Synthesis of Variable Dancing Styles Based on A Compact Spatiotemporal Representation of Dance

João Lobato Oliveira† and Luiz Naveda‡ and Fabien Gouyon* and Marc Leman‡ and Luis Paulo Reis+

Abstract— **Dance as a complex expressive form of motion is able to convey emotion, meaning and social idiosyncrasies that opens channels for non-verbal communication, and promotes rich cross-modal interactions with music and the environment. As such, realistic dancing characters may incorporate cross-modal information and variability of the dance forms through compact representations that may describe the movement structure in terms of its spatial and temporal organization. In this paper, we propose a novel method for synthesizing beat-synchronous dancing motions based on a compact topological model of dance styles, previously captured with a motion capture system. The model was based on the Topological Gesture Analysis (TGA) which conveys a discrete three-dimensional point-cloud representation of the dance, by describing the spatiotemporal variability of its gestural trajectories into uniform spherical distributions, according to classes of the musical meter. The methodology for synthesizing the modeled dance traces back the topological representations, constrained with definable metrical and spatial parameters, into complete dance instances whose variability is controlled by stochastic processes that considers both TGA distributions and the kinematic constraints of the body morphology. In order to assess the relevance and flexibility of each parameter into feasibly reproducing the style of the captured dance, we correlated both captured and synthesized trajectories of samba dancing sequences in relation to the level of compression of the used model, and report on a subjective evaluation over a set of six tests. The achieved results validated our approach, suggesting that a periodic dancing style, and its musical synchrony, can be feasibly reproduced from a suitably parametrized discrete spatiotemporal representation of the gestural motion trajectories, with a notable degree of compression.**

I. INTRODUCTION

The process of generating human-like motions plays a key role in robotics, computer graphics, computer games and virtual reality systems. On other hand, the success in reproducing natural human body motions may be highly improved by introducing expressiveness and style which is able to convey emotion, meaning and social idiosyncrasies. Dance movements form a complex class of human motions that offer infinite forms of expressiveness, modes of non-verbal communication, diverse cultural vocabularies and a rich use of multimodal interactions with music and other modalities. It imposes fascinating challenges to robotics and outstanding opportunities to deepen our understanding about the phenomenon of dance.

State of the art applications in robotics/computer animation and dance often manipulate captured instances of dance performances as temporal sequences that are wrapped [1], [2], [3], merged [1], [2], [4] and segmented [5], [6], [7]. Such applications are specially relevant in the universe of popular dances, which often exhibit a limited vocabulary of idiosyncratic gestures and frequently exhibit periodic spatiotemporal characteristics. However, the manipulation of static temporal instances of dance performances are not sufficient to provide the necessary quality attached to human movement and the proper balance between variability and predictability. The expressiveness of human movement appears to be influenced by stochastic processes and cross-modal links that affect the dancer's reasoning about space and all the kinematic and kinetic constraints of his/her body in space. Realistic dancing characters may benefit from methods that manage cross-modal information and variability and incorporate perception-action loops verified in reality. Methods that incorporate *generative models* of dance forms may offer more quality and less processing load than fixed *temporal instantiations* of dance performances.

In this study, we propose an original method for generating beat-synchronous dancing sequences based on a topological model of a dance style. The model builds on TGA (Topological Gesture Analysis) method [8], which offers a formalized description that conveys both spatiotemporal variability and cross-modal characteristics of dance gestures. This method relies in the projection of cues (e.g. musical cues) onto the gestural trajectories, which generates spatiotemporal point cloud representations in the three-dimensional space. Point clouds are filtered, discriminated and then interpreted as distributions or topologies in space. The characteristics of these clouds inform about the space and the variability of the dance in relation to the musical cue. We used metrical cues (e.g.: bar, beat, half beat points) to find these distributions of gestural space according to musical meter. Our methodology for generating dance sequences traces back the topological representations into complete dance instances whose variability is controlled by stochastic processes and complementary parameters manipulated in accordance to the interactive model. This process involves (i) a solution of kinematic constraints, (ii) a stochastic mechanism to find choices within the representation of variability and a (iii) heuristics for interpolation of key-poses. We used a selection of these parameters to generate sequences of samba dances

This work was supported by Fundação para a Ciência e Tecnologia (FCT), Portugal and by Fonds Wetenschappelijk Onderzoek (FWO), Belgium.

†J. L. Oliveira is with Institute for Systems and Computer Engineering of Porto (INESC Porto) and Artificial Intelligence and Computer Science Laboratory (LIACC), FEUP, Porto University, Portugal jmso@inescporto.pt

‡L. Naveda and M. Leman are with Institute for Psychoacoustics and Electronic Music (IPEM), Ghent University, Belgium luiznaveda@gmail.com, Marc.Leman@Ugent.be

*F. Gouyon is with INESC Porto fgouyon@inescporto.pt

+L. P. Reis is with LIACC lpreis@fe.up.pt

modeled from a motion capture recording of a professional samba dancer. In order to assess the relevance and flexibility of each parameter into feasibly reproducing the style of the captured dance, we correlated both captured and synthesized trajectories of samba dancing sequences in relation to the level of compression of the used model, and report on a subjective evaluation over a set of six tests. Such tests consider different temporal and spatial hypotheses of interaction between the movement and music, regarding the chosen metrical resolution and the integration of spatial variability.

The paper is structured as follows: in the Section II we review related work on the synthesis of beat-synchronous dancing and rhythmic motions from mocap data. Section III specify the details of modeling, representation and re-synthesis of dance sequences. Section IV describes our evaluation method and discusses the main results concerning the level of similarity and style representation of a set of synthesized dancing sequences over captured samba dance. Finally, in the Section V we conclude this paper and present some paths for future work.

II. Related Work

The generation of human motion data from scratch has proven to be time expensive and complex, even when techniques such as key-framing and physical simulation are available. A common solution for this problem is to capture human motion data using motion capture systems (e.g.: [9], [10], [11], [12], [6], [13]), which became increasingly accessible in the last years.

In the attempt to bring expressiveness to movement sequences, researchers have often opted to manipulate or transform of human motion recordings rather than deal with the biomechanical, choreological and musicological network and reasoning behind the gestures. For example, early computer animation researchers developed techniques for processing [9] and retargetting [14] motion capture data in order to automatically generate new character animations. Later approaches explored the rearrangement and blending of mocap clips, recurring to motion graphs [15], [11] and statistical models [10] for synthesizing smooth motion transitions, while flexibly and on-the-fly [16] satisfying user-defined constraints. Others dealt with the compression and representation of mocap data recurring to Bezier curves and clustered Principal Component Analysis (PCA) [17], or by exploring the redundancy of motion patterns [18], to reduce dimensionality of the data with minor loss.

A more comprehensive approach to expressive human movements, and in special dance movements, requires analysis and representation of the movement structure. Relevant information of dance gestures seem to be encoded in the two main dimensions of the dance gesture, namely, *space*, which is considered the medium for the deployment of movement, and *time*, which is considered the medium for segmentation and synchronization of movement. Dance gestures often deploy gestural forms through synchronization with musical time, which can be structured by hierarchies of musical meter. The question is how both space and musical time can be articulated together. Most approaches so far have focused on the spatial deployment of gesture using a temporal grid for the time dimension that impregnate the majority of dance representation. The literatures spans from early attempts, such as [19], to more contemporary score-like notations, such as [20] or computer based representations (see [21] for more detail). Yet, in robotics and computer animation most researchers generated dancing motions upon symbolic dance representations made up of primitive motions that are further synched with music. These motions constitute essential postures, characteristic of a given style [6], [13]. Besides, researchers had to overcome the kinematic and dynamic constraints resultant from mapping captured dancing motions onto different humanoid morphologies [12], [6], [13].

Working with synthesis of Japanese folk dances, [5] and [6] segmented captured dancing sequences, according to the minimum velocities of the end-effectors' (hands and feet) trajectories. The resulting key-poses were clustered and interpolated for generating variations of the original dance. Similarly, [4] extracted motion key-poses in terms of motion rhythm and intensity, calculated from local minimums of Laban's "weight effort" [22] (stop motions), which were respectively correlated with musical rhythm and intensity features for modeling musical synchronization. [1] and [2] generated rhythmic motion patterns, such as dancing and locomotion, by clustering and interpolating unlabeled mocap segments in terms of motion beats, corresponding to moments of rapid change in the motion signal, given by zero-crossings of the second derivative of all joints orientation. [7] calculated points for the feet and found key-poses by correlating them with indicators of extreme positions of arm swings given by Kinematic Centroid Segmentation (KCS). After retrieving motion features and the corresponding musical cues (such as beats, pitch, intensity and chord progression), mostly matched music and dance streams by relying on signal alignment and optimization techniques such as time-warping [1], [3], dynamic programming [7], and genetic algorithms [2].

On other hand, by using a cross modal algorithm for periodicity analysis, [23], [24] developed a methodology for the representation and analysis of periodic gestures in popular dances.

Although these methodologies build on the concept of time and space, the representation of time seem to be always fixed in a linear concept of time and a strictly deterministic concept of gesture. In other words, time is always represented linearly and sequentially and gestures tend to be represented as poses or patterns that tend to a precise point in space. In order to cope with these questions [25] proposed a method that projects classes of musical features in the space of dance gestures. The TGA method (Topological gestures analysis) cluster temporal information of the same class in a region or "topology" in the space, which gives rise to representations that are not linear in time but encode the variability of the original system. A better description of this method will be presented in Section III-B.

III. METHODOLOGY

In this section we describe the methodology for analysis and synthesis of beat-synchronous dancing styles based on captured dancing sequences. Our methodology includes four stages: (1) data acquisition, (2) representation of the dance style, (3) synthesis of dancing sequences, (4) visualization of dancing sequences, and (5) evaluation of the synthesized dances. The architecture of the process is displayed in Fig. 1.

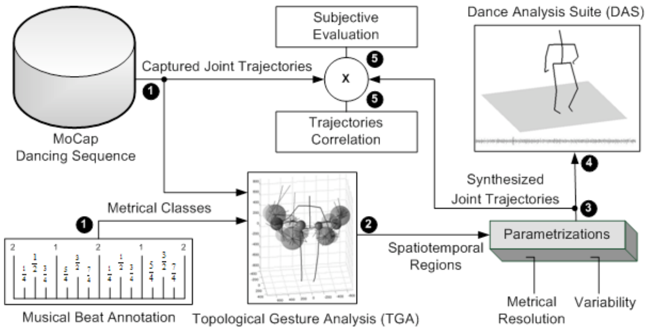

Fig. 1. Methodology workflow for dancing representation and synthesis.

A. Data acquisition

1) Motion Capture data: The dance recordings were realized in Brazil with a motion capture system (Optitrack / Natural Point) that consisted of 8 cameras positioned around a the dancer. The dance movements were recorded at a frame rate of 60 Hz, interpolated to 100Hz in the editing phase. The motion recordings were synchronized with audio in the editing phase. The dance sequences were also normalized at each frame, in relation to the centroid of the body. This process subtracts the effect of the movement of the whole body on the trajectories of the limbs. The sequences were imported into Matlab by using the Mocap toolbox [26]. The calculation of body basic joint positions, the filtering of raw vectors, the normalization and part of the visualization functions were also based on the Mocap toolbox.

2) Annotation data: The manual annotation of metrical points of the audio sequences (see Fig. 2) were realized by specialists, using Sonic Vizualizer[27]. From the beat annotation we derived both macro level (2 beat cycles, derived by mathematical multiplication) and micro levels of the musical meter (half-beat and quarter-beat levels, derived by mathematical subdivision). These levels encompass the resolution of the metrical parameters that will be used in the syntheses. See a schematic description of these levels in the time domain in Fig. 2 (”Metric levels”).

3) Procedures: The dances were performed by a professional female dancer, specialized in Afro-Brazilian dances. We asked the dancer to perform simple dance gestures in *Samba-no-pé* style, which is the most recognizable and popular sub-style of the Afro-Brazilian samba dances. After a few trial runs without any limitation, the dancers were instructed to dance the standard steps of the styles, without exhibiting improvisations, turns or embellishments.

B. Representation of the Dance Style

The Topological Gesture Analysis (TGA) [8] is a method that maps the use of space of musical gestures. It relies in a simple projection of musical cues onto spatial trajectories, which generates a visual representation of points in space. If the gesture in space is organized according to the music cues it is likely that the projection of points in space generate clusters, or point clouds. Point clouds can be interpreted as topologies or spatial regions equipped with musical qualities (see [28]), which informs us about the relationships between gesture, space and music. In the case of repetitive dances, as exemplified in [8], the method describes the space occupied by the dancer at the time they were synchronized with classes of musical meter (1 beat, half-beat, 2 beats, etc.). Each of these regions can be further parametrized by assuming certain characteristics for each point cloud. Fig. 2 describes one of these parameterizations which assumes an homogeneous spherical distribution of the point clouds around the average. The implementation of the method in this paper involve the definition of musical cues (manual annotation), projection of these cues onto spatial trajectories and discrimination of the point cloud regions by means of Linear Discriminant Analysis (see [8], for more information). The description of the TGA parameters convey one mean value (3 dimensions) for each body part (20) and the radius of the spherical distribution.

The process of TGA projection results in point clouds distributed in regions in space. These regions can be treated as topological spaces equipped with musical qualities. There are many ways to represent and treat these points. First we discriminated these regions by using a linear discriminant analysis over the classes of metrical cues. From the discriminated points, we opted to assumed spherical distributions whose radius is defined by the mean of the euclidean distances of all points to the centroid of the distribution. This very compact description of the distributions (composed of a centroid (x,y,z) and a radius) offer an simple and effective representation to compare the parameters of the syntheses. Fig. 3 illustrates the final spherical distributions for the hands of the samba dancer.

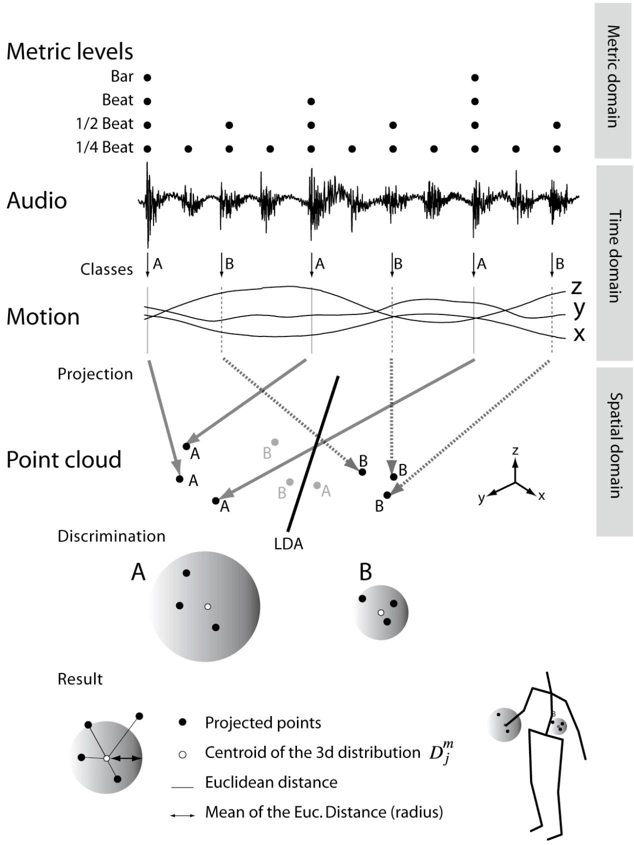

Fig. 2. Process of projection of musical cues (metrical classes) onto the dance trajectories. First, the metrical structure of the music synchronized with the MoCap recording is annotated. Then these cues are projected onto the movement vectors (in the example, right hand movements). The metrical points are projected as different classes (e.g.: 1st beat, 2nd beat, etc.), here described as A and B. Finally, the point clouds are discriminated using LDA analysis. This results in distributions in space, which can be represented by topologies or regions in space. In this study we simplified these distributions as spheres containing

C. Synthesis of dancing sequences

Based on the former representation, the actual dancing motion is synthesized by generating and propagating constrained stochastic variations of the dancing pattern described by the TGA spatial distributions for every joint of our 20-joints body model (see Fig. 3b)). As illustrated in Fig. 3c), the dancing pattern described by each joint is temporally represented by a closed-loop metrical cycle segmented in discrete categories at a given resolution. The chosen metrical resolution (see Fig. 2) represents the discrete classes of the musical meter (musical cues) considered by the TGA representation, which additionally constrain key-pose regions in the movement (space) occurring at specific key-frames in the music (time). Such key-poses constitute pseudo-unlimited

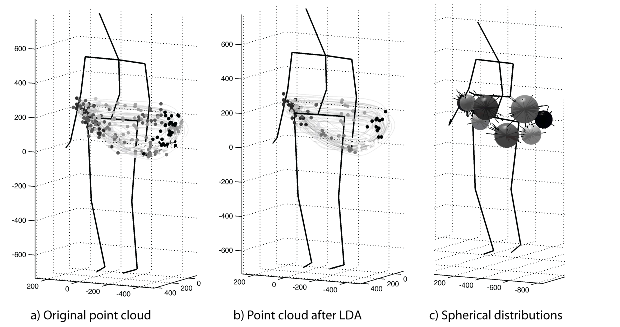

Fig. 3. a) Point cloud representation for quarter-beat classes within a two-beat metrical cycle, left hand b) Point cloud after LDA analysis. Note that classes of points are visually and linearly discriminated from each other. c) Representation of point clouds as homogeneous spherical distributions around the mean of the left hand gestural trajectories.

variations of the fundamental postures characteristic of the given dancing style, which are built upon full-body joint-positions. These joint-coordinates are stochastically generated within every metrical class distribution while satisfying the kinematic constraints, for keeping the body morphology and the represented spatial variability.

1) Inducing variability in the key-poses: The motion variability was generated by a stochastic process which takes into consideration both TGA distributions and the kinematic constraints of the body in space. The described process was repeated for every considered metrical classes.

Initially, the used body model was split into 5 kinematic chains, derived from 2 anchor joints – see Fig. 4. Both anchor joint coordinates were assigned as the mean values of their respective TGA spherical distributions.

For the given metrical class m, each joint chain was processed independently by stochastically calculating every of its joint coordinates. Starting from the anchor joint p_0^m until the chain extremity, each joint coordinates were calculated based on the calculated former joint position p_{j-1}^m, the length of the body segment delimited by both joints $l_{j-1,j}$, and the TGA spherical distribution of the considered joint D_j^m. As depicted in Fig. 4, the current joint position p_j^m is therefore generated as a random point on the surface of the spherical cap C_j^m resultant from intersecting D_j^m with a sphere S_{j-1}^m centered in p_{j-1}^m and radius equal to the segment length $l_{j-1,j}$:

$$\begin{cases} p_j^m = rand_p : p \in C_j^m \\ C_j^m = D_j^m \cap S_{j-1}^m \end{cases} \quad , p, C_j^m, D_j^m, S_{j-1}^m \in \mathbb{R}^3. \quad (1)$$

Within a considered kinematic chain and metrical class, the calculation of a joint coordinates depends on the previous joints positions (all stochastically determined). Therefore, the process was iteratively computed until successfully managing to calculate all joints while satisfying the propagated kinematic constraints.

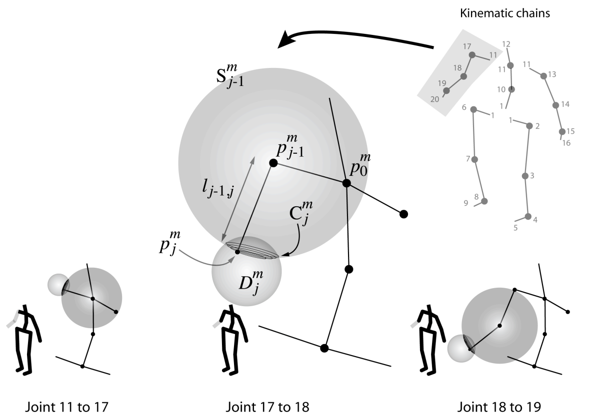

Fig. 4. Example of solution of the kinematic constrains for one section of the kinematic chain (joints of the right arm, metrical class m). Starting from anchor p_0^m at joint 11 to chain extremity at joint 20, the variability is computed by stochastically calculating all joint positions p_j^m inside the intersection C_j^m between the considered joint's TGA distribution at such class D_j^m and a sphere S_{j-1}^m centered on the last joint position p_{j-1}^m with radius equal to the segment length $l_{j-1,j}$.

2) Motion Interpolation between key-poses: The order of transition between the synthesized key-poses was intrinsically defined by the TGA topology and its cross-modal incorporation, which implicitly assured musical synchronization.

The dancing motion trajectories were finally synthesized by orderly interpolating key-poses constituted by full sets of joint key-positions at each metrical class, along all metrical cycles of the dancing sequence. The interpolation between postures, within all metrical classes, was generated by interpolating each joint independently. As such, all joint coordinates p_j were interpolated along all k intervals $[t_m, t_{m+1}]$ between consecutive pair of key-frames t (interpolation knots) pointed by such classes, by means of a piecewise cubic interpolant $I(j_d)$ over each joint coordinate dimension j_d (x, y, z) and class m, given by:

$$I(j_d) = [I_0, I_1, \cdots, I_{k-1}] : [(t_0, t_1), \cdots, (t_{k-1}, t_k)] \to \mathbb{R} \tag{2}$$

where

$$\begin{cases} \quad I_m(j_d) = c_0 + c_1(j_d - p_{j_d}^m) + c_2(j_d - p_{j_d}^m)^2 + \\ \quad +c_3(j_d - p_{j_d}^d)^3 : [t_m, t_{m+1}], m = 0, \cdots, k-1 \\ I_m(j_d) = I_{m-1}(j_d) \\ I'_m(j_d) = I'_{m-1}(j_d) \\ I''_m(j_d) = I''_{m-1}(j_d) \\ I''_0(j_d) = I''_{k-1}(j_d) = 0 \end{cases} \tag{3}$$

D. Visualization of dance sequences

For visualizing and animating our body model with the captured and synthesized joint trajectories, in synchrony with the considered musical input, we developed an interface based on the DAS (Dance Analysis Suite) software [29] – see Fig. 5.

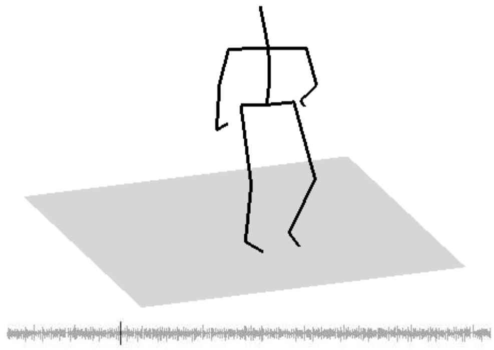

Fig. 5. DAS visualization of synthesized samba dancing.

E. Temporal and Spatial Hypotheses

By considering different hypotheses of interaction between movement and music, we assessed the flexibility of our dancing representation towards feasibly and naturally synthesizing the gestures of the original dancing style. For such we constrained our representation model with specific parameters in order to evaluate the following temporal (metrical) and spatial hypotheses:

1) **Metrical Resolution:** Evaluate the influence of selecting metrical classes at different tempo resolutions. For such we generated dancing sequences, as a series of periodic metrical cycles, by selecting, and interpolating, key-poses at time-points given by distinct hierarchical structures of the musical meter (metrical levels) – at bar, beat, half-beat, and at all quarter-beat sub-divisions – see Fig. 2.
2) **Variability:** Assess the impact of introducing spatial variability within sequent metrical cycles. For such we compared synthesized dancing with repetitions of the same rhythmic pattern, by assuming the centroids of the TGA distributions for each joint and metrical class, against other with controlled random variations of it, calculated as described above.

Such hypotheses were tested by constraining our representation and motion synthesis models with specific parameters concerning the metrical resolution and consideration of spatial variability. All tests were numerically and subjectively evaluated as described below.

IV. Evaluation and Results

This section describes our method for evaluating the reliability of our model towards synthesizing dancing sequences representative of the captured dancing style. The tests were performed over samba dance, which was recorded and preprocessed as described in section III-A.

A. Experimental Tests

In order to evaluate and compare the proposed hypotheses, the evaluation was performed over a set of 6 tests considering $30s$ dancing sequences synthesized with different parameterizations. In each of the 6 tests an excerpt, with also $30s$, of the original captured dancing was compared with each of the following dancing sequences:

- **original:** another excerpt of the captured dancing sequence, for delimiting the evaluation – theoretical best;
- **randgaps+2:** equal to "variability+2" but with random gaps between half of the considered metrical classes, for delimiting the evaluation – theoretical worst;
- **variability+1:** synthesized dancing sequence with variability and lowest metrical resolution (beat);
- **variability+2:** synthesized dancing sequence with variability and medium metrical resolution (half-beat);
- **variability+4:** synthesized dancing sequence with variability and highest metrical resolution (quarter-beat);
- **centroids+4:** synthesized dancing sequence without variability and highest metrical resolution (quarter-beat);

B. Level of Similarity and Compression

In order to evaluate the level of similarity Sim_i between the tested $test_i$ and the captured dancing sequences *capt*, we computed the correlation at zero lag between both joint trajectories, normalized by the autocorrelation of the captured sequence, in relation to the dimensionality and level of compression of the respective representation models:

$$\begin{cases} Sim_i = \frac{\sum_n \sum_j (capt[n,j] * test_i[n,j])}{\sum_n \sum_j (capt[n,j]^2)} * 100(\%) \\ n = [1, nFrames], j = [1, nJoints * 3], i = [1, 6] \end{cases} \quad (4)$$

where $nFrames$ is the number of frames (3000 for $30s$ at $100fps$) of the dancing sequences and $nJoints$ is the number of joints (20) of the considered body model.

The dimensionality was measured by the spatiotemporal dimension of the used representation model beyond the synthesized trajectories, in terms of all body joints (20) – *Body*, the 3-dimensional coordinates of the mean (centroid) of the TGA distributions and their radius (when considering variability) – *Space*, and the number of considered metrical classes – *Time*. The level of compression of every model was measured by comparing its dimensionality with the dimension of the captured dancing, dependent on the size (in frames) of the synthesized sequence. All results are presented in Table I. Fig. 6 presents a comparison of the captured joint trajectories with the ones synthesized by "variability+4" and"centroids+4" parametrization, for the right hand joint (19).

Test	Sim (%)	Dim (BodyxSpacexTime)	Compression
original	86.9	20x3xnFrames = 60xnFrames	0
randgaps+2	43.8	20x(3+1)x2 = 160	0.38xnFrames
variability+1	27.3	20x(3+1)x2 = 160	0.38xnFrames
variability+2	79.6	20x(3+1)x4 = 320	0.19xnFrames
variability+4	80.6	20x(3+1)x8 = 640	0.09xnFrames
centroids+4	81.2	20x3x8 = 480	0.13xnFrames

TABLE I

Correlation of the joint trajectories between the tested and captured dancing sequences, in relation to the spatiotemporal dimensionality and level of compression of the respective representation models.

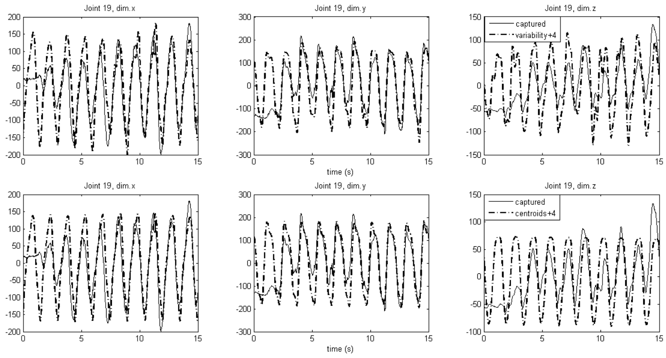

Fig. 6. Captured vs synthesized trajectories for right hand joint (19), for different synthesis parameterizations: a)"variability+4"; b) "centroids+4".

C. Level of Dancing Style Representation

In order to evaluate the level of samba style representation of the synthesized dances and assess every proposed hypotheses, we run a subjective evaluation over each test. For such we inquired 15 subjects, 7 of them Brazilian and acculturated with samba dancing. We started by showing each of them instructive videos about samba and previously submitted them to two runs of the 6 tests. For each test we showed them, on DAS, an excerpt of the captured dancing and one of the dancing sequences described in section IV-A, randomly ordered among the subjects to raise the level of confidence. After each test we asked them to point which of the two sequences they considered to be the original and to grade, from 1 to 5, the level of dance style representation of the considered synthesized dance over the original. For the first question 98.33% of the answers correctly classified the original sequence over the synthesized one. From the 3 misses, 2 erroneously chose the "centroids+4" and 1 the "variability+4". A box plot with the statistical results over the level of style representation evaluation is presented in Fig. 7.

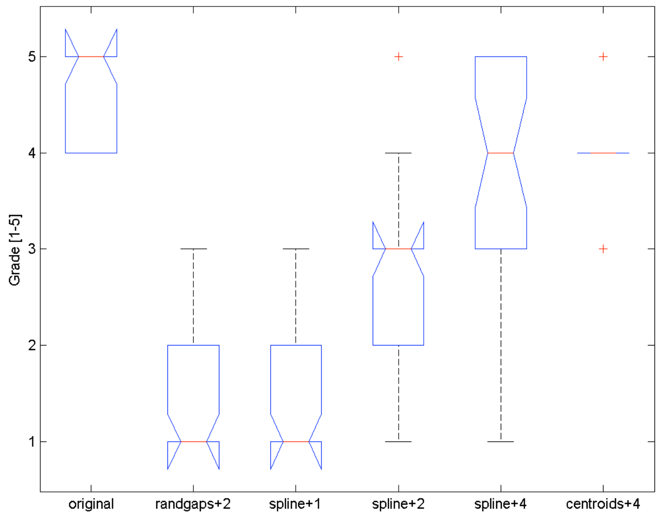

Fig. 7. Subjective statistical results over level of style representation of samba dancing.

D. Discussion

1) Similarity: By considering both numerical and subjective evaluations, there is an overall agreement on that "centroids+4" and "variability+4" feasibly reproduced and synthesized the captured dancing in terms of similarity and style representation, outperforming all other parameterizations. This result is enforced by their competitive results with the ones achieved when comparing different excerpts of the same captured dancing sequence ("original"). The difference is given by 6.3% to 5.7% in terms of trajectories similarity and 1 point in terms of dancing style representation given by the mean of the subjects' responses.

The results seem to suggest that the original dancing style may be feasibly reproduced by the TGA model at quarter-beat resolutions which conveys a compact representation of the original dancing spatiotemporal structure. Table I indicates that the TGA model offers a virtually unlimited reduction of the original dancing, by allowing the synthesis of endlessly variable dancing sequences, with a compression in the order of $\frac{1}{11}$ multiplied by the size of the generated sequence.

2) Resolution: When comparing the tests' results in terms of their used metrical resolution, we observed that the chosen metrical level pays a fundamental role on completely describing and representing the original dance, which may additionally depend on the considered style. For samba dancing, Table I suggests that the use of all half-beat subdivisions ("variability+2"), assures the reproduction of the original trajectories with a similarity of 79.6%, only surpassed by 1% by doubling that resolution to quarter-beat ("variability+4"), with the trade-off of halving the level of compression. Yet the subjective results in Fig. 7 reveal a statistical outperformance of using a quarter-beat resolution for reliably reproducing the original samba dancing style. This seem to suggests that subjective reasoning may play an important factor while evaluating similarity between dance patterns. These factors could be related with the focus on specific body parts in determining the style or the influence of the non-ecological elements of the set up of the experiment (use of stick figure, backgrounds and computer simulations). On its hand, the strict consideration of full-beat points, at "variability+1", revealed the worst performance, even providing a worse trajectory description than "randgaps+2" in terms of similarity to the original motion. It seem to suggest a non-linear relation between resolution of the process and the realism of sequence: when the synthesis drops to a certain threshold of numerical resolution (in the whole process) it may dramatically decrease the perceived similarity as a dance sequence.

3) Variability: The effect of variability in the system was verified by comparing the use of stochastic and fixed points for the definition of key-poses, respectively defined by the "variability+4" and "centroids+4" tests. The results indicate that the variability imposed by the process is not sufficient to impose an effect in terms of similarity with the original. The subjective evaluation of the "centroids+4" sequence (which displays very repeating patters, as depicted in Fig. 6b)) was consistently less divergent than the "variability+4", (which was sequenced using stochastic processes – see Fig. 6a)), suggesting a negative effect of the variability on reproducing the original dance style. An explanation for such result may rely on the repetitive nature of the captured dance, which may imply that periodicity would be considered by the subjects as a key factor for their assessment. In addition, assuming an homogeneous spherical distribution for the stochastic process may impose random combinations of movements that are perceived as non-realistic. More studies in this field are necessary in order to uncover this relationship between variability and dance expressiveness.

V. Conclusions and Future Work

In this study we proposed and evaluated a method for synthesizing dance movements from a very compressed representation of dance gestures. The process starts from information of the original dance recorded with a motion capture system combined with musical information packed in the TGA representation. This representation was re-synthesized into dance sequences, and tested against the captured dance. The results shown that quarter-beat representations offer a proper level of similarity while offering a great compression of the original signal. Smaller resolutions offer a decreasing reproduction of the original dance, but keep an increasing compression ratio. There were no significant positive effects on inducting variability, as suggested by both evaluations.

The overall results seem to validate the TGA representation as a reversible form that may be applied for both analysis and synthesis of periodic dancing styles in animated characters. In addition, the topological structure of the concept offers new perspectives to further manipulation and syntheses of these topologies. Such representation may offer means for satisfying the kinematic constraints imposed by using different robotic humanoid morphologies, by allowing a flexible transformation of the dancing gestures' geometry while keeping the topology of the movement structure and shape, fundamental to describe a given dancing style. More studies are needed in order to verify the role of the variability and importance of body parts in the perception of expression in popular dance styles.

VI. Acknowledgments

This work was funded by a PhD scholarship endorsed by FCT, with ref. SFRH/BD/43704/2008, and by FWO. We would like to acknowledge all subjects who participated on the reported subjective evaluation.

References

[1] T.-H. Kim, S. I. Park, and S. Y. Shin, "Rhythmic-motion synthesis based on motion-beat analysis," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 392–401, 2003.

[2] G. Alankus, A. A. Bayazit, and O. B. Bayazit, "Automated motion synthesis for dancing characters," *Journal of Visualization and Computer Animation*, vol. 16, no. 3-4, pp. 259–271, 2005.

[3] G. Kim, Y. Wang, and H. Seo, "Motion control of a dancing character with music," in *ACIS-ICIS*, pp. 930–936, 2007.

[4] T. Shiratori, A. Nakazawa, and K. Ikeuchi, "Dancing-to-music character animation," *Comput. Graph. Forum*, vol. 25, no. 3, pp. 449–458, 2006.

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

[6] S. Nakaoka, A. Nakazawa, K. Yokoi, H. Hirukawa, and K. Ikeuchi, "Generating whole body motions for a biped humanoid robot from captured human dances," in *ICRA*, pp. 3905–3910, 2003.

[7] H.-C. Lee and I.-K. Lee, "Automatic synchronization of background music and motion in computer animation," *Comput. Graph. Forum*, vol. 24, no. 3, pp. 353–362, 2005.

[8] L. Naveda and M. Leman, "The representation of spatiotemporal music gestures, using topological gesture analysis (tga)," *submitted to Music Perception*, 2010.

[9] A. Bruderlin and L. Williams, "Motion signal processing," in *SIGGRAPH*, pp. 97–104, 1995.

[10] M. Brand, M. Brand, A. Hertzmann, and A. Hertzmann, "Style machines," in *Proceedings of SIGGRAPH 2000*, pp. 183–192, 2000.

[11] L. Kovar, M. Gleicher, and F. H. Pighin, "Motion graphs," in *SIGGRAPH*, pp. 473–482, 2002.

[12] N. S. Pollard, J. K. Hodgins, M. J. Riley, and C. G. Atkeson, "Adapting human motion for the control of a humanoid robot," in *Proceedings of International Conference on Robotics and Automation*, pp. 1390–1397, 2002.

[13] T. Shiratori, S. Kudoh, S. Nakaoka, and K. Ikeuchi, "Temporal scaling of upper body motion for sound feedback system of a dancing humanoid robot," in *IROS*, pp. 3251–3257, 2007.

[14] M. Gleicher, "Retargeting motion to new characters," in *SIGGRAPH*, pp. 33–42, 1998.

[15] O. Arikan and D. A. Forsyth, "Interactive motion generation from examples," in *SIGGRAPH*, pp. 483–490, 2002.

[16] J. Lee, J. Chai, P. S. A. Reitsma, J. K. Hodgins, and N. S. Pollard, "Interactive control of avatars animated with human motion data," in *SIGGRAPH*, pp. 491–500, 2002.

[17] O. Arikan, "Compression of motion capture databases," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 890–897, 2006.

[18] Q. Gu, J. Peng, and Z. Deng, "Compression of human motion capture data using motion pattern indexing," *Comput. Graph. Forum*, vol. 28, no. 1, pp. 1–12, 2009.

[19] K. Tomlinson, *The Art of Dancing Explained by Reading and Figures, repr.* 1735.

[20] R. Laban and L. Ullmann, *Choreutics*. London: MacDonald and Evans, 1966.

[21] "Digital Representations of Human Movement," *ACM Computing Surveys (CSUR)*, vol. 11, no. 1, pp. 19–38, 1979.

[22] R. Laban and L. Ullmann, *Mastery of Movement*. Princeton Book Company Publishers, 1960.

[23] L. Naveda and M. Leman, "A Cross-modal Heuristic for Periodic Pattern Analysis of Samba Music and Dance," *Journal of New Music Research*, vol. 38, no. 3, pp. 255–283, 2009.

[24] M. Leman and L. Naveda, "Basic gestures as spatiotemporal reference frames for repetitive dance/music patterns in Samba and Charleston," *Music Perception (Accepted)*, 2010.

[25] L. Naveda and M. Leman, "The spatiotemporal representation of dance and music gestures using Topological Gesture Analysis (TGA)," *Music Perception (In Press)*, 2010.

[26] P. Toiviainen and B. Burger, "MoCap Toolbox Manual," 2008.

[27] C. Cannam, C. Landone, M. Sandler, and J. P. Bello, "The sonic visualiser: A visualisation platform for semantic descriptors from musical signals," 2006.

[28] G. Carlsson, "Topology and data," *Journal: Bull. Amer. Math. Soc*, vol. 46, pp. 255–308, 2009.

[29] L. Naveda, M. Leman, and D. Mota, "DAS - Dance Analysis Suite," 2010.