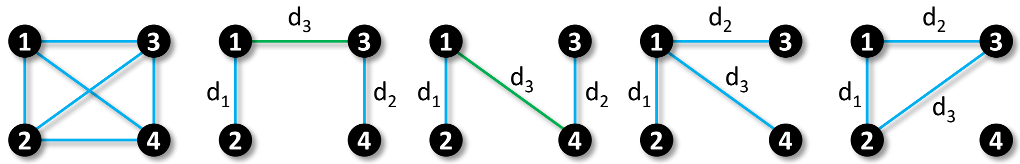

Fig. 5. (a) A fully connected graph with six edges. After removing the three edges with the largest distance, there are several possible configurations. For instance, (b) and (c) are potential reliable quartets since d_3 is a bridge connecting the two pairs, while the (d) and (e) configurations are discarded since the removal of d_3 do not separate the four vertices into two pairs.

5.1 Defining quartets

The first step towards building the categorization tree is to select appropriate and reliable quartets. To form quartets, we compute the distance between signatures of the dances (both motion and/or style signatures), as use it as the dis-similarity value among the dances; the clearer the topological structure of this separation is, the more reliable the quartet is. The similarity value used in our experiments can be manually tuned to weight the influence of each of the descriptor features and change the sensitivity toward posture or stylistic correlation. The process of acquiring a reliable quarter is:

(1) Select any four different dances (A, B, C, D) from $\mathcal{D}$. For each pair of dances in the collection, calculate the distances between their signatures (geometric, stylistic, or combine), and construct a full 4-node connected graph. Each node of the graph represents a dance sequence and each edge is associated with the distance between the two connected nodes.
(2) Sort the six edges according to their distance similarity, and remove the three edges with the largest distance. Check whether the four nodes are still connected, and if not, discard this quartet. If all nodes are connected, let the largest edge of the three remaining edges be d_3 and the other two edges be d_1 and d_2. Check if d_3 is a bridge, where its removal separates the four nodes into two pairs, and if not, discard this quartet; see Figure 5.
(3) Compute the ratio between the distance values of the edges: if $d_3/d_1 > R$ and $d_3/d_2 > R$, where R is a user defined threshold, then this quartet is consider to be a reliable quartet. Thus, the topological constraints ensure that the elements of each pair of dances which are connected by the edges d_1 and d_2, respectively, are close, while the two pairs are far apart.

We used the $k = 20$ nearest neighbors while searching for quartet candidates. The threshold R, that is the ratio between the value of inner to inter pair distances in a quartet, controls the number of reliable quartets. Setting a small value for R creates a large number of quartets, at the cost of having less reliable quartets, while setting the value of R too large, only a small number of conservative quarters pass through the reliability test, which again reduces the overall accuracy of the tree. In this work, we adjusted the value of $R = 1.4$ so as to allow the minimum number of quarters that contain all the dances in the collection.

5.2 Building the categorization tree

After obtaining a set of reliable quartets, and similarly to [49], we adopt the Quartets MaxCut algorithm (QMC) [91] to construct the categorization tree. Every quartet defines a topological relation between two pairs of dances that must be maintained by the categorization tree. The QMC algorithm approximates the categorization tree (which is unrooted tree), that maximally preserves the topological constrains of the dance sequences defined by the quartets (see Figure 6). Each quartet becomes a sub-tree and each leaf in the tree represents a dance motion. If two leafs have the same parent node, it indicates that these two leafs are very similar to each other. If two non-leaf nodes have the same parent node, it indicates their sub-trees are similar to each other to some degree.

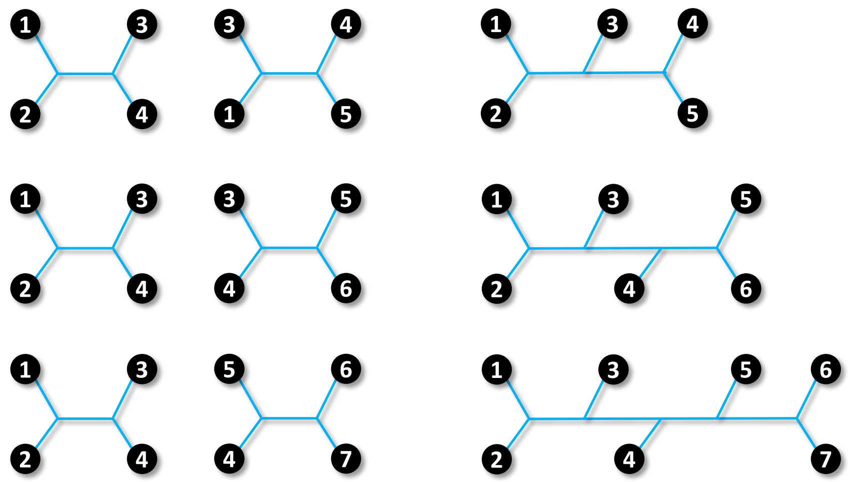

Fig. 6. Building a tree based on different quartets using the MaxCut algorithm. In each row, the input is a set of two quartets, and the algorithm creates a tree respecting these quartets topology. More details can be found in [91]

The degree of separation of any two dances in the collection can be represented by the number of edge hops between their corresponding leaves in the tree. Figure 14 shows the categorization tree built from our dance motion collection.

5.3 Set the hierarchical structure

To put dances in a chronological or geographical order, apart from the distance matrix that gives the similarity between all dances, we also need to define *semantic links* from metadata information to link dances hierarchically, and establish parental-to-children or siblings relationships. The semantic links use metadata information and are meant to assist the hierarchical structuring of the tree with regard to the similarity between the dance sequences, the date they first appeared, and their origin. In this way, we can study or visualize the evolution of dance, and unveil different variations of the same dance in neighboring nations and areas, or at different time periods. We use the following metadata to establish our semantic links: (a) date/year, (b) origin, (c) male/female/mix, and (d) solo/group. The date/years information gives the parental-to-child relation, while the gender, origin and type, allow to refine, prune, group and/or separate the dances (e.g., brake all connections between dances performed by actors of the opposite gender).

6 RESULTS

In this section, we provide several experiments to evaluate the performance of our motion and style analysis, and illustrations that demonstrate the effectiveness of our method.

6.1 Implementation Details

We have implemented our system in Matlab R2018b, the Siamese network was implemented in torch, while trees were generated in C++. All experiments, including the network training, were run on a six-core PC with Intel i7-6850K at 3.6GHz, 32GB RAM, and with nVIDIA Titan XP GPU. We used the trained triplet network of Aristidou *et al.* [5], that has been trained using a large number of motion words from various types of motion, including dance motions. It took approximately seven hours to embed the motion words into a high dimensional

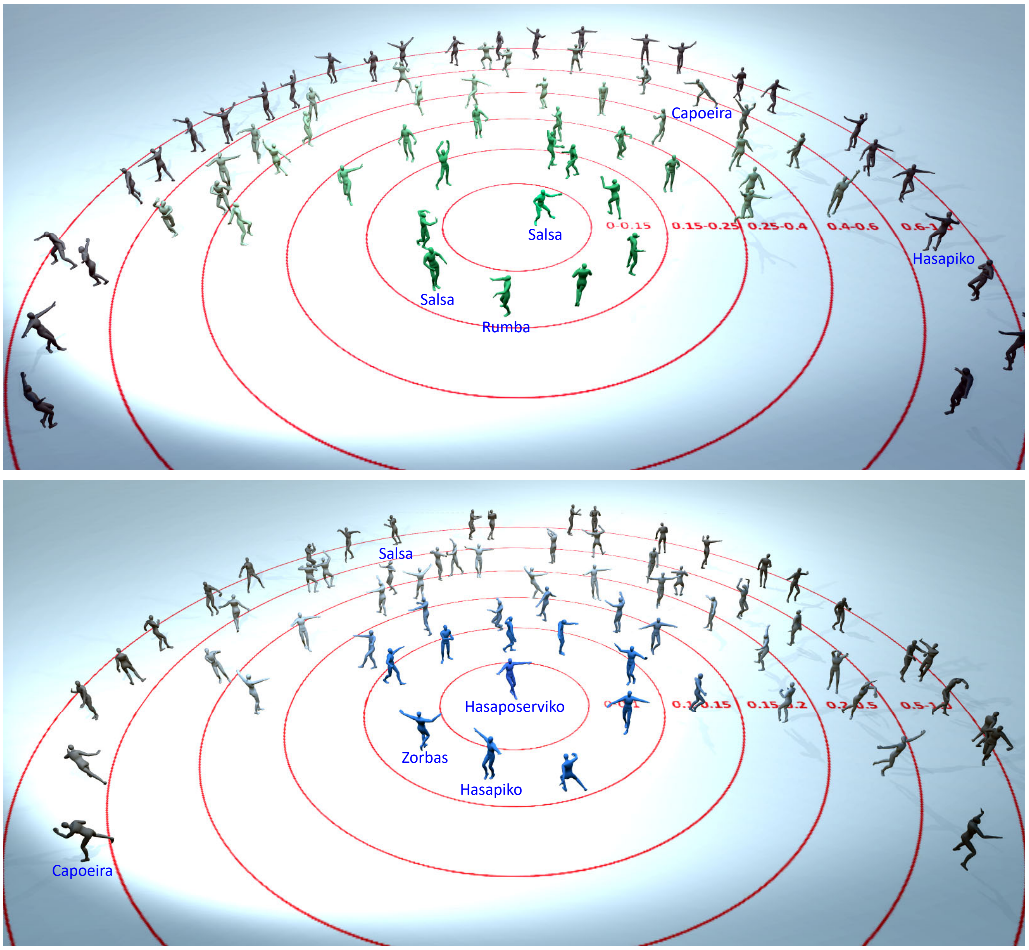

Fig. 7. This figure demonstrates, in a circular partition, the degree of similarity of the *Salsa* (top) and *Hasaposerviko* (bottom) dances to other dances in the collection. Similar dances to the query dance are placed closer to the center circle. The similarity is also illustrated by different shades of green (for Salsa) and blue (for Hasaposerviko); the numbers in red indicate the degree of dissimilarity for that partition. Refer to the supplementary video for an animated visualization of this demonstration.

feature space for our dance motion database $\mathcal{D}$. For style words, it only took less than a couple of hours to create the embedding. Clustering words (approximately 25K motion and 25K style words) into mutually common clusters requires approximately 2 hours for each method, creating motion and style signatures, that is the distribution of these words in a motion sequence. Motion retrieval and comparison is then achieved in real-time.

6.2 Evaluation

Figure 14 presents the categorization tree where dance motions are organized, producing reliable neighbors for each dance motion. Its performance is demonstrated by its effectiveness in selecting reliable quartets as topological constraints, indicating that the use of motion and style signatures can be a reliable motion representation. We evaluate the importance of our geometric and stylistic measurements by constructing trees using only motion

Fig. 8. Our method finds similarities in the dance culture of neighboring countries in the Caribbean.

or style signatures as the distance metric, and conclude that using an equally weighted metric matches better the recommendations of our associated dance professionals. More specifically, we asked our collaborators (five dance experts) to group our dance collection (by indicating the 5-Nearest Neighbors for each dance from those dances in our database), and compare our results with their suggestions. The accuracy of our method in finding the 5 contextually closest dances, for a given dance, matches the suggestions made by our dance specialists by approximately 87% (76% when using only the motion signatures, and 71% when the style signatures were used); the accuracy increases to 90% when comparing the 3 closest dances, and 93% for the single most related dance. For further analysis purposes, we also portray our results using some other popular clustering and ordering methods, such as the Multi-Dimensional Scaling (MDS) [86] in Figure 11, where dances are ordered into 2 dimensional space and similar objects are near and dissimilar are far apart, the t-Distributed Stochastic Neighbor Embedding (t-SNE) [99] in Figure 12, and the hierarchical clustering [72] in Figure 13, that gathers similar dances into groups.

Organizing a dance collection into a tree or an embedding allows fast and intuitive exploration. In this work, we additionally visualize the similarity of a given dance to the collection in a two-dimensional neighbor map. The distance between the given dance and the collection is computed by equally weighting the EMD distance of their motion and style signatures. The 2D map facilitates the exploration of the collection, placing the nearest neighbors to the given dance closer to the center of the 2D map, and the dissimilar dances further away. Moreover, the closer the dances are to the center of the 2D map, the shade of color is more similar to the given dance, while as we move further from the circle center, the color fades and becomes gray. Figure 7 illustrates two examples where the dance collection has been reordered around a given dance by their degree of similarity in circular partitions. The top image shows the relation of the *Salsa* dance (green color) to the collection (similar dances are placed closer to the given dance and colored in green shade), and the bottom image the corresponding relation of the *Hasaposerviko* dance (blue color) to the collection (similarity is also highlighted in different shades of blue). For instance, note how the *Zorbas* and *Hasapiko* dances can be found close to the input dance (*Hasaposerviko*), while dances with Latin or Carribean origin (e.g., *Capoeira*, *Salsa*) are placed at the outer partition, meaning that they have a large degree of dissimilarity to the input dance. This is also illustrated in Figures 11, 12, 13, and 14.

All of these methods give you the sense of similarity or the differences between dances, but no method offers a visualization to semantically perceive the homogeneity of dances, their evolution in time, and their inter or meta-influences. To illustrate the chronological and geographical evolution of dance, we use a constrained version of the well-known Motion Graphs method [62], that allows finding transitions points at similar poses. The use of motion and style signatures add contextual constraints, allowing to connect only dances with similar content (their degree of dissimilarity is less than a threshold) that go beyond the body's postural configuration. The addition of semantic information in the process ensures that the selected transition is made with chronological or geographical order. This is achieved by embedding a contextual assessment that prunes incoherent transitions. In particular, we compare the postural and stylistic contents of the input and the candidate motions near the transition frame (anchored on the transition words), by using a window of $n = 12$ motion and style words (that is a 2.5 seconds of motion), and discard the transitions that have large degree of dissimilarity. For instance, to illustrate the chronological evolution of a given dance (e.g., the Cypriot *1st Antikristos*), we build a graph with

Fig. 9. Our method reveals some unexpected correlations, such as those between the Chinese *Xin-Jiang* (shown in red) and the Egyptian *Belly* dance (shown in yellow); both dances have Oriental roots and influences.

all possible transitions that are contextually similar (above a predefined threshold) to the given dance, and are initialized in a later time. This may result in numerous paths; the selection of the optimal or ideal path depends on the case. Here, we select the longest path with the lowest mean contextual cost. Figure 1 illustrates an example of the evolution over time in the Cypriot folk dancing, while Figure 8 shows an example of related dances with similar geographical orientation. For an animated version, please refer to our supplementary video.

6.3 Discussion

Our method, apart from finding similar dances (both ethnically, and geographically) and placing them near each other, can also help in revealing some interesting observations. More specifically, there are motion motifs that appear in different types of dances and create unexpected connections; for instance, we found similarities in the motion and style signatures of the Serbian (e.g., *Kolo*) and Greek folk dances (e.g., *Podaraki*, *Rasopoulos*). All these are group dances where the performers are forming a circle, holding each other's hands or having their hands around each other's waists. Another interesting association is that between the Chinese *Xin-Jiang* dance (taken from the NUS [77]) and the Egyptian dance (*Belly* dance). The dances at the Xin-Jiang region in China have Oriental (Middle East) influences due to some ethnic minority groups, e.g., the Turkic Uyghur people. Both aforementioned connections are also confirmed by our associated folklore specialists. Figure 9 shows a short sequence of the Chinese *Xin-Jiang* and *Belly* dance, where the similarities in motion between the two dances is obvious. Please refer to the supplementary video for an animated version.

The importance of using both the motion and style signatures is demonstrated when comparing similar dances e.g., the *Tsamikos* and the *Syrtos sta tria*, or the *Hasapiko* and *Zorbas* dances. Both pairs have analogous distribution of similar motions (the geometry of their movements), but their style (e.g., the speed) is different. Our equally weighted metric allows to find the similarities in motion, but also highlights the differences in rhythm between the two dances. Figure 10 shows snapshots of the *Tsamikos* and the *Syrtos sta tria* dances, and their corresponding motion and style signatures. An animated example is given in the supplementary video.

It is important to note that the correlations found in our experiments are highly depended on the limited size of our database. For instance, *modern* dance is not contextually related to *Capoeira* (see Figures 11, 12, 13, and 14), but is the most similar dance to it from those in our database. As our dataset will grow and become enriched with a larger number of dances, we believe that more observations, connections and links could be found between dances. We look forward to algorithmically unveil more unexpected correlations and insights among the dance cultural heritage of many neighboring countries.

7 CONCLUSIONS

In this paper we present a descriptive representation for dance motion sequences, based on the BoW principles, to contextually analyze and organize large dance motion libraries. We use motion capture to acquire a large number of folk, modern, and other dances. We then divide motions into short temporal descriptors, taking into account both the dancer's body geometry and style, creating a high-dimensionality embedding in feature space. The descriptors are then clustered into groups based on their similarity, and their distribution along a dance sequence

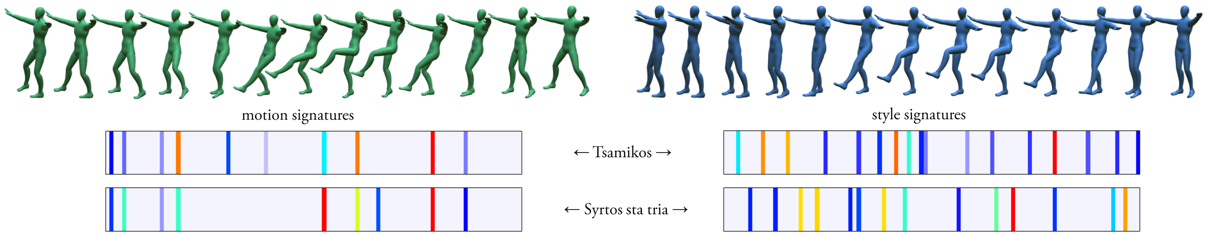

Fig. 10. The *Tsamikos* (shown in green) and the *Syrtos sta tria* (shown in blue) dances have similar distribution of motion words but they differ in style (rhythm), as illustrated by their corresponding motion and style signatures.

defines signatures that characterize the content of that movement. Signatures offer a succinct but descriptive representation of motion sequences, that is time-scale and temporal-order invariant. Similar dance sequences have a similar distribution of their motion and style descriptors. We demonstrated the use of our dance motion representation as an efficient tool for organizing dances by assembling a categorization tree, while the use of semantic links inferred from metadata information helped in structuring dances in a chronological order.

Our work, however, has some limitations. First, only a small subset of dances has been considered, with the structure of the organizing tree not being adequate. Forming the genealogical tree of dances is an ultimate target, but a much more complete database is required. We look forward to enrich the DMCD with a larger number of dances, from many different countries, aiming to put the foundations for creating the first virtual museum of dances. The categorization tree is scalable so new dance data can be added and organized as they became available. The DMCD should also be expanded to store other metadata information to enhance the dance collection, such as clothes, facial painting, hair fashion etc. Having no access to the original version of the dance, as it was performed in the past, our dance motion data may be a good approximation but differs from its original version; unfortunately, some aspects of our ICH may be forever lost. We also aim at capturing various versions of each dance, from multiple performers, so the performer's personal experience and talent will be factored out of our observations, as well as recorded as stylistic element. Second, the structuring of our categorization tree still relies on the quality of our computational methods (i.e., our motion and style signatures), which may create conflicting quartets. Finally, our trees are not structured taking into account the sociopolitical influences, the music, facial expressions, or the story-telling of dances, but is based only on the geometric and stylistic characteristics. Future work will see the introduction and evaluation of other similarity links (e.g., semiological links) that take into consideration the social and cultural aspects of dance, the clothes, the music and/or other metadata information.

ACKNOWLEDGMENTS

The authors would like to thank the Cyprus Folklore Research Association, the dancers and actors (in alphabetical order) Fanie Alexandrou, Vaso Aristeidou, Anna Charalambous, Anna Cortesi, Andreas de la Vara, Clio de la Vara, Maritsa Elia, Stefanos Koullapis, Olivia Kyriakides, Nicos Michael, Ioanna Tacha, Stefanos Theodorou, and Theodora Tsiakka who performed in our lab. This work has been supported by the RESTART 2016-2020 Programmes for Technological Development and Innovation, through the Cyprus Research Promotion Foundation, with protocol number P2P/JPICH_DH/0417/0052. It has also been partly supported by the project that has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 739578 (RISE-Call: H2020-WIDESPREAD-01-2016-2017-TeamingPhase2) and the Government of the Republic of Cyprus. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

REFERENCES

[1] Dimitrios S. Alexiadis and Petros Daras. 2014. Quaternionic signal processing techniques for automatic evaluation of dance performances from MoCap data. *IEEE Transactions on Multimedia* 16, 5 (Aug. 2014), 1391–1406.

[2] Dimitrios S. Alexiadis, Philip Kelly, Petros Daras, Noel E. O'Connor, Tamy Boubekeur, and Maher Ben Moussa. 2011. Evaluating a Dancer's Performance Using Kinect-based Skeleton Tracking. In *Proceedings of the 19th ACM International Conference on Multimedia (MM '11)*. ACM, New York, NY, USA, 659–662.

[3] Okan Arikan, David A. Forsyth, and James F. O'Brien. 2003. Motion Synthesis from Annotations. *ACM Trans. Graph.* 22, 3 (July 2003), 402–408.

[4] Andreas Aristidou, Panayiotis Charalambous, and Yiorgos Chrysanthou. 2015. Emotion Analysis and Classification: Understanding the Performers' Emotions Using the LMA Entities. *Comput. Graph. Forum* 34, 6 (Sept. 2015), 262–276.

[5] Andreas Aristidou, Daniel Cohen-Or, Jessica K. Hodgins, Yiorgos Chrysanthou, and Ariel Shamir. 2018. Deep Motifs and Motion Signatures. *ACM Trans. Graph.* 38, 6, Article 187 (Nov. 2018), 187:1–187:13 pages.

[6] Andreas Aristidou, Daniel Cohen-Or, Jessica K. Hodgins, and Ariel Shamir. 2018. Self-similarity Analysis for Motion Capture Cleaning. *Comput. Graph. Forum* 37, 2 (May 2018), 297–309.

[7] Andreas Aristidou, Efstathios Stavrakis, Panayiotis Charalambous, Yiorgos Chrysanthou, and Stephania Loizidou-Himona. 2015. Folk Dance Evaluation Using Laban Movement Analysis. *J. Comput. Cult. Herit.* 8, 4 (Aug. 2015), 20:1–20:19.

[8] Andreas Aristidou, Efstathios Stavrakis, and Yiorgos Chrysanthou. 2014. LMA-Based Motion Retrieval for Folk Dance Cultural Heritage. In *5th International Conference on Cultural Heritage* (November 3 - 8) *(EuroMed'14)*, Marinos Ioannides, Nadia Magnenat-Thalmann, Eleanor Fink, Roko Žarnić, Alex-Yianing Yen, and Ewald Quak (Eds.), Vol. LNCS, 8740. Springer International Publishing, 404–413.

[9] Andreas Aristidou, Efstathios Stavrakis, Margarita Papaefthimiou, George Papagiannakis, and Yiorgos Chrysanthou. 2017. Style-based Motion Analysis for Dance Composition. *The Visual Computer* (2017).

[10] Andreas Aristidou, Efstathios Stavrakis, Margarita Papaefthimiou, George Papagiannakis, and Yiorgos Chrysanthou. 2018. Style-based Motion Analysis for Dance Composition. *Vis. Comput.* 34, 12 (2018), 1725–1737. https://doi.org/10.1007/s00371-017-1452-z

[11] Andreas Aristidou, Qiong Zeng, Efstathios Stavrakis, Kangkang Yin, Daniel Cohen-Or, Yiorgos Chrysanthou, and Baoquan Chen. 2017. Emotion Control of Unstructured Dance Movements. In *Proceedings of the ACM SIGGRAPH / Eurographics Symposium on Computer Animation (SCA '17)*. ACM, New York, NY, USA, 9:1–9:10.

[12] Murtha Baca. 2014. *Introduction to Metadata* (third edition ed.).

[13] Philippe Beaudoin, Stelian Coros, Michiel van de Panne, and Pierre Poulin. 2008. Motion-motif Graphs. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '08)*. Eurographics Association, 117–126.

[14] Jürgen Bernard, Eduard Dobermann, Anna Vögele, Björn Krüger, Jörn Kohlhammer, and Dieter Fellner. 2017. Visual-Interactive Semi-Supervised Labeling of Human Motion Capture Data. In *Proceedings of Visualization and Data Analysis (VDA '17)*. Society for Imaging Science and Technology, 34–45.

[15] Jürgen Bernard, Nils Wilhelm, Björn Krüger, Thorsten May, Tobias Schreck, and Jörn Kohlhammer. 2013. MotionExplorer: Exploratory Search in Human Motion Capture Data Based on Hierarchical Aggregation. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Dec. 2013), 2257–2266.

[16] Leonid Blouvshtein and Daniel Cohen-Or. 2018. Outlier Detection for Robust Multi-dimensional Scaling. *IEEE Trans. Pattern Anal. Mach. Intell.* (2018).

[17] Jinxiang Chai and Jessica K. Hodgins. 2005. Performance Animation from Low-dimensional Control Signals. *ACM Trans. Graph.* 24, 3 (July 2005), 686–696.

[18] Jacky C.P. Chan, Howard Leung, Jeff K.T. Tang, and Taku Komura. 2011. A Virtual Reality Dance Training System Using Motion Capture Technology. *IEEE Transactions on Learning Technologies* 4, 2 (April 2011), 187–195.

[19] Giannis Chantas, Sotiris Karavarsamis, Spiros Nikolopoulos, and Yiannis Kompatsiaris. 2018. A Probabilistic, Ontological Framework for Safeguarding the Intangible Cultural Heritage. *J. Comput. Cult. Herit.* 11 (2018).

[20] Min-Wen Chao, Chao-Hung Lin, Jackie Assa, and Tong-Yee Lee. 2012. Human Motion Retrieval from Hand-Drawn Sketch. *IEEE Transactions on Visualization and Computer Graphics* 18, 5 (May 2012), 729–740.

[21] Songle Chen, Zhengxing Sun, and Yan Zhang. 2015. Scalable Organization of Collections of Motion Capture Data via Quantitative and Qualitative Analysis. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval (ICMR '15)*. 411–418.

[22] Suepphong Chernbumroong, Pradorn Sureephong, Karim Tabia, and Yoeven D. Khemlani. 2018. Thai Dance Training Game-Based Model. In *Proceedings of the 3rd International Conference on Digital Arts, Media and Technology (ICDAMT'18)*. IEEE, 194–197.

[23] Diane Chi, Monica Costa, Liwei Zhao, and Norman Badler. 2000. The EMOTE Model for Effort and Shape. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '00)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 173–182.

[24] Myung G. Choi, Kyungyong Yang, Takeo Igarashi, Jun Mitani, and Jehee Lee. 2012. Retrieval and Visualization of Human Motion Data via Stick Figures. *Comput. Graph. Forum* 31, 7 (2012), 2057–2065.

[25] CMU. 2019. Carnegie Mellon University MoCap Database: http://mocap.cs.cmu.edu/. (2019).
[26] Bruce Croft, Donald Metzler, and Trevor Strohman. 2009. *Search Engines: Information Retrieval in Practice* (1st ed.). Addison-Wesley Publishing Company, USA.
[27] Liqun Deng, Howard Leung, Naijie Gu, and Yang Yang. 2011. Real-time mocap dance recognition for an interactive dancing game. *Comput. Animat. Virtual Worlds* 22, 2-3 (2011), 229–237.
[28] Zhigang Deng, Qin Gu, and Qing Li. 2009. Perceptually Consistent Example-based Human Motion Retrieval. In *Proceedings of the Symposium on Interactive 3D Graphics and Games (I3D '09)*. 191–198.
[29] DMCD. 2019. Dance Motion Capture Database: http://dancedb.cs.ucy.ac.cy/. (2019).
[30] Anastasios Doulamis, Athanasios Voulodimos, Nikolaos Doulamis, Sofia Soile, and Anastasios Lampropoulos. 2017. Transforming Intangible Folkloric Performing Arts into Tangible Choreographic Digital Objects: The Terpsichore Approach. In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 5, (VISAPP)*. INSTICC, SciTePress, 451–460.
[31] Katerina El Raheb and Yannis Ioannidis. 2012. A Labanotation Based Ontology for Representing Dance Movement. In *Proceedings of the 9th International Conference on Gesture and Sign Language in Human-Computer Interaction and Embodied Communication (GW'11)*. Springer-Verlag, Berlin, Heidelberg, 106–117.
[32] Katerina El Raheb and Yannis Ioannidis. 2013. Dance in the World of Data and Objects. In *Information Technologies for Performing Arts, Media Access, and Entertainment*, Paolo Nesi and Raffaella Santucci (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 192–204.
[33] Katerina El Raheb, Aristotelis Kasomoulis, Akrivi Katifori, Marianna Rezkalla, and Yannis Ioannidis. 2018. A Web-based System for Annotation of Dance Multimodal Recordings by Dance Practitioners and Experts. In *Proceedings of the 5th International Conference on Movement and Computing (MOCO '18)*. ACM, New York, NY, USA, Article 8, 8:1–8:8 pages.
[34] Katerina El Raheb, Nicolas Papapetrou, Vivi Katifori, and Yannis Ioannidis. 2016. BalOnSe: Ballet Ontology for Annotating and Searching Video Performances. In *Proceedings of the 3rd International Symposium on Movement and Computing (MOCO '16)*. ACM, New York, NY, USA, 5:1–5:8.
[35] Katerina El Raheb, George Tsampounaris, Akrivi Katifori, and Yannis Ioannidis. 2018. Choreomorphy: A Whole-body Interaction Experience for Dance Improvisation and Visual Experimentation. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces (AVI '18)*. ACM, New York, NY, USA, 27:1–27:9.
[36] Slim Essid, Dimitrios Alexiadis, Robin Tournemenne, Marc Gowing, Philip Kelly, David Monaghan, Petros Daras, Angelique Drémeau, and Noel E. O'Connor. 2012. An advanced virtual dance performance evaluator. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2269–2272.
[37] Brian S. Everitt, Sabine Landau, and Morven Leese. 2009. *Cluster Analysis* (4th ed.). Wiley Publishing.
[38] Kevin Forbes and Eugene Fiume. 2005. An Efficient Search Algorithm for Motion Data Using Weighted PCA. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '05)*. 67–76.
[39] Eftychia Fotiadou, Ioannis Kapsouras, Nikos Nikolaidis, and Anastasios Tefas. 2016. A Bag of Words Approach for Recognition of Greek Folk Dances. In *Proceedings of the 9th Hellenic Conference on Artificial Intelligence (SETN '16)*. ACM, New York, NY, USA, 21:1–21:4.
[40] Brendan J. Frey and Delbert Dueck. 2007. Clustering by passing messages between data points. *Science* 315 (2007), 972–976.
[41] Satoru Fukayama and Masataka Goto. 2014. Automated Choreography Synthesis Using a Gaussian Process Leveraging Consumer-generated Dance Motions. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology (ACE '14)*. ACM, New York, NY, USA, 23:1–23:6.
[42] Stamatios Giannoulakis, Nicolas Tsapatsoulis, and Nikos Grammalidis. 2018. Metadata for Intangible Cultural Heritage - The Case of Folk Dances. In *13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP)*, Francisco H. Imai, Alain Trèmeau, and Josè Braz (Eds.). SciTePress, 634–645.
[43] Michael Good. 2001. MusicXML for Notation and Analysis. *The virtual score: representation, retrieval, restoration.* 12 (2001), 113–124.
[44] Nikolaos Grammalidis, Kosmas Dimitropoulos, Filareti Tsalakanidou, Alexandros Kitsikidis, Pierre Roussel, Bruce Denby, P. Chawah, Lise Buchman, Stèphane Dupont, Sohaib Laraba, Benjamin Picart, Mickaël Tits, Joëlle Tilmanne, Stelios Hadjidimitriou, L. Hadjileontiadis, V. Charisis, Christina Volioti, A. Stergiaki, Athanasios Manitsaris, Odysseas Bouzos, and Sotiris Manitsaris. 2016. The i-Treasures Intangible Cultural Heritage Dataset. In *Proceedings of the 3rd International Symposium on Movement and Computing (MOCO '16)*. ACM, New York, NY, USA, 23:1–23:8.
[45] Ann Hutchinson Guest. 2013. *Labanotation: The System of Analyzing and Recording Movement* (forth edition ed.).
[46] Joseph F. Hair Jr., Rolph E. Anderson, Ronald L. Tatham, and William C. Black. 1995. *Multivariate Data Analysis (4th Ed.): With Readings*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
[47] Jonathan Hatol. 2006. *MovementXML: A representation of semantics of human movement based on Labanotation.* Master's thesis. Interactive Arts and Technology - Simon Fraser University.
[48] Diane I. Hillmann, Rhonda Marker, and Chris Brady. 2008. Metadata Standards and Applications. *The Serials Librarian* 54 (2008), 7–21. Issue 1-2.

[49] 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 (July 2013), 71:1–71:10 pages.

[50] Minsu Jang, D. Kim, Y. Kim, and Jaehong Kim. 2017. Automated dance motion evaluation using dynamic time warping and Laban movement analysis. In *Proceedings of the IEEE International Conference on Consumer Electronics (ICCE'17)*. 141–142.

[51] Yu Jin, Xiaoxiang Hu, and GangShan Wu. 2012. A Tai Chi Training System Based on Fast Skeleton Matching Algorithm. In *Proceedings of the 12th International Conference on Computer Vision - Volume Part III (ECCV'12)*. Springer-Verlag, Berlin, Heidelberg, 667–670.

[52] Adrienne L. Kaeppler. 2000. Dance Ethnology and the Anthropology of Dance. *Dance Research Journal* 32, 1 (2000), 116–125.

[53] Rajkumar Kannan, Frederic Andres, and Christian Guetl. 2010. DanVideo: An MPEG-7 Authoring and Retrieval System for Dance Videos. *Multimedia Tools Appl.* 46, 2-3 (Jan. 2010), 545–572.

[54] Mubbasir Kapadia, I-kao Chiang, Tiju Thomas, Norman I. Badler, and Joseph T. Kider, Jr. 2013. Efficient Motion Retrieval in Large Motion Databases. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D '13)*. 19–28.

[55] Ioannis Kapsouras and Nikos Nikolaidis. 2014. Action Recognition on Motion Capture Data Using a Dynemes and Forward Differences Representation. *J. Vis. Comun. Image Represent.* 25, 6 (Aug. 2014), 1432–1445.

[56] Eamonn Keogh, Themistoklis Palpanas, Victor B. Zordan, Dimitrios Gunopulos, and Marc Cardle. 2004. Indexing Large Human-motion Databases. In *Proceedings of the International Conference on Very Large Data Bases (VLDB '04)*. 780–791.

[57] 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).

[58] Eugenia S. Kim. 2011. ChoreoSave: A digital dance preservation system prototype. *Proceedings of the American Society for Information Science and Technology* 48, 1 (2011), 1–10.

[59] Alexandros Kitsikidis, Kosmas Dimitropoulos, Stella Douka, and Nikos Grammalidis. 2014. Dance analysis using multiple Kinect sensors. In *Proceedings of the International Conference on Computer Vision Theory and Applications (VISAPP)*, Vol. 2. 789–795.

[60] Alexandros Kitsikidis, Kosmas Dimitropoulos, Deniz Uğurca, Can Bayçay, Erdal Yilmaz, Filareti Tsalakanidou, Stella Douka, and Nikos Grammalidis. 2015. A Game-like Application for Dance Learning Using a Natural Human Computer Interface. In *Universal Access in Human-Computer Interaction. Access to Learning, Health and Well-Being*, Margherita Antona and Constantine Stephanidis (Eds.). Springer International Publishing, Cham, 472–482.

[61] Lucas Kovar and Michael Gleicher. 2004. Automated Extraction and Parameterization of Motions in Large Data Sets. *ACM Trans. Graph.* 23, 3 (Aug. 2004), 559–568.

[62] Lucas Kovar, Michael Gleicher, and Frédéric Pighin. 2002. Motion Graphs. *ACM Trans. Graph.* 21, 3 (July 2002), 473–482.

[63] Björn Krüger, Jochen Tautges, Andreas Weber, and Arno Zinke. 2010. Fast Local and Global Similarity Searches in Large Motion Capture Databases. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '10)*. Eurographics Assoc., 1–10.

[64] Matthew Kyan, Guoyu Sun, Haiyan Li, Ling Zhong, Paisarn Muneesawang, Nan Dong, Bruce Elder, and Ling Guan. 2015. An Approach to Ballet Dance Training Through MS Kinect and Visualization in a CAVE Virtual Reality Environment. *ACM Trans. Intell. Syst. Technol.* 6, 2 (March 2015), 23:1–23:37.

[65] Sohaib Laraba and Joëlle Tilmanne. 2016. Dance Performance Evaluation Using Hidden Markov Models. *Comput. Animat. Virtual Worlds* 27, 3-4 (May 2016), 321–329.

[66] Jehee Lee, Jinxiang Chai, Paul S. A. Reitsma, Jessica K. Hodgins, and Nancy S. Pollard. 2002. Interactive Control of Avatars Animated with Human Motion Data. *ACM Trans. Graph.* 21, 3 (July 2002), 491–500.

[67] Jongwook Lee and Boncheol Goo. 2014. A study on the e-learning for Korean traditional dance. In *Proceedings of the International Conference on Virtual Systems Multimedia (VSMM)*. 240–245.

[68] Guodong Liu, Jingdan Zhang, Wei Wang, and Leonard McMillan. 2005. A System for Analyzing and Indexing Human-motion Databases. In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '05)*. 924–926.

[69] Irene Lourdi, Christos Papatheodorou, and Mara Nikolaidou. 2007. A Multi-layer Metadata Schema for Digital Folklore Collections. *J. Inf. Sci.* 33, 2 (April 2007), 197–213.

[70] Nadia Magnenat-Thalmann, Dimitrios Protopsaltou, and Evangelia Kavakli. 2008. Learning How to Dance Using a Web 3D Platform. In *Proceedings of the 6th International Conference on Advances in Web Based Learning (ICWL'07)*. Springer-Verlag, Berlin, Heidelberg, 1–12.

[71] Anupama Mallik, Santanu Chaudhury, and Hiranmay Ghosh. 2011. Nrityakosha: Preserving the Intangible Heritage of Indian Classical Dance. *J. Comput. Cult. Herit.* 4, 3 (Dec. 2011), 11:1–11:25.

[72] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval.* Cambridge University Press, New York, NY, USA.

[73] Meinard Müller, Andreas Baak, and Hans-Peter Seidel. 2009. Efficient and Robust Annotation of Motion Capture Data. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '09)*. 17–26.

[74] Meinard Müller and Tido Röder. 2006. Motion Templates for Automatic Classification and Retrieval of Motion Capture Data. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '06)*. Eurographics Association, 137–146.

[75] Meinard Müller, Tido Röder, and Michael Clausen. 2005. Efficient Content-based Retrieval of Motion Capture Data. *ACM Trans. Graph.* 24, 3 (July 2005), 677–685.
[76] Minako Nakamura and Kozaburo Hachimura. 2006. An XML representation of labanotation, LabanXML, and its implementation on the notation editor labaneditor2. *Review of the National Center for Digitization* 9 (2006), 47–51.
[77] NUS. 2019. National University of Singapore MoCap Database: http://animation.comp.nus.edu.sg/nusmocap.html. (2019).
[78] PhaseSpace Inc. 2018. Optical Motion Capture Systems: http://www.phasespace.com, Accessed: 01/2019. (2018).
[79] Thanawin Rakthanmanon, Bilson Campana, Abdullah Mueen, Gustavo Batista, Brandon Westover, Qiang Zhu, Jesin Zakaria, and Eamonn Keogh. 2012. Searching and Mining Trillions of Time Series Subsequences Under Dynamic Time Warping. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '12)*. ACM, New York, NY, USA, 262–270.
[80] Ioannis Rallis, Nikolaos Doulamis, Anastasios Doulamis, Athanasios Voulodimos, and Vassilios Vescoukis. 2018. Spatio-temporal summarization of dance choreographies. *Computers & Graphics* 73 (2018), 88–101.
[81] Michalis Raptis, Darko Kirovski, and Hugues Hoppe. 2011. Real-time Classification of Dance Gestures from Skeleton Animation. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '11)*. 147–156.
[82] Qaiser Riaz, Björn Krüger, and Andreas Weber. 2016. Relational Databases for Motion Data. *Int. J. Innov. Comput. Appl.* 7, 3 (Jan. 2016), 119–134.
[83] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas. 2000. The Earth Mover's Distance As a Metric for Image Retrieval. *Int. J. Comput. Vision* 40, 2 (Nov. 2000), 99–121.
[84] Anindhya Sankhla, Vinanti Kalangutkar, Himadri B. G. S. Bhuyan, Tanwi Mallick, Vivek Nautiyal, Partha Pratim Das, and Arun Kumar Majumdar. 2017. Automated Translation of Human Postures from Kinect Data to Labanotation. In *Proceedings of the 6th National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics (NCVPRIPG'17)*. 494–505.
[85] Florian Schroff, Dmitry Kalenichenko, and James Philbin. 2015. FaceNet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '15)*. 815–823.
[86] George Arthur Frederick Seber. 2008. *Multivariate Observations*. John Wiley & Sons, Inc.
[87] Simon Senecal, Louis Cuel, Andreas Aristidou, and Nadia Magnenat-Thalmann. 2016. Continuous body emotion recognition system during theater performances. *Comput. Animat. Virtual Worlds* 27, 3-4 (May 2016), 311–320.
[88] Takaaki Shiratori and Katsushi Ikeuchi. 2008. Synthesis of Dance Performance Based on Analyses of Human Motion and Music. *Transactions on Computer Vision and Image Media* 1, 1 (June 2008), 34–46.
[89] Takaaki Shiratori, Atsushi Nakazawa, and Katsushi Ikeuchi. 2006. Dancing-to-Music Character Animation. *Comput. Graph. Forum* 25, 3 (September 2006), 449–458.
[90] Miguel-Angel Sicilia. 2014. *Handbook of Metadata*. World Scientific.
[91] Sagi Snir and Satish Rao. 2010. Quartets MaxCut: A Divide and Conquer Quartets Algorithm. *IEEE/ACM Trans. Comput. Biol. Bioinformatics* 7, 4 (Oct. 2010), 704–718.
[92] Efstathios Stavrakis, Andreas Aristidou, Maria Savva, Stephania Loizidou Himona, and Yiorgos Chrysanthou. 2012. Digitization of Cypriot Folk Dances. In *Proceedings of the 4th International Conference on Progress in Cultural Heritage Preservation (EuroMed'12)*. Springer-Verlag, Berlin, Heidelberg, 404–413.
[93] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '15)*. 1–9.
[94] Jeff K. T. Tang, Jacky C. P. Chan, and Howard Leung. 2011. Interactive Dancing Game with Real-time Recognition of Continuous Dance Moves from 3D Human Motion Capture. In *Proceedings of the International Conference on Ubiquitous Information Management and Communication (ICUIMC '11)*. ACM, New York, NY, USA, 50:1–50:9.
[95] Matthew Thorne, David Burke, and Michiel van de Panne. 2004. Motion Doodles: An Interface for Sketching Character Motion. *ACM Trans. Graph.* 23, 3 (Aug. 2004), 424–431.
[96] Mickaël Tits, Joëlle Tilmanne, and Thierry Dutoit. 2017. Morphology Independent Feature Engineering in Motion Capture Database for Gesture Evaluation. In *Proceedings of the International Conference on Movement Computing (MOCO '17)*. ACM, NY, USA, 26:1–26:8.
[97] Georgios Tsampounaris, Katerina El Raheb, Vivi Katifori, and Yannis Ioannidis. 2016. Exploring Visualizations in Real-time Motion Capture for Dance Education. In *Proceedings of the Pan-Hellenic Conference on Informatics (PCI '16)*. ACM, New York, NY, USA, 76:1–76:6.
[98] Unesco. 2003. Convention for the Safeguarding of the Intangible Cultural Heritage. (2003).
[99] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing Data using t-SNE. *J. Mach. Learn. Research* 9 (2008), 2579–2605.
[100] Yingying Wang and Michael Neff. 2015. Deep Signatures for Indexing and Retrieval in Large Motion Databases. In *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games (MIG '15)*. ACM, New York, NY, USA, 37–45.
[101] Yang Yang, Howard Leung, Lihua Yue, and Liqun Deng. 2013. Generating a two-phase lesson for guiding beginners to learn basic dance movements. *Computers & Education* 61 (2013), 1–20.
[102] Peng Ye and Yaolin Zhou. 2013. The Framework of Chinese Intangible Cultural Heritages Metadata Standards. In *Proceedings of the 2013 International Workshop on Computer Science in Sports (IWCSS'18)*. Atlantic Press, 194–197.

APPENDIX

A Dance Collection

The dance motion capture collection used in our experiments consists of 72 dances from 17 different countries. All these dances are grouped by their country of origin and listed in Table 2.

Table 2. The dance metadata information used in our experiments. Note that, the dates for some dances (e.g., the folk dances in Greece, Cyprus and Serbia) are indicative and approximate; folk dances are in general experiential, expressing the sociopolitical and cultural aspects of the local community, thus there is not enough data regarding the exact date they have been initialized.

	Performer's Gender	Dance Type	Origin		Date
			Country	Region	
Tango	Male	Couple	Argentina	River Plate	1880's
Tango	Female	Couple	Argentina	River Plate	1880's
Capoeira	Male	Solo	Brazil		16th century
Capoeira	Male	Solo	Brazil		16th century
XinJiang Dance	Female	Solo	China		10th century
XinJiang Dance	Female	Solo	China		10th century
Zumba	Female	Solo	Colombia		1990's
Rumba	Male	Couple	Cuba		1930's
Rumba	Male	Couple	Cuba		1930's
Rumba	Female	Couple	Cuba		1930's
Rumba	Female	Couple	Cuba		1930's
Salsa	Male	Couple	Cuba		1970's
Salsa	Male	Couple	Cuba		1970's
Salsa	Female	Couple	Cuba		1970's
Salsa	Female	Couple	Cuba		1970's
1st Antikristos	Male	Solo	Cyprus		17th century
2nd Antikristos	Male	Solo	Cyprus		17th century
3rd Antikristos	Male	Solo	Cyprus		17th century
Tatsia	Male	Solo	Cyprus		17th century
Zeibekiko	Male	Solo	Cyprus		18th century
Zeibekiko - Modern	Male	Solo	Cyprus		20th century
Zeibekiko - Modern	Female	Solo	Cyprus		20th century
Bachata	Male	Solo	Dominican Rep.		1960's
Bachata	Female	Solo	Dominican Rep.		1960's
Belly Dance	Female	Solo	Egypt		5th century
Belly Dance	Female	Solo	Egypt		5th century
Modern - Excited	Female	Solo	E.U.		1900's
Modern - Excited	Female	Solo	E.U.		1900's
Modern - Happy	Female	Solo	E.U.		1900's
Modern - Happy	Female	Solo	E.U.		1900's
Modern - Sad	Female	Solo	E.U.		1900's
Modern - Sad	Female	Solo	E.U.		1900's
Modern - Tired	Female	Solo	E.U.		1900's
Modern - Tired	Female	Solo	E.U.		1900's
Ballet	Female	Solo	France		19th century

Ballet	Female	Solo	France		19th century
Waltz	Male	Couple	Germany		16th century
Waltz	Female	Couple	Germany		16th century
Waltz	Female	Couple	Germany		16th century
Dimitroula	Female	Group	Greece	Macedonia	15th century
Haniotikos	Female	Group	Greece	Crete	10th century
Hasapiko	Female	Group	Greece	Thrace	12th century
Hasaposerviko	Male	Group	Greece	Thrace	15th century
Laziotikos	Female	Group	Greece	Crete	15th century
Maleviziotikos	Male	Group	Greece	Crete	15th century
Maleviziotikos	Female	Group	Greece	Crete	15th century
Outsai	Female	Group	Greece	Pontus	10th century
Pentozali	Male	Group	Greece	Crete	19th century
Podaraki	Female	Group	Greece	Pontus	10th century
Poustseno	Male	Group	Greece	Macedonia	18th century
Rasopoulos	Female	Group	Greece	Pontus	10th century
Roditikos	Female	Group	Greece	Rhodes	10th century
Syrtos sta tria	Male	Group	Greece	Epirus	5th century
Tsamiko	Male	Group	Greece	Epirus	18th century
Zonaradiko	Male	Group	Greece	Thrace	12th century
Zaloggo	Male	Group	Greece	Epirus	19th century
Zeibekiko - Modern	Male	Solo	Greece		20th century
Zeibekiko - Modern	Male	Solo	Greece		20th century
Zeibekiko - Fast	Male	Solo	Greece		20th century
Zorbas	Female	Group	Greece		1960’s
Bollywood	Female	Solo	India		1950’s
Bollywood	Female	Solo	India		1950’s
Bollywood	Female	Solo	India		1950’s
Bollywood	Female	Solo	India		1950’s
Reggaeton	Female	Solo	Puerto Rico		1990’s
Reggaeton	Female	Solo	Puerto Rico		1990’s
Kolo	Female	Group	Serbia		18th century
Pastirske	Female	Group	Serbia		18th century
Flamenco	Female	Couple	Spain	Andalusia	18th century
Hip-Hop	Female	Solo	U.S.A		1970’s
Salsa	Male	Solo	U.S.A.	New York	1970’s
Salsa	Female	Solo	U.S.A.	New York	1970’s

B Comparison with other methods

We have compared our Categorization Tree (Figure 14) with other clustering and ordering methods, such as the 2D embedding from the Multidimensionality Scaling (Figure 11), the t-Distributed Stochastic Neighbor Embedding (Figure 12) methods, and the dendrogram of the Agglomerative hierarchical clustering (Figure 13). Please zoom in the electronic version to see details.

Received July 9, 2019

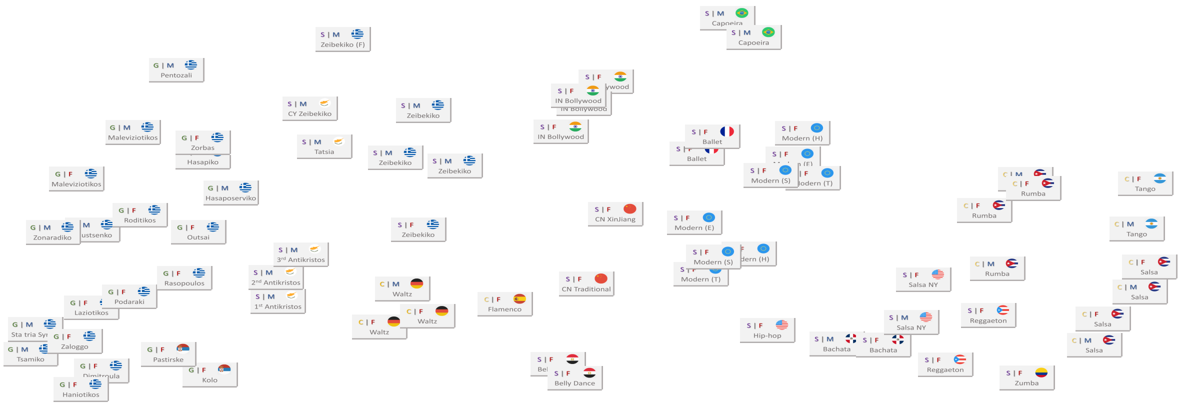

Fig. 11. 2D embedding of our dance motion collection using MDS.

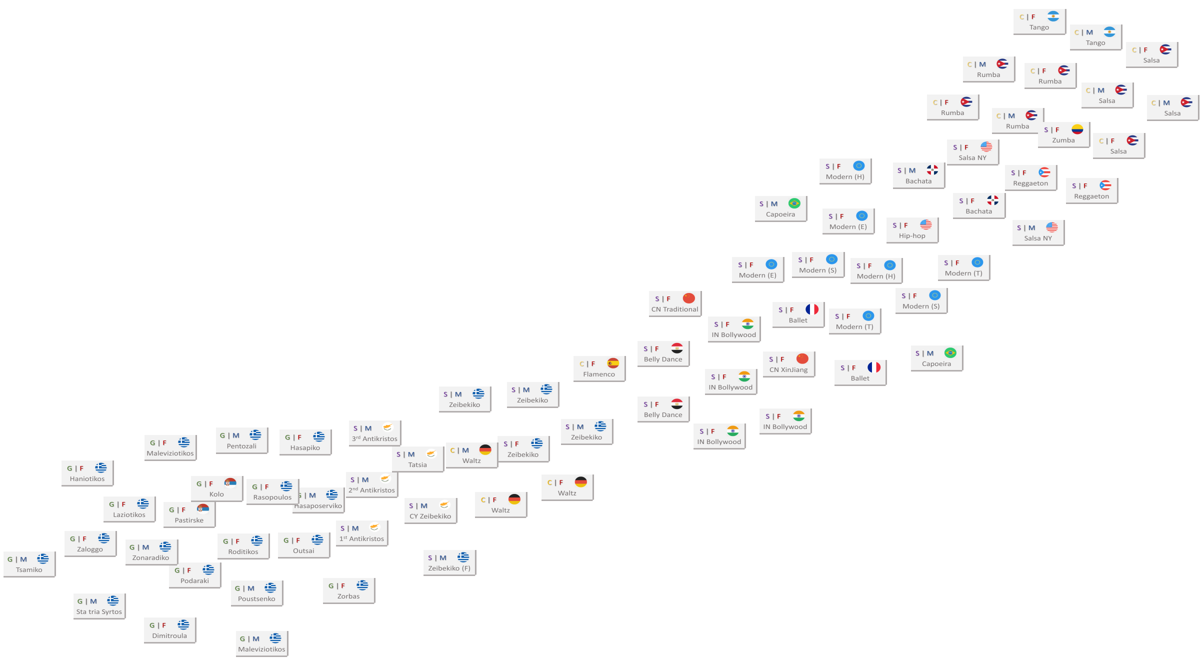

Fig. 12. 2D embedding of our dance motion collection using t-SNE.

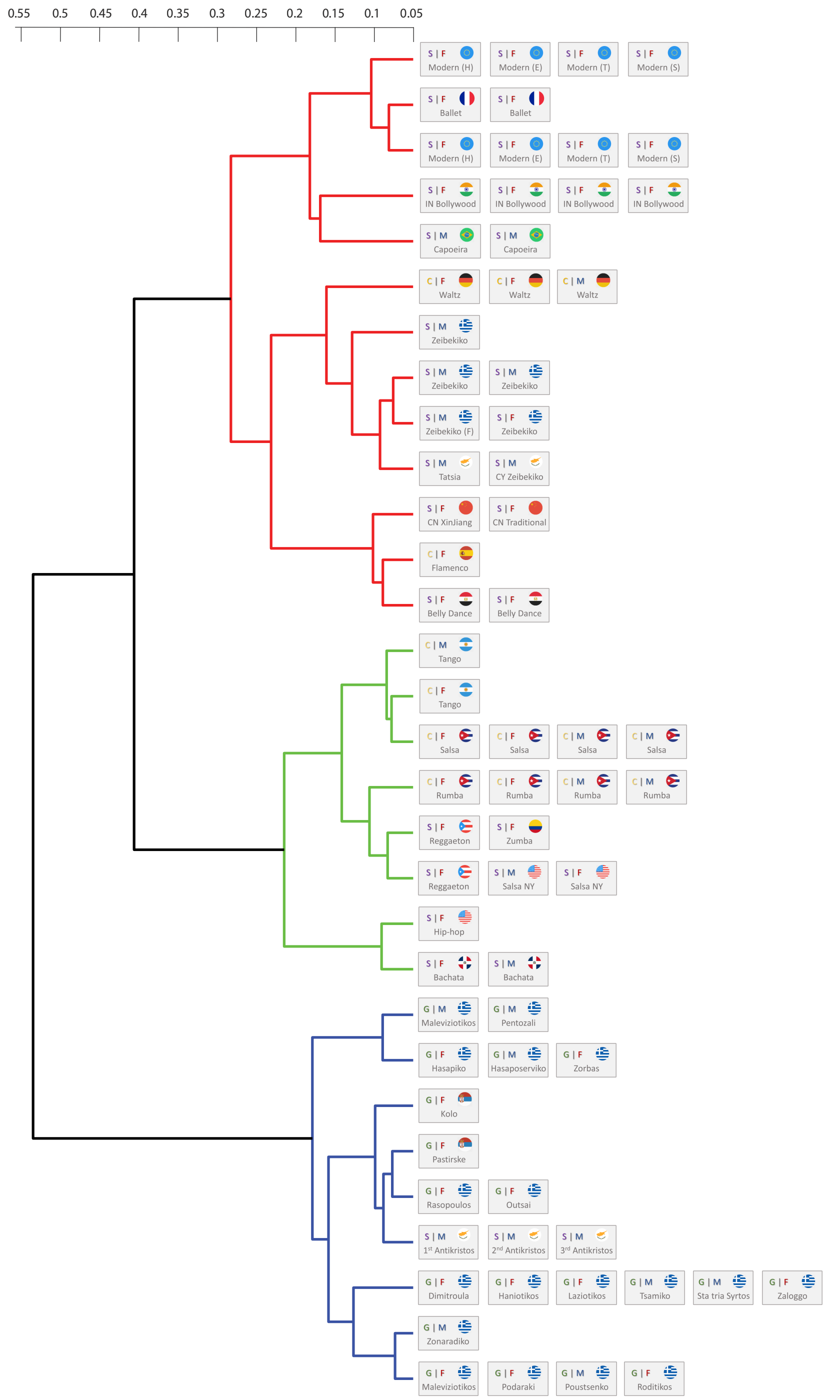

Fig. 13. Agglomerative hierarchical clustering and its dendrogram visualization. The ruler at the top of the dendrogram shows the distance between the dances.

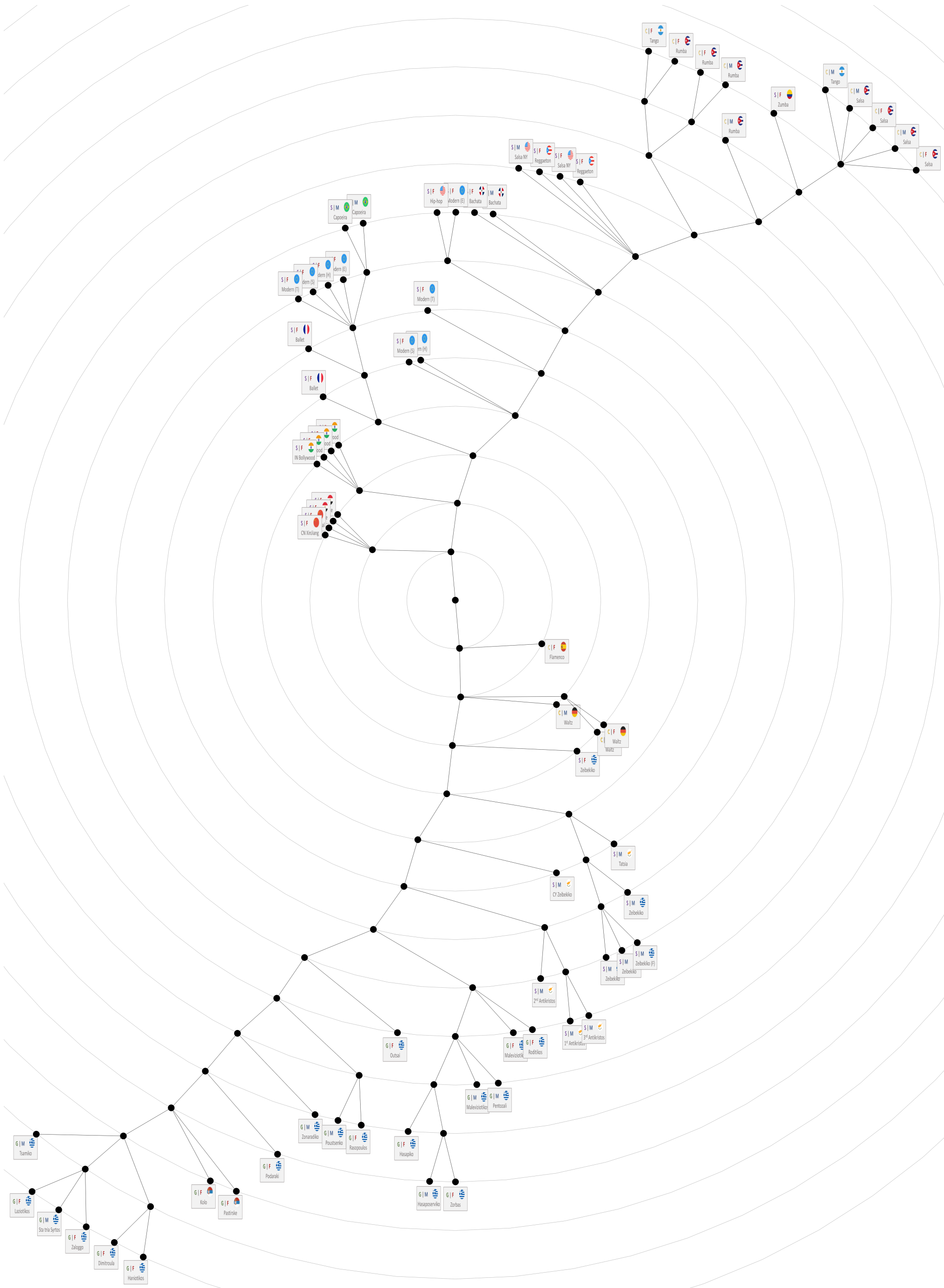

Fig. 14. The Categorization Tree computed using [49].