Eurographics/ ACM SIGGRAPH Symposium on Computer Animation (2006), pp. 1–2
M.-P. Cani, J. O'Brien (Editors)

Posters and Demos

Multilinear Motion Synthesis Using Geostatistics

Tomohiko Mukai and Shigeru Kuriyama

Toyohashi University of Technology, Aichi, Japan

Abstract

The trend in interactive entertainment is towards scenes with massive numbers of characters, and requiring huge amounts of motion data, which must be compactly and efficiently stored without sacrificing quality or controllability of the motions. Multilinear algebra is a powerful tool for efficiently representing multivariate data, including human motion data, through the analysis of multimodal correlations. The multilinear model, however, often suffers from undesirable artifacts when motion data are sparsely and non-uniformly sampled in a high-dimensional control space. For overcoming this defect, we introduce a geostatistical interpolation to the multilinear model by formulating it to fit into the motion representation with tensor approximation. The advantages of this approach are demonstrated by the motion synthesis in a high-dimensional control space and by a level of detail control. This technique provides practical tools for implementing interactive animations of many characters while ensuring accurate and flexible controls with a small amount of storage.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-dimensional Graphics and RealismAnimation

1. Introduction

Data-driven motion generation techniques are widely used for creating natural human animations. Several techniques, such as motion editing/retargeting and reassembling, have been proposed for dealing with the motion capture data through intuitive operations. Motion interpolation is a simple and efficient technique for synthesizing a desired motion by blending multiple motion samples with some control parameters that correspond to kinematical, physical, and emotional attributes [MK05]. Most methods separately interpolate a rotational and positional element at each time and joint. However, such a naive interpolation is redundant because human motion data usually indicates a multimodal correlation of joints, time series, and control parameters.

This paper proposes a practical method for interpolating motion samples with the multilinear model. Our method first identifies a subspace of the given motion samples by analyzing the multimodal correlations of joints, time series, and control parameters, by which the redundancy of motion samples is fully eliminated. A new motion is then synthesized by interpolating the principal features of the samples using geostatistics. Our motion representation within a multilinear subspace can reduce the storage and computational costs of geostatistical motion interpolation while preserving the prediction accuracy. Adaptive selection of the subspace (the number of principal features) can control the level of detail of motion synthesis.

2. Algorithm

Our method utilizes a high order singular value decomposition (HOSVD) that is a high order generalization of principal component analysis (PCA) [WWS*05]. Using HOSVD, a high order tensor is decomposed into a core tensor and basis matrices with reduced ranks through the analysis of multimodal correlation factors. A new data is efficiently generated by interpolating the principal components which the redundancy is fully eliminated.

Given motion samples, a motion tensor is constructed and decomposed in three steps. First, motion samples are temporally aligned with the same duration via dynamic time-warping. Secondly, the motion tensor is composed by lining up all samples in one dimension. HOSVD is then applied to the motion tensor for simultaneously analyzing three mode correlation: joint correlation, time correlation, and correlation between motion samples. The motion tensor is finally decomposed into a core tensor and three basis matrices with reduced ranks.

New motions are synthesized by interpolating the reduced principal components. Given target time frame and control parameters, a pose vector is synthesized by a product of the core tensor, joint basis matrix, and time and control coefficient vectors. The time coefficient vector is computed by linearly interpolating two column vectors of the time basis matrix. However, the simple linear interpolation cannot provide the accurate control coefficient vector due to the high dimensionality of the control space. Therefore, we introduce universal kriging to accurately estimate the control coefficient vector.

Universal kriging is a practical technique of geostatistics for estimating the continuous data distribution in a high dimensional space by the weighted average of the scattered samples. The weight function is optimized through the correlation analysis between the spatial distance and dissimilarity of the samples. Previous methods compute the positional distance between every joint pair as the dissimilarity of poses. We simply compute the dissimilarity of entire motions by Euclid distance between the core tensors. Though this simplification does not take the visual differences into account, we have experimentally confirmed that the resulting motions can sufficiently preserve the accuracy.

3. Experimental Result

We take 92 samples of reaching motions for evaluating the accuracy and performance of our method. The sample dataset consists of the motions of three performers, reaching their right hand to different locations while standing or sitting. The motion capture data is sampled at 120 Hz with 94 DOFs, where all motions are temporally aligned with 100 frames. The motion samples are parameterized in the 4D control space with the height of the performer, the hip height at the initial pose, and the turning angle from the forward direction and the height of the corresponding target position. We use the motion synthesized by naive kriging as the ground truth. The approximation error is evaluated by the sum of the positional distance between corresponding joints over the motion. About two thousand control parameters are uniformly sampled for computing the root mean squared error. Table 1 shows the compression ratio required for reducing the RMS error less than a threshold, and the average computational times for the preprocess and the pose synthesis. This shows that our method achieves a high compression ratio while preserving the accuracy. The preprocessing time and runtime speed is also faster than that in the naive interpolation.

We demonstrate a gait animation of hundreds of characters whose motions are generated using our method, where the root position of the character is individually controlled. Since the cost of motion synthesis linearly increases with the number of characters in a scene, we control the compression ratio as in geometric level of detail control, by switching the four compression levels. The compression level is automatically determined for each character by the distance from the viewpoint so that the higher level is assigned to the more distant character. The average and maximum computational time for the motion synthesis are 270 and 360 msec per frame, respectively. Without the control of compression levels, the average computational times with the lowest compression and the higher compression are 510 and 190 msec per frame, respectively.

Table 1: *Compression ratio and computational times required for reducing RMS error to less than a threshold.*

RMSE threshold	Data size	Precomputation (sec)	Runtime (msec/pose)
Naive	$92 \times 100 \times 92$	119.8	2.76
500	$22 \times 6 \times 24$	6.0	0.84
750	$18 \times 6 \times 28$	6.4	0.81
1000	$20 \times 6 \times 20$	6.3	0.83

4. Conclusions

This paper has introduced geostatistical motion interpolations into motion representation within a multilinear subspace. Multilinear representation provides a flexible control of the subspace dimension, depending on the desirable accuracy of the resulting motions. On the other hand, geostatistical interpolation enhances the prediction accuracy of synthesized motions. As a result, our integration of two numerical methodologies is more accurate than the simple restorations of a multilinear model, and requires less computational cost and storage than naive interpolations.

Multilinear analysis has the potential to be applied to other motion editing techniques such as motion transition and warping. Moreover, the synthesized motions are not limited to captured human motions; any type of motion data could be managed if they satisfy the statistical preconditions. The verification of this potential is the subject of our future work. From a practical viewpoint, the size of the tensors should be automatically optimized by estimating the visual effect of the approximation on the motions. Moreover, the implementation of multilinear motion synthesis on GPU could enhance its usability.

References

[MK05] MUKAI T., KURIYAMA S.: Geostatistical motion interpolation. *ACM Transactions on Graphics 24*, 3 (2005), 1062–1070.

[WWS*05] WANG H., WU Q., SHI L., YU Y., AHUJA N.: Out-of-core tesor approximation of multidimensional matrices of visual data. *ACM Transactions on Graphics 24*, 3 (2005), 527–535.