

A VECTOR-SPACE REPRESENTATION OF MOTION DATA FOR EXAMPLE-BASED MOTION SYNTHESIS *

Ik Soo Lim[†] and Daniel Thalmann

Computer Graphics Lab (LIG), Swiss Federal Institute of Technology (EPFL), CH-1015 Lausanne, Switzerland

Keywords: example-based motion synthesis, vector space representation, principal components

Abstract We approach the problem of example-based motion synthesis by transforming motion data into a vector space representation. This allows many techniques successful for stationary object synthesis applicable to that of motion. Especially, by separating generation of motion into a time-consuming preprocess and a fast process, it lets on-the-fly motion synthesis able to use a rich set of exemplar motions and handle motion attributes invariant for each individual, both of which are difficult to be addressed by previous approaches based on interpolation.

1. INTRODUCTION

Motion control of articulated figures such as humans has been a challenging task in computer animation [BPW93]. Once an acceptable motion segment has been created, either from key-framing, motion capture or physical simulations, reuse of it is important.

This article describes example-based motion synthesis with parametric manipulation of motion attributes. This is done by transforming motion data into a vector space representation based on a linear combination of prototypical motions in full correspondence/alignment. This representation allows those techniques successful for the synthesis of stationary objects such as 2D images and 3D shapes applicable to motion.

* This work is supported in part by PAVR under the EU Training and Mobility of Researchers program.

[†] Thanks to Norman Badler for helpful discussions.

By separating motion generation into a time-consuming preprocess and a fast process, it also fits well to real-time animation for applications such as game and virtual reality while taking advantage of a rich set of exemplar motions.

In the following sections, we first describe related work of editing and reuse of motion data. A vector-space representation of stationary objects is briefly reviewed. We present its extension to motion data. Taking advantage of the representation, manipulation of motion attributes is introduced. First results of our approach with hand-crafted motion data follow it. We conclude with discussions comparing ours with some of previous works.

2. RELATED WORK

Much of the recent research in computer animation has been directed towards editing and reuse of existing motion data. Stylistic variations are learned from a training set of very long unsegmented motion-capture sequences [BH00]. An interactive multi-resolution motion editing is proposed for fast and fine-scale control of the motion [LS99]. Whereas most of other methods may produce results violating the laws of mechanics [WP95], an editing method maintaining physical validity is suggested [PW99]. Motion editing is also done in frequency-domains [BW95] [UAT95]. Interpolation of existing motion data is employed for the on-the-fly synthesis [RCB98][WH97].

3. LINEAR COMBINATIONS OF STATIONARY OBJECTS

The basic idea of the vector space representation by a linear combination of stationary objects can be described as follows, first proposed by Ulman and Basri [UB91] and followed up for 2D images [BP96][VP97] and 3D geometries [BV99][She00]. It is based on a data set of stationary objects in a same class. All of these exemplar objects are assumed in full correspondence, which can be done using techniques based on optic flow algorithms [BP96][BV99][VP97]. Given a set of m exemplar objects in full correspondence, characterized by feature vectors $\mathbf{X}_1, \dots, \mathbf{X}_m$ such as pixels for 2D images or vertices for 3D geometries, a linear combination of them produces a new object in the same class:

$$\mathbf{X} = \sum_{i=1}^m w_i \mathbf{X}_i$$

This linear combination is meaningful or valid since all the examples are in full correspondence: unless the 2D images align pixel-to-pixel, for ex-

ample, a simple linear combination of them would look like a transparent superposition of different images rather than a new image in the same class. The vectors \mathbf{X}_i comprise the basis of a linear vector space. The method parametrizes a continuous class of objects and the weight vector $\vec{w} = (w_1, \dots, w_m)$ characterizes each object of this class in a compact way. Object transformations, like view point changes in 2D images of an object [UB91] or attribute manipulation in 3D faces [BV99], can be expressed in terms of changes in the weight vector.

4. LINEAR COMBINATIONS OF MOTIONS

Motion can be described by a set of motion curves each giving the value of one of the model's parameters as a function of time, e.g. joint angles over time for articulated objects such as human figures.

For instructional purposes, we start with motion data of just a single curve $\theta(t)$. We represent the motion data with a shape-vector $\mathbf{S} = (\theta_1, \dots, \theta_n) \in \mathbb{R}^n$ and a timing-vector $\mathbf{T} = (t_1, \dots, t_n) \in \mathbb{R}^n$ where θ_j giving the value of θ at time t_j , i.e. $\theta_j = \theta(t_j)$. Given m exemplar motions, each represented by its shape-vector and timing-vector, an arbitrary motion among them is chosen as a reference motion, \mathbf{S}_{ref} and \mathbf{T}_{ref} . Correspondence between all other motions and this reference is computed, which can be done automatically by a time-warp algorithm [BW95]: each shape-vector is warped into the reference \mathbf{S}_{ref} and its corresponding timing-vector is accordingly recomputed. New shapes \mathbf{S}_{mod} and new timings \mathbf{T}_{mod} can be expressed as a linear combination of the shapes and timings of the m exemplar motions in full correspondence:

$$\mathbf{S}_{mod} = \sum_{i=1}^m a_i \mathbf{S}_i, \quad \mathbf{T}_{mod} = \sum_{i=1}^m b_i \mathbf{T}_i.$$

For motion data consisting of a set of motion curves $\theta^{(1)}(t), \dots, \theta^{(k)}(t)$ for articulated bodies such as human figures, a shape-vector and a timing-vector can be straightforwardly extended by concatenating those of the single curves such as $\mathbf{S} = (\theta_1^{(1)}, \dots, \theta_n^{(1)}, \dots, \theta_1^{(k)}, \dots, \theta_n^{(k)}) \in \mathbb{R}^{nk}$ and $\mathbf{T} = (t_1^{(1)}, \dots, t_n^{(1)}, \dots, t_1^{(k)}, \dots, t_n^{(k)}) \in \mathbb{R}^{nk}$. This representation parametrizes a continuous class of motions by the weight vectors $\vec{a} = (a_1, \dots, a_m)$ and $\vec{b} = (b_1, \dots, b_m)$ in a compact way and motion manipulation / transformation can be expressed in terms of changes in the weight vectors, analogous to that of stationary objects as above.

Principal Component Analysis (PCA) [Jac91] can be employed for a basis transformation to an orthogonal coordinate system with the eigenvectors of the covariance matrices computed over the shape and timing

differences with their averages $\bar{\mathbf{S}}$ and $\bar{\mathbf{T}}$, respectively. This will lead to further data compression and level-of-details as explored for stationary objects [AM00][BH00][PW89].

5. MOTION ATTRIBUTES

For the functional relationship between the weight vectors and the continuous parameters of interest, we adopt the technique used for geometric models of 3D faces [BV99]: radial basis functions network [Bis95] [GJP95] are often used for the similar purposes [BP96][She00]. Based on a set of motions $(\mathbf{S}_i, \mathbf{T}_i)$ with labels μ_i describing the markedness of the attribute, we compute the weighted sums

$$\Delta\mathbf{S} = \sum_{i=1}^m \mu_i (\mathbf{S}_i - \bar{\mathbf{S}}), \quad \Delta\mathbf{T} = \sum_{i=1}^m \mu_i (\mathbf{T}_i - \bar{\mathbf{T}}).$$

Multiples of $(\Delta\mathbf{S}, \Delta\mathbf{T})$ can now be added to or subtracted from any individual motion generated by the motion model, which will manipulate a specific attribute while keeping all other attributes as constant as possible. Motion caricature is also possible, analogous to that of a face [Bre85][BV99]. Individual motions are caricatured by increasing their distance from the average motion.

6. RESULTS

We performed experiments with hand-crafted motion data of ten examples similar but distinct (Figure 1). The key-framed data were re-sampled densely, 64 samplings over 2 seconds or so for each of 75 degree-of-freedom as if simulated/captured data: a vector of 4800 dimension. These examples went through a basis transformation by PCA (Figure 2). To illustrate the manipulation of motion attributes, two attributes vectors were extracted and applied: one for the height of a human figure's right hand, the other for energy (Figure 3). Motion caricatures were also generated with different degree of distinctiveness (Figure 4).

7. DISCUSSIONS AND CONCLUSIONS

A vector space representation of motion data based on a linear combination of prototypes is presented for example-based motion synthesis analogous to that of stationary objects. Due to the representation, many techniques successful for synthesis of stationary objects become applicable to motion synthesis and are done in a compact way with the low dimensional weight vector of the linear combination.

Especially noticeable, among them, is the extraction of motion attribute vectors which, when added to or subtracted from a motion, will manipulate a specific attribute of motion while keeping all other attributes as constant as possible: adopted from a technique to handle attributes such as weight, age, and gender of faces [BV99], the motion attribute vectors can deal with those invariant for each individual. This separates generation of motion into a preprocess that may require complex and time-consuming computations, and a process that is fast and tolerant of various types of complexity: one for the extraction of the motion attribute vectors, the other for the addition/subtraction of them. Even on-the-fly motion synthesis can, hence, take advantage of a rich set of example motions while only a small set of examples are interpolated for the motion synthesis in previous works [RCB98][WH97], hardly handling those attributes invariant for each individual.

The basis transformation to one formed by the eigenvectors of the covariance matrix is often considered for dimension reduction or level-of-details [AM00][BV99][PW89]. Characterizing the variation between the examples, the eigenvector representation also serves well for efficient generation of motion variations which should be especially useful for applications such as crowd motions [MGT99]: altering the weights is more efficient and of better quality than directly perturbing the motion curves in a data-independent way [Per95].

The linear combination of stationary objects is also suggested for computer animation [AM00][BV99][EP00][PW89]. This linear combination in space may be suitable for a single stream of motion data. For a set of motions, however, a linear combination in space-time as done here fits better in terms of computation and storage.

Lately, in computer vision, a similar notion of the vector space representation is proposed for synthesis and analysis of motion patterns [GP00]. Handling video sequences directly without any 3D model, its motion synthesis is limited to, for example, that of similar view angles and the motion correspondence becomes complicated, being a spatio-temporal correspondence problem. In our approach, however, 3D model-based representations such as joint angles are used so that it is suitable for 3D animation and its motion correspondence problem is reduced to that of temporal correspondence for which a time-warp algorithm can be used [BW95].

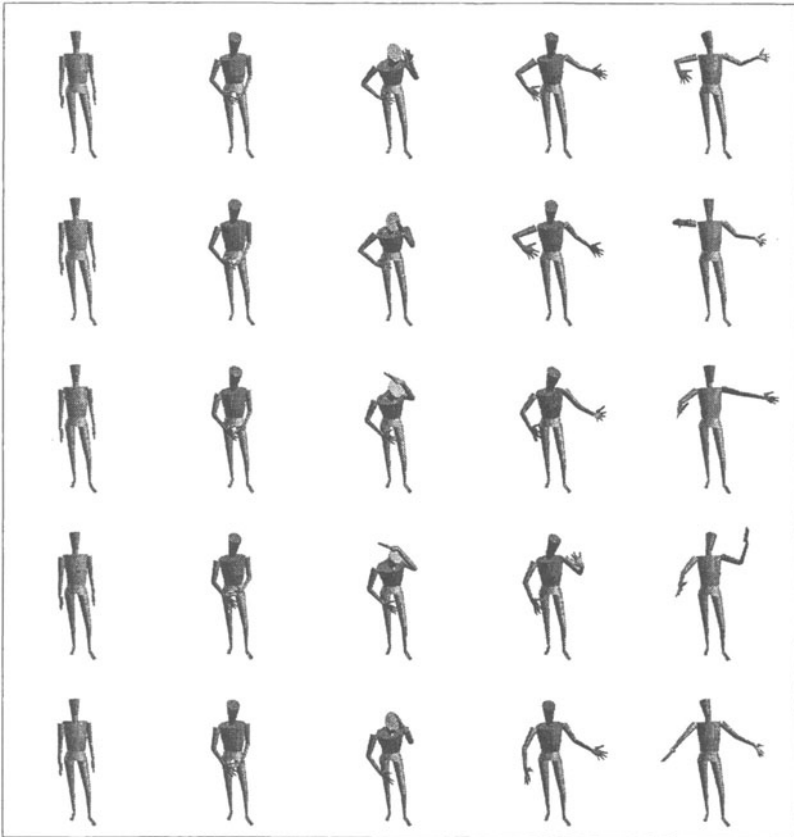


Figure 1 Five of ten similar but distinct example motions used for the motion synthesis.

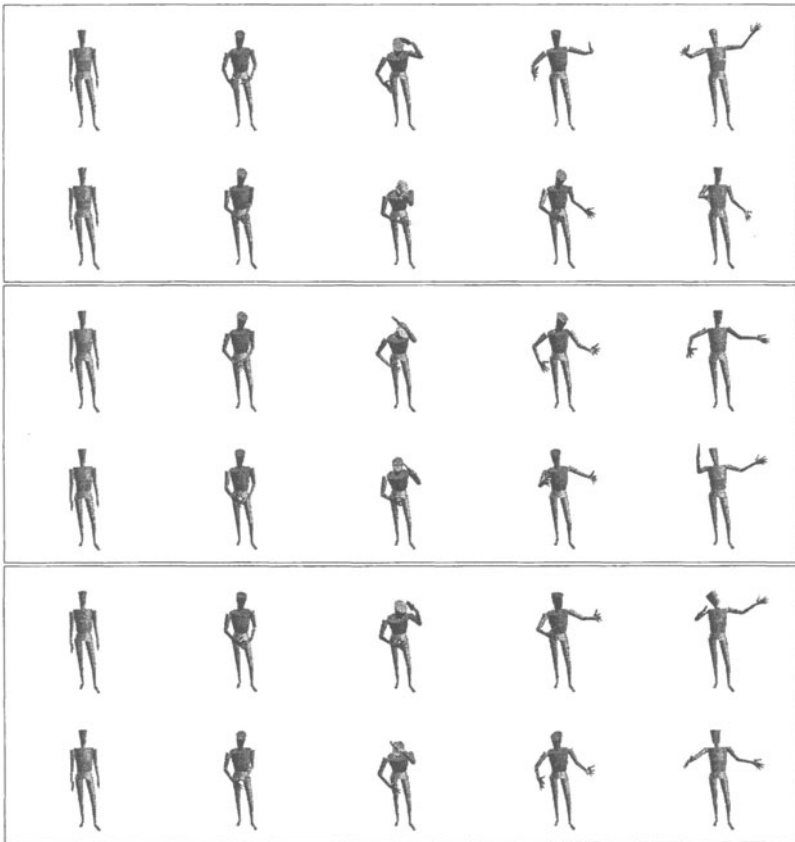


Figure 2 Three of the eigenvectors scaled by \pm constants are added to the average motion.

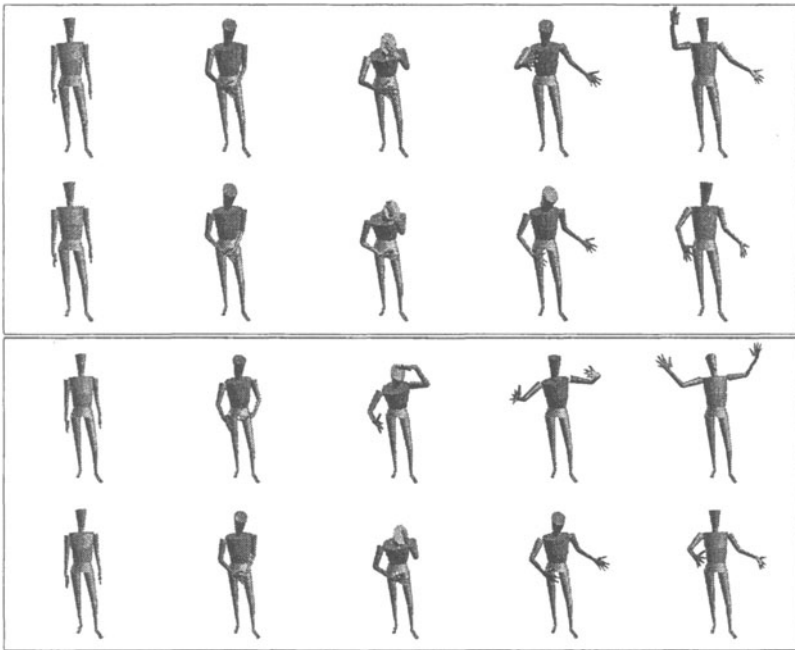


Figure 3 Motion Attributes. (Upper) the height of the human figure's right hand, high/low and (Lower) the energy of the human figure, high/low.

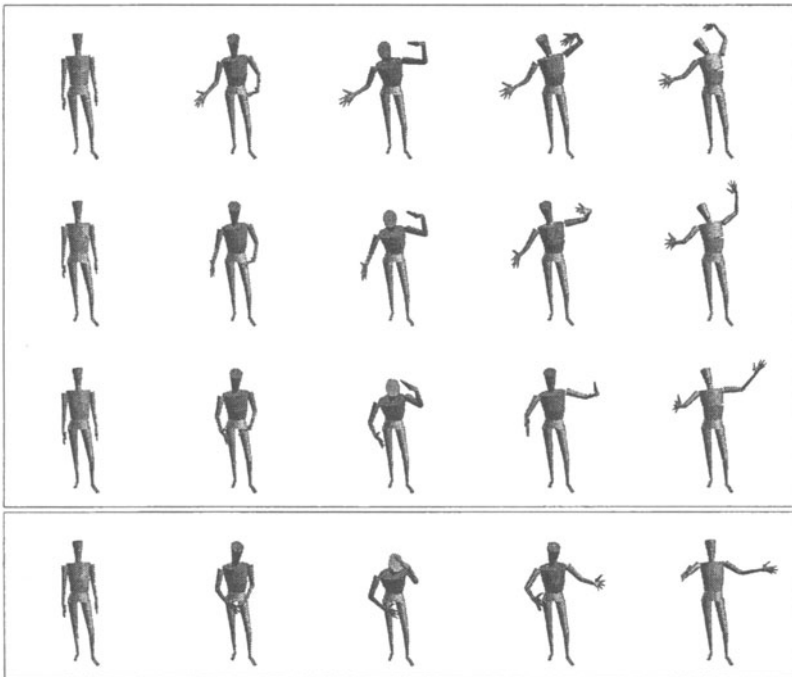


Figure 4 Motion Caricature. (Starting from the bottom) the average motion, the original motion and two of its caricatures with different degrees of distinctiveness.

References

- [AM00] M. Alexa and W. Muller. Representing animations by principal components. In *Proceedings of EUROGRAPHICS 2000*, 2000.
- [BH00] Matthew Brand and Aaron Hertzmann. Style machines. In *Proceedings of SIGGRAPH 2000*, pages 183 – 192, 2000.
- [Bis95] C. M. Bishop. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1995.
- [BP96] D. Beymer and T. Poggio. Image representations for visual learning. *Science*, 272(5270):1905 – 1909, 1996.
- [BPW93] N. Badler, C. Phillips, and B. Webber. *Simulating Humans: Computer Graphics, Animation, and Control*. Oxford University Press, 1993.
- [Bre85] S. E. Brennan. The caricature generator. *Leonardo*, 18:170 – 178, 1985.
- [BV99] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In *Proceedings of SIGGRAPH 1999*, pages 187 – 194, 1999.
- [BW95] A. Bruderlin and L. Williams. Motion signal processing. In *Proceedings of SIGGRAPH '95*, pages 97 – 104, 1995.
- [EP00] T. Ezzat and T. Poggio. Visual speech synthesis by morphing visemes. *International Journal of Computer Vision*, 38(1):45 – 57, 2000.
- [GJP95] F. Girosi, M. Jones, and T. Poggio. Regularization theory and neural network architecture. *Neural Computation*, 7:219 – 269, 1995.
- [GP00] M. Giese and T. Poggio. Morphable models for the analysis and synthesis of complex motion patterns. *International Journal of Computer Vision*, 38(1):59 – 73, 2000.
- [Jac91] J. E. Jackson. *A User's Guide to Principal Components*. John Wiley & Sons, New York, 1991.
- [LS99] J. Lee and S. Y. Shin. A hierarchical approach to interactive motion editing for human-like figures. In *Proceedings of SIGGRAPH 1999*, pages 39 – 48, 1999.
- [MGT99] S. R. Musse, F. Garat, and D. Thalmann. Guiding and interacting with virtual crowds in real-time. In *Proceedings of Computer Animation and Simulation 1999*, pages 23 – 34, 1999.

- [Per95] K. Perlin. Real-time responsive animation with personality. *IEEE Transactions on Visualization and Computer Graphics*, 1(1):5 – 15, 1995.
- [PW89] A. P. Pentland and J. Williams. Good vibrations: Modal dynamics for graphics and animation. In *Proceedings of SIGGRAPH '89*, pages 215 – 222, 1989.
- [PW99] Z. Popovic and A. Witkin. Physically based motion transformation. In *Proceedings of SIGGRAPH 1999*, pages 11 – 20, 1999.
- [RCB98] C. Rose, M. Cohen, and B. Bodenheimer. Verbs and adverbs: Multidimensional motion interpolation. *IEEE Computer Graphics and Applications*, 18(5):32 – 48, September/October 1998.
- [She00] C. Shelton. Morphable surface models. *International Journal of Computer Vision*, 38(1):75 – 91, 2000.
- [UAT95] M. Unuma, K. Anjyo, and R. Takeuchi. Fourier principles for emotion-based human figure animation. In *Proceedings of SIGGRAPH '95*, pages 91 – 96, 1995.
- [UB91] S. Ulman and R. Basri. Recognition by linear combination of models. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 13:992 – 1006, 1991.
- [VP97] T. Vetter and T. Poggio. Linear object classes and image synthesis from a single example image. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 19(7):733 – 742, 1997.
- [WH97] D. Wiley and J. Hahn. Interpolation synthesis of articulated figure motion. *IEEE Computer Graphics and Applications*, 17(6):39 – 45, November/December 1997.
- [WP95] A. Witkin and Z. Popovic. Motion warping. In *Proceedings of SIGGRAPH '95*, pages 105 – 108, 1995.