PCA-based Walking Engine using Motion Capture Data

Article · May 2004

DOI: 10.1109/CGI.2004.73 · Source: CiteSeer

PCA-based Walking Engine using Motion Capture Data

Pascal Glardon, Ronan Boulic and Daniel Thalmann
Swiss Federal Institute of Technology
Virtual Reality Lab
1015 Lausanne, Switzerland
{Pascal.Glardon, Ronan.Boulic, Daniel.Thalmann}@epfl.ch

Abstract

This paper aims to propose a novel approach to generate new generic human walking patterns using motion-captured data, leading to a real-time engine intended for virtual humans animation. The method applies the PCA (Principal Component Analysis) technique on motion data acquired by an optical system to yield a reduced dimension space where not only interpolation, but also extrapolation are possible, controlled by quantitative speed parameter values. Moreover, with proper normalization and time warping methods, the generic presented engine can produce walking motions with continuously varying human height and speed with real-time reactivity.

1. Introduction

In computer graphics, virtual humans animation is still a challenging topic due to the complexity to obtain an acceptable level of physical plausibility and controllability.

A major computer animation technique is procedural modelling that applies established an explicit analytic formulation linking high-level parameters to the animation parameters (i.e the skeleton joint values). This method requires a careful analysis of several real patterns performed with various sets of qualitative and quantitative motion characterizations.

An alternate approach aims to re-use an existing database of motions by editing them to produce new motions driven by parameters such as speed or style of motion. For example, statistical and probabilistic methods, coupled with interpolation techniques, can be applied to all database motions to generate similar motions according to database singularities such as the height of the human or the locomotion speed.

Both described techniques can benefit from motion capture, at present widely used for data acquisition [15]. Indeed, complex motions can be recorded and afterwards analyzed or re-used, so highlighting one of the main weaknesses of motion capture, namely to adapt the raw motion data to another human.

We here focus on the synthesis of real-time reactive animation using motion-captured data. Both procedural and statistical methods are powerful tools in contexts where many characters need to be animated simultaneously. However, to avoid a fastidious analysis of the high amount of large data, the second method is more appropriate, organizing the data in a manner that guarantees a low computational cost.

When applying this approach, it is important to drive the interpolation with one or more high-level parameters that not only modify it qualitatively (i.e with adverbs, ”more” or ”less” fast for example), but also quantitatively (i.e with exact scalar values). Indeed animators in general prefer to use animation engines providing a fine control over the motion.

Our goal is to develop an engine for human walking motion, one of the most common locomotion movements, which successfully interpolates and extrapolates motion-captured data with quantitative high-level parameters to generate new motion in real-time, adaptable to any kind of virtual human. We use the statistical-based PCA (Principal Component Analysis) technique to reduce the dimension of the input data, so allowing an efficient and precise extraction of the quantitative speed parameter. Moreover, this space allows to interpolate between movements at lower computational cost.

The remainder of this paper is organized as follows. First, related work is presented. Then, section 3 focuses on the pre-processing tasks necessary for our model. Next, we explain the method for generating a new walking motion applicable to a human of any height. In section 5, results and discussions are provided, and finally future work is presented.

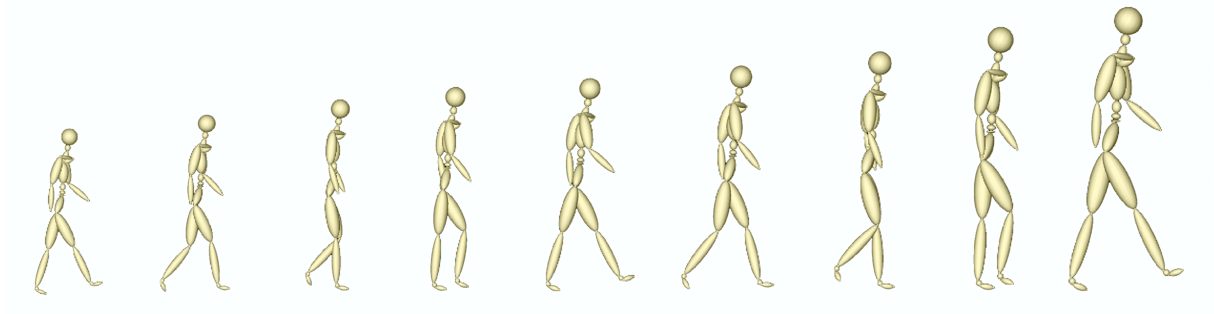

Figure 1. A walking cycle with continuously varying human height

2. Related Work

Walking motion synthesis is a central problem due to its strategic importance in virtual animation. Three main research directions, as presented in [17], can be identified among the various methods to synthesize walking movements.

Kinematics consist of specifying at each key time the corresponding key positions of the joints angles, according to biomechanical data information. This method is exploited in [5] to develop a walking engine. In spite of the efficient solution, there is no guarantee of physical realism, leading to overly mechanical movements.

The second direction aims at solving dynamical equations of motion [10, 9]. The goal is to find the correct forces and torques to apply to joints in order to reach a determined position. This approach can produce smooth results, but sometimes however impossible according to the joint limits of the human skeleton. To tackle this problem, controllers have been implemented in order to check and eventually correct the resulting posture.

The last research direction encompasses techniques to edit motion capture data. Often, interpolation methods are applied.

Ashraf and Wong [6] interpolate walking and running motions in 3D space where the axes correspond to significant parameters of a locomotion cycle. The method computes, via bilinear interpolation, a new motion between four given motions having different values for each of the three dimensions. However, this approach is limited to a small number of input motions and parameters.

In [20], multivariate interpolation on motion capture data is performed. Motions (verbs) are classified manually according to characteristics, leading to a parameter vector for each motion. Similar verbs are time-normalized using a time warping process that structurally aligns motions. A new motion is generated by applying polynomial and RBF (Radial Basis Function) interpolation between the B-Spline coefficients defining the verbs, according to a given parameter vector. Similar work is presented in [19].

Kovar and Gleicher [13] introduce a data structure, called ”registration curves”, which contains all information needed to perform blending operations on motions: timing, coordinates of the root and constraints. The transitions are achieved by simple linear interpolation between two motions.

A serious drawback of all these methods is that no extrapolation is possible. Moreover it is difficult for the user to set some parameters, as they rather qualify a motion than quantify it (for example, in the case of speed: faster, slower but not effective speed value). The method we present in this paper not only allows extrapolation, but also the quantification of parameters.

In addition to interpolation, PCA [12] has recently been introduced in animation synthesis. It is used either to compress the data [3] or to emphasize similarities between input data [4, 14] in order to generate motion according to control parameters such as age or gender. Unfortunately, these parameters do not characterize the motion in a precise way.

More recently, Troje [21] presented an approach more similar to ours, where 3D motion capture data is used and each walking sequence is decomposed into a PCA space. This produces a reduced dimension space where discriminant functions are determined to compute corresponding coefficients for a given parameter (male/female, happy/sad, relaxed/nervous). However, a disadvantage resides in that changing a stylistic parameter can modify locomotion speed. Moreover, as the data are computed in global 3D space, no retargeting on humans of different size is possible.

To avoid the drawbacks of this approach, we propose a PCA space built from angular input data, where modifying a parameter does not influence other motion characteristics.

3. Data Pre-processing

As we present a method that uses motion capture data organized in a PCA space, we first describe the acquisition process of motion data and then the creation of a reduced dimension space composed of these data.

3.1. Motion capture data process

We used a commercial optical motion capture system [22] with a set of 37 markers, illustrated in Figure 2. Regarding the convention of skeleton modelling, we use the H-ANIM standard [7] that defines a radical common denominator for the joint orientations (all the joints have the same orientation in the default posture) and a flexible topology for the hierarchical structure (any sub-set respecting the parent-child relation is valid).

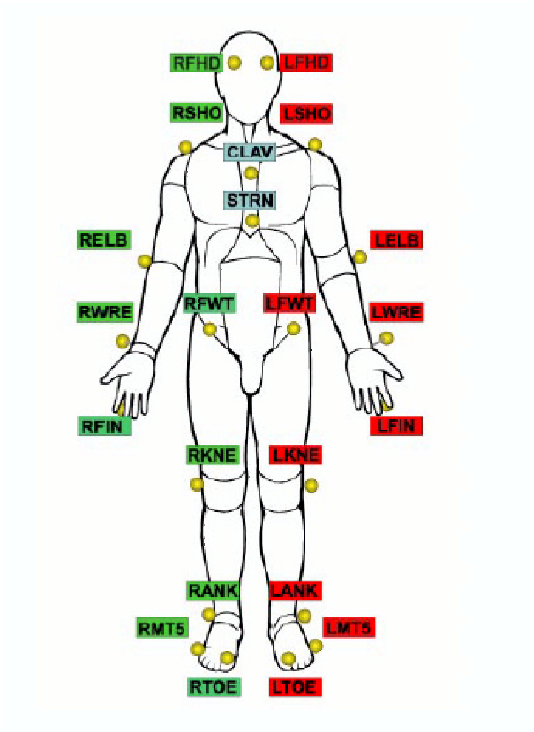

Figure 2. Front view of the marker set used

The mapping of marker positions to joint angles is achieved applying the method described in [16]. However, this conversion process has two major sources of errors:

- In the *skeleton fitting* stage the joint centers and associated local frames are constructed using vectors built from the marker positions in a reference pose featuring all the degrees of mobility [8]. The strong assumption of a marker being rigidly linked to such a local frame biases skeleton fitting and thus the deduced joint angles. A systematic analysis of local skin deformations as a function of all mobility parameters is needed to establish a complete marker model.

- The mechanical model of the skeleton is generally simplified with respect to biomechanics to ease the design and manipulation of motion. Secondary mobilities like small translations at the joint level are ignored thus introducing an error in the local position of markers. Here a joint model that better reflects anatomic shapes and deformations is needed to establish a more robust model of the relationship between the joint states and skin deformation.

Eight cameras recorded the performance of five subjects (two women and three men) walking on a treadmill. The speed of the various sequences varies from 3.0 km/h to 7.0 km/h, by increments of 0.5 km/h. The sequences were then manually segmented into cycles (one cycle includes two steps, starting at right heel strike) and converted to joint angle space and finally normalized, so that each sequence is represented by the same number of samples. To complete the database, a standing position motion sequence has been inserted to define a lower bound of the motion space.

In practice, the posture of a person in a given keyframe can be defined by the position and orientation of a root node and a vector of joint angles coded in axis-angle notation. As we capture walking motion on a treadmill, the root node contains only local oscillations of the motion. A motion of our database is then represented by an angular *motion vector* θ, which is a set of such joint angle vectors measured at regularly sampled intervals.

3.2. Creation of the PCA space

We want to benefit from the PCA technique [12] that transforms a set of variables into a smaller set of its linear combination, so as to account for most of the variance of an original dataset. Therefore, PCA spaces composed of various walking motion sequences that only differ in speed values are established, leading to one space per subject.

A PCA space is computed with a motion matrix $\mathbf{M}$ composed of various walking motion vectors θ from a specific subject v. To place the origin of the space at the center of the data set, we define θ_0 as the average vector of all n motion vectors of $\mathbf{M}$. The basis vectors describing this space are the m first orthogonal PC's (Principal Components) necessary to calculate an approximation of the original data. Let $\overrightarrow{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_m)$ be a coefficient vector and $\mathbf{E} = (\overrightarrow{e_1}, \overrightarrow{e_2}, \dots, \overrightarrow{e_m})$ a vector matrix of the first PC's (or eigenvectors) of $\mathbf{M}$, a walking motion θ can be expressed as:

$$\theta \cong \theta_0 + \sum_{i=1}^{m} \alpha_i \overrightarrow{e_i} = \theta_0 + \overrightarrow{\alpha}\mathbf{E} \qquad (1)$$

4. Walking model

Our presented walking engine is driven by two high-level parameters: speed and human size. First, interpolation and extrapolation are achieved in the PCA spaces of each subject to generate new motion according to a speed value. Secondly, a time warping method allows to handle the human height parameter.

4.1. Motion interpolation and extrapolation

As the walking motions that compose a PCA space differ only at the speed parameter level, its PC's tend to express the most variance between slow and fast motions. Therefore, a relationship between a coefficient vector $\overrightarrow{\alpha}$ and its corresponding speed value is determined in order to allow motion interpolation and extrapolation.

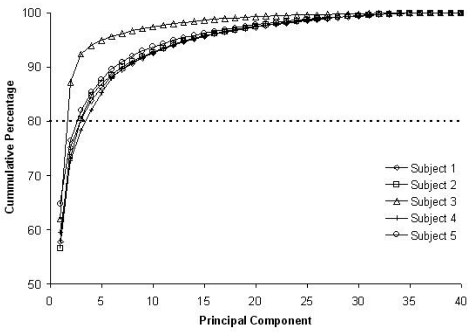

Figure 3. The cumulative percentage of the PC's for the five subjects of our database

First the number m of PC's that significantly influence the motion is specified by taking the first PC's representing 80% of the motion information. Indeed, as Figure 3 illustrates, the contribution of the PC's beyond this percentage value is small compared to the first m, and probably does not provide a relevant relation to speed values. Moreover, this percentage represents an important dimension reduction factor of the space, appropriated for our real-time specifics.

Then, the various captured speed values are compared to their corresponding coefficient vectors $\overrightarrow{\alpha}$, for each dimension of the PCA space. The resulting graphs, depicted in Figure 4 and Figure 5, clearly show a linear relationship between motions having speed values greater than zero, whereas the standing posture cancels this linearity. The fitting of a polynomial curve to these motions is unfortunately unadapted in practice, especially for motion extrapolation. Hence two linear functions are constructed and attached to each dimension.

The first function is a linear least-squares fit performed on the pairs of coefficient vectors $\overrightarrow{\alpha}$ and their corresponding speed values S, excluding the speed value zero. Thus, for a given i-PC, a linear approximation function $A_i(S) =$

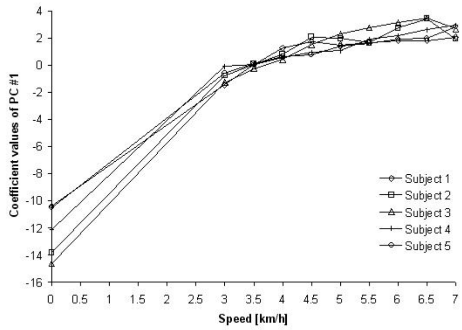

Figure 4. Comparison between the speed values and the coefficient values of the first PC, for five subjects

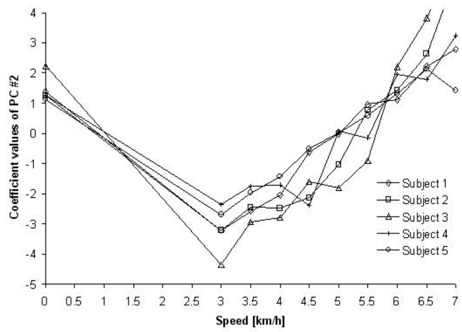

Figure 5. Comparison between the speed values and the coefficient values of the second PC, for five subjects

$A_i = m_i S + b_i$ over the j increasing speed values is computed by minimizing the sum of the square distances between the actual coefficient values α_{ij} and the approximated coefficient value d_{ij}, as the following equation describes:

$$\sum_{j=1}^{nbS}(\alpha_{ij} - dij)^2 = \sum_{j=1}^{nbS}(\alpha_{ij} - m_i S_i - b_i)^2 \qquad (2)$$

where nbS is the number of speed values. The second function is a simple linear interpolation between the coefficient vector at null speed value α_{i1} and the function A_i evaluated at the minimal captured speed value ($j = 2$). Figure 6 illustrates the result of the two fitting functions on a subject.

Therefore, applying the two described approximation functions, it is now possible not only to interpolate, but also to extrapolate walking motions for a given speed value. Actually, these functions return a coefficient vector $\overrightarrow{\alpha}$ that has to be inserted into Eq. 1 to compute the new generated walking motion.

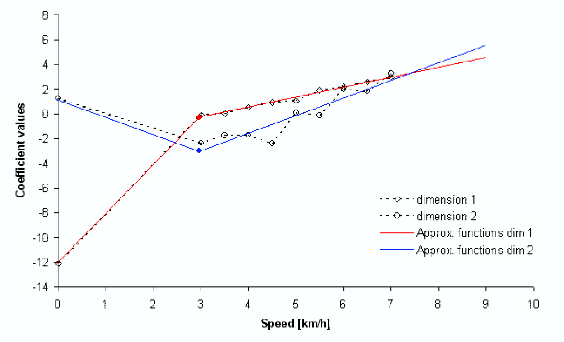

Figure 6. The approximation functions (linear interpolation and A_i) relative to the first two PC

Linear interpolation between coefficient vectors $\overrightarrow{\alpha}$ implies linear interpolation between axis-angles, which is incorrect in most of the case. Alexa in [2] solves this problem for general cases by a subdivision of the transformation matrices that have to be linear combined. In our situation, we do not need to apply this subdivision scheme, as the posture (i.e the transformation matrices) in the different walking motions are very near. For the standing posture, which is quite different from other postures, their transformation matrices are the identity matrices and therefore allows linear interpolation.

4.2. Normalization and time warping

The generalization of the heterogeneous input data we used is an important aspect to produce animation adaptable to any kind of virtual humans. Indeed, we capture various subjects, not only with differences in the style of motion, but also, and more importantly, differences in height. Murray [18] has shown that all the leg-relative angles in the sagittal plane (hip, knee, ankle) show very similar trajectories for all adult men for the same value of normalized speed V, obtained by dividing the walking velocity v (in m/s) by the hip joint height H (i.e. the leg length in meters), as Eq. 3 describes:

$$V = \frac{v}{H} \tag{3}$$

The normalization of motion is achieved simultaneously when the time warping method is performed. Every input locomotion cycle and also each generated cycle contain a fixed number of frames due to the normalization step during pre-processing. The induced time warp is handled using a walking cycle frequency function that links a given normalized speed to a cycle frequency [11]. This function is called the Inman law and is defined by Eq. 4.

$$f = 0.743\sqrt{V} \tag{4}$$

We adapted this law to our observations, performed on a treadmill. We fit the data to an approximation function of the form ax^b, similar to the Inman law. Figure 7 illustrates the evolution of the frequencies with respect to normalized speed, for the five subjects captured during a walking motion. The resulting frequency function is described in Eq. 6.

Therefore, the animation engine is able to continuously vary the speed and compute the phase update as in the walking engine in [5]. In short, the phase variation $\Delta\varphi$ for a given duration Δt (update in seconds) is given by Eq. 5:

$$\Delta\varphi = \frac{\Delta t}{cycleDuration} = \Delta t f \tag{5}$$

where f is obtained from the instantaneous normalized velocity through Eq. 6. Hence, our time warping method is a function of speed value, normalized according to a given human height.

$$f = 0.85x^{0.4} \tag{6}$$

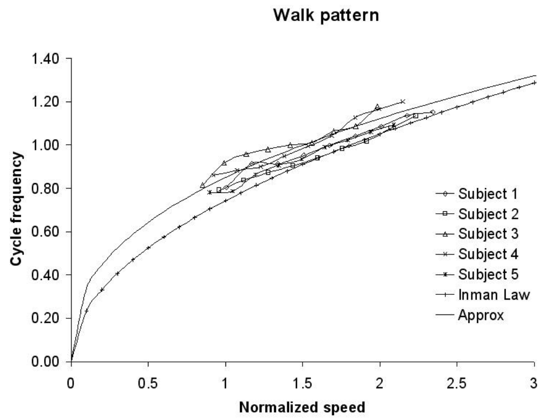

Figure 7. Frequency functions

5. Results and discussion

We create PCA spaces as described in section 3.2 for the five subjects we captured. The input data enclose 40 sequences (nine different walking motions and one standing posture, in four exemplars) recorded at 120 Hz and normalized, using SLERP based on quaternion interpolation, to 100 frames containing animations of 78 DOF's. Therefore, each PCA space has maximal 40 PC's and represents up to 3MB of input data.

To represent 80% of the total information data, between two and four PC's were necessary, depending on the subject as summarized in Table 1. Additionally, approximation

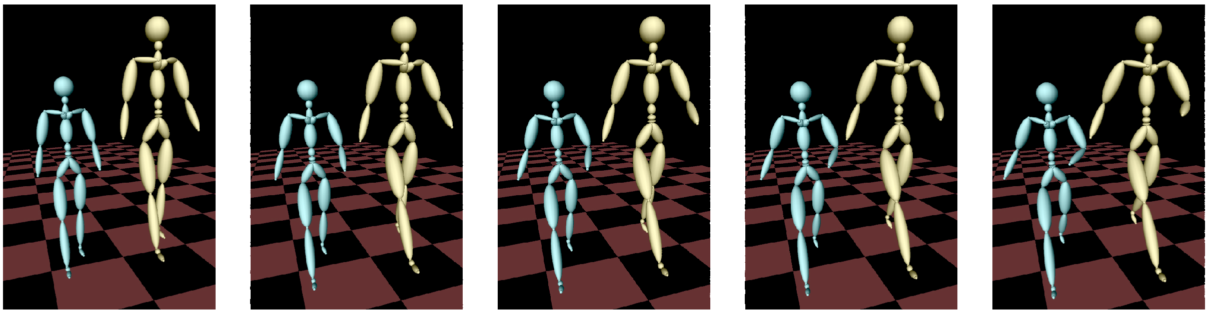

Figure 8. Postures at identical moment in the walk cycle, with two different skeleton sizes. From left to right: 2.0, 4.0, 6.0, 8.0 and 10.0 [km/h]

functions that fit speed to a coefficient vector are computed, as well as the frequency functions. All the above described operations are computed in a pre-processing phase.

For a given speed, our method computes the corresponding entire walking cycle which will be looped. Despite of the fact that no cyclification method [1] has been performed on our captured sequences, the animation is smooth because of two important statements in our methodology. First, the approximation functions return a coefficient vector for a given speed representing a mean motion over all captured motions at this speed. Secondly, the PCA algorithm has a filter effect on the data, as only the most important variation of the motions has been taken into account.

The global translation of the root node is added according to the desired speed, and the time un-warping is performed using the normalized speed with the given leg length, and by applying the frequency function described in Eq. 6

Table 1. PCA space information for each subject

Subject ID	#PCS's for 80% of whole data	Reduction factor
1	3	13
2	3	13
3	2	20
4	4	10
5	3	13

In the executing phase, a human skeleton based on the H-ANIM hierarchy is animated according to the high-level parameter values that the user can change in real-time: speed and leg length. The updating of the motion is interactive, with a computation time of 0.8 milliseconds on average on a CPU 1.8 GHz machine.

We can extrapolate and interpolate realistic and pleasant walking motions from 0 km/h up to 10 km/h, as shown in Figure 8, in spite of the fact that motions are only represented by few dimensions. However, beyond 10 km/h, undesired behavior occurs at the skeleton level, especially unrealistic movements at the level of the leg. Another effect is that the double support phase is no longer ensured.

Scaling the human directly influences the cycle frequency in order to preserve motion properties. In the case of two humans of different heights, the smaller one walks with a higher frequency for an identical motion speed, as shown in the video attached to this paper.

The method we developed in this paper presents many advantages. First of all, we use the full potential of PCA, by reducing data dimension and by the fact that the first PCs induce the most changes in the dataset. Our method does not require lengthy pre-processing (10 seconds for five subjects on a CPU 1.8 GHz machine). Then, applying approximation functions, new generated motions can be interpolated or extrapolated with a quantitative speed parameter. Finally, our generic method is applicable to any human height that can be modified in real-time. As any update of these high-level parameters is performed in real-time (0.8 milliseconds), the method allows to drive multiple virtual humans.

6. Future work

The method presented in this paper describes the creation of a walking engine for a specific captured subject. We want to integrate all subjects within a global structure allowing to parameterize the style of the motion.

An important difficulty in the development of human animation patterns resides in the sliding effect of the feet. Our implementation tries to limit this effect by constructing adapted frequency functions, but still presents some artifacts when the speed value nears null or takes on values well beyond those measured during motion capture. Therefore, we intend to refine our frequency functions, by introducing linearity when the speed value exceeds a boundary value that we have yet to determine. Moreover, animation of walking pattern at low speed can be improved by applying other functions as linear interpolation between the standing posture and the walking sequence with minimal speed.

Finally we also want to apply our method to other locomotion types, and especially by increasing our database with running motions to observe the adjustments needed for the time warping process and for motion transitions.

7. Acknowledgment

This research has been funded by the Federal Office for Education and Science in the framework of the European project MELIES (IST-2000-28700).

We wish to thank Tom Molet for motion capture support and Olivier Renault for video post-production.

References

[1] A. Ahmed, F. Mokhtarian, and A. Hilton. Cyclification of human motion for animation synthesis. In *Short Paper Proceedings of Eurographics 2003*, sept 2003.

[2] M. Alexa. Linear combination of transformations. In S. Spencer, editor, *Proceedings of the 29th Conference on Computer Graphics and Interactive Techniques (SIGGRAPH-02)*, volume 21, 3 of *ACM Transactions on Graphics*, pages 380–387, New York, July 21–25 2002. ACM Press.

[3] M. Alexa and W. Müller. Representing animations by principal components. In M. Gross and F. R. A. Hopgood, editors, *Computer Graphics Forum (Eurographics 2000)*, volume 19(3), pages 411–426, 2000.

[4] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In A. Rockwood, editor, *Proceedings of the Conference on Computer Graphics (SIGGRAPH99)*, pages 187–194, N.Y., Aug. 1999. ACM Press.

[5] R. Boulic, N. M. Thalmann, and D. Thalmann. A global human walking model with real-time kinematic personification. *The Visual Computer*, 6(6):344–358, Dec. 1990.

[6] A. Golam and K. C. Wong. Dynamic time warp based framespace interpolation for motion editing. In *Graphics Interface*, pages 45–52, May 2000.

[7] H-ANIM. Humanoid animation working group. www.hanim.org, 2004.

[8] L. Herda, P. Fua, R. Plänkers, D., R. Boulic, and D. Thalmann. Using skeleton-based tracking to increase the reliability of optical motion capture. *Human Movement Science*, 20:313–341, 2001.

[9] J. K. Hodgins. Three-dimensional human running. In *Proceeding of the IEEE Conference on Robotics and Automation*, April 1996.

[10] J. K. Hodgins, W. L. Wooten, D. C. Brogan, and J. F. O'Brien. Animating human athletics. In R. Cook, editor, *SIGGRAPH 95 Conference Proceedings*, Annual Conference Series, pages 71–78. ACM SIGGRAPH, Addison Wesley, Aug. 1995. held in Los Angeles, California, 06-11 August 1995.

[11] V. Inman, H. Ralston, and F. Todd. *Human Walking*. Wiliams & Witkins, Baltimore, 1981.

[12] I. T. Jolliffe. *Principal Component Analysis*. Springer series in statistics. Springer-Verlag, 1986.

[13] L. Kovar and M. Gleicher. Flexible automatic motion blending with registration curves. In D. Breen and M. Lin, editors, *Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA-03)*, pages 214–224, Aire-la-Ville, July 2003. Eurographics Association.

[14] I. S. Lim and D. Thalmann. Construction of animation models out of captured data. In *Proc. IEEE Conf. Multimedia and Expo*, Aug. 2002.

[15] A. Menache. *Understanding Motion Capture for Computer Animation and Video Games*. Academic Press, San Diego, 2000.

[16] T. Molet, R. Boulic, and D. Thalmann. Human motion capture driven by orientation measurements. In *Presence, MIT*, volume 8, pages 187–203, 1999.

[17] F. Multon, L. France, M.-P. Cani-Gascuel, and G. Debunne. Computer animation of human walking: a survey. *The Journal of Visualization and Computer Animation*, 10(1):39–54, 1999.

[18] M. P. Murray. Gait as a total pattern of movement. *Am J Phys Med*, 46(1):290–333, 1967.

[19] S. I. Park, H. J. Shin, and S. Y. Shin. On-line locomotion generation based on motion blending. In *ACM Symposium on Computer Animation 2002*, 2002.

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

[21] N. Troje. Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, (2):371–387, 2002.

[22] Vicon^{TM}. Vicon motion systems. www.vicon.com, 2004.