EURASIP Journal on Audio, Speech, and Music Processing

SpringerOpen

A parameterizable spatiotemporal representation of popular dance styles for humanoid dancing characters

EURASIP Journal on Audio, Speech, and Music Processing 2012,
2012:18 doi:10.1186/1687-4722-2012-18

João Lobato Oliveira (jmso@inescporto.pt)
Luiz Naveda (luiznaveda@gmail.com)
Fabien Gouyon (fgouyon@inescporto.pt)
Luis Paulo Reis (lpreis@fe.up.pt)
Paulo Sousa (paulosousa12.2@gmail.com)
Marc Leman (marc.leman@ugent.be)

ISSN 1687-4722

Article type Research

Submission date 15 April 2011

Acceptance date 5 May 2012

Publication date 19 June 2012

Article URL http://asmp.eurasipjournals.com/content/2012/1/18

A parameterizable spatiotemporal representation of popular dance styles for humanoid dancing characters

João Lobato Oliveira*[1,2,3], Luiz Naveda[4], Fabien Gouyon[1,3],

Luis Paulo Reis[1,2], Paulo Sousa[1] and Marc Leman[4]

[1]Faculty of Engineering of the University of Porto (FEUP), Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal

[2]Artificial Intelligence and Computer Science Laboratory (LIACC), Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal

[3]Institute for Systems and Computer Engineering of Porto, Rua Dr. Roberto Frias, 378, 4200-465 Porto, Portugal

[4]Institute for Psychoacoustics and Electronic Music (IPEM), Ghent University, Blandijnberg 2, Rozier 44, 9000 Ghent, Belgium

*Corresponding author: joao.lobato.oliveira@fe.up.pt; jmso@inescporto.pt

LV: luiznaveda@gmail.com

FG: fgouyon@inescporto.pt

LPR: lpreis@fe.up.pt

PS: paulosousa12.2@gmail.com

ML: marc.leman@ugent.be

Abstract

Dance movements are a complex class of human behavior which convey forms of non-verbal and subjective communication that are performed as cultural vocabularies in all human cultures. The singularity of dance forms imposes fascinating challenges to computer animation and robotics, which in turn presents outstanding opportunities to deepen our understanding about the phenomenon of dance by means of developing models, analyses and syntheses of motion patterns. In this article, we formalize a model for the analysis and representation of popular dance styles of repetitive gestures by specifying the parameters and validation procedures necessary to describe the spatiotemporal elements of the dance movement in relation to its music temporal structure (musical meter). Our representation model is able to precisely describe the structure of dance gestures according to the structure of musical meter, at different temporal resolutions, and is flexible enough to convey the variability of the spatiotemporal relation between music structure and movement in space. It results in a compact and discrete mid-level representation of the dance that can be further applied to algorithms for the generation of movements in different humanoid dancing characters. The validation of our representation model relies upon two hypotheses: (i) the impact of metric resolution and (ii) the impact of variability towards fully and naturally representing a particular dance style of repetitive gestures. We numerically and subjectively assess these hypotheses by analyzing solo dance sequences of Afro-Brazilian samba and American Charleston, captured with a MoCap (Motion Capture) system. From these analyses, we build a set of dance representations modeled with different parameters, and re-synthesize motion sequence variations of the represented dance styles. For specifically assessing the metric hypothesis, we compare the captured dance sequences with repetitive sequences of a fixed dance motion pattern, synthesized at different metric

resolutions for both dance styles. In order to evaluate the hypothesis of variability, we compare the same repetitive sequences with others synthesized with variability, by generating and concatenating stochastic variations of the represented dance pattern. The observed results validate the proposition that different dance styles of repetitive gestures might require a minimum and sufficient metric resolution to be fully represented by the proposed representation model. Yet, these also suggest that additional information may be required to synthesize variability in the dance sequences while assuring the naturalness of the performance. Nevertheless, we found evidence that supports the use of the proposed dance representation for flexibly modeling and synthesizing dance sequences from different popular dance styles, with potential developments for the generation of expressive and natural movement profiles onto humanoid dancing characters.

---

1 Introduction

The process of generating human-like motions plays a key role in robotics, computer graphics, computer games and virtual reality systems. The modeling and generation of expressive and natural forms of human motion has an impact on our knowledge about human behaviors and on the application of this knowledge in science and technology. Dance movements are a complex class of human motions that offer infinite forms of expressiveness and modes of non-verbal communication that are distributed in cultural vocabularies enriched with interactions with music and other modalities. These characteristics impose fascinating challenges to

robotics and outstanding opportunities to deepen our understanding about the phenomenon of dance.

In [1], Naveda and Leman proposed a topological spatiotemporal method to analyze the relationships between gesture, space and music in popular dance styles characterized by repetitive movement patterns in synchrony with temporal regularities in music. In [2], we explored this method of analysis to model a mid-level dance representation[a] and synthesize beat-synchronous dance sequences of Afro-Brazilian samba. This article extends the latter by making use of the proposed representation model to investigate two hypotheses: that *(i) there is a minimum and sufficient temporal metric resolution required to represent a style of repetitive dance gestures*, and that *(ii) spatiotemporal variability*[b] *is an essential quality that relates to expressiveness in dance, and consequently to more naturalness in a dance display.* In addition, this article proposes an alternative method to synthesize the spatiotemporal variability observed in the recorded dance performance from an improved formalization of the representation. By also proposing improved quantitative metrics of similarity and variance we compare the real dance performances against synthesized dance sequences of the analyzed dance style, using different parameters as independent variables (i.e., different *parameterizations*). Finally, this article introduces strategies to manipulate our representation model in order to reproduce these dance styles onto different humanoid dancing characters. The method was applied to two different popular dance styles characterized by repetitive gestures, in particular the Afro-Brazilian samba and the American Charleston dances.

The article is structured as follows: the remainder of this section describes the concept

of musical meter, refers to previous spatiotemporal representations of dance gestures and introduces methods in recent literature used for modeling and synthesizing dance movements onto robotic and computer animation characters based on captured dance movements. Additionally, it summarizes the proposed method for analyzing and representing dance sequences of popular dance styles and the evaluation methods used for assessing the proposed representation model according to the stated hypotheses. Section 2 specifies the details of the recording, analysis and representation of popular dance motion data tested on Afro-Brazilian samba and American Charleston. It additionally describes means to parameterize our representation model and re-synthesize variations of the analyzed dances from it, also offering a solution to synthesize the same level of variability observed in the recorded dances. Section 3 describes our evaluation method as the procedures undertaken for assessing and validating our representation model according to the stated hypotheses. Section 4 presents and discusses the achieved results in accordance with these hypotheses, and presents some paths for future work, namely introducing strategies to manipulate our representation model towards generating dance sequences onto different humanoid dancing characters. Finally, Section 5 summarizes and concludes this article.

1.1 Related work

A number of studies have used recordings of human movement in an attempt to investigate how expressiveness and meaning can be attached to artificial motion profiles of robotic and computer animated characters (e.g., [3–6]). However, the manipulation of pre-recorded

sequences of movements is time-consuming and highly dependent on the context in which the movement was recorded, which narrows the range of applications and interactions.

From a psychophysical perspective, a great part of the experience of motion can be described by the dimensions of *space*, which is considered the medium for the deployment of movement, and by the *time*, which is considered the medium for segmentation and synchronization of movement [1]. One could then attempt to model and generate dance movements by means of generative algorithms, but modeling expressiveness depends on deeper knowledge on the nature and structure of dance behaviors. This kind of knowledge would involve models for biomechanics, kinematic representations of dance displays, and multimodal interactions, which are not easily formalized from an algorithmic perspective nor easily implemented from the viewpoint of applied robotic applications.

State-of-the-art applications in robotics and computer animation frequently use symbolic dance representations made of primitive motions synchronized with music [7]. Primitive motions represent characteristic postures (key-poses) of a given dance style. These are typically selected from the movement by identifying sudden trajectory changes in the motion profile. For example, [3, 8] segmented movement sequences of real Japanese folk dancers according to the minimum velocities of their end-effectors' (hands and feet) trajectories. The resulting key-poses were clustered and interpolated a posteriori for generating variations of the captured dance. Similarly, [9] extracted motion key-poses by means of motion rhythm and intensity features calculated from local minima in the motion signal (stop motions), which were based on the Laban's concept of "weight effort" [10]. On the other hand, [4, 11]

generated rhythmic motion patterns, such as dancing and locomotion, by clustering and interpolating unlabeled MoCap segments as "motion beats". Motion beats corresponded to moments of rapid change in the motion signal, given by zero-crossings of the second derivative of all joints' orientation. After retrieving motion features and different musical cues, e.g., beats, pitch, intensity and chord progression, methods for matching and aligning the dance movement with music typically apply signal alignment and optimization techniques, such as time-warping [4, 6], dynamic programming [5], and genetic algorithms [11].

The majority of these studies seem to be mainly centered on a linear concept of time and a strictly deterministic concept of gesture (e.g., fixed key-poses). In other words, time is always represented as a monochronic sequence and movements are often represented as fixed poses in space. Such a concept of time might contrast with the structure of musical time, which is usually based on concurrent temporal regularities (see the concept of *meter* in the next section) and has a strong influence on the structure of traditional and popular dance styles. Literature in the field of dance ethnology and gesture/movement analysis points out that the universe of dance and movement extrapolate the notion of precision in space [1, 12] and that reasoning in dance is much more diverse than the key-pose paradigm [13]. As such, a more comprehensive modeling of dance should involve more flexible (i.e., variable) representations of the use of space while manipulating time according to temporal cues of the musical rhythm. In great part of the popular or traditional dances, gestures are often deployed through synchronization with events in music, which are traditionally organized by the *musical meter*. The question is how both space and musical meter can be articulated in a

compact and parameterizable representation of a dance performance situated in a particular dance style, and how to re-synthesize dance sequences from the latter representation model. These processes should induce the observed level of naturalness, expressiveness and musical synchronization of the original movement, while keeping the overall spatiotemporal structure of the analyzed dance style.

1.2 Spatiotemporal representation of musical meter and dance

A significant part of dance styles depends on the structure of *musical meter*, which organizes dance choreographies, the timing of the gestures, and the music structure itself. The concept of musical meter captures the idea that rhythm and temporal regularities are organized as hierarchical structures in music, resembling a hierarchical structure of beats and metric levels depicted as a grid, as represented in Figure 1a. In this representation, the establishment of temporal regularities caused by past and present events (metric accents) reinforce or conflict with the metric structure [14, 15].

The concept of meter proposed by Lerdahl et al. [14] is expressed in the structure seen in Figure 1a. It indicates that meter is organized in hierarchies composed of layers (vertical axis) of periodic and symmetric metric accents distributed through time (horizontal axis). However, when dance gestures are synchronized with musical meter, it can be said that the meter becomes integrated with dance in the *spatiotemporal domain*. Because dance and music share the same time domain, regular events in the musical tessiture are reflected as regularities in the use of space. Figure 1b illustrates the process in which metric accents are

projected onto the dance trajectories in the spatiotemporal domain.

Spatiotemporal representations of dance are not new and several forms of representation have been proposed so far. Figure 2 shows a chronological prospect of some of these representations, which denote a long term effort to represent dances. Note that, given the complexity of dance engagement, none of the approaches managed to provide a complete solution for a representation of the dance structure. For example, the lack of systematic representation of all body articulations in time (e.g., the first and third graphs [16, 17] in Figure 2), the lack of representation of cross modal interactions with music and other modalities (e.g., fourth and fifth graphs [18, 19] in Figure 2), the absence of representations of the variability of the dance gesture (e.g., the first, second and third graphs [16, 17, 20] in Figure 2) and the lack of structural models, at some of the representations in Figure 2, indicate how the representation of dance is often complex. Thus, in order to render expressive beat-synchronous dance movements we needed to extend the existing representations into a novel spatiotemporal model (see Section 2.2) that would consider all of the following: *(i)* the possibility of re-synthesis of original motions; *(ii)* encoding motion trajectories and musical time (meter) in the same representation; *(iii)* accounting for the variability of dance sequences; *(iv)* support of multi-modal parameterizations for assessing different hypotheses.

1.3 Method

The method, depicted in Figure 3 involves analysis and representation of dance sequences and the evaluation procedures necessary to validate the proposed representation model. The

analysis and representation are applied to motion capture recordings of dance performances of popular dance styles and the process of validation compares different parameters applied to the representation model according to a set of evaluation criteria. The parameterization of the representation model, related to the chosen metric resolution and representation of variability, compose our set of independent variables used to test hypotheses that assess the feasibility of our model towards fully and naturally representing a dance style of repetitive gestures.

The processes of analysis and representation include four stages: *(i)* data acquisition, *(ii)* data analysis, *(iii)* parameterization of the representation model, and *(iv)* the final representation of the dance style. The process of validation involves the synthesis of dance sequences from different choices of parameters, and the comparison against the captured dance sequence through a set of evaluation criteria. Our evaluation method consists of numerical and subjective assessments. In the numerical evaluation we consider the degree of variance and correlation of both synthesized and captured joint trajectories in relation to the consequent dimensionality and level of reduction provided by the representation model. In the subjective assessment we evaluated the degree of similarity between captured and synthesized dance sequences by asking fifteen subjects to measure their subjective similarity. The dance performances of Afro-Brazilian samba and American Charleston were synchronized to their respective music styles. These dances were recorded with a MoCap system and the musical pieces were manually annotated by experts.

2 Dance movement analysis and parameterizable spatiotemporal representation

2.1 Recording procedures

The recorded dances were performed by two professional female dancers, one specialized in Afro-Brazilian dances and other in old/traditional dances. The first dancer performed simple dance gestures in *samba-no-pé* style, which is the most recognizable and popular sub-style of the Afro-Brazilian samba dances. The second dancer performed dance gestures in the basic American Charleston style. After a few trial runs without any limitations, the dancers were instructed to dance the standard steps of the style without exhibiting improvisations, turns or embellishments.

The musical stimulus used in the samba recording was composed of looped samples of a samba percussion ensemble (*surdo*, *tamborim* and *caxixi*) sequenced at 80 BPM (beats-per-minute). The musical stimulus used in the Charleston recording was composed of phrases of Charleston music exhibiting a mean tempo of 111 BPM.

2.1.1 Motion capture data acquisition

The dance recordings of samba and Charleston were respectively performed in Brazil and Belgium, both with an 8-camera MoCap system setup (Optitrack/Natural Point [21]). The dance movements were recorded at a frame rate of 60 Hz and upsampled to 100 Hz in the editing phase. The motion data was synchronized with the musical stimulus used in the recording and the motion trajectories of each recorded dance sequence were normalized in

relation to the centroid of the body, frame per frame. This process subtracted the effect of the movement of the whole body on the trajectories of the limbs. The sequences were imported into Matlab and edited using the MoCap Toolbox [22]. This process resulted in one dance sequence of samba and one other of Charleston, both synchronized to music, to be further analyzed.

2.1.2 Annotation data

The musical sequences were manually annotated by experts and all metric accents (here described as time points and classes of musical levels) classified using Sonic Visualizer [23]. From the beat annotation we derived both a macro level (by downsampling it into bars of 2 beats) and micro levels of the musical meter (by upsampling it into half-beat, quarter-beat, and eighth-beat levels). These levels encompass the resolution of the metric parameters that are used to parameterize our dance representation model, considering bar levels (i.e., the size of the metric cycles in which the meter is decomposed) of 2 beats for samba and 4 beats for Charleston. Previous knowledge about the Jazz (which includes the Charleston dance styles) and Afro-Brazilian culture (which includes the samba dance styles)[c] indicate that their couple music forms have the metrical characteristics indicated in this study, more specifically the organization of the subdivisions of the beat in 1/4th beat divisions in both styles and the metrical properties mentioned before. A schematic description of these metric levels (hereafter named *metric classes*) in the time/metric domain is shown in Figure 1 and Figure 4.

2.2 Analysis and parameterizable spatiotemporal representation of the dance movement

Our representation model is build upon a method that analyzes the spatiotemporal relationships between music and use of space in popular dance styles [1]. This method, denominated Topological Gesture Analysis (TGA), maps the structure of musical gestures into topological spaces.

As depicted in Figure 4, the TGA method relies on a projection of musical cues (here, metrical classes—see Figure 4a) onto trajectories (see Figure 4c), which, by definition, generates a combined spatiotemporal representation using musical and choreographical information. Considering that the use of space in dance is organized according (or in synchrony) to the projected musical cues it is likely that the projection of points generate clusters in space, or point clouds (Figure 4d), which can be clustered, discriminated and organized in different geometries or representations [24]. Since we assumed that the modalities of audio and movement are intrinsically interdependent and synchronized, we projected a set of W annotated metric classes, decomposed into C metric cycles and matching $Q = W * C$ time points in the audio, onto the 3D trajectories of dance. This resulted in a sequence of metric classes, $M = [m_1, \ldots, m_W, \ldots, m_1, \ldots, m_W, \ldots]$, that match a set of time-points in the audio given by $T = [t_1, \ldots, t_Q]$. Subsequently, they are projected onto a set of Q 3D points, P, given by the 3d coordinates of the motion trajectories, Z, occurring at the time-points of T, such as $P = [[Z_x(t_1), Z_y(t_1), Z_z(t_1)], \ldots, [Z_x(t_Q), Z_y(t_Q), Z_z(t_Q)]]$.

The set of points P results in point clouds that charge the dance space with musical

qualities, defining geometries that we call gesture topologies (see Figure 4e). In short, they inform how the relationships between the gesture of the dancer and the respective musical characteristics are performed in space. Since we are interested in how the dancer uses the space in relation to classes of the musical meter (half-beat, 1 beat, 2 beats, etc.), i.e., how the dancer performs in beat-synchrony, we discriminated the 3d points of P into W point clouds (i.e., one point cloud X_m per metric class m), $X = [X_{m_1}, \ldots, X_{m_W}]$, by clustering the points according to their represented metric class, such as $X = [P(t \equiv m_1), \ldots, P(t \equiv m_W)]$; where $P(t \equiv m)$ represents all 3d points whose time occurrence match the considered metric class m. To improve the discrimination of these regions (see Figure 4d) we used linear discriminant analysis (LDA) [25] which guaranteed higher separation between classes of point-clouds by calculating the between class variability through the sample covariance of the class means. From the separation of the classes we discarded the set of points, $L = [L_{m_1}, \ldots, L_{m_W}]$, from X that could not be discriminated in the LDA, such as $X' = X - L = [X'_{m_1}, \ldots, X'_{m_{W'}}]$, where W' corresponds to the number of classes of W that are represented by at least one point of X after discrimination. Ultimately, from the discriminated point clouds, X', we delimited W' topological regions (i.e., W' topologies) given by uniform spherical distributions. The radius of each spherical distribution, V_m, is defined by the mean of the Euclidean distances, $E_{x'_{m,i},\mu_m}$, of all the I_m points represented in the given point cloud of class m, $X'_m = [x'_{m,1}, \ldots, x'_{m,I_m}]$, to the 3d centroid (i.e., center of mass μ_m) of its distribution (see Figure 4e):

$$V_m = \frac{4}{3}\pi \left(\frac{1}{I_m} \sum_{i=1}^{I_m} E_{x'_{m,i},\mu_m} \right). \tag{1}$$

The described process was replicated for each of the dancer's joint motion trajectories such

that the complete TGA representation conveys one mean 3d value and the radius of the spherical distribution for each metric class and each of the 20 joints of the considered body model (see the considered body model description on Figure 5a.

As described, this dance model offers a compact representation of the dance movement in relation to the musical meter, being at the same time able to describe the dance according to different levels of the musical meter (different temporal resolutions), and flexible to convey variability of the gestures in space. In addition, the model can be parameterized in different ways since it is able to provide different variations of the same dance representation, which may specifically differ in terms of the considered metric resolution (i.e., by the number of considered metric classes—see Figure 4a) and in the consideration and discrimination of spatiotemporal variability (e.g., Figure 4e), where spherical distributions are considered to represent variability). Figure 6 illustrates the final spherical distributions for the left hand of a dancer provided by a spatiotemporal representation model of samba dance parameterized with quarter-beat resolution and with variability.

2.3 Synthesizing dance sequences

Our TGA model of the dance performances represents a set of discrete metric classes which intrinsically delineate likelihoods of key-poses in space, describing pseudo-unlimited variations of the fundamental postures characteristic of the analyzed dance style. The process of synthesis consists in generating and concatenating closed-loop cycles of the key-poses underlying in the TGA model, with or without variability, and interpolating them according

to the represented musical meter, at the chosen metric resolution. Consider the example illustrated in Figure 7: while a representation parameterized with a 1-beat resolution ("beat resolution") enables the synthesis of dance sequences with one key-pose interpolated within the time-interval, Δt, of one musical beat (i.e., two different key-poses per metric cycle of two beats), a resolution of quarter-beat provides four key-poses for the same duration (i.e., eight different key-poses per two-beat metric cycle). The musical beat-synchrony is therefore implicitly projected into the gesture itself by assigning each synthesized key-pose to specific key-frames in the music, respectively representing each of the annotated metric classes in the time domain.

Alternatively, the representation can be parameterized with or without variability, by respectively considering or ignoring the radii of the spherical distributions represented in the TGA model (see Figure 4e). Therefore, the synthesis of dance sequences without variability is built by concatenating repeated sequences of a fixed dance pattern, composed by the same set of key-poses. In this case, the set of repetitive key-poses is build on the centroids of the TGA distributions for each joint at each metric class.The method of synthesizing full-body key-poses with variability is described in detail in Section 2.3.1. For both alternatives, the synthesized key-poses are interpolated in complete dance instances as described in Section 2.3.2.

2.3.1 Synthesizing key-poses with variability

The question of synthesizing key-poses with variability from our representation model can be potentially solved by, at every metric class, stochastically generating joint coordinates/rotations that would satisfy both kinematic constraints and their respective spherical distributions; preserving both body morphology and the represented spatiotemporal variability.

Contrarily to [2], the proposed solution is formulated in quaternion algebra to be directly applied into robotic and/or computer animated humanoid characters, enabling an easier and more reliable manipulation of the dance representation to be applicable onto different humanoid body models. This process (depicted in Figure 5) involves an initial decomposition of the MoCap representation of the human body into five kinematic joint chains, derived from two anchor joints (hip, at joint 1, and neck, at joint 11). At this stage every kinematic chain is processed independently by calculating random joint rotations confined by the represented TGA distributions. The correspondent body segments are synthesized as the norm of the given unity vectors according to the original body model (as illustrated in Figure 5c). This process is iteratively computed until all joints of each key-pose can be successfully calculated while satisfying the propagated kinematic constraints.

In order to ensure that the fixed geometry of the human body is not violated in the process at any given metric class, if one segment does not fit both TGA distributions at its joint extremities the algorithm points all the following joint rotations (up to the chain extremity) towards their respective TGA's centroids. This occurs when the choice of a random joint position at one segment extremity makes it impossible for the segment, with

fixed length, to reach the other extremity's distribution. This restriction is mostly caused by the discrimination of the topologies at the dance analysis stage by the use of the LDA. In this eventuality, the algorithm retries the whole process from the beginning (i.e., from the anchor joint of the considered kinematic chain and metric class) and keeps trying to accomplish the whole chain's constraints and the represented spatiotemporal variability for a maximum of 25 times (limited due to the computational overload), saving the resulting key-poses' joint positions of each trial. This process is realized for every metric cycle of the desired dance sequence, in order to ensure the highest degree of variability among them. Each cycle is therefore built by the group of key-poses which, when summed, required the lowest amount of forced joint positions (i.e., the lowest amount of joint rotations set towards the centroid of their respective TGA distributions).

The process of randomly calculating all joint's rotations (and consequently their 3d coordinates) inside each body kinematic chain is described in detail as follows. As illustrated in Figure 5c, for every metric class, m, the first step consists of determining the possible variations of the quaternion, qv, (i.e., the 3d rotation of a target unity vector, $\vec{v}'$, around its base unity vector, $\vec{v}$) defined between every two body segments. Every two body segments are defined by the current segment, s_j^m, which links the formerly assigned joint position, p_{j-1}^m, to the current joint coordinates, p_j^m, to be randomly calculated, and the previously processed segment, s_{j-1}^m, which links p_{j-1}^m to the preceding joint position, p_{j-2}^m, in that same kinematic chain. p_j^m is generated from a rotation quaternion, $qv_{s_j^m}$, randomly assigned inside the intersection cap, C_j^m, between the considered TGA distribution, T_j^m, and a sphere, J_{j-1}^m,

centered on p^m_{j-1} with radius equal to the current segment length, $l_{j-1,j}$.

Initially, the base vector for calculating the orientation of the spine segment that connects the two anchor joints of the used body model (joint 1 to joint 10 in Figure 5a) is considered to be fixed in space at $\vec{v}_{s^m_0} = (0, -1, 0)$. This anchor segment, s^m_0, is then considered as the initial base vector of all kinematic chains. From this point on, every generated target vector is used as the base vector of the following segment, in a recursive process, up to the extremity segment of the considered kinematic chain. As such, starting from s^m_0, the possible variations of each segment rotational quaternion, $qv_{s^m_j}$, are constrained by the former calculated joint position, p^m_{j-1}, the former segment unity vector, $\vec{v}'_{s^m_{j-1}}$, (i.e., the current base vector, $\vec{v}_{s^m_j}$), and C^m_j.

The current joint rotation, $qv_{s^m_j}$, is therefore randomly selected inside a spatial range confined by six extremity quaternions, $qv^i_{s^m_j}$, (one maximum and one minimum for each spatial dimension, $d = \{x, y, z\}$). These six $qv^i_{s^m_j}$ are indicated by the rotation of the current segment, s^m_j, around its base segment vector, $\vec{v}_{s^m_j}$, towards each dimensional extremity, $C^{i\ m}_{\text{ext}\,j}$, of C^m_j, as follows (note that for simplification we omitted the m index from all variables in Equations (2) and (3), although all calculations are relative to a specific metric class):

$$\begin{cases} qv^i_{s_j} = \cos\left(\alpha^i_{s_{j-1},s_j}/2\right) + \vec{u}^i_{s_j} * \sin\left(\alpha^i_{s_{j-1},s_j}/2\right) \\ \vec{v}'^i_{s_j} = \vec{v}^i_{s_{j+1}} = C^i_{\text{ext}\,j} - p_{j-1} : \\ C^i_{\text{ext}\,j} = \left\{\min_d\left(C_{\text{ext}\,j}\right) \cup \max_d\left(C_{\text{ext}\,j}\right)\right\}; \quad i = 1, \ldots, 6 \end{cases}, \tag{2}$$

where $\vec{u}^i_{s_j}$ is the unity vector representing the 3d axis of rotation between both segments, s_{j-1} and s_j, towards one of the $C^i_{\text{ext}\,j}$ extremities, and $\alpha^i_{s_{j-1},s_j}$ is the correspondent rotation

angle.

The second step consists of calculating a random quaternion, $qv_{s_j^m}$, inside the spatial range described by the six extremity quaternions, $qv^i_{s_j^m}$, (calculated in Equation (2)), as follows:

$$\begin{cases} qv'_{s_j} = \overline{qv}_{s_j} \pm \left[\max_i \left(\left|\overline{qv}_{s_j}\right| - \left|qv^i_{s_j}\right|\right) * \text{rand}[0,1]\right] \\ qv_{s_j} = \frac{qv'_{s_j}}{\|qv'_{s_j}\|} \end{cases}, \tag{3}$$

where $\overline{qv}_{s_j}$ is the mean quaternion, representing a rotation from the last calculated joint position to the center of the current spherical cap.

The third and final step consists of calculating the current joint position, p_j^m, based on the obtained target rotation vector, $\vec{v}'_{s_j^m}$, the former calculated joint position, p^m_{j-1}, and the current segment length, $l_{j-1,j} = \|\vec{v}'_{s_j^m}\|$, as follows:

$$\begin{cases} p_j^m = p^m_{j-1} + l_{j-1,j} * \vec{v}'_{s_j^m} : p_j^m \in T_j^m \\ \vec{v}'_{s_j} = \left(\frac{qv_{s_j}}{\|qv_{s_j}\|}\right) * \vec{v}_{s_j} * \left(\frac{qv_{s_j}}{\|qv_{s_j}\|}\right)^{-1} = \vec{v}_{s_{j+1}} \end{cases}. \tag{4}$$

2.3.2 Motion interpolation between key-poses

In order to synthesize complete dance instances from the synthesized key-poses, we generated continuous joint trajectories by recurring to motion interpolation techniques. Motion interpolation (or blending) is a highly discussed topic in computer animation and robotics literature. Interpolation functions typically blend point-to-point positions (e.g., key-poses) or motion primitives into continuous trajectories according to a set of kinematic and/or dynamic constraints. These functions can be applied both in the time [11] or frequency [26]

domains, to joint/links coordinates or rotations, and assume various forms, ranging from slerps [11] to hierarchical B-splines [27], B-spline wavelets [28], and piecewise [5] spline functions.

Considering our application, we selected a linear point-to-point joint coordinates interpolation between key-poses. Although this method does not ensure that the geometry of the humanoid body model is fixed along the whole synthesized motion sequence (see Section 3.2 for the imposed body error), it is computational inexpensive and provides the required reliability to validate our dance representation model. Yet, the use of this representation model for the generation of dance movements onto computer animated or robotic characters would imply the use of more sophisticated interpolation functions. (Further discussion on this topic is outside the scope of this article, and left for future work—see Section 4.5.)

In detail, continuous dance motion trajectories are synthesized by interpolating all synthesized key-poses in the order of the represented metric structure, along all metric cycles of the dance sequence at the defined resolution. In such a way, the beat-synchrony observed in the recorded dance is implicitly translated into the interpolated dance sequence.

The motion transition between postures, within all W metric classes, is generated by interpolating each joint independently. As such, all joint coordinates, $p_{j_{x,y,z}}$, are interpolated between W consecutive pairs of key-frames, $[\{t_0, t_1\}, \ldots, \{t_m, t_{m+1}\}, \ldots, \{t_W, t_0\}]$, (the interpolation knots) pointed by consecutive pairs of metric classes, m, by means of a piecewise cubic spline interpolant, I, over each joint coordinate dimension, j_d, given by (follow

Figure 8):

$$I = [I_k(j_x), I_k(j_y), I_k(j_z)], \tag{5}$$

where $m = 0, \ldots, W-1; k = 0, \ldots, W;\ d = \{x, y, z\};\ j_d \in [\{t_0, t_1\}, \ldots, \{t_m, t_{m+1}\}, \ldots, \{t_W, t_0\}];$

and

$$\begin{cases} I_m(j_d) = c_0 + c_1(j_d - p^m_{j_d}) + c_2(j_d - p^m_{j_d})^2 + c_3(j_d - p^m_{j_d})^3 \\ I_m(j_d) = I_{m-1}(j_d) \\ I'_m(j_d) = I'_{m-1}(j_d) \\ I''_m(j_d) = I''_{m-1}(j_d) \\ I''_0(j_d) = I''_W(j_d) = 0 \end{cases} . \tag{6}$$

3 Experiments and validation procedures

This section describes the evaluation method for validating the proposed representation model using recordings of samba and Charleston dance, which were recorded and pre-processed as described in Section 2. The proposed evaluation consists of three sections: *(i)* experimental setup, and *(ii)* numerical and *(iii)* subjective evaluations.

3.1 Experimental setup

The experiments consist of *numerical* and *subjective* assessments that evaluate the capacity of the TGA model to represent repetitive displays of popular dance styles according to the proposed hypotheses. The *numerical* evaluation includes measures of *similarity*, *variance*, level of *reduction*, and *dimensionality* that aim to describe how dance sequences, synthesized

from differently parameterized representations and of distinct dance styles (i.e., samba and Charleston), differ from the captured data and among each other, and furthermore what gain can be obtained in terms of data compression by the use of the proposed representation model. Ultimately, it measures the overall body size error imposed by our simplistic interpolation method. The subjective evaluation consist of subjects' assessment over the visual similarity between the synthesized and captured dances.

Both these processes aim to investigate the optimal set of parameters (i.e., the optimal parameterization) necessary to represent each dance style, and consequently to identify the minimum amount of information necessary to reliably describe them by means of a compact spatiotemporal representation, thus validating our model in respect to the proposed hypotheses.

3.1.1 Hypotheses

In order to validate the proposed representation model we relied upon two hypotheses: *metric resolution* and *variability*. The remainder of this article addresses the validity of these hypotheses by proposing a set of parameterizations and evaluation criteria to assess our representation model.

The confirmation of the hypothesis of *metric resolution* should imply that the quantity or density of metric classes has a positive impact on a full and natural description of the represented dance. In other words, there should be a minimum and sufficient temporal metric resolution required to satisfactorily represent the dance style, leading to the optimal

similarity between synthesized and captured dances. In order to test this hypothesis, we varied the metric resolutions (independent variable) of the synthesized dance sequences in four levels: beat, half-beat, quarter-beat, and eighth-beat resolutions.

The confirmation of the hypothesis of *variability* should imply that spatiotemporal variability in the system lead to more perceived naturalness and, consequently, more similarity with the captured dance. We accessed the impact of this hypothesis by comparing dance sequences synthesized from representations parameterized with spatiotemporal variability (as described in Section 2.3.1) with others built of repetitive sequences of a fixed pattern (by assuming the centroid of the TGA distributions for each joint and metric class).

3.1.2 Assessed parameterizations

Table 1 shows the eight parameterizations applied to the proposed representation model for validating it in respect to the stated hypotheses. These different parameterizations were individually applied to the TGA representation model of each dance style and respectively synthesized into dance sequences to be numerically and subjectively evaluated against their respective captured dances. In addition, we also included an excerpt of the captured dance sequence of each dance style off-set by one metric cycle (i.e., by one bar), hereafter denominated "original" sequence. To ensure that all synthesized sequences are also aligned with the captured dance sequence the initial frame of these sequences is mapped to the first metric class of a metric cycle.

The numerical evaluation assesses dance sequences of 15 s synthesized from all the eight

parameterizations described in Table 1 applied to the TGA representation model, plus the "original" sequence, against the captured sequence of each dance style. Due to time constraints the subjective assessment only considers the most relevant sequences for measuring the effect of inducing spatiotemporal variability in the dance representation model. These consist of 30 s dance sequences synthesized from the four parameterizations presented in bold in Table 1 applied to the TGA representation model, plus the "original" sequence.

3.2 Numerical evaluation

3.2.1 Level of similarity

In order to evaluate the *level of similarity* between the captured and the synthesized dance sequences we looked into the literature for measures of interdependence (synchrony) between signals [29]. From the studied metrics we selected the correlation coefficient, r_{s_1,s_2}, which quantifies the linear time-domain correlation between two signals. Between two motion trajectory signals, s_1 and s_2, it can be formulated as follows:

$$r_{s_1,s_2} = \frac{\sum_{n=1,j=1,d=1}^{N,J,D} \left[\left(s_1(n,j,d) - \bar{s_1}(j,d)\right)\left(s_2(n,j,d) - \bar{s_2}(j,d)\right)\right]}{\sqrt{\sum_{n=1,j=1,d=1}^{N,J,D} \left(s_1(n,j,d) - \bar{s_1}(j,d)\right)^2 \sum_{n=1,j=1,d=1}^{N,J,D} \left(s_2(n,j,d) - \bar{s_2}(j,d)\right)^2}}, \qquad (7)$$

where N is the length of the signals (set to 1500 frames – corresponding to 15 s sequences at 100 fps), J is the total number of joints of the considered body model ($J = 20$), D is the number of considered spatial dimensions ($D = 3$, for the 3d space), and $\bar{s_1}$ and $\bar{s_2}$ are the mean frames across all J and D for s_1 and s_2, respectively. This metric translates both period and phase interdependence between s_1 and s_2, resulting in a maximum of $r_{s_1,s_2} = 1$ in the presence of identical signals.

3.2.2 Degree of variability

In order to measure the *degree of variability* observed in each dance sequence we looked for the spatiotemporal variability observed between the motion trajectories of each individual metric cycle composing the whole dance sequence. Therefore, we split each dance sequence into several excerpts corresponding to individual metric cycles. The 15 s samba sequences, at 80 BPM, were split into ten complete metric cycles of 2 beats each, whereas the 15 s Charleston sequences, at 111 BPM, were split into seven complete metric cycles of 4 beats each. To measure the spatiotemporal variability between the motion trajectories, s_c, delimited by each metric cycle, c, we calculated the mean variance, $\bar{v}_s$, among all joints, J, and spatial dimensions, D, of s_c. $\bar{v}_s$ is measured in square millimeters (mm^2) and calculated across all frames, N_c, of each s_c between all metric cycles of the considered sequence's signal, s, as follows:

$$\bar{v_s}(mm^2) = \frac{\sum_{n=1,j=1,d=1}^{N_c,J,D} \sqrt{\frac{1}{C-1} \sum_{i=1}^{C} \left(s_c(n,j,d) - \bar{s_c}(n,j,d)\right)^2}}{J \cdot D \cdot N_c}, \tag{8}$$

where $\bar{s_c}$ is the mean value of the considered dimension, $\{n, j, d\}$, across all metric cycles of s, and C is the total number of metric cycles described in s.

3.2.3 Dimensionality

The *dimensionality*, *Dim(J,S,T)*, of each parameterized representation model was measured as the number of spatiotemporal arguments used to describe the full-body 3d trajectories of the whole dance sequence, according to the defined parameterization. It is described in terms of the number of joints, J, considered in the used body model, the number of spatial

arguments, S, needed to represent the dance motion (in the case of the TGA spherical representation it implies a 3d centroid and the radius, if emulating variability, for each distribution), and the used temporal resolution, W, (i.e., number of metric classes used in the TGA representation):

$$\mathrm{Dim}(J, S, W) = J \cdot S \cdot W. \tag{9}$$

3.2.4 Reduction

The consequent *Reduction* of each synthesized dance sequence measures the degree of data compression of the used representation model, by comparing the dimensionality of the synthesized sequence, $\mathrm{Dim}_s(J, S_s, W_s)$, with that of the captured dance, $\mathrm{Dim}_o(J, S_o, W_o)$. This is always dependent on the length, N, (in frames) of the synthesized sequence, as follows:

$$\mathrm{Reduction} = \frac{\mathrm{Dim}_o(J, S_o, W_o)}{\mathrm{Dim}_s(J, S_s, W_s)} \cdot N. \tag{10}$$

This criterion represents a measure of efficiency and compactness of our representation model under the different applied parameterizations.

3.2.5 Interpolation error

We measure the overall *interpolation error* imposed by our interpolation method in terms of mean body size differences between the synthesized and captured body models. It is calculated as follows:

$$e_i(\%) = \left(1 - \frac{\frac{1}{N} * \sum_{n=1}^{N} \sum_{j_s=1}^{J-1} |p_s^n - p_{j_s+1}^n|}{\sum_{j_b=1}^{J-1} |p_{j_b}^n - p_{j_b+1}^n|}\right) \cdot 100, \tag{11}$$

where $p^n_{j_s}$ are the 3d coordinates of the given joint, j_s, for the considered frame number, n, of the synthesized sequence, s, and $p^n_{j_o}$ are the same 3d joint coordinates in the original (i.e., captured) body model, b.

3.3 Subjective evaluation

In the subjective assessment we asked fifteen subjects (seven Brazilians and eight non-Brazilians) to evaluate dance sequences of samba dance only. The restriction to samba on the subjective evaluation was meant to avoid bias in the evaluation, since the reliance on different cultural backgrounds could lead to uncontrollable bias on the comparison between samba and Charleston. The subjective assessment over Charleston, and other dance styles, will be considered in future work.

In the training phase of the inquiry we described the experiment using a training example and a demonstration of human samba. In the assessment, the subjects were presented to two series of the five dance sequences (i.e., ten trials) described in Section 3.1.2. These included four sequences, each synthesized from one of the four parameterizations displayed in bold in Table 1 applied to our representation model, plus the "original" sequence. In order to evaluate the degree of subjective similarity between the five assessed dance sequences and the captured dance, we run a user-oriented evaluation over each selected parameterization by randomly displaying the captured dance sequence followed by one of its synthesized versions (or the "original" sequence) or vice-versa. After each trial we asked the subjects (1) to indicate which of the two sequences they considered to be the captured sequence and (2) to

grade, from 1 to 5, the level of similarity between the considered captured sequence and the synthesized or the "original" one.

All dance sequences were displayed through a graphic animation of the dance movement synchronized with the used musical stimulus by using an interface based on the dance analysis suite (DAS) software [30]. The visual representation of the human body, displayed in Figure 9, contains a stick figure description of the body model and a clean graphical environment.

4 Results, discussion and future work

In this section we present and discuss both numerical and subjective results (from Table 2 and Figure 10, respectively) according to the accuracy of our representation model on reliably representing a particular dance style in respect to the proposed hypotheses.

4.1 Numeric results

Every dance sequence was synthesized ten times from the same parameterized representation model, and for both dance styles. This aimed to improve the results' precision given the stochastic elements present in our algorithm, which is responsible for injecting spatiotemporal variability into the syntheses. The mean results across the ten dance sequences synthesized from each differently parameterized representation are presented in Table 2. Figure 11 exemplifies the comparison between an excerpt of the captured sequence of Charleston (straight line), and synthesized sequences of Charleston with (dashed line) or without (dotted line)

variability, extracted from the trajectories of the right hand joint (joint 19), for all considered metric resolutions (see Table 1). The "original" sequence is unique per dance style and thus only evaluated once for each style .

Note that the small error imposed by our interpolation method, with a mean of 1.54% among all assessed sequences (see Table 2), ensures a good approximation of the synthesized body with the captured, especially at higher metric resolutions. Its reliability is additionally supported by the high correlation verified between the synthesized and captured dance sequences, even surpassing the results of the "original" sequence. Such results validate the application of our simplistic interpolation method for evaluation purposes.

The fact that all the dance sequences synthesized without variability ("fixed" parameterizations) present some degree of variability among metric cycles (although all sets of synthesized key-poses are the same for all concatenated metric cycles) can be explained by the slightly different time-length of each metric cycle, which results in slightly different interpolated motion trajectories. This effect is greater at lower metric resolutions because the interpolation function has less knots (i.e., less key-poses) which increase the variation of the synthesized motion trajectories among metric cycles with different lengths. The same outcome is intrinsically present in the degree of variability measured in the "original" and all "variability" synthesized sequences, since all dance sequences of each style share the same meter in the time-domain.

4.2 Subjective results

A box plot with the overall statistical results of question (2) (see Section 3.3) for samba is presented in Figure 10. A box plot provides a graphic visualization of the statistical properties of the data [31]. The explanations in the graph indicate that the "the notches surrounding the medians provide a measure of the rough significance of differences between the values. Specifically, if the notches about two medians do not overlap in this display, the medians are, roughly, significantly different, at about a 95% confidence level." [32].

The depicted results are discussed in detail in the following Section 4.3 and Section 4.4.

4.3 The impact of metric resolution

When comparing the results of Table 2 in terms of metric resolution, we observed that the metric level considered in the representation model plays a fundamental role in describing and representing the analyzed dance, which seems to not vary according to the dance style, as observed by the similar trend of $r_{s,o}$ among the synthesized sequences of both samba and Charleston. For synthesized dance sequences of both dance styles we observed a non-linear relationship between resolution and similarity with the captured dance which indicates that when the representation model drops to a certain threshold of numerical resolution (in the whole process) it compromises the geometry and shape of the dance motion trajectories. As observed in the $r_{s,o}$ results of Table 2 (in bold), this saturation threshold seems to be defined by a quarter-beat resolution for both samba and Charleston.

For both samba and Charleston dances, there is an overall agreement between numerical

and subjective evaluations that a correct parameterization of our representation model feasibly reproduces the captured dance in terms of similarity. From a numerical point of view (see Table 2), we could synthesize dance sequences of each particular dance style, with or without variability, with an average accuracy of 0.89 ± 0.01 correlation points. These even outperformed the similarity between excerpts of the same captured dance sequence (i.e., the captured *vs* the "original" sequence), by a maximum difference of 0.03 points. Yet, this was contradicted by the subjects' responses over samba (see Figure 10), by attributing to the "original" sequence the maximum similarity, of 5 points, with the captured dance, and outperforming the best synthesized dance sequence by, on average, 1 point. This suggests that subjective reasoning may play an important factor while evaluating similarity between dance patterns. These factors could be related with the cognitive attention to specific body parts for determining the dance style or with the influence of the non-ecological elements of the set up of the experiment (e.g., the use of a stick figure, backgrounds and computer simulations).

Although we achieved similar optimal correlations with dance sequences synthesized from representations of both dance styles, it seems that there is a minimum temporal metric resolution required to fully represent each dance that does not seem to depend on the analyzed style, at least among samba and Charleston. As presented in Table 2, for both samba and Charleston the optimal solution was achieved with a quarter-beat resolution (i.e., by the "fixed-4" and "variability-4" parameterizations). Since we observed no positive effect on increasing the resolution above quarter-beat for both dance styles (see Table 2) it seems that

there is also a sufficient (i.e., maximum) metric resolution required to fully, and consequently naturally, represent a particular style of repetitive dance gestures.

As a final remark, we verify that metric resolution has a direct impact on the dimensionality of the proposed representation model by proportionally decreasing the compactness of our representation with the increase of the resolution. The results of Table 2 suggest that the numerical structure and subjective impact of the analyzed dance style may be feasibly reproduced by a compact representation model, with a reduction of information in the order of 13% or 6% the size of the dance sequence, for samba and Charleston respectively. The differences in reduction (by a factor of two) is due to the segmentation of samba in compasses of two beats and of Charleston in compasses of four beats, which means that a representation of Charleston requires two times the number of metric classes required by samba for encompassing the same metric resolution.

4.4 The impact of variability

The impact of introducing variability in the proposed representation model towards improving the naturalness and similarity of the represented dance style with the captured dance was specifically measured by the numerical correlation and degree of variance, presented in Table 2, and the subjective reasoning of the inquired subjects, presented in Figure 10. The numerical results indicate that the variability imposed by the proposed stochastic method negatively affected the similarity between the synthesized joint trajectories and the captured ones, with an average decrease in correlation of 0.02 ± 0.02 points against dance sequences

synthesized without variability, among all metric resolutions and for both dance styles. The slight outperformance of the dance sequences synthesized from representations parameterized without variability can be justified by the use of repeated sequences of a fixed movement pattern representing the mean joint trajectories among metric cycles of the analyzed dance, which minimizes the difference against it. Yet, although the correlation with the captured dances had been slightly compromised by introducing variability in our representation model, we observed a significant increase in variance in comparison to the dance sequences synthesized without it. The mean variance ratio between the dance sequences synthesized with variability and the ones without it is in the order of $22.59 \pm 12.87\%$, among all metric resolutions and dance styles. Nevertheless, the results of Table 2 suggest that the induced variance is linearly disproportional to the parameterized metric resolution, which reveals a trade-off between the degree of variability and the degree of similarity with the analyzed dance. Therefore, at the optimal metric resolution for each dance style, we could only induce 20.33% and 2.34% of the variance observed in the captured dance, respectively for samba and Charleston. This can be justified by the use of a rough representation of the observed spatiotemporal variability through the use of homogeneous spherical distributions in the TGA analysis. As such, the validity of using a spherical approximation for representing the variability of the observed joint trajectories depends on the uniformity of the analyzed dance style, justifying the much lower proportion of variability induced in the synthesized dance sequences of Charleston than of samba (which is proportional to the difference in variance between the captured sequence of each dance style). As observed in Table 2, the captured Charleston's

dance sequence exhibits approximately six times the variance of samba's whereas the optimally synthesized dance sequence of Charleston exhibits approximately nine times less the variance of the optimally synthesized dance sequence of samba.

Ultimately, regarding the subjective assessment over samba (see Figure 10), we observed that the evaluation of the "fixed-4" parameterization was consistently less divergent than the one of "variability-4", enforcing the negative effect of the induced variability on feasibly representing the analyzed dance style. An explanation for this result may rely on the repetitive nature of the captured dance, which might imply that periodicity is considered by the subjects as a key factor in their assessment. Another justification may be the reliance on an incomplete representation of the observed variability by the lack of relative information among the represented topologies. This factor, combined with the use of uniform spherical distributions, could potentially lead to random combinations of movements that are perceived as unrealistic.

Nevertheless, although the proposed representation of variability was not convincing, and therefore the hypothesis of *variability* could not be fully confirmed both in numeric and subjective terms, there are enough considerations to support the notion that variability may be a fundamental quality to represent expressiveness of movement and consequently the naturalness observed in the performance of any particular dance style of repetitive gestures. By looking into Table 2 and Figure 10, this can be supported by the correlation "ceiling" at around 0.90 points for dance sequences of both dance styles synthesized without variability, the great differences in the variances measured in the synthesized and captured dance se-

quences, and the 1 point subjective similarity difference between the "original" and "fixed-4" sequences of samba against its captured dance sequence.

4.5 Towards humanoid robot dancing

The topological map provided by the TGA concept in the proposed dance representation model offers new perspectives for further manipulation of the dance gesture structure demanded by different motion retargeting requirements, without compromising the spatiotemporal structure of the original dance style. Such a parameterizable representation, in combination with the use of a motion synthesis method based on rotational quaternions and the use of a proper rotational joint interpolation method (e.g., slerps), offers a means for retargeting the captured dance trajectories onto different humanoid morphologies while overcoming the kinematic constraints imposed by their body models. Such an application can take advantage of the kinematic malleability of the TGA representation and the flexibility of quaternion algebra for synthesizing equivalent motion profiles adjusted to the new body segments' dimensions, and to the verified kinematic constraints, in terms of degrees-of-freedom and rotational limitations. A first approach towards retargeting beat-synchronous samba dance movements onto a simulated humanoid robot was described in [33]. The presented method manipulates and adapts the represented TGA topologies according to the target humanoid morphology, in terms of segment lengths, the number of joints, and the joints' degrees-of-freedom. From this morphologically adjusted dance representation we synthesized closed-loop sets of the represented key-poses (i.e., one set per metric cycle), and interpolated

them, using a sine interpolation function, according to the original musical meter in order to replicate the beat-synchrony of the analyzed dance.

A full implementation in a real humanoid robot requires further considerations that cannot be inferred from the proposed representation model. These include offline/online optimization (e.g., [34]) and/or dynamic control techniques (*e.g*, [35]) for refining the generated robot dance motion in order to ensure the humanoid's biped balance, avoid self-collisions, and overcome additional kinematic/dynamic constraints. Since the used dance representation is fully integrated with a formalized description of the music structure, an autonomous beat-synchronous robot dancing system will also require a real-time beat tracker (already developed in [36]) for synchronizing the generated dance behaviors on-the-fly to live musical stimuli. This beat-synchrony can be reproduced at different resolutions according to the metric parameterization of the TGA model. A design for improving the real-time beat tracking performance in the presence of ego-motion noise of a dancing robot was already proposed and evaluated on [37]. A first approach towards synchronizing humanoid robot dancing movements to online musical stimuli was also already implemented on [38].

5 Conclusions

In this study we proposed a parameterizable spatiotemporal representation of human dance movements applicable for the generation of expressive dance movements onto different humanoid dancing characters. The proposed dance representation model was assessed according to two hypotheses, namely the impact of metric resolution and the impact of variability

towards fully and naturally representing a particular popular dance style built on repetitive gestures. The overall results validate the use of the TGA model as a reversible form of data representation, and consequently compression, which indicates that it can be applied for motion analysis and synthesis of musically-driven dance styles for humanoid dancing characters.

The proposed method starts from information of the captured dance, recorded with a motion capture system combined with musical information, which is packed into a spatiotemporal representation of the captured dance movements in the form of a topological model (TGA). This representation was re-synthesized into dance sequences using different parameterizations and compared against MoCap recordings of real performances of popular dance styles, namely of samba and Charleston. The results seem to confirm the hypothesis that there is a minimum and sufficient temporal metric resolution required to fully represent a particular popular dance style of repetitive gestures. Specifically, for the analyzed dance styles of samba and Charleston, quarter-beat representations offered both a sufficient level of similarity to the captured dance while consequently offering a great compression of the captured signal. Smaller resolutions offer a decreasing reproduction of the analyzed dance with the trade-off of an increased compression ratio. Concerning the impact of variability, both numeric and subjective evaluations pointed to no positive effects on considering spatiotemporal variability into our representation model, and that the proposed representation of variability, at the optimal metric resolution, offers only some extent of the variance observed in the analyzed dances. This can be justified by the use of a rough and incom-

plete representation of the observed spatiotemporal variability, by the use of homogeneous spherical distributions in the TGA model, and by missing relative information among the represented topologies. These could lead to random combinations of movements that are perceived as unnatural and might generate discrepant motion trajectories.

Further studies are needed in order to clarify the role of spatiotemporal variability and the importance of specific body parts in the perception of expressiveness in popular dance styles. In the future we should also verify the applicability of the proposed representation model and hypotheses on other popular dance styles, with different metrical structures (e.g., dances at the 3-beat bar level of the Waltz music forms).

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was partly supported by the COST-SID Action, with an STSM hosted by Prof. Marc Leman at IPEM, Ghent University (Belgium), with ref. COST-STSM-IC0601-4790, a PhD scholarship endorsed by the Portuguese Foundation for Science and Technology (FCT) Portugal, with ref. SFRH/BD/43704/2008, and by FWO (Belgium). It was also partly supported by the European Commission, FP7 (Seventh Framework Programme), ICT-2011.1.5 Networked Media and Search Systems, grant agreement No 287711; and the European Regional Development Fund through the Programme COMPETE and by Na-

tional Funds through the FCT, within projects ref. PTDC/EAT-MMU/112255/2009 and PTDC/EIA-CCO/111050/2009.

Endnotes

[a]Dance representation and dance representation model are used indistinctively throughout the article and refer to a formalized description or "visualization" of the dance by means of a systematic analysis of its spatiotemporal structure. [b]The expression spatiotemporal variability refers to the distribution of the positions in space where the limbs of a dancer hit specific music cues in time (thus, spatiotemporal variability in dance). It is well known that dancers and musicians do not perform repetitive movements or events at the precise time points or positions. Such variation is claimed to be related to perceived expressiveness, naturalness and expertise and are ubiquitous in human performances (see [39–43]). [c]For examples of previous studies that support this assumption in Jazz see [44–46]; for studies in Afro-Brazilian Music see [39, 47, 48]

References

1. L Naveda, M Leman, The spatiotemporal representation of dance and music gestures using Topological Gesture Analysis (TGA). Music Percept. **28**(1), 93–111 (2010)

2. JL Oliveira, L Naveda, F Gouyon, M Leman, LP Reis, Synthesis of dancing motions based on a compact topological representation of dance styles, in *Workshop on Robots and Musical*

Expressions (WRME) at IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS), Taipei, Taiwan, 2010, p. 8

3. A Nakazawa, S Nakaoka, K Ikeuchi, K Yokoi, Imitating human dance motions through motion structure analysis, in *In Proc. of International Conference on Intelligent Robots and Systems*, EPFL, Switzerland, 2002, pp. 2539–2544

4. T-H Kim, SI Park, SY Shin, Rhythmic-motion synthesis based on motion-beat analysis. ACM Trans. Graph. **22**(3), 392–401 (2003)

5. H-C Lee, I-K Lee, Automatic synchronization of background music and motion in computer animation. Comput. Graph. Forum **24**(3), 353–362 (2005)

6. G Kim, Y Wang, H Seo, Motion control of a dancing character with music, in *ACIS-ICIS*, Melbourne, Australia, 2007, pp. 930–936

7. T Shiratori, A Nakazawa, K Ikeuchi, Detecting dance motion structure through music analysis, in *IEEE International Conference on Automatic Face and Gesture Recognition (FGR)*, Seoul, Korea, 2004, pp. 857–862

8. S Nakaoka, A Nakazawa, K Yokoi, H Hirukawa, K Ikeuchi, Generating whole body motions for a biped humanoid robot from captured human dances, in *IEEE International Conference on Robotics and Automation (ICRA)*, Taipei, Taiwan, 2003, pp. 3905–3910

9. T Shiratori, A Nakazawa, K Ikeuchi, Dancing-to-music character animation. Comput. Graph. Forum **25**(3), 449–458 (2006)

10. R Laban, L Ullmann, *Mastery of Movement* (Princeton Book Company Publishers, 1960)

11. G Alankus, AA Bayazit, OB Bayazit, Automated motion synthesis for dancing characters. J. Vis. Comput. Anim. **16**(3–4), 259–271 (2005)

12. N Stergiou, *Innovative Analyses of Human Movement* (Human Kinetics Publishers, 2004)

13. JL Hanna, RD Abrahams, NR Crumrine, R Dirks, RV Gizycki, P Heyer, A Shapiro, Y Ikegami, AL Kaeppler, JW Kealiinohomoku, Movements toward understanding humans through the anthropological study of dance. Curr. Anthropol. **20**(2), 313–339 (1979)

14. F Lerdahl, R Jackendoff, R Jackendoff, *A Generative Theory of Tonal Music* (The MIT Press, 1996)

15. HC Longuet-Higgins, CS Lee, The rhythmic interpretation of monophonic music. Music Percept. **1**(4), 424–441 (1984)

16. K Tomlinson, *The Art of Dancing Explained by Reading and Figures, repr* (Westmead Gregg International, London, 1735)

17. R Laban, FC Lawrence, *Effort* (Macdonald and Evans, London, 1947)

18. AR Jensenius, Using motiongrams in the study of musical gestures, in *In Proceedings of the 2006 International Computer Music Conference* (Scholarly Publishing Office, University of Michigan University Library, New Orleans, USA, 2006), pp. 6–11

19. M Palazzi, NZ Shaw, Synchronous objects for one flat thing, reproduced, in *SIGGRAPH 2009: Talks* (ACM, New Orleans, Louisiana, USA, 2009) p. 2

20. A Saint-Léon, F Pappacena, *La Sténochorégraphie, 1852* (Libreria Musicale Italiana, 2006)

21. NaturalPoint, Optitrack, http://www.naturalpoint.com/optitrack. Accessed 13 Jan 2012

22. P Toiviainen, B Burger, MoCap Toolbox Manual, Jyväskylä, Finland. (2008), http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual. Accessed 13 Jan 2012

23. C Cannam, C Landone, M Sandler, JP Bello, The sonic visualiser: a visualisation platform for semantic descriptors from musical signals, in *Proceedings of the International Conference on Music Information Retrieval*, Victoria, Canada, 2006, pp. 324–327

24. G Carlsson, Topology and data. J. Bull. Am. Math. Soc. **46**, 255–308 (2009)

25. DG Morrison, On the interpretation of discriminant analysis. J. Market Res. **6**(2), 156–163 (1969)

26. A Bruderlin, L Williams, Motion signal processing, in *SIGGRAPH '95 Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 1995, pp. 97–104

27. T Shiratori, K Ikeuchi, Synthesis of dance performance based on analyses of human motion and music. IPSJ Online Trans. **1**, 80–93 (2008)

28. A Ude, CG Atkeson, M Riley, Planning of joint trajectories for humanoid robots using B-spline wavelets, in *IEEE International Conference on Robotics and Automation (ICRA)*, San Francisco, California, USA, 2000, pp. 2223–2228

29. J Dauwels, F Vialatte, T Musha, A Cichocki, A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. NeuroImage **49**(1), 668–93 (2010)

30. L Naveda, M Leman, D Mota, DAS—Dance Analysis Suite, Ghent (2010), http://navedacodex.tumblr.com/das. Accessed 13 Jan 2012

31. L Nelson, Evaluating overlapping confidence intervals. J. Qual. Technol. **21**(2), 140–141 (1989)

32. RA McCarthy, AF Blackwell, S DeLahunta, AM Wing, K Hollands, PJ Barnard, I Nimmo-Smith, AJ Marcel, Bodies meet minds: choreography and cognition. Leonardo **39**(5), 475–477 (2006)

33. P Sousa, JL Oliveira, LP Reis, F Gouyon, Humanized robot dancing: humanoid motion retargeting based in a metrical representation of human dance styles, in *Progress in Artificial Intelligence, 15th Portuguese Conference on Artificial Intelligence, EPIA*, ed. by L Antunes, HS Pinto, Lecture Notes in Computer Science, vol. 7026 (Springer, Lisbon, Portugal, 2011), pp. 392–406

34. L Cruz, LP Reis, L Rei, Generic optimization of humanoid robots' behaviours, in *15th Portuguese Conference on Artificial Intelligence, EPIA*, Lisbon, Portugal, Oct 2011, pp. 385–397

35. E Yoshida, C Esteves, I Belousov, J Laumond, T Sakaguchi, K Yokoi, Planning 3-d collision-free dynamic robotic motion through iterative reshaping. IEEE Trans. Robot. **24**(5), 1186–1198 (2008)

36. JL Oliveira, F Gouyon, LG Martins, LP Reis, IBT: A real-time tempo and beat tracking system, in *International Society for Music Information Retrieval (ISMIR)*, Utrecht, The Netherlands, 2010, pp. 291–296

37. JL Oliveira, G Ince, K Nakamura, K Nakadai, Online audio beat tracking for a dancing robot in the presence of ego-motion noise in a real environment, in *IEEE International Conference on Robotics and Automation (ICRA)*, Minnesota, USA, May 2012, pp. 403–408

38. CB Santiago, J Oliveira, L Reis, A Sousa, Autonomous robot dancing synchronized to musical rhythmic stimuli, in *First Workshop in Information Systems for Interactive Spaces (WISIS), 6th Iberian Conference on Information Systems and Technologies (CISTI)*, Castelo Branco, Portugal, June 2011, pp. 1002–1007

39. L Naveda, F Gouyon, C Guedes, M Leman, Microtiming patterns and interactions with musical properties in samba music. J. New Music Res. **40**(3), 223–236 (2011)

40. P Desain, H Honing, Tempo curves considered harmful. Contemp. Music Rev. **7**(2), 123–138 (1993)

41. R Chaffin, AF Lemieux, C Chen, It is different each time I play: variability in highly prepared musical performance. Music Percept. **24**(5), 455–472 (2007)

42. D Krasnow, M Wilmerding, S Stecyk, M Wyon, Y Koutedakis, Biomechanical research in dance: a literature review. Medi. Probl. Perform. Art. **26**(1), 3 (2011)

43. ML Latash, JP Scholz, G Schöner, Motor control strategies revealed in the structure of motor variability. Exer. Sport Sci. Rev. **30**(1), 26 (2002)

44. JA Prögler, Searching for swing: participatory discrepancies in the jazz rhythm section. Ethnomusicology **39**(1), 21–54 (1995)

45. RP Dodge, N Jazz, More than Rhythm The Charleston, *Culture Makers: Urban Performance and Literature in The 1920s*, 2009, p. 64

46. F Benadon, Slicing the beat: Jazz eighth-notes as expressive microrhythm. Ethnomusicology **50**(1), 73–98 (2006)

47. M Sodré, *Samba, O Dono do Corpo* (Codecri, Rio de Janeiro, 1979)

48. C Sandroni, *Feitiço Decente: Transformações do samba no Rio de Janeiro, 1917–1933* (Jorge Zahar, Rio de Janeiro, 2001)

49. R Laban, Schrifttanz [Writing Dance] (Universal Edition, Vienna, Austria, 1928)

Figure 1. Spatiotemporal representation of musical meter and dance: **(a)** Hierarchical representation of the structure of meter (based on [14]), with a period of 2 beats. From top to bottom, each hierarchical metric level is subdivided or grouped in other levels. **(b)** Spatiotemporal representation of metric accents in a dance gesture.

Figure 2. Five different spatiotemporal representations of the dance gesture: **(a)** Tomlinson [16] proposed representations that guide steps distributed in the dance floor; **(b)** Saint-Leon and Pappacena [20] developed a mixture of musical score and figurative descriptions of key-poses to represent music and dance in the same process; **(c)** Laban [49] developed the *labanotation* method, perhaps the most disseminated form of dance notation; **(d)** Jensenius [18] developed a representation based on video recordings whose pixels are collapsed and inform about movement in time; **(e)** Palazzi and Shaw [19] used videogrammetry to create a set of 3D video representations of dance.

Figure 3. Workflow of the method: **(a)** dance analysis and representation, and **(b)** validation of the proposed representation model according to different hypotheses.

Figure 4. Projection of musical cues (metric classes) onto the dance trajectories. Firstly, **(a)** the annotation of metric structure of the **(b)** music is synchronized with the MoCap recording. These cues are projected onto **(c)** the movement vectors (in the example, right hand movements) as different classes of points (e.g., 1st beat, 2nd beat—respectively described as 1 and 2 in the figure). Finally, **(d)** the point clouds are discriminated using LDA analysis which guarantees the separation of point-clouds into **(e)** topologies. In this study we assumed a spherical distribution for the point clouds whose radius is defined by the average of the Euclidean distances from all points of the class to the mean.

Figure 5. Process of key-pose synthesis with variability, from kinematic chains decomposition to the stochastic calculation of the joints' rotations. The top graph shows **(a)** the decomposition of the body model into five kinematic chains, and **(b)** the sequence of propagation of stochastic processes along the kinematic chain of the character's right arm. The bottom graph shows the proposed solution for generating the key-pose of metric class m with variability, by replicating the same variability observed in the recorded dance. This process can be implemented by **(c)** randomly calculating all key-pose's joint rotations: starting from the anchor segment, s_0^m, at the spine, which links joint 1 to joint 10, to the chain extremity at joint 20, each joint position, p_j^m, is randomly calculated inside its respective, C_j^m, by selecting a random quaternion, $qv_{s_j^m}$, that describes a possible rotation of that joint segment, s_j^m, around its base unity vector, $\vec{v}_{s_j^m}$, (given by the last segment target vector, $\vec{v}'_{s_{j-1}^m}$), circumscribed by C_j^m.

Figure 6. Spatiotemporal dance representation model of samba, parameterized with quarter-beat resolution and variability, within two-beat metric cycles (i.e., dance represented by the spherical distributions of eight metric classes, which correspond to $\frac{1}{4}$resolution $*$ 2beats): **(a)** point cloud representation of the dance gesture of the left hand; **(b)** point cloud after LDA analysis. Note that classes of points are visually and linearly discriminated from each other; **(c)** representation of point clouds as homogeneous spherical distributions around the mean trajectories of the left hand.

Figure 7. Synthesis of dance sequences from representation models parameterized with different metric resolutions, within two-beat metric cycles: concatenating closed-loop cycles of the represented key-poses (i.e., one different key-pose per metric class and one full set of key-poses per metric cycle), and interpolating them according to the represented musical meter at different metric resolutions.

Figure 8. Generating one movement cycle of the right hand joint by orderly interpolating the discrete joint positions calculated for all defined metric classes at different resolutions.

Figure 9. DAS visualization of a synthesized samba dance sequence synchronized to music.

Figure 10. Overall statistical results for the subjective evaluation over the level of similarity of each synthesized dance sequence, plus the "original" sequence, of samba to the captured dance sequence.

Figure 11. Captured (straight) versus synthesized trajectories of the right hand joint (joint 19) for four metric cycles (delimited by the vertical lines) of dance sequences of Charleston. The synthesized trajectories are generated from representations parameterized with (dashed) or without (dotted) variability, at the following metric resolutions: **(a)** beat—"variability-1"/"fixed-1"; **(b)** half-beat—"variability-2"/"fixed-2"; **(c)** quarter-beat—"variability-4"/"fixed-4"; **(d)** eighth-beat—"variability-8"/"fixed-8".

Table 1. Assessed parameterizations for validating the proposed representation model in respect to the stated hypotheses

Parameterization	**Metric resolution**	**Spatiotemporal variability**
Fixed-1	Beat	None
Fixed-2	Half-beat	None
Fixed-4	**quarter-beat**	**None**
Fixed-8	Eighth-beat	None
Variability-1	**Beat**	**Spherical distribution**
Variability-2	**Half-beat**	**Spherical distribution**
Variability-4	**quarter-beat**	**Spherical distribution**
Variability-8	Eighth-beat	Spherical distribution

The numerical evaluation considers all the present parameterizations whereas the subjective assessment only considers the parameterizations in bold.

Table 2. Correlation coefficient, $r_{s,o}$, between the joint trajectories of the assessed dance Sequence, s, and the captured dance sequence, o, and mean variance, $\bar{v_s}$, among the metric cycles composing the assessed sequence in relation to the dimensionality, *Dim*, and level of reduction, *Reduction*, of its respective representation model

Style	**Sequence**	$\mathbf{r_{s,o}}$	$\bar{\mathbf{v}}_\mathbf{s}(mm^2)$	**Dim**(J,S,W)	**Reduction**	$\mathbf{e_i}(\%)$
Samba						
	Original	0.86	646.84 (846.86)	20×3×N	0	NA
	Fixed-1	0.46 (0.00)	133.12 (527.68)	20×3×2 = 120	0.50×N	2.11
	Fixed-2	0.81 (0.00)	60.45 (314.44)	20×3×4 = 240	0.25×N	1.08
	Fixed-4	**0.89 (0.00)**	**24.56 (66.47)**	**20×3×8 = 480**	**0.13×N**	**0.31**
	Fixed-8	0.88 (0.00)	13.73 (31.96)	20×3×16 = 960	0.06×N	0.62
	Variability-1	0.41 (0.01)	539.59 (683.64)	20×(3+1)×2 = 160	0.38×N	1.61
	Variability-2	0.77 (0.01)	258.50 (406.60)	20×(3+1)×4 = 320	0.19×N	0.23
	Variability-4	**0.87 (0.00)**	**131.52 (166.68)**	**20×(3+1)×8 = 640**	**0.09×N**	**0.09**
	Variability-8	0.87 (0.00)	64.94 (91.46)	20×(3+1)×16 = 1280	0.05×N	0.10

Table 2. continues

Style	Sequence	$\mathbf{r_{s,o}}$	$\bar{\mathbf{v}}_\mathbf{s}(mm^2)$	**Dim**(J,S,W)	**Reduction**	$\mathbf{e_i}$(%)
Charleston						
	Original	0.87	3772.72 (5793.63)	20×3×N	0	NA
	Fixed-1	0.74 (0.00)	157.14 (540.80)	20×3×4 = 240	0.25×N	5.82
	Fixed-2	0.87 (0.00)	21.27 (77.07)	20×3×8 = 480	0.13×N	2.31
	Fixed-4	**0.90 (0.00)**	**6.89 (16.33)**	**20×3×16 = 960**	**0.06×N**	**1.93**
	Fixed-8	0.89 (0.00)	7.07 (18.18)	20×3×32 = 1920	0.03×N	2.53
	Variability-1	0.73 (0.01)	319.27 (336.48)	20×(3+1)×4 = 320	0.19×N	4.60
	Variability-2	0.86 (0.00)	241.19 (275.32)	20×(3+1)×8 = 640	0.09×N	1.11
	Variability-4	**0.89 (0.00)**	**88.40 (156.87)**	**20×(3+1)×16 = 1280**	**0.05×N**	**0.19**
	Variability-8	0.89 (0.00)	26.16 (65.19)	20×(3+1)×32 = 2560	0.03×N	0.05

The mean error caused by the used interpolation method in each synthesized dance sequence is given by e_i. The numbers in parentheses refer to the standard deviation of, respectively, the mean $r_{s,o}$ and the mean $\bar{v_s}$ across the ten synthesized dance sequences for each applied parameterization.

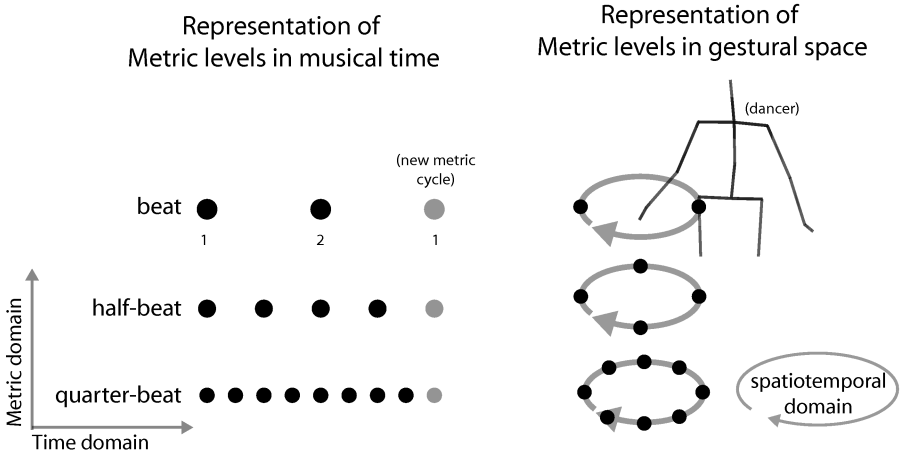
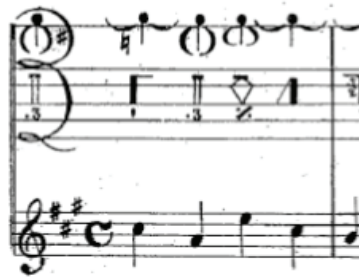
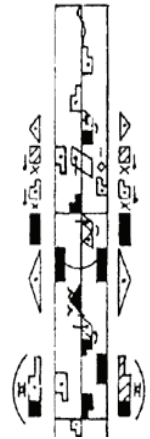

Representation of
Metric levels in musical time
(new metric
cycle)
beat
1
2
1
half-beat
quarter-beat
Metric domain
Time domain
a
Representation of
Metric levels in gestural space
(dancer)
spatiotemporal
domain
b

Figure 1

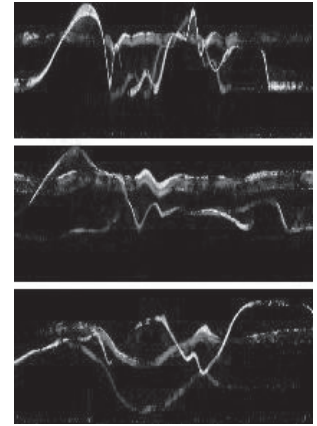
Tomlinson
1795



Saint-Léon
1852



Laban
1947



Jensenius
2006

Palazzi & Shaw
2009

Figure 2

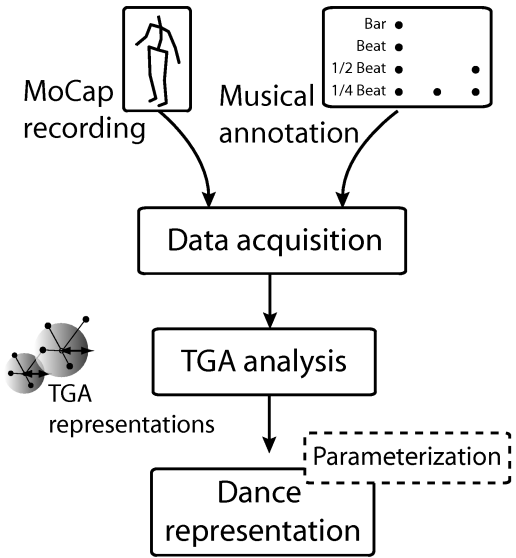

Analysis - Synthesis
MoCap recording
Musical annotation
Bar
Beat
1/2 Beat
1/4 Beat
Data acquisition
TGA analysis
TGA representations
Parameterization
Dance representation
Assessment - Validation
Parameterization
Dance representation
Synthesized dance sequences
MoCap recording
versus
Numerical evaluation
Subjective evaluation
Hypotheses

Figure 3

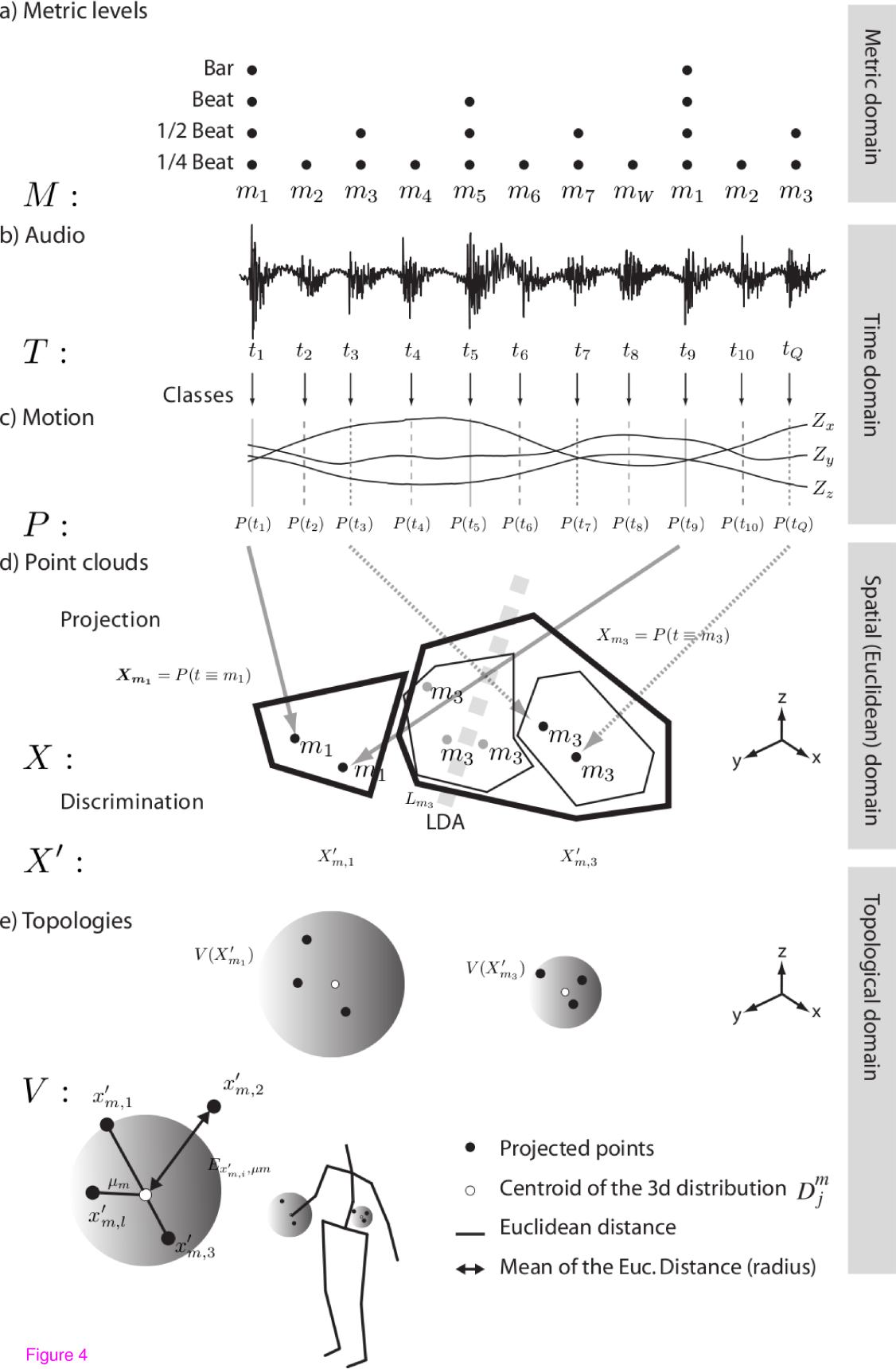

a) Metric levels
Bar
Beat
1/2 Beat
1/4 Beat
M :
m_1 m_2 m_3 m_4 m_5 m_6 m_7 m_W m_1 m_2 m_3
b) Audio
T :
t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_Q
Classes
c) Motion
Z_x
Z_y
Z_z
P :
$P(t_1)$ $P(t_2)$ $P(t_3)$ $P(t_4)$ $P(t_5)$ $P(t_6)$ $P(t_7)$ $P(t_8)$ $P(t_9)$ $P(t_{10})$ $P(t_Q)$
d) Point clouds
Projection
$X_{m_1} = P(t \equiv m_1)$
$X_{m_3} = P(t \equiv m_3)$
X :
Discrimination
L_{m_3}
LDA
X' :
$X'_{m,1}$
$X'_{m,3}$
e) Topologies
$V(X'_{m_1})$
$V(X'_{m_3})$
z
y
x
V :
$x'_{m,1}$
$x'_{m,2}$
$x'_{m,3}$
$x'_{m,l}$
μ_m
$E_{x'_{m,i},\mu m}$
Projected points
Centroid of the 3d distribution D_j^m
Euclidean distance
Mean of the Euc. Distance (radius)
Metric domain
Time domain
Spatial (Euclidean) domain
Topological domain

Figure 4

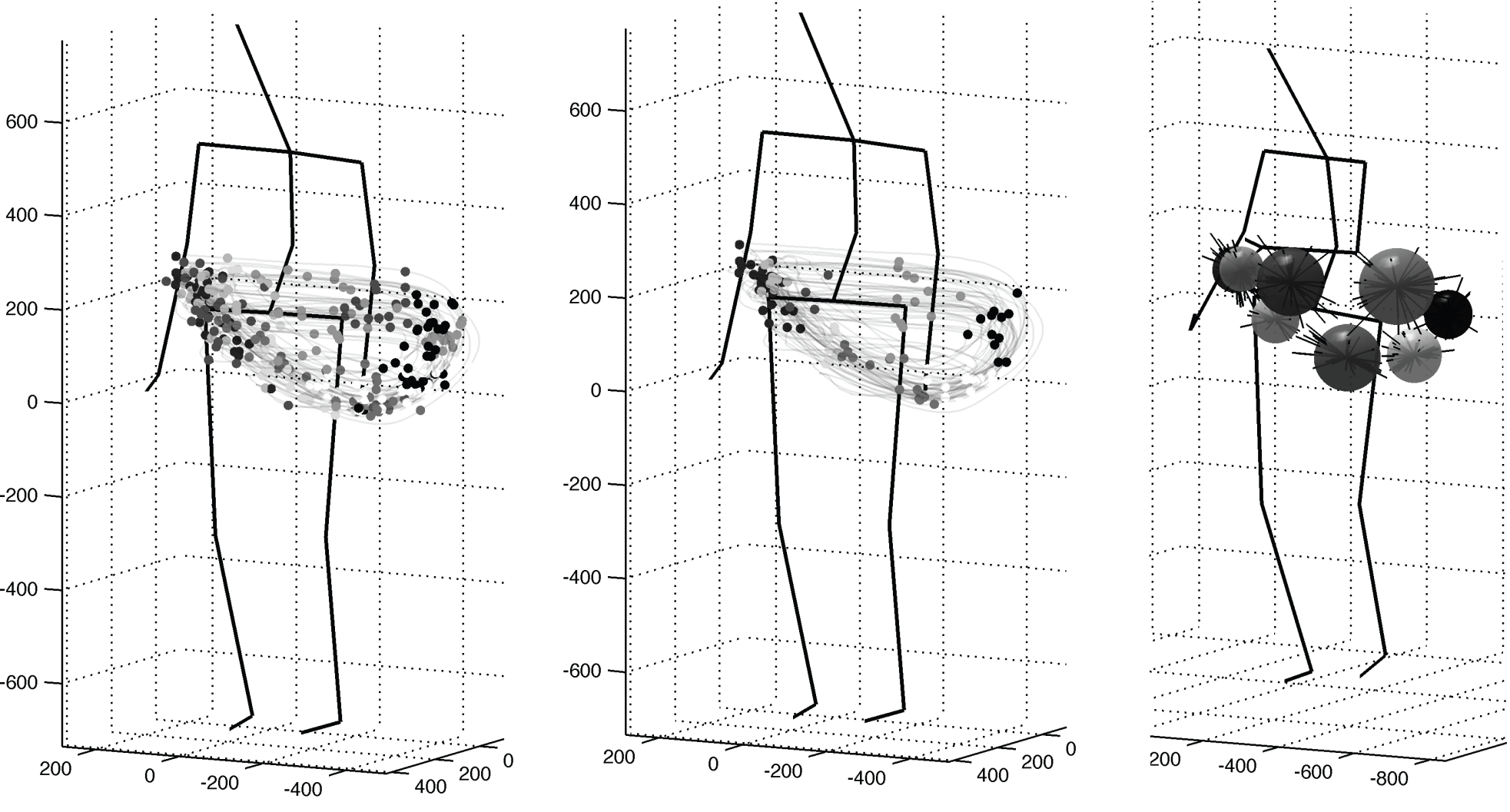


Figure 5

a) Original point cloud

b) Point cloud after LDA

c) Spherical distributions

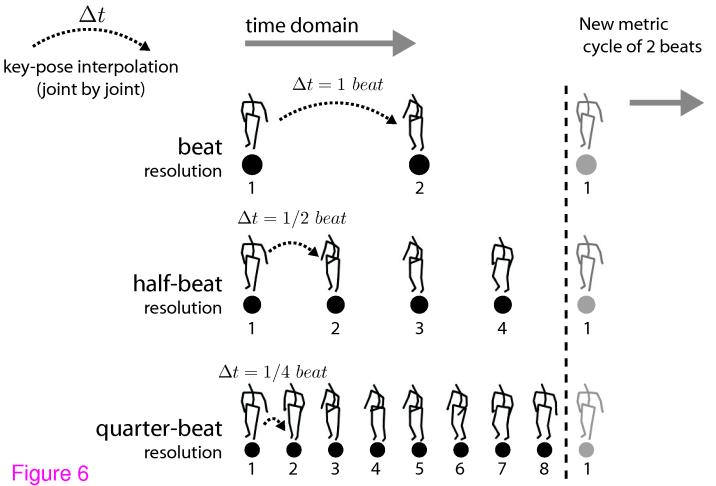

Δt
key-pose interpolation
(joint by joint)
time domain
New metric
cycle of 2 beats
$\Delta t = 1\ beat$
beat
resolution
1
2
1
$\Delta t = 1/2\ beat$
half-beat
resolution
1
2
3
4
1
$\Delta t = 1/4\ beat$
quarter-beat
resolution
1
2
3
4
5
6
7
8
1

Figure 6

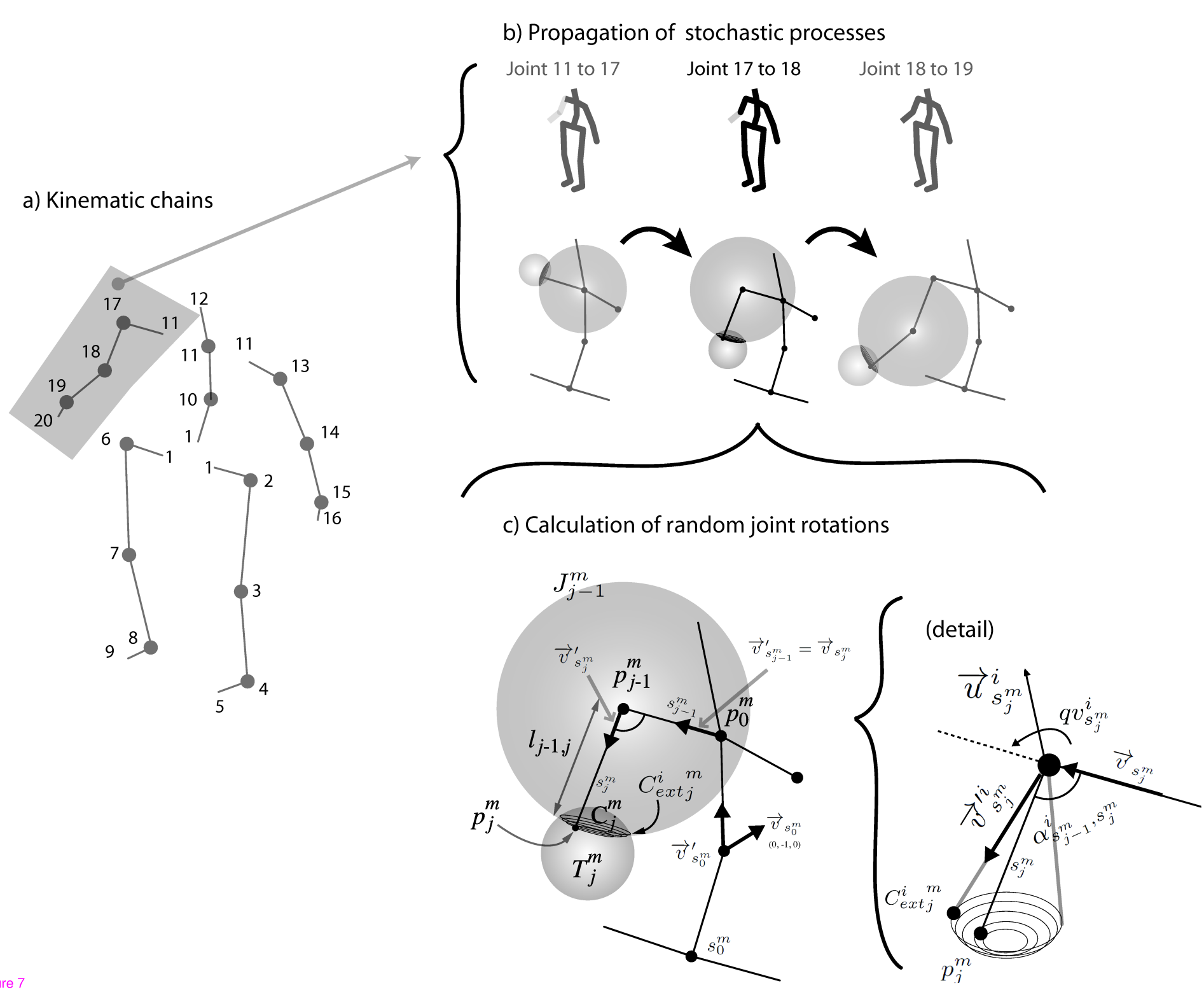
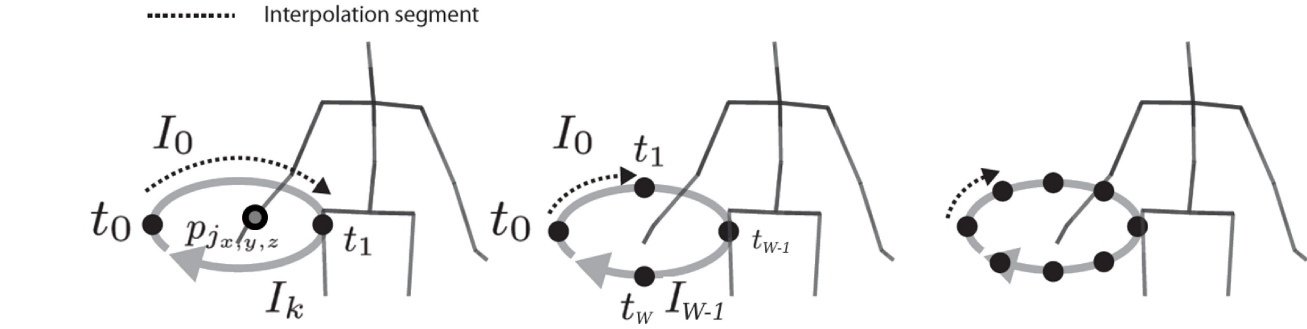

a) Kinematic chains
b) Propagation of stochastic processes
Joint 11 to 17
Joint 17 to 18
Joint 18 to 19
c) Calculation of random joint rotations
(detail)

Figure 7



Resolution 1 beat half-beat quarter-beat

Figure 3

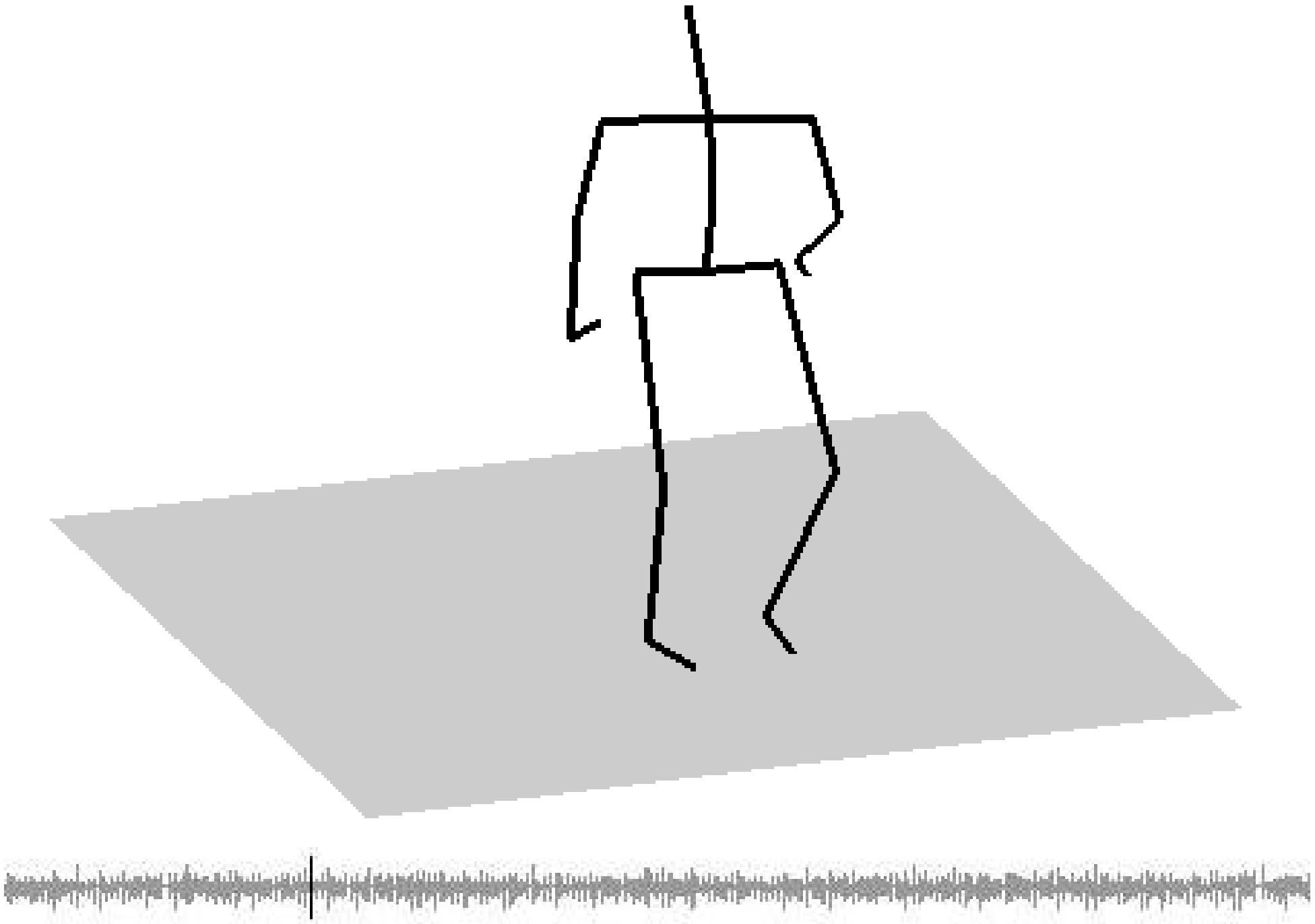


Figure 9

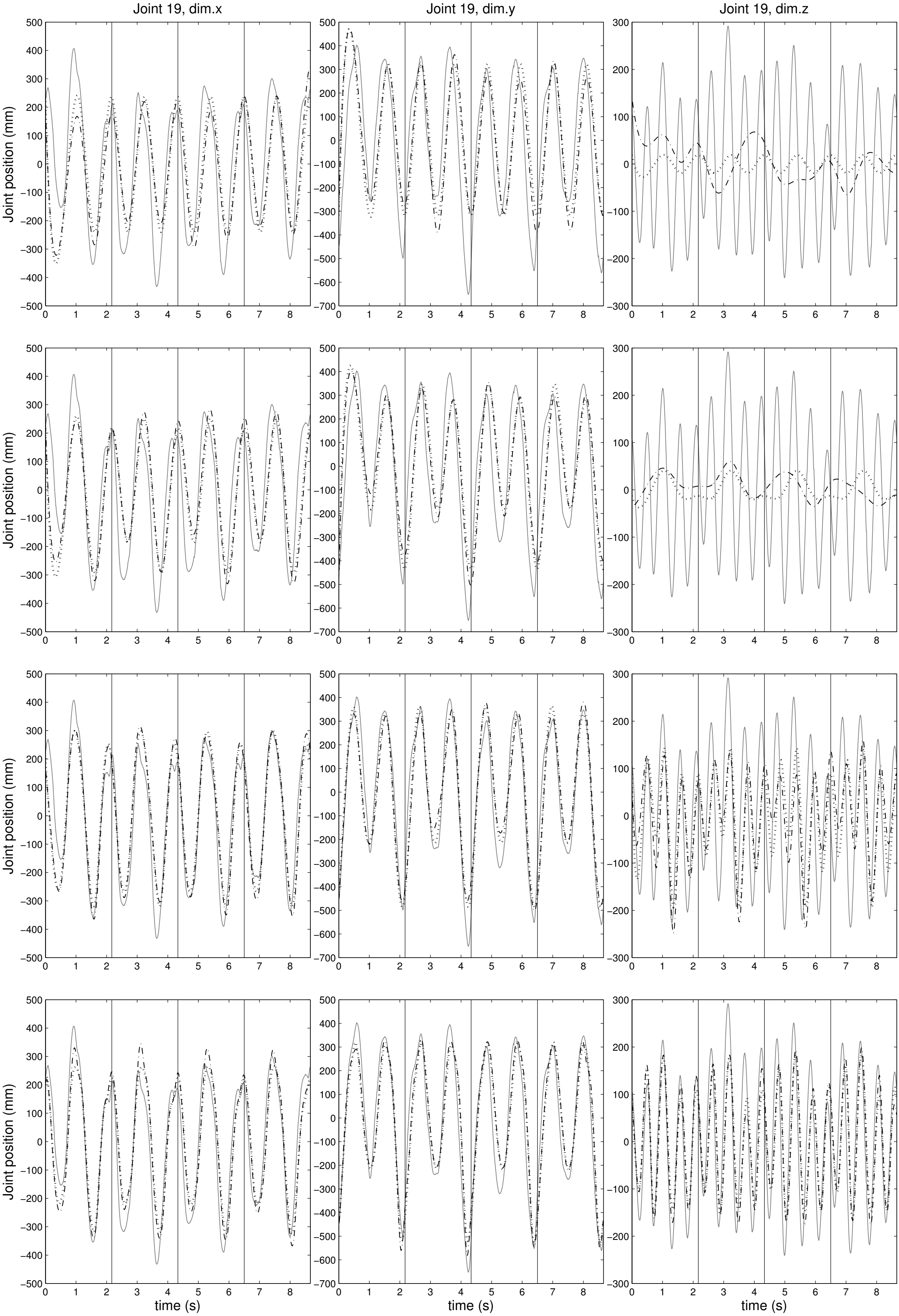
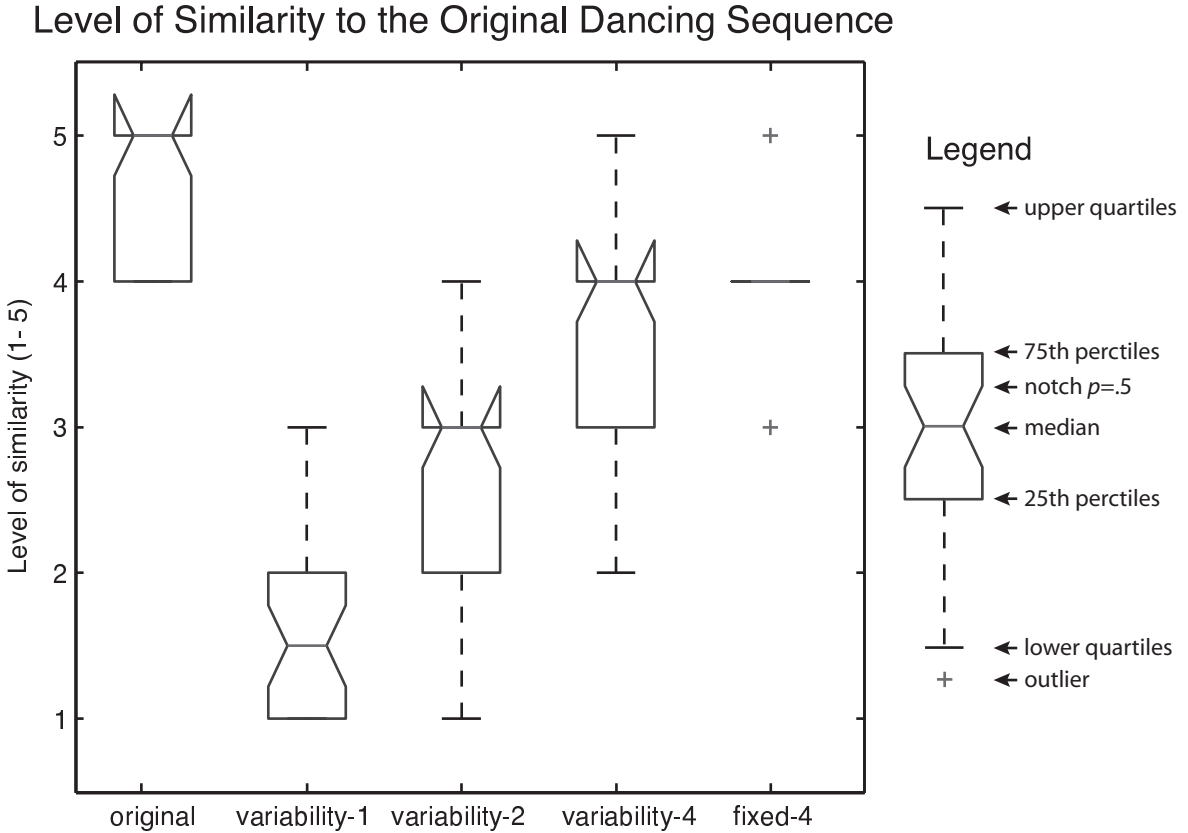

Joint 19, dim.x
Joint 19, dim.y
Joint 19, dim.z
Joint position (mm)
time (s)

Figure 10


Level of Similarity to the Original Dancing Sequence
Level of similarity (1- 5)
5
4
3
2
1
original
variability-1
variability-2
variability-4
fixed-4
Legend
upper quartiles
75th perctiles
notch p=.5
median
25th perctiles
lower quartiles
outlier

igure 11