Style Subspaces for Character Animation

Gengdai Liu, Zhigeng Pan, Zuoyan Lin

State Key Lab of CAD&CG - Zhejiang University

Hangzhou, 310027, P.R. China

Tel. (+81)571 88206683 401 Fax. (+81)571 88206680

email: liugengdai@cad.zju.edu.cn

Abstract

In this paper, we present a novel method for editing stylistic human motions. We represent styles as differences between stylistic motions and introduced neutral motions, including timing differences and spatial differences. Timing differences are defined as time alignment curves, while spatial differences are found by a machine learning technique: Independent Feature Subspaces Analysis, which is the combination of Multidimensional Independent Component Analysis and Invariant Feature Subspaces. This technique is used to decompose two motions into several subspaces. One of these subspaces can be defined as style subspace that describes the style aspects of the stylistic motion. In order to find the style subspace, we compare norms of the projections of two motions on each subspace. Once the time alignment curves and style subspaces of several motion clips are obtained, animators can tune, transfer and merge the style subspaces to synthesize new motion clips with various styles. Our method is easy to use since manual manipulations and large training data sets are not necessary.

Keywords: motion capture, style, independent feature subspaces analysis

Introduction

Nowadays, motion capture is widely used in many communities, including film industry, game development and robotics. However, motion capture is an expensive technique and lack of flexibility. Therefore, reuse of motion capture data becomes a promising research area where many researchers have contributed. Our method in this paper can fall into this category, but it mainly focuses on how to synthesize new human motions with various styles from existing motion data.

Style is an abstract concept and hard to express quantitatively. But generally, it can be regarded as subtle variations on the basic motions. Therefore, two assumptions are proposed in this paper. Firstly, the style can be defined as the differences between the stylistic motions and neutral motions, including spatial differences and timing differences [1, 2, 3, 4]. Secondly, the style and content of the motions are independent mutually and can be separated [16].

According to our assumptions, the spatial differences are expressed as a subspace of the motion data in this paper. This subspace that contains the stylistic aspects of the motion data is called *style subspace*. In order to find the style subspace of a single stylistic motion clip, a neutral motion clip with no style must be employed. We use independent feature subspaces analysis to find the style subspace of the stylistic motion. Once the style subspaces of several motion clips are obtained, animators can not only change the degree of style of the original motion but also transfer and merge styles between two motions to generate new motions with various styles. Our method is easy to use for animators since manual manipulation and large data sets are not necessary.

Figure 1 shows the overview of our method. At decomposition stage, several subspaces are obtained by using independent feature subspace analysis and one of them is defined as style subspace. At synthesis stage, new motions with various styles can be generated in real time by tuning, transferring and merging style subspaces.

Related works

The idea of defining the style of a stylistic motion as differences between the stylistic motion and a neutral motion is similar to that of our method. Amaya et al [1] think emotion of a motion lies in speed and spatial amplitude. Their method can extract emotion by comparing neutral and stylistic motions. Then the differences between two motions are added to another motion to generate new style. Kawasaki et al [2] extract the differences in angular velocity and torque of the joints between standard and stylistic motions and transfer the differences to the third motion. Terasaki et al [3] extract style features from differences between a base motion and a reference motion and apply the style features to other motions. The style features in this paper consist of postural and temporal differences between two motions and are defined in an abstract form. Hsu et al [4] learn Linear Time Invariant models by comparing the input and output motions to perform style translation. All the works above inspired us to extract style by comparing a neutral motion and a non-neutral motion. And we adopt a distinct technique to find the differences.

Since a motion can be regarded as a time-varying signal, some researchers applied signal processing techniques to edit stylistic motion. Unuma et al [5] use Fourier techniques to change the style of human gaits in Fourier domain. Using their method, the Fourier characteristics that describe the style of human gaits can be extracted. Bruderlin and Williams [6] edit the stylistic motions by varying the frequency bands of the signal. Perlin [7] adds rhythmic and stochastic noise functions to joints to drive computer generated puppets with various personalities. Our method is unrelated to signal processing, though we use band pass filter at pre-processing stage to reduce noise and facilitate estimation of independent feature subspaces. We do not assume that the style lies in high frequency, but think it can be found by analyzing the statistical characteristics of a single motion.

Rose et al [8] present multidimensional motion interpolation, a method to generate a style space by interpolating between sample motions using RBF. Our method differs in that we do not interpolate between two motions, but extract style subspaces of stylistic motions and edit the subspaces.

Neff et al [9] present an approach and prototype system for generating stylistic character animation. Their system provides animator with a set of edit modes and bridges the gap between artistic and technical communities.

In recent years, more and more researchers applied machine learning to character animation. Urtasun et al [10] use PCA to train large sets of locomotion data, and use PCA coefficients to synthesize new motions with different heights and speeds. Brand and Hertzmann [11] use Hidden Markov Models to capture the style of training motions. These styles can be reused to other motions. Torresani et al [12] introduce a motion blending based motion controller. The blending weights are learned from large sets of motions, whose styles are labelled by some specialists. Liu et al [13] construct a physical model and use optimization to generate new motions with learned physical parameters that contain the style aspects. We do not consider physical aspects in our algorithm, which is a limitation of our method. However, we use a balance filter to correct some poses that break the physical rules at post-processing stage. Elgammal and Lee [14] take the style of a human motion as a time-invariant parameter, and learn a decomposable generative model that explicitly decomposes the style from a walk motion video. Cao et al [15] use ICA to find the emotion aspects of facial motions automatically for the purpose of style editing. All the research works above can produce good results, but large sets of training motion data are required. Our method differs in that it works effectively if we want to extract the style from only a single stylistic motion with the aid of an introduced neutral motion.

Shapiro et al [16] use ICA to decompose a single motion into many components. The animators have to select one of them as style component manually. Our method is the modification and extension of Shapiro's. We can extract style automatically by introducing a neutral motion and using Independent Feature Subspaces Analysis.

Data pre-processing

Motion data representations

All the motion capture data used in this paper come from CMU Motion Capture Database [17]. But we simplify the skeleton structure and remove some negligible joints and links as shown in Figure 2.

Our method can work on either Euclidean coordinate of joints, Euler angles or quaternions of links. Euler angle is a poor choice due to Gimbal lock. Quaternion is not suitable for motion decomposition because we can not separate motions in quaternion space meaningfully [16]. Therefore, we adopt Euclidean coordinate representation of motion. Since all data in CMU Motion Capture Database are represented with Euler angles, we have to convert the motion capture data into global Euclidean coordinate point representation at first.

It is worth mentioning that the global translation should be removed from the motion firstly because we assume that these values have no contribution to the style of a motion. Then the motion $m(t)$ is represented by the trajectories of joints $p_i(t), 1 \le i \le K$, where K is the number of its joints. In our method, $K = 25$. Our algorithm works on a matrix, whose rows represent different DOFs of the motion and whose columns represent all individual frames involved.

Time alignment

In order to find spatial differences between a neutral motion and a stylistic motion, we have to compare the two motions frame to frame. Therefore a time alignment process has to be taken.

Let $d(p,q)$ be the distance between motion 1 at frame p and motion 2 at frame q. To perform time alignment, we need to find a time path $\lambda(n) = (p(n), q(n)), n = 1, 2, ..., L$ to make sum of distance between every corresponding frame pair along the time path minimum, where L is the length of the time path, and $p(n)$ and $q(n)$ is the corresponding frame of motion 1 and motion 2 at n st point on the time path, which can be defined as:

$$D(M_1, M_2) = \min_w \sum_{n=1}^{L} d(\lambda(n)) \quad (1)$$

where M_1 and M_2 represent the two motions aligned. The distance metric used originally by Kovar [18] is chosen for $d(p,q)$.

After computing the distances between every corresponding frame pair, we can draw a distance image where larger distance corresponds to whiter pixel and vice versa, as shown in Figure 3.

We use Kovar's method [19] to figure out the time path $\lambda(n)$ that is a curve on the distance image. In order to obtain this curve correctly, three constraints must be taken into account: continuity, monotonicity and slope limit (it is set to 2). These constraints can be described as:

$$\textit{if } \lambda(n) = (p,q) \textit{ then } \lambda(n-1) \in \{(p-1,q),(p-1,q-1),(p,q-1)\} \quad (2)$$

$$\begin{aligned} &\textit{if } \lambda(n) = (p,q) \textit{ and } \lambda(n-1) = (p-1,q) \textit{ then } \lambda(n+1) \in \{(p+1,q+1),(p,q+1)\} \\ &\textit{or if } \lambda(n) = (p,q) \textit{ and } \lambda(n-1) = (p,q-1) \textit{ then } \lambda(n+1) \in \{(p+1,q),(p+1,q+1)\} \end{aligned} \quad (3)$$

Our time alignment system allows animators to select the start and end of the curve interactively. Then the system solves Equation (1) for $\lambda(n)$ subject to constraints above using dynamic programming. In Fig.3, a neutral walk and a stagger are aligned by a red curve whose start and end is selected manually.

After obtain $\lambda(n)$, we can resample the two motions along the curve to produce synchronized versions of them with new timings. Suppose the index $n_1,...,n_{T_1}$ are chosen and T_1 is the number of frames in aligned motion 1 such that $p(n_i) = i$. Then the motion 1 will be replayed with its original timing, i.e. the motion 2 is aligned to motion 1. Considering the fact that independent feature subspaces analysis can only find the spatial differences between motions and that the timing of original motion which can also be regarded as a part of style (e.g. slow-in and slow-out) will be lost if we wrapped stylistic motions to neutral motions, we align neutral motions to stylistic motions so that the timing of stylistic motions can be preserved.

Filtering and whitening

Before we apply our algorithm the aligned motion data has to go through two phases: filtering and whitening.

The reason why we have to use filtered motion data to train the independent feature subspace model is to obtain more mutually independent subspaces. It has been proven true that ICA model (and its extensions) is still valid after the data is filtered in time domain [20]. The filter we adopt for our method is a bandpass filter. We use Bruderlin's method of multiresolution filtering [6] to perform bandpass filtering. After the Laplacian bandpass pyramid is constructed, the motion can be expressed as sum of DC value and all the bandpass bands. Then the motion is filtered by decreasing slightly the signal in highest frequency-band and signal in some lower frequency-bands. The purpose of decreasing a little high frequency signal is to reduce noise that makes human skeleton quivering rapidly and will deteriorate the final results. The purpose of eliminating some low frequency signal is to construct a high-pass filter to turn the original motion data into an approximate innovation process that is likely to lead to much better estimates of the mixing matrix [21].

Whitening is a common pre-processing technique for ICA and its extensions [20]. Before whitening, the data should be centred. That is to shift the data towards its mean so that the resulting variables have zero mean. Whitening transforms the input data linearly so that we can obtain new data whose components are uncorrelated and whose variances equal unity. PCA can achieve decorrelation, so it is commonly used to perform this transformation. Whitening can also reduce the dimension of the motion data and accelerate convergence.

Style subspaces

According to our assumptions, we can extract style by finding the differences between neutral and non-neutral motions and apply extracted style to other motions. Moreover, we assume style of a motion has some invariant features, i.e. the style can keep invariant even though the motion changes partly. Considering this, we use independent feature subspace model [22], which is a modification of ICA and a combination of the technique of multidimensional independent component analysis and principle of invariant feature subspaces to extract styles.

In this section, we introduce the concept of independent feature subspaces and the method on how to apply this technique to human motions to extract style subspaces that describe the style aspects of stylistic motions.

Independent feature subspaces

Multidimensional independent component analysis is a linear generative model. Compared with ICA, the components s_i are not supposed to be all mutually independent. Instead, it is assumed that the s_i can be divided into n-tuples, such that s_i in a given n-tuples can be dependent on each other, but dependencies among different n-tuples are prohibited. The principle of invariant feature subspaces is an approach to represent features with some invariances. An invariant feature can be considered as a linear subspace in a feature space. The value of the invariant, high-order feature is given by the square of the norm of the projection of the given data on the subspace. The value can be described as:

$$F_j(t) = \sum_{i=1}^{k} (b_i^T x(t))^2, i \in S_j \qquad (4)$$

Independent feature subspace is a technique combined by the two techniques introduced above. In this model, the likelihood of observed data $x(t)$ can be described as:

$$\log L(x(t), t=1,\cdots,T; b_i, i=1,\cdots,n) = \sum_{t=1}^{T}\sum_{j=1}^{J}\log p(\sum_{i\in S_j}(b_i^T x(t))^2) + T\log|\det B| \qquad (5)$$

where b_i is the basis vector of subspaces, J is the number of independent feature subspaces and S_j, $j=1,\cdots,J$ represents the set of the indices of the s_i belonging to the subspace of index j. B is a matrix with b_i as its columns. $p(\cdot)$ gives the probability density inside the jth n-tuple of s_i. We can assume the norm of the projection of data on any subspace has a supergaussian distribution.

4.2 Learning a style subspace

In fact, independent feature subspace analysis is a technique not only for feature extraction, but also for signal decomposition. Therefore, a subspace expresses the style while it is a part of the original data. For animators, it is an intuitive model for motion editing.

After concatenating pre-processed stylistic motion $m^S(t)$ and neutral motion $m^N(t)$ end to end, we obtain a new motion $m(t)$. We replace $x(t)$ in Equation (4) and (5) with $m(t)$ and estimate independent feature subspaces by maximizing the Equation (5) using a stochastic gradient ascent algorithm. In Equation (5), we use the following probability distribution that is supergaussian:

$$\log p(u) = -\alpha u^{1/2} + \beta \qquad (6)$$

In order to find the style subspace, a metric is defined to describe the dissimilarity between corresponding subspaces of stylistic and neutral motions. It is described as:

$$d_j = (\frac{1}{T}(\sum_{t=1}^{T}(F_j^S(t) - F_j^N(t))^2))^{1/2} \qquad (7)$$

where $F_j^S(t)$ and $F_j^N(t)$ represent the value of feature of the stylistic motion and neutral motion respectively and T is number of frames of aligned motions. We define the style subspace as one of these subspaces that makes d_j maximum, which is described as:

$$j^s = \arg\max d_j, j \in \{1,\cdots J\} \qquad (8)$$

In Figure 4, an example of decomposition of two motions is illustrated. We concatenate a neutral walk motion and a stride motion to construct $m(t)$. $F^S(t)$ is represented by blue curves and $F^N(t)$ is represented by red curves. It is obvious that the dissimilarity between $F^S(t)$ and $F^N(t)$ of the third subspace is the largest. Therefore this subspace is specified as style subspace. It is amazing that for the most cyclic motions we can always obtain a subspace where the dissimilarity between features of two motions is clearly different from others. This is why we choose this technique to decompose motions.

Edit of stylistic human motions

Due to orthonormality of the basis vectors of all subspaces, the motion can be expressed as:

$$m(t) = E\{m(t)\} + P\tilde{A}s(t) \qquad (9)$$

where $E\{\cdot\}$ represents the means of input data, and P is the PCA matrix related to whitening. $\tilde{A}$ is the mixing matrix defining the mapping from subspaces to the whitened data space. It is the inverse matrix of B in Equation (5) which is estimated by independent feature subspace analysis. $s(t)$ is projection of whitened original data on all basis vectors at given time t. Recall previous section, matrix P and $\tilde{A}$ is estimated using filtered motion. But in Equation (9), $s(t)$ is the projection of whitened original motion data but not projection of whitened filtered data on the subspaces. At decomposition stage, our method uses filtered data to estimate the independent feature subspaces model in order to obtain much better estimation. At synthesis stage, our method uses projection of original data on each subspace to reconstruct the motion $m(t)$. Therefore, our method is essentially unrelated to signal processing. We do not model style in high frequency signal, but extract style by decomposing motion statistically.

Based on Equation (9), we proposed three editing modes similar to those in [15] that originally worked on facial animation: style tuning, style transfer and style merging.

Style tuning

Once the style subspace of a stylistic motion is obtained, the style is parameterized by the norm of style subspaces. Therefore, we can scale projection of data on the style subspace to change the value of style. This editing mode can be mathematically expressed as:

$$m(t) = E\{m(t)\} + P\tilde{A}(s(t) + (\alpha \sum_{i=1}^{k} (s^T(t) e_i^{j_s}) e_i^{j_s})^T) \quad (10)$$

where α is the coefficient to control the degree of style $e_i^{j_s}$ is the unit basis corresponding to the basis vector of the style subspace. We can obtain new motions with different degrees of style by tuning only one parameter. Essentially, this editing mode is an interpolation between the stylistic motion and the neutral motion. However, this kind of interpolation is done in the low-dimensional style subspace.

Style transfer

Other than editing on the original motion, we can also transfer style from the original motion to other motions. It is notable that motion 1 and 2 must be aligned before editing. Considering timing is also a part of style of motion data, we need align motion 1 to motion 2 in order to preserve the timing of motion 2. This editing mode can be expressed as:

$$m(t) = E\{m(t)\} + P\tilde{A}(s_1(t) + (\sum_{i=1}^{k}((s_2(t) - s_1(t))^T e_i^{j_s})e_i^{j_s})^T) \qquad (11)$$

Equation (11) shows how to transfer style of motion 2 to motion 1 mathematically. If motion 1 and 2 are regarded as neutral motion and stylistic motion respectively, the style of motion 2 can be transferred to motion 1 directly using Equation (11). But if motion 1 is a third motion, it has to be projected on the style subspace of motion 2 at first. The projection can be expressed as:

$$s_1(t) = (P\tilde{A})_2^{+}(m_1(t) - E\{m_1(t)\}) \qquad (12)$$

where $+$ represents pseudo-inverse of a matrix.

Style merging

Another interesting editing mode is to merge styles of multiple motions. It can be expressed as:

$$\begin{aligned} m(t) = E\{m(t)\} &+ (P\tilde{A})_1(s_1(t) - (\sum_{i=1}^{k}(s_1^T(t)e_i^{j_{S1}})e_i^{j_{S1}})^T) + \alpha(P\tilde{A})_1(\sum_{i=1}^{k}(s_1^T(t)e_i^{j_{S1}})e_i^{j_{S1}})^T \\ &+ (1-\alpha)(P\tilde{A})_2(\sum_{i=1}^{k}(s_2^T(t)e_i^{j_{S2}})e_i^{j_{S2}})^T \end{aligned} \qquad (13)$$

Equation (13) shows how to merge styles of two motions with weight α . Obviously, this editing mode preserves the style of motion 1 while style transfer replaces the style of motion 1 with that of motion 2. When the coefficient α is changed, the degrees of two styles vary with it. When $\alpha = 1$, the synthesized motion is motion 1 itself. When $\alpha = 0$, the style of motion 1 is removed clearly and the style of motion 2 is added entirely. Style merging is a kind of space-time interpolation between two aligned motion described by Torresani et al [12], but the interpolation is implemented only in two low-dimensional style subspaces. As for timing, we can either align motion 1 to motion 2 or vice versa, which depends on which motion’s timing we wish to preserve. Moreover, we find that style merging is very effective when we merge two styles, but when we merge three or more styles, some unnatural and undesirable motion clips may be produced.

Post-processing

After editing, the motion data has to be post-processed. Firstly, the global translation removed at the pre-processing stage must be added again.

Secondly, since our method that works on global joint positions can not guarantee that the limb lengths of the character keep unchangeable, the limb lengths have to be restored. In order to do that, we represent each limb as a vector in global coordinate, and convert these vectors into quaternions. Then the pose at each frame can be reconstructed according to lengths of all limbs. Subsequently, the motion data is expressed by quaternions, so that the limb lengths are corrected but the joint angles are preserved.

Finally, some artifacts must be removed. It is obvious that our method does not guarantee physical and kinematical correction, which is a common weakness of most statistical methods. Therefore we use an IK solver to clean foot sliding on the ground. In addition, we develop a motion balance filter using existing technique to correct the poses that break the basic physical rules [23]. The filter examines the ZMP for each frame and corrects the faulty poses when the corresponding ZMP goes outside the support area of foot.

Results and discussions

There are 25 joints in the skeleton model used in this paper, thus the dimension of motion data is 75. After whitening, the dimension is reduced to 10, but over 99% of variation of original data is preserved. We employ independent feature subspace analysis to obtain five 2-D subspaces, one of which is the style subspace.

In Figure 5, a sneak walking motion is edited by tuning the scalar α. It is obvious that when the norm of projection of motion data on the style subspace is adjusted, the degree of sneaking changes.

In Figure 6, the style of a stride motion is transferred to a jogging motion. The newly synthesized motion is a run at a stride motion where the timing of stride is preserved. We can find that it is striding more than jogging because part of the jogging style is replaced by striding style.

In Figure 7, the style of a brisk walk motion and the style of a soldier march motion are merged. The newly synthesized motion is a brisk soldier march motion. Both brisk walk style and soldier march style are preserved. When the coefficient α is tuned, the ratio of each style in the synthesized motion varies. In this case, the timing of brisk walk is preserved because it is crucial to express the style of brisk.

In [16], Obviously, more precise differences between the neutral motion and non-neutral motion can be found by using IFSA than by using ICA. The reason why IFSA can find the differences more precisely is that this technique considers some high-order invariant features that describe the style more precisely. IFSA takes advantage of the residual dependences between independent components in ICA model (In fact, decomposing a vector into entirely independent components is impossible). Recall Equation (5), where the feature value is the norm of projection of motion data on each subspace, so each subspace is spherically symmetric. IFSA does not discriminate projections of input data on the different basis vectors of a subspace, but considers only the norm of projection, i.e. the components of a subspace are not all independent mutually. This characteristic enables the features to keep invariant even though some parts of input data change. Therefore, we can obtain more independent stylistic aspects which are influenced less by the content of motion data. Although the detailed meaning of invariance in this model is not specified clearly, we can expect it is more suitable for style extraction, transfer and merging than ICA, which has been proven reasonable by the results.

Another point worth noticing is that the results produced by style transfer and style merging are similar sometimes though they are expressed by different formulas. For example, transfer style of stride to jogging and merging styles of stride and jogging can give similar results that are both jogging at a stride. However, the former looks more like triple jump than jogging. This is caused by the fact that only a part of style of jogging remains after the data of motion 1 is projected on the style subspace of motion 2.

Although some good results can be obtained by using our method, there are still several limitations.

Our method is more suitable for cyclic motions (e.g. locomotion such as walking and jogging) than acyclic motions. There are two reasons. Firstly, our method depends highly on the statistical characteristics of given motions that cyclic motions provide more than acyclic motions. Secondly, it is difficult to align an acyclic motion with a neutral motion. However, if we choose a proper neutral motion for certain acyclic motions, we can also obtain satisfactory results. As shown in Figure 8, we can generate an old man playing the violin motion by merging an old man waiting motion and a playing the violin motion that is an acyclic motion. An idle standing motion is chosen as the neutral motion in this case.

Compared with some previous works [1, 2, 4, 10], our method can not synthesize motions at different speeds due to time alignment at pre-processing stage. ISFA on a stylistic motion and a neutral motion can only find the spatial differences between them. However, we can preserve timing of a stylistic motion that is also regarded as a part of style by aligning neutral motions to stylistic motions.

Another limitation is that the speed of estimation of subspaces for our method is lower than that of ICA. Our method estimates the subspaces by using stochastic gradient algorithm while ICA method uses fast fixed-point algorithm which has faster convergence speed. Moreover, our method has to estimate more orthogonal vectors spanning the subspaces. For the motion data used in this paper (about 200 frames), it takes us about 30 seconds to obtain 5 subspaces spanned by 2 vectors running on the machine with 2.4GHz CPU and 1GB memory. Therefore our method is not suitable for real-time interactive editing as described in [16], but synthesizing new motions with various styles when style subspaces have been obtained can be implemented in real time.

Conclusion

In this paper, a novel method to edit stylistic human motions is presented. Instead of learning a statistical or linear model from large motion data sets, our method decomposes a single stylistic motion into several subspaces to find the style aspects. By using Independent Feature Subspace Analysis, we can get the correlations between DOFs of a motion that express some essential features of this motion. After comparing the values of features between a stylistic motion and a neutral motion, our method finds the style subspace automatically. Based on the proposed decomposition method, a set of editing modes that can not only change the style of the original motion but also transfer and merge styles between two motions are given.

Although we consider the kinematics and dynamic aspect of a motion and remove the artifacts at post-processing stage, the statistics-based method is not a perfect approach to edit stylistic motions. How to combine statistics, kinematics and dynamics naturally to synthesize more realistic stylistic motion is our future work.

Acknowledgements

Many thanks to CMU Graphics Lab for providing free motion capture data. This work is co-supported by 863 Project (grant no: 2006AA01Z335), NSFC project (grant no: 60533080) and “On-line Expo” project by MOST.

References

[1] K. Amaya, A. Bruderlin and T. Calvert. Emotion from motion. In Proceeding of Graphics Interface 1996, pages 222-229, May 1996.

[2] R. Kawasaki, Y. Kitamura and Fumio Kishino. Extraction of motion individuality in sports and its application to motion of characters with different figures. In Proceeding of Computer Graphics International 2003.

[3] T. Terasaki, M. Oshita, Motion Style Transformation by Extracting and Applying Motion Feature, International Conference on Computer Animation and Social Agents 2006 (CASA 06), pp. 243-252, Geneva, Switzerland, July 2006.

[4] E. Hsu, K. Pulli and J. Popovic. Style translation for human motion. ACM Transactions on Graphics, 24(3):1082-1089, 2005.

[5] M. Unuma, K. Anjyo and R. Takeuchi. Fourier principles for emotion-based human figure animation. In Proceedings of ACM SIGGRAPH, pages 91.96, New York, NY, USA, 1995.ACM Press.

[6] A. Bruderlin and L. Williams. Motion signal processing. In Proceedings of ACM SIGGRAPH: Computer Graphics, pages 97-104, August 1995.

[7] K. Perlin. Real time responsive animation with personality. IEEE Transaction on Visualization and Computer Graphics, vol.1, no.1, pp.5-15, 1995.

[8] C. Rose, M. F. Cohen and B. Bodenheimer. Verbs and adverbs: Multidimensional motion interpolation. IEEE Computer Graphics and Applications, 18(5):32-41, 1998.

[9] M. Neff, E. Fiume, AER: Aesthetic Exploration and Refinement for Expressive Character Animation, In Proceeding of Eurographics/SIGGRAPH Symposium on Computer Animation 2005.

[10] R. Urtasun, P. Glardon, R. Boulic, D. Thalmann, and P. Fua. Style based Motion Synthesis. Computer Graphics Forum, 23(4):1-14, 2004.

[11] M. Brand and A. Hertzmann. Style machines. In Proceeding of Siggraph 2000, pages 183-192. ACM Press / ACM SIGGRAPH / Addison Wesley Longman, 2000.

[12] L. Torresani, P. Hackney and C. Bregler. Learning motion style synthesis from perceptual observations. In Proceeding of NIPS 19, 2007.

[13] C. K. Liu, A. Hertzmann, and Z. Popovic. Learning physics-based motion style with nonlinear inverse optimization. ACM Transactions on Graphics, 24(3):1071-1081, 2005.

[14] A. Elgammal and C.S. Lee. Separating Style and Content on Nonlinear Manifold. In Proceeding of CVPR 2004.

[15] Y.Cao, P.Faloutsos and F.Pighin. Unsupervised learning for speech motion editing. In Proceeding of Eurographics/SIGGRAPH Symposium on Computer Animation 2003.

[16] A. Shapiro, Y.Cao and P. Faloutsos, Style Components, In Proceeding of Graphics Interface 2006.

[17] CMU Graphics Lab Motion Capture Database, http://mocap.cs.cmu.edu/, 2008.

[18] L. Kovar, M.Gleicher and F.Pighin. Motion graphs. In Proceeding of ACM SIGGRAPH 2002, pp.473-482.

[19] L. Kovar and M. Gleicher. Flexible Automatic motion blending with registration curves. In Proceeding of Eurographics/SIGGRAPH Symposium on Computer Animation 2003, 2003, pp. 214-224.

[20] A. Hyvarinen, J. Karhunen, and E. Oja. Independent Component Analysis, John Wiley & Sons, 2001.
[21] A. Hyvarinen, Independent Component Analysis for Time-dependent Stochastic Processes, in Proceeding of International Conference on Artificial Neural Networks'98, 1998.
[22] A.Hyvarinen, P.Hoyer, Emergence of phase- and shift-invariant features by decomposition of natural images into independent feature subspaces, Neural Computation, pp.1705-1720, 2000.
[23] S. Tak, O.Y. Song, and H, S, Ko. Motion Balance Filtering, Eurographics 2000.

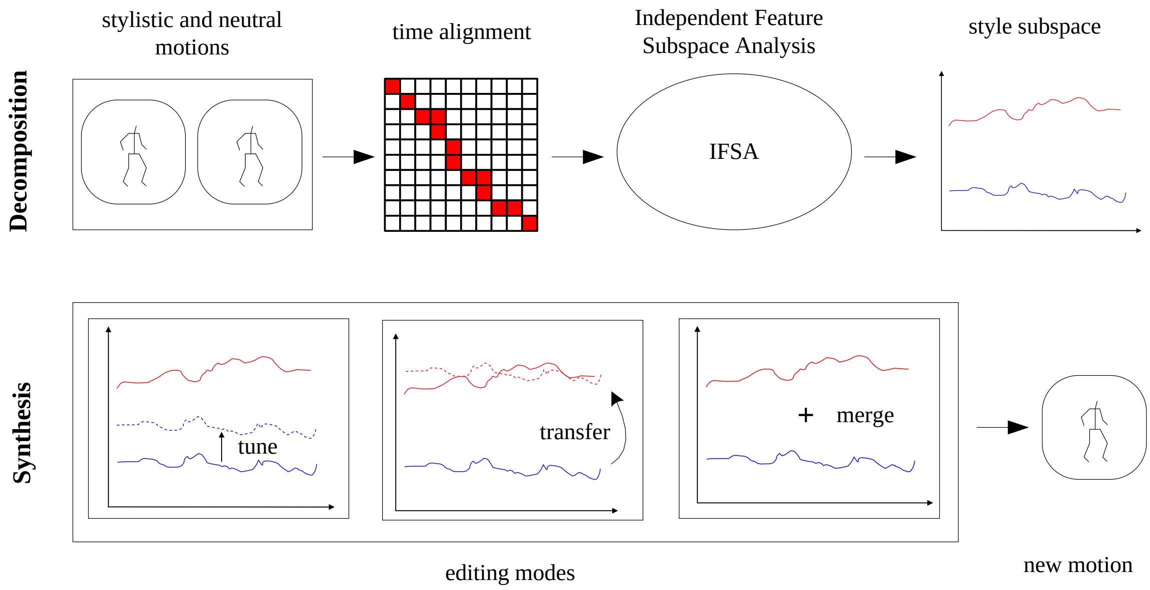


Figure 1 : Overview of our method

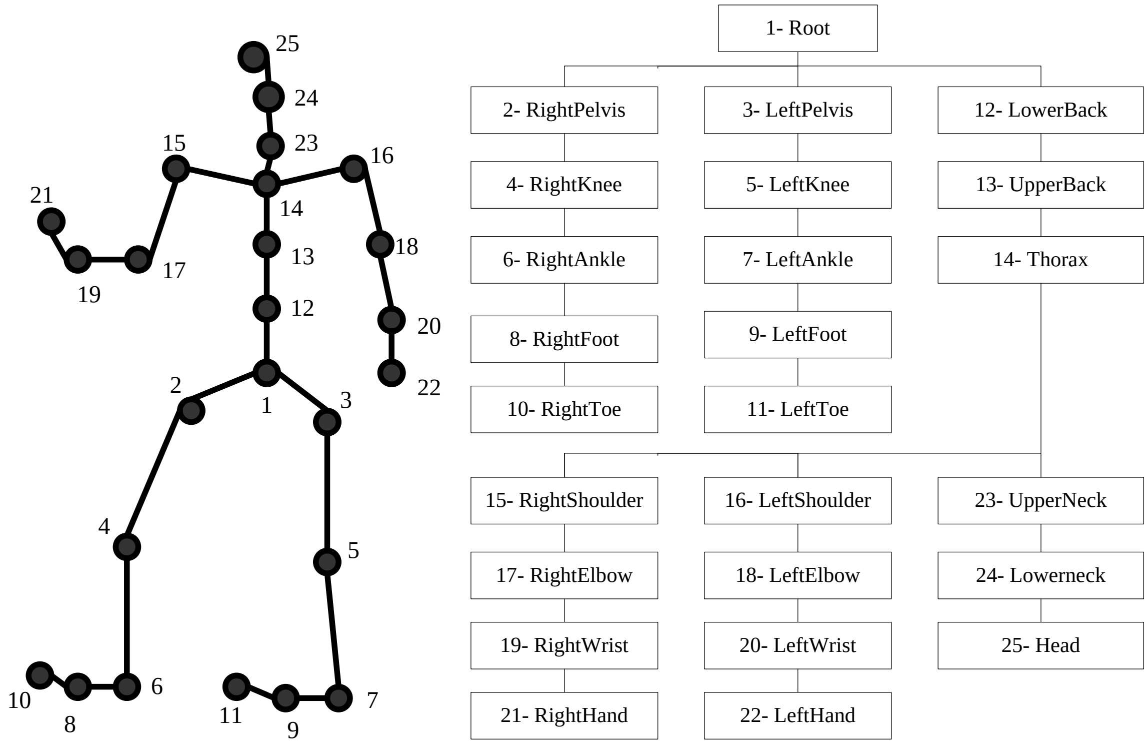


Figure 2 : Skeleton model and its hierarchical structure

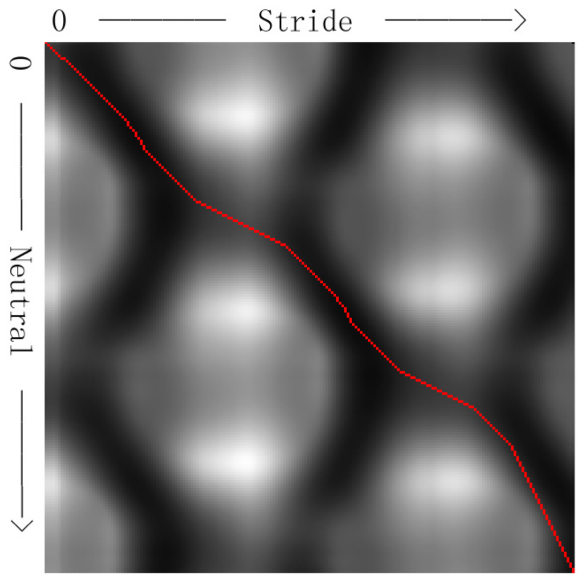


Figure 3 : An example of time alignment

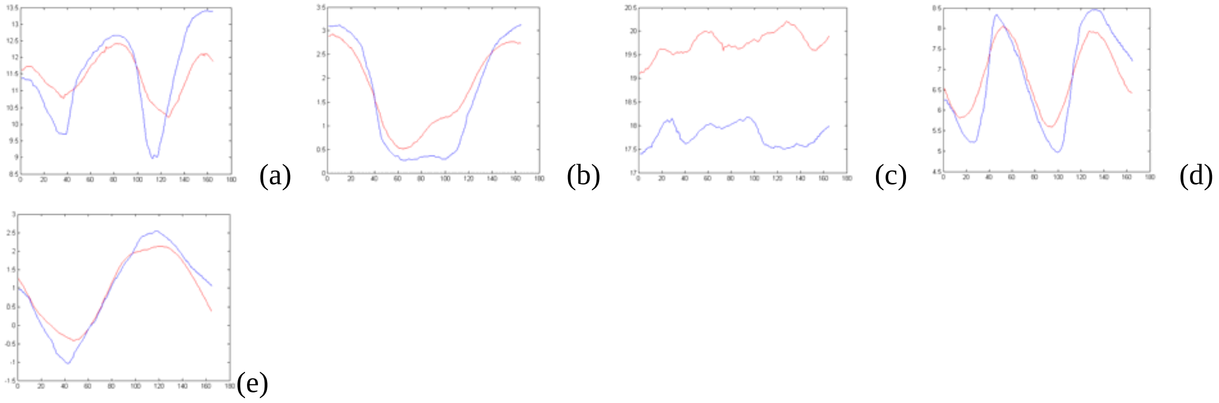


Figure 4 : The norm of projection of joined motion data on five subspaces.

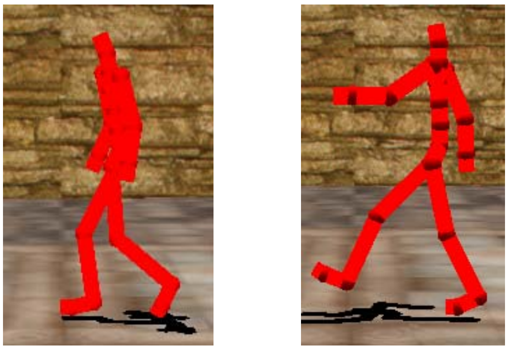
(a) original motion

(b) $\alpha = 0.15$

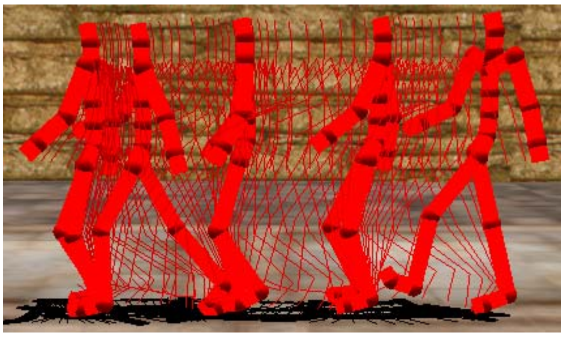
Figure 5 : Style tuning for sneak

(a) jogging and stride (b) run at a stride

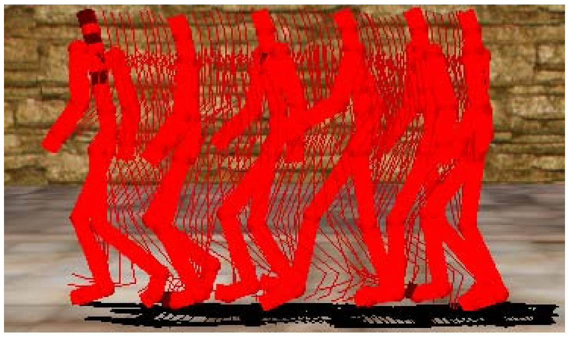
Figure 6 : Style transfer



a. brisk walk and soldier march



b. brisk soldier march ($\alpha = 0.3$)



c. brisk soldier march ($\alpha = 0.7$)

Figure 7 : Style merging

Figure 8 : Generating an old man playing the violin (right) by merging playing the violin (left) and old man waiting (middle) ($\alpha = 0.5$)