Towards a Live Dance Improvisation between an Avatar and a Human Dancer

Alexander Berman
alex@timebend.net

Valencia James
valenciajamesprojects@gmail.com

ABSTRACT

This paper presents an attempt to generate novel dance movement based on motion captured human dance. Captured movements are analyzed statistically using nonlinear principal component analysis in order to create a "map" of observed poses. A method for automatically exploring the map by generating random trajectories is then presented, constituting a kind of improvisation. The result is an animated avatar that exhibits creative and novel movements in the style of its teacher, and paves the way towards a fully improvised live performance between an avatar and a human dancer.

Categories and Subject Descriptors

J.5 [**Arts and humanities**]: Performing arts;
I.2.10 [**Artificial intelligence**]: Vision and Scene Understanding—*Motion*

1. INTRODUCTION

Improvisation is a universal phenomenon we all exercise in daily life. As in music, improvisation is also a specialized performance genre in dance as well as a tool for accessing creativity in its purest form during the choreographic process. For a dancer the main challenge of improvisation is the "getting stuck" and consequently becoming "trapped" into habitual movement sequences. These habits could constitute the unique style of the dancer, but the lack of understanding about the nature of those tendencies undermines the dancer's freedom of choice in a performance setting [5]. From this challenge sprung the idea of using artificial intelligence (AI) to not only give insight into the nature of a dancer's habits but also to use those tendencies to derive new movements otherwise unimaginable by the dancer and hence to inspire her own movement choices.

This paper presents the preliminary results of a research into the possibilities of generating novel dance movement based on motion captured human dance. The research is an initial step in the interdisciplinary project AI_am whose long-term goal is to imitate human dance and creativity by developing software that reproduces and extends the human dancer's style, ultimately enabling a fully improvised live performance where a human dancer and an avatar learn from and inspire each other.

MOCO'14, June 16-17 2014, Paris, France.
ACM 978-1-4503-2814-2/14/06 $15.00.
http://dx.doi.org/10.1145/2617995.2618026.

Our current approach utilizes a small database of captured movements that were selected and performed by a dancer in the team. The moves were derived from poses (distinct but dynamic bodily shapes) that relate to certain spatial points, and the movement that arises from moving back and forth between these poses. This approach was directly influenced by Rudolf Laban's Choreutics or Space Harmony method [13]. Ten directions were randomly chosen out of the basic 27 possible directions of the cube and dance poses were set to spatially reflect each direction. The movement derived by moving between each of these positions served as the training input for the AI.

The captured movements were analyzed statistically using Principal Component Analysis (PCA) in order to create a "map" of observed poses. A method for automatically trajecting territories of this map was then developed and evaluated. In essence, this method constitutes a kind of "improvisation" by which the AI explores both known and unknown regions of the space of possible poses. By manipulating parameters manually, we can move the AI between "cautious" and "brave" modes of improvisation. The preliminary results of the method indicate its potential usefulness as an inspiring virtual dance partner, and as a basis for a live dance duet between a human and an avatar.

2. RELATED WORK

Several attempts at modelling live improvisation have been carried out in the field of music. For example, [15] describes a system that generates novel music and adapts to the reaction of a human co-performer in real time. Much fewer examples of artificial improvisation have been reported within dance. One exception is [1] who describes a robot that moves according to distinct categories in Laban Movement Analysis. Its improvisation consists in switching between predefined modes such as copy, follow-copy and oppose. Software has also been used as a tool for generating human-computer choreographies [7] and to offer choreographic material to dancers as part of their creative process [2]. In contrast, the current study focuses on learning and extending human movement qualities and does not constrain the AI to a predefined movement repertoire.

In terms of the specific methods employed in our approach, dimensionality reduction of motion capture data has

Figure 1: Example of a motion sequence from the training set.

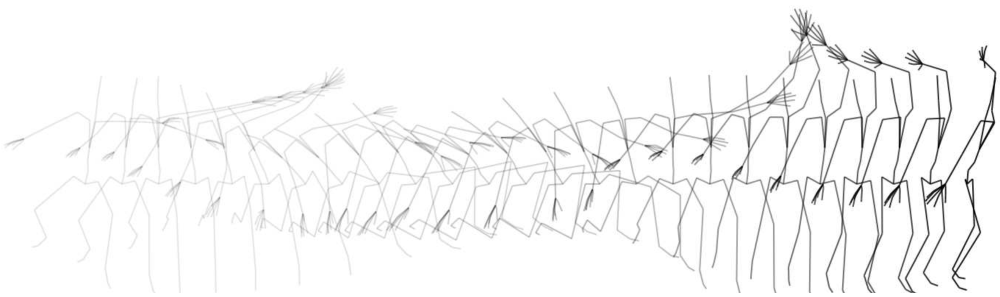

Figure 2: Example of a synthesized motion sequence, extending beyond movements in the training set.

previously been used for generating movement sequences. For example, [6] used PCA and quaternions in order to enhance expressiveness of motions in avatars. Other related approaches include [10] who generated novel trajectories in PCA-reduced space using Hidden Markov Models, with the purpose of synthesizing realistic movements between arbitrary keyframes in the captured database. With similar goals, Gaussian process latent variable models have been used instead of PCA [14, 4]. In contrast to these examples, the current project focuses on dance improvisation and more specifically on trying to generate creative and previously unseen movements that may surprise even the teacher. The main contribution of the presented work lies in the automated exploration of a pose map, rather than in the specifics of the chosen motion processing.

3. MOVEMENT ANALYSIS

The dancer's movements were recorded and represented using a skeleton model with 52 joints (see figure 1). In each captured frame, each joint contains information about its angular orientation, represented as Euler angles, while the root joint (in our case the hip) also encodes 3d translation (position of the whole body in space). In other words, each input frame is represented by a vector of size $3+52\times3 = 159$.

The high dimensionality of the input space makes it unsuitable for direct exploration. It also carries a large amount of redundancy, since orientations represented orthogonally are in fact in most cases highly correlated. For instance, the orientation between the left foot and shinbone correlates strongly with the one between the shinbone and the thigh. Additionally, physical body constraints and the nature of real human movements also limit the actual orientational range of individual joints. For example, only a few joints are physically able to rotate 360 degrees along any axis.

High dimensionality and large redundancy motivate the use of dimensionality reduction, whereby a two-way mapping between a high-dimensional input space and a low-dimensional map space is achieved. In the current approach, nonlinear Kernel Principal Component Analysis (KPCA) [8] was used for this purpose.

3.1 Preprocessing and dimensionality reduction

Since Euler angles suffer from singularities, they are not suitable for motion analysis and synthesis. Therefore, orientations in the motion capture data were transformed to unit quaternions [9] which overcome this disadvantage. However, reducing the dimensionality of orientation data is not straightforward, regardless of representation form. Most standard dimensionality reduction techniques assume Euclidean data, while 3d orientations are elements of the rotation group SO(3), which is not a vector space. Although kernel PCA is nonlinear, our implementation does not generalize to SO(3). Instead, the problem was worked around by assuming local linearity in the training data, which sufficed for our motion data.[1] A more general solution could employ exponential maps [3] or nonlinear manifold learning such as principal geodesics analysis [12].

For the purpose of normalization between translation and orientation, the translation vector was scaled to unit magnitude and then multiplied by a factor representing the weight of translational movement versus angular orientations. This translation weight could then be used as an experimental parameter.

Since quaternions have 4 components, each preprocessed frame of motion data is represented by a normalized vector of size $3 + 52 \times 4 = 211$. Using PCA, we found that the input space for the motion database could be reduced to a

[1] The so called double covering of quaternions was handled in the same way: it simply did not pose a problem with our data. Other authors such as [6] apply hemispherization to address double covering.

7-dimensional map space accounting for 85% of the variance of the input. Figure 3 shows the distribution in map space of all recorded poses in the database.

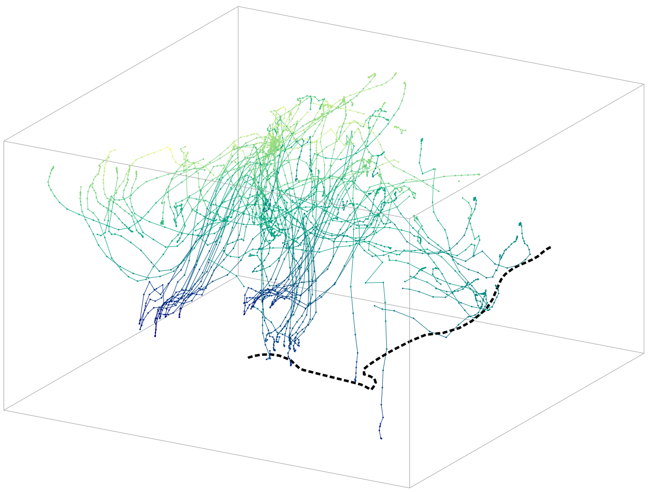

Figure 3: Movements in the training set visualised as trajectories across the map. The dashed path was generated using the improvisation algorithm. Note that while the figure illustrates a 3d pose map, a 7d map was used for the actual study.

4. MOVEMENT SYNTHESIS AND EXPLORATION

With preprocessing and dimensionality reduction in place, observed sequences of human movement can be analyzed as trajectories in map space. Reversibly, by applying inverse KPCA transform and inverse preprocessing, map trajectories can be interpreted as movement and generated graphically in the form of an avatar. However, special care needs to be taken when interpreting the quaternion values reconstructed from the map. Since PCA treats the quaternion coefficients independently, magnitude is unconstrained. Simple renormalization to unit magnitude after each inverse transform was found to yield satisfactory results.

4.1 Eliminating foot skating

A common artifact resulting from motion compression is "foot skating", caused by a reliance on statistical methods and a lack of physical and physiological modeling. This limitation was overcome using a highly simplified model of friction. In this model, the bottom-most joint is assumed to support the rest of the body and is not allowed to move, while the rest of the joints are centered around the supporting point, similar to the technique described in [11].

4.2 Improvisation: sequencing novel movements

A simple "improvisation" mechanism was developed for automatic exploration of the pose map. The method generates trajectories across the map as smoothly interpolated curves in the vicinity of observed areas, and has the following logic:

- Select a random destination q with a distance of at least E from the departure p, and with a distance of N to the nearest observed pose (N is a "novelty" parameter)
- Choose some intermediate points between p and q, where each intermediate point lies between a point on the straight line between p and q and its nearest observation in the map (closer to the nearest observation for lower N values)
- Smooth the resulting path using spline interpolation
- Create the next trajectory by treating the current destination as a departure and repeating the steps

When $N = 0$, generated trajectories are found in the immediate vicinity of observed areas. As N increases, so does the distance between generated trajectories and observations. The "extension" parameter E controls the length of generated trajectories. An example of a generated trajectory is shown in figure 3.

An additional "dynamics" parameter controls the speed with which trajectories are followed when performed, ranging from flat (constant speed) to smooth (constant plus sinusoidal).

4.3 Choice of mapping

Experiments were performed with linear PCA as well as nonlinear kernel PCA. The most interesting results were obtained with a polynomial kernel of degree 3. Compared to the linear version, the nonlinear mapping gave rise to much more "daring" and "creative" output, with maintained accuracy in the observed areas.

5. RESULTS

An evaluation of the prototype was performed by letting the dancer in the team watch videos of generated motion with different combinations of parameter values. The result of the evaluation is summarized below.

5.1 General observations

For low values of novelty and extension, the movements seem like tentative attempts. This cautious style is characterised by a restrained range of motion and a hesitant, disjunct flow. When extension is increased, the hesitancy gives place for a more fluid and expansive behaviour.

When novelty is increased, the avatar begins to excel in innovative movements. Some motions appear virtuosic and fluid as limbs start to extend further in space and spirals are more sustained as the body rotates around one fixed joint at times. The general flow is still controlled but more free.

When novelty is at its maximum, the avatar's movement quality evolves significantly and the motion seems at its bravest. Its limbs rotate and bend to extreme positions, with unnatural coordination and weight placement. Many times it demonstrates humanly impossible movements such as a 360-degree rotation on a leg extended away from the vertical axis, without impetus, or an unsupported head stand starting from an upright stance. However one can sometimes very briefly identify expressive pedestrian gestures such as a shrug of the shoulders with hands splayed.

While the training set focused entirely on Laban's Space Harmony, the avatar also exhibits other motion factors that constitute Laban's Effort theory such as flow and time. The resulting movement qualities are at times more motivating than its direction and shape.

5.2 Inspiration

In general, the avatar produces surprising and interesting improvisations. All parameter combinations produce varying qualities of movement that inspire the desire to experiment and embody them.

Generally, extreme parameter values cause less useful behaviour. Low values of novelty and extension yield abrupt and partial movement, lacking intuitive development or full embodiment of the movements. On the other hand, the cautious style is easier to imitate and its nuances are clearer. In contrast, maximum novelty sometimes causes shapes that are almost impossible to imitate.

6. EXAMPLES

An example of a generated sequence is presented in figure 2. Videos of the motion capture database used for training as well as examples of automated improvisation can be found at `http://timebend.net/AI_am/MOCO14/`

7. DISCUSSION AND FUTURE WORK

We have described a method for generating novel and previously unseen dance movements based on a statistical analysis of motion captured human dance and automatic exploration of the statistical model. Some early results of these experiments have also been presented. An evaluation indicates that the avatar exhibits novel, surprising and inspiring movements. As initially suspected, the superior analytical capabilities of the AI facilitate the discovery of movement otherwise not possible with traditional methods. Exploring unobserved regions in the pose map produces more inconceivable results than relying exclusively on a captured dataset. Such novelty is highly stimulating during an improvisation and encourages its further development and innovation by the human dancer.

Parameters governing the style of generated motions (novelty, extension and dynamics) have so far been set manually. As a future option, their values could be guided by feedback from the human dancer. Alternatively, the extension and dynamics parameters could be replaced with more sophisticated temporal processing, attending to the actual dynamics in the training set.

The presented approach is fundamentally statistical in nature, lacking higher-order levels of analysis such as segmentation of observed sequences into "moves", or any semantic processing. It is interesting to note that relevant and useful results can be achieved even in the absence of segmentation and classification. Nevertheless, the strive for a meaningful interaction with a human dancer could drive future development towards higher-order levels of processing.

While the results presented here are preliminary, we believe that they form a basis for a fully improvised live performance between a human dancer and the avatar. As a next step, we plan to implement the presented method as a component in a framework for real-time interaction between human and machine. This will enable explorations of various means of interaction such as live feedback from the human dancer, serving as reinforcement of generated movements and guiding the avatar's improvisation towards trajectories that stimulate response in the human dancer. Segmentation and classification could also enable "games" in which the participants exchange known and novel moves with each other.

8. REFERENCES

[1] J. Bryden, D. Hogg, S. Popat, and M. Wallis. Building artificial personalities: expressive communication channels based on an interlingua for a human-robot dance. In S. Bullock, J. Noble, R. Watson, and M. A. Bedau, editors, *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*, pages 80–87. MIT Press, Cambridge, MA, 2008.

[2] S. deLahunta. The choreographic language agent. In C. Stock, editor, *Conference Proceedings of the 2008 World Dance Alliance Global Summit*, 2009.

[3] F. S. Grassia. Practical parameterization of rotations using the exponential map. *J. Graph. Tools*, 3(3):29–48, Mar. 1998.

[4] K. Grochow, S. L. Martin, A. Hertzmann, and Z. Popović. Style-based inverse kinematics. *ACM Trans. Graph.*, 23(3):522–531, Aug. 2004.

[5] I. Hagendoorn. Cognitive dance improvisation: How study of the motor system can inspire dance (and vice versa). *Leonardo*, 36(3):221–228, Jun 2003.

[6] M. P. Johnson. *Exploiting Quaternions to Support Expressive Interactive Character Motion*. PhD thesis, MIT Media Lab, 2003.

[7] F.-J. Lapointe and M. Époque. The dancing genome project: Generation of a human-computer choreography using a genetic algorithm. In *Proceedings of the 13th Annual ACM International Conference on Multimedia*, MULTIMEDIA '05, pages 555–558, New York, NY, USA, 2005. ACM.

[8] B. Schölkopf, A. Smola, and K.-R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.*, 10(5):1299–1319, July 1998.

[9] K. Shoemake. Animating rotation with quaternion curves. *SIGGRAPH Comput. Graph.*, 19(3):245–254, July 1985.

[10] L. M. Tanco and A. Hilton. Realistic synthesis of novel human movements from a database of motion capture examples. In *Proceedings of the Workshop on Human Motion (HUMO'00)*, HUMO '00, pages 137–, Washington, DC, USA, 2000. IEEE Computer Society.

[11] J. Tilmanne and T. Dutoit. Expressive gait synthesis using pca and gaussian modeling. In *Proceedings of the Third International Conference on Motion in Games*, MIG'10, pages 363–374, Berlin, Heidelberg, 2010. Springer-Verlag.

[12] M. Tournier, X. Wu, N. Courty, E. Arnaud, and L. Reveret. Motion Compression using Principal Geodesics Analysis. *Computer Graphics Forum*, 28(2):355–364, Apr. 2009.

[13] R. von Laban. *The Language of Movement: A guidebook to choreutics*. Plays, Inc, Boston, 1966.

[14] J. M. Wang, D. J. Fleet, and A. Hertzmann. Gaussian process dynamical models for human motion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(2):283–298, Feb. 2008.

[15] M. Young. Nn music: Improvising with a 'living' computer. In R. Kronland-Martinet, S. Ystad, and K. Jensen, editors, *CMMR*, volume 4969 of *Lecture Notes in Computer Science*, pages 337–350. Springer, 2007.