Human motion and emotion parameterization

Michal Kraus*
Department of Computer Science
Czech Technical University
Prague / Czech Republic

Abstract

This paper describes the methods and data structures used for generating new human motions out of existing motion data. These existing animation data are derived from live motion capture or produced through the traditional animation tools. We present a survey of techniques for generating new animations by physical and emotional parameterizations. Discussed new animations are generated by Fourier transformations, transformations of velocity and amplitude and scattered data interpolation in this report.

Keywords: Human animations, motion capture, motion control, motion parameterization, motion synthesis

1 Introduction

As we know, virtual human characters use has been spreading widely in recent years. These characters should be moving and we want them to move realistically. Real human motion is usually quite dependent on the emotional state of the person, performing the move. Although the influence of emotion is obviously projected to the kinematic parameters (speed, range) of the motion, it is quite difficult if not impossible to describe it in terms of these parameters. Thus, we can decide to consider the human emotional state (happiness, tiredness) to be just another parameter.

This decision becomes important if we want to apply some motion synthesis technique. Virtual humans can be animated using various approaches from keyframing to motion capture, but in each case, we have to create (record) each particular motion. The need for wide variety of animations of the same class (varying step-length of character walking in uneven terrain) leads to the situation when motion capturing or animating is just a waste of resources and another way is sought after. The motion synthesis (or parameterization) is a process of creating new animation varying in a specific parameter. This parameter can be the above mentioned step-length as well as emotional influence elusive by nature. In motion parameterization process we typically create new animations out of some basic set of animations representing typical examples of our desired parameters. If these parameters are of the emotional kind, this set should be a set of recorded motions that we label as influenced by particular emotion.

This paper is a beginning of our master thesis research aimed at the motion parameterization. We have done a survey for the available methods with focus at the specifics of parameterizing the emotional parameters. When we mention an emotion, we are interested in the effect of emotional state on human motion. Therefore, we have neglected the facial expression that definitely indicates an emotional state, but has no significant effect on motion parameters. In Section 2 and 3, we have introduced the prerequisites for any motion synthesis, data representation of animation and humanoid. Section 4 describes motion synthesis problem in general. Section 5, mentions briefly related sub problems of animation data reduction. A few special selected methods of motion synthesis are presented in section 6. In section 7, the methods mentioned in section 6 are compared. Section 8 refers to our future work.

2 Representation of humanoid

Virtual humanoids made by segments are used in our discussed animations. These segments are connected together by joints. These joints are arranged to a special form of a hierarchy. The joint, which contains position of the humanoid in the global coordinate system, is a top of the hierarchy. In the different animations, a different number of the joints is used. The number of joints depends on the complexity of animation.

3 Motion data

Animation can be described by a set of motion curves [1]. These curves give the value of the human's position and rotation of all its joints represented as a function of time. Animation curves are represented by a vector $\theta^{(1)}(t) \ldots \theta^{(k)}(t)$. In case of the motion described by one motion curve (for example rotation of one joint), we can define this motion by its shape vector $S = (\theta_1, \ldots, \theta_n) \in R^n$ and time vector $T = (t_1, \ldots, t_n) \in R^n$. Generally, the animation is represented by concatenating single curves represented as vectors $S = (\theta_1^{(1)}, \ldots, \theta_n^{(1)}, \ldots, \theta_1^{(k)}, \ldots, \theta_n^{(k)}) \in R^{nk}$ and $T = (t_1^{(1)}, \ldots, t_n^{(1)}, \ldots, t_1^{(k)}, \ldots, t_n^{(k)}) \in R^{nk}$.

4 Motion synthesis

Given a set of distinct example motions of the same class, we can synthesize new motions under following conditions. The example motions must be of the same

* krausm1@fel.cvut.cz

time scale and must have the same number of keyframes. This means it is necessary to compute the missing frames in the source animation. The joint structure must be the same (same joints and same number of them). This means that the motion data are represented by a vector of the same dimension. Each of the example motions must be labeled with a parameter specific for it. For instance, if our class of motions is one step of human gait, we can have a step that is labeled as an example of long step, or a low step, etc.

We can synthesize new motions with different value of the selected parameter by interpolating between the example and average motion. We can also create linear combinations of such motions to express the combination of parameters. A good example of this scattered data interpolation approach can be found at [5][6]. This approach is discussed in section 6.3.

If we label our examples with emotional parameters ("happy" or "tired" walk) we can do the same, thus obtaining the animations of "happier" or "more tired" walk. We can do more with emotional parameters. We can obtain the "essence" of the parameter that can be applied to another motion of different class. Several research papers address this problem [3][4]. Some of these approaches are presented in closer detail in sections 6.1 and 6.2.

5 Reduction of motion data

Data reduction methods, which are mentioned in this section, are not necessary for our discussed parameterizations. However, with respect to a time-consuming processing of the proposed transformations, it is very useful to make reduction of animation data. This leads to smaller storage and faster computation, especially in the real-time applications.

We can do it by a curve simplification on the level of single animation [2] or by algorithms, which choose representative set of motion data [8][9][10].

5.1 Curve simplification

Given a curve as a chain of line segments, a simplification algorithm generates an approximation of it represented by a smaller number of vertices (Figure 1).

For this approximation, we can choose Lowe's algorithm [7]. The algorithm approximates the chain by selecting vertices, which can be set out of this chain. In the beginning, the algorithm connects two endpoints of the chain. Then this approximation is tested by a distance criterion. If the maximum deviation of the farthest point is greater than the distance criterion, the line segment is sub-divided into two segments at the curve point most distant from the line segment. This step is recursively repeated until the distance criterion is satisfied.

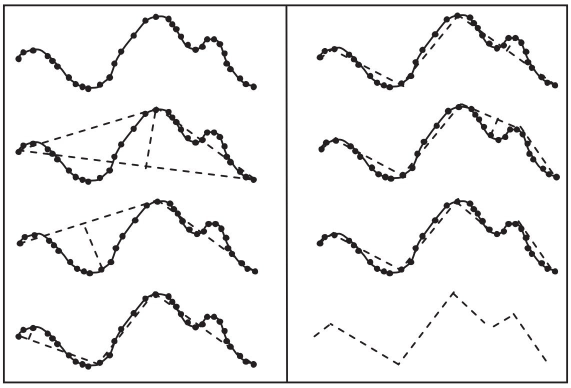


Figure 1: Curve simplification

5.2 Principal component analysis

By Principal component analysis (PCA)[8][9][10] we can choose representative set of the motion data. Basic idea is that PCA takes average (M_{avg}) of motion data instances M_i, i = 1..k and then counts differences out of the motion data instances and average. The instances, which are quite similar to average, we can leave out(1).

$$(1)\quad M_{avg} = \frac{1}{k}\sum_{i=1}^{k} M_i \qquad \Delta M_i = M_i - M_{avg}$$

, where k is a number of instances.

If $\Delta M_i < M_{\lim it}$ then ΔM instance can be vanished. M_{limit} is a distance criterion.

6 Methods for controlling motions by parameters

Let us to introduce methods for generating and controlling motions in this section. With respect to a fact that we are going to be interested in emotional parameterizations of animations in our future work, we have selected methods, which provide possibilities of application of extracted emotional part of the motion. In section 6.3 more general method of parameterization is described.

6.1 Fourier principles for human emotion-based motions

Unuma, Anjyo a Takeuchi, have introduced method [3], which uses Fourier principles in order to make new emotion-based motions. This method is based on Fourier series application to animation data using interpolations and extrapolations in order to make a new animation. Authors have experimented with emotions, step length and others during walk.

6.1.1 Fourier series

The series, which are expanded to sine or cosine functions, are called Fourier series (2).

(2) $f(t)=\frac{a_0}{2}+\sum_{k=1}^{\infty}(a_k\cos k\omega t+b_k\sin k\omega t)$,

where a_k, b_k are real or complex numbers, $\omega>0$, t is time

$$a_k=\frac{2}{T}\int_0^T f(t)\cdot\cos k\omega t dt$$

$$b_k=\frac{2}{T}\int_0^T f(t)\cdot\sin k\omega t\, dt \quad , \text{ where } k = 0,1 \ \ldots$$

$T=\frac{2\pi}{\omega}$, where T is a period of motion

In method, which is proposed in this section, the special case of Fourier series is used. Equation 3 defines this special case.

(3) $f(t)=\sum_{k=1}^{\infty}b_k\sin k\omega t$

$$b_k=\frac{4}{T}\int_0^{\frac{T}{2}} f(t)\cdot\sin k\omega t \quad \text{and} \quad a_k = 0$$

Motion data are represented in the same way as in the section 3. The rotation of the m-th joint is represented by vector $\theta^m{}_x(i\Delta t)$, $\theta^m{}_y(i\Delta t)$, $\theta^m{}_z(i\Delta t,)) \mid i = 1..n\}$, where Δt is a measuring interval of the motion capture system.

6.1.2 Fourier approximation

Based on the discrete data of the m-th joint angle, let us first construct a functional model with continuous parameter of time. This functional model describes motion of each joint in time. Then the motion (for example "normal" walk) is normalized to 2π period. Thus, we get the following expression (4), which we call a rescaled Fourier functional model:

(4) $\theta_m(t)=A_{m0}+\sum_{n\geq 1}A_{mn}\sin(nt+\phi_{mn})$

Based on the discrete animation data of the joint angles for another periodic locomotion (for example "tired" walk), we get expression (5):

(5) $\prod_m(t)=B_{m0}+\sum_{n\geq 1}B_{mn}\sin(nt+\psi_{mn})$

6.1.3 Interpolation and extrapolation

New animation is created by interpolation and extrapolation between the motions θ_m (t) a $\prod_m$ (t) ("normal" and "tired" walk). This new animation is defined by equation number 6. There is an important variable s, which defines a smooth transition between the both motions. If 0<s<1 we see "a little tired walk" (Figure 2c). In this case, the motion is interpolated. If s>1, the motion is amplified and extrapolated (Figure 2d). On the other hand, if s < 0, the model looks rather brisk (Figure 2e).

(6)

$$\Xi_m(s,t)=\{(1-s)\cdot A_{m0}+s\cdot B_{m0}\ \}+$$
$$\sum_{n\geq 1}\{(1-s)\cdot A_{m0}+s\cdot B_{m0}\ \}\sin\{nt+(1-s)\phi_{mn}+s\psi_{mn}\ \}$$

Examples:

a) s = 0 => $\Xi_m(0,t)=\theta_m(t)$ => "normal walk"
b) s = 0.5 => "a little tired" walk
c) s = 1 => $\Xi_m(1,t)=\Pi_m(t)$ => "tired walk"
d) s > 1 => "a very tired walk"
e) s<0 => "walk seems to be a little brisk"

Using Fourier approximations, we can also control a kinematical aspects of the human locomotion, such as a step length, a speed etc. (Figure 3)

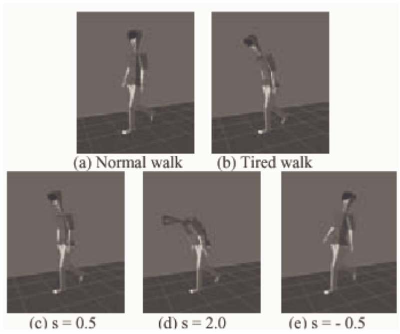


Figure 2: Interpolation and extrapolation of "normal" and "tired" walk (from [3])

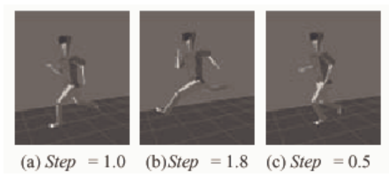


Figure 3: "Step" effect (from [3])

6.1.4 Separated emotional part application and superposition of human behaviors

Another possibility how to make use of Fourier principles is explained in this section. This method separates emotional part of the motion in order to apply it to the another motion. The source motions are described by the rescaled Fourier models (4)(5). Then the emotion is defined by Fourier coefficients (7), which are extracted from the source motions (4)(5). We extract "tiredness" in this case. Now, we can create rescaled Fourier model out of Fourier coefficients (8).

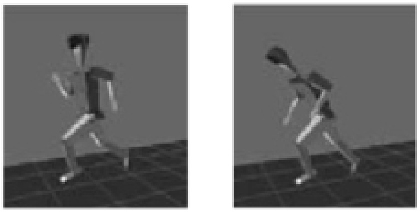
(7) $C_{mn} = A_{mn} - B_{mn}$

$$\lambda_{mn}=\phi_{mn}-\psi_{mn}$$

$$\Psi_m(t) = C_{m0} + \sum_{n \geq 1} C_{mn} \sin(nt + \lambda_{mn}) \quad (8)$$

$$\Omega_m(t) = D_{m0} + \sum_{n \geq 1} D_{mn} \sin(nt + \upsilon_{mn}) \quad (9)$$

We have two Fourier characteristic functions $\Psi_m(t)$ in (8) and $\Omega_m(t)$ in (9). $\Omega_m(t)$ was obtained out of the new animation data set (for example "normal" run). Using interpolation and extrapolation of $\Psi_m(t)$ and $\Omega_m(t)$, as we have already mentioned in section 5.1.5, we obtain a superposition of human behaviors. In this case, we obtain "tired run". Figure 4 shows examples of this technique.



a) a „normal" run b) a „tired" run

Figure 4: Result of separated emotional part application (from [3])

This technique shows possibilities of creating emotional-based motions out of animation data set. This is performed by using the interpolation and the extrapolation of the exemplary motions, which are represented by Fourier series.

6.2 Transformation of separated emotional part

Bruderin, Amaya and Calvert [4] have proposed another method, which describes separating of the emotional part of the human locomotion. This separated emotion is obtained as a difference between "neutral" and emotional-based animation. Finally, the emotion can be applied to an existing "neutral" animation of different class of motion. It was discovered by measurement that the emotion is characterized by two parameters. These parameters are the intensity and the velocity, where the intensity is characterized by amplitude and the velocity is specified by the time (Figure 5).

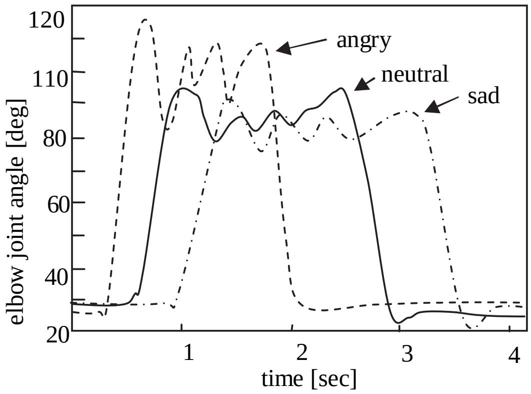


Figure 5: Knocking on the door with the different emotions

6.2.1 Transformation of speed

In this section we explain a method for speed transformation that is demonstrated on an example motion (emotional and normal) performed by actor when drinking a beverage. With respect to a fact, that we are using different motions, it is necessary to subdivide both the neutral and emotional motion into units of motion, called "basic periods". For example, we can subdivide such a motion into periods of "hand to cup", "cup to mouth", "cup down" and "hand back" (Figure 6). The first step to obtain the speed transformation is to get the absolute speed of the end effector. Then the data are normalized. This ensures the same duration for all animations. The next step is to get the distribution of frames along the trajectory (Figure 7). Then, the speed transformation is applied to another, neutral movement (possibly even of different class) by calculation of distribution of its frames, followed by division into basic period. For all basic periods of each motion the steps mentioned above are performed. The distribution of frames of an original and a new emotion animation is used as a correspondence between the original neutral and new emotional movement.

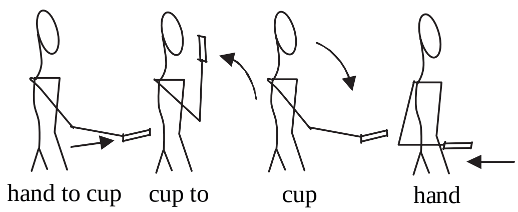


Figure 6: Drinking motion

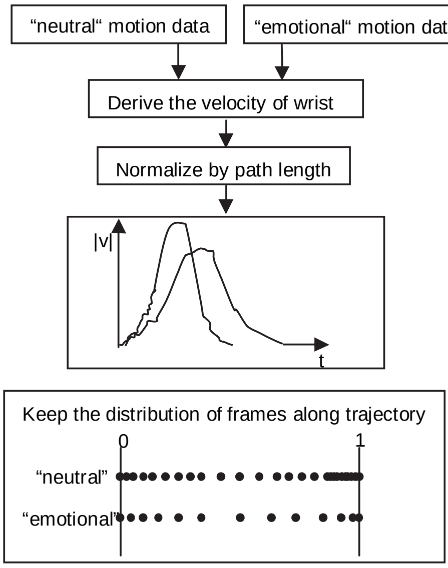


Figure 7: Algorithm to obtain speed transformation

6.2.2 Transformation of amplitude

At first, the joints are divided into four categories according to the levels of hierarchy of humanoid. This division is necessary because the ranges of motion for the joints are different in each category (Figure 8).

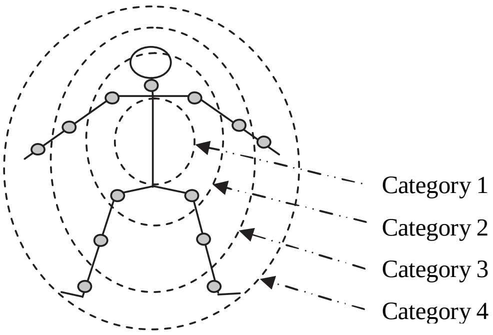


Figure 8: Joint categories

Each of these categories represents a multidimensional space, Θ^{i}_{NorE} that is defined by its joint angles and time t. The subscript N or E denotes “normal” or “emotional” motion. The extraction of the amplitude is provided by the factors d_N^{i} and d_E^{i}. Each of these factors represent the maximum distance from the straight line, which connects the start and the end of the selected basic period , and the curve that represents the time-joint angles space of each category (Figure 9).

The next step is to get ratio of the distance $d^i = d^i_E / d^i_N$. Then the ratio of the distance is applied to a new neutral motion $\Theta_{new}^{ij}(t)$, where j is corresponding basic period (Figure 10). The result is denoted as $\Theta_{gen}^{ij}(t)$.

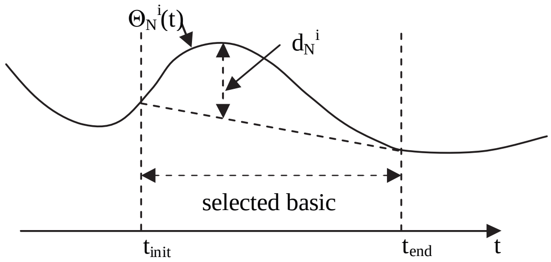


Figure 9: Intensity factor of spatial amplitude

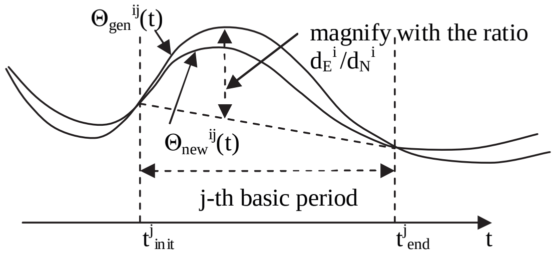


Figure 10: Transformation of spatial amplitude

The basic idea of this method is an extraction of velocity and amplitudes out of the captured neutral and motional movement. Then these separated parts are applied to a new neutral motion. The results of this method provide Figure 11.

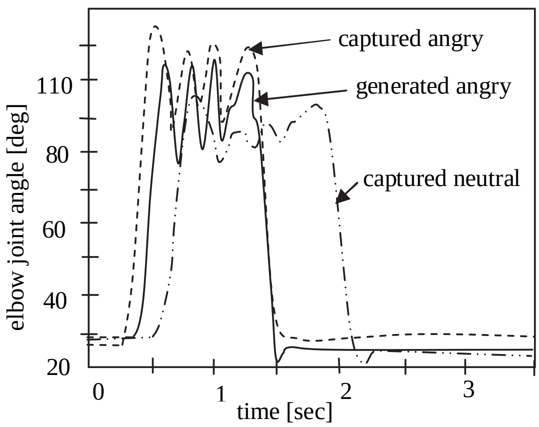


Figure 11: Comparison between generated and captured elbow joint-angles

6.3 Parameterization using scattered interpolation

In this part a method for constructing parametric animations out of captured motion and skeleton data is described. This construction is provided by scattered data interpolation.

6.3.1 Motion data

A set of example motions M_i from different individuals are used to construct a parametric animation. Each of example motion is characterized by its parameter p_i. This parameter provides a set of properties of motion M_i. Mentioned parameter includes age, gender, height and weight, whose are invariant for each individual. Moreover, it also includes physical and emotional states, whose are varying for each individual. Thus complete input motion data are represented by vectors $\{(M_i, p_i, t_i)\}^n_{i=1}$, where each of motion data M_i consists a human's position and a set of each joint's orientation in time. Unit quaternions are used to represent orientations of joints. Vector t_i represents key-times and n is a number of instances. With respect to a fact that the set of input motion data represents a multidimensional space it is suitable to make reduction of these example data. This can by done by PCA, which we have already mentioned in section 5.2.
Because of the motion data are not structurally similar it is necessary to perform a time reparameterization. This means for example, in walking, that the time when the left foot strikes the ground is different in the exemplar data set. Thus, each set of key-times t_i is transformed to a generic time $t' \in [0,1]$.

6.3.2 Scattered data interpolation

The construction of parametric animation is formulated as a scattered interpolation problem in this method. For this interpolation problem we can use radial basis function (RBF) [11] or a network consisting of Gaussians [12].
Finally, new animations are generated by interpolations functions, which map parameters from multi-dimensional parameter space of age, gender, height etc. to space of motion and skeleton data.

Thalmann and Rose [5][6] have provided a general method of motion parameterization. This method is specific in the fact that it uses not only motional data parameters, but skeleton parameters such as weight, age etc. are also used.

7 Conclusion

We summarized methods for generating new realistic animations from captured motion data in this paper. We can create these animations by parameterization of the separated parameters of motion or by interpolations among the example motions. The motion syntheses, which use separated parameters, are represented by Fourier principles and transformation of velocity and amplitude in this report. Although, these methods seem to be quite similar, there are some differences. The main difference is that the method, which uses Fourier principles, is suitable for generating only periodic movements. The reason is that Fourier series are represented by periodic functions sine and cosine. The other difference is that the technique, which makes transformation of amplitude and velocity works with discrete motion data whereas Fourier principles work only with a functional model of animation data with a continuous parameter of time.
We also discussed the more general method in this paper. This method creates new animations by interpolations among the examples motions in the parametric space. This technique was formulated as a scattered interpolation problem. The method provides general parametric model. There are no separated parameters in contradistinction to methods mentioned above in this method. The main idea of this method is that the existing motions are mapped to space of parameters. We also discussed methods of the data representation and reduction in this report. The data reduction methods lead to smaller data storage and a faster computation of mentioned parameterizations.

8 Future work

This survey of the human motion and emotion parameterization methods was performed as a beginning of our master thesis research. Now, we are going to use acquired knowledge from this work in implementation of our own approach. This approach will be very close to method, which we have already mentioned in the section 6.3. The main idea is that the labeled animations of the same class of the movement can be treated as scattered points in the multidimensional parametric space. Then the new animations are generated by linear combinations within the parametric space. The main advantage is that we can use both, the emotional and kinematic parameters, in this approach.
In particular, we might possibly apply this idea in virtual teaching of tennis [13] in the future. For example, we will have an animation set of strokes of forehand. Each of these forehand animations will be labeled as an example of certain feature of stroke (typical mistake, special technique). Finally, we will simulate particular trainee's performance as linear combinations of these examples. Other applications that would consider emotional parameters are certainly possible.
The question that appears is how to measure quality of our generated animations. We are going to judge the visual quality by subjective appreciation of observers.

Acknowledgements

I would like to thank to Vladimír Štěpán for his patience and useful advice, which helped me to make this paper.

References

[1] S. Lim and D. Thalmann, *A vector-space representation of motion data for example-based motion synthesis*, In Proceedings of the IFIP TC5/WG5.10 DEFORM'2000 Workshop and AVATARS'2000 Workshop on Deformable Avatars citation, pages 169-179, 2000

[2] S. Lim and D. Thalmann, *Key-posture extraction out of human motion data by curve simplification,*, In Proceedings of 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBC 2001, pages 1167-1169, 2001

[3] M. Unuma and K. Anjyo, *Fourier Principles for Emotion-based Human Figure Animation*, In Proceedings of SIGGRAPH 1995, pages 91-96, 1995

[4] K. Amaya T. Calvert and A. Bruderlin, *Emotion from Motion*, In Proceedings of SIGGRAPH 1995, pages 97-104, 1995

[5] Ch. Rose, B. Bodenheimer and M.F.Cohen, *Multidimensional Motion Interpolation Using Radial Basis Functions*, IEEE Computer Graphics and Applications, pages 32-40, 1998

[6] S. Lim and D. Thalmann, *Construction of animation models out of captured data*, In Proceedings of IEEE International Conference on Multimedia and Expo (ICME 2002), pages 829-832, 2002

[7] D. G.Lowe, *Three-dimensional object recognition from single two-dimensional images*, Artificial intelligence, pages 355-397, 1987

[8] M. Alexa and W. Muller, *Representing animations by principal components*, In Proceedings of EUROGRAPHICS 2000, pages 411-418, 2000

[9] V. Blanz and T. Vetter, *A morphable model for the synthesis of 3d faces*, In Proceedings of SIGGRAPH 1999, pages 187-194, 2000

[10] A.P. Pentland and J. Williams, *Good vibrations: Model dynamics for graphics and animation*, In Proceedings of SIGGRAPH 1989, pages 215-222, 1989

[11] J.C.Carr, *Reconstruction and representation of 3d objects with radial basis functions*, In Proceedings of SIGGRAPH 2001, pages 67 - 76, 2001

[12] F. Grassia, *Practical parameterization of rotations using exponential map*, Journal of Graphics Tools, pages 29 – 48, 1998

[13] V. Štěpán and J. Žára, *Teaching tennis in virtual environment*, In Proceedings of SCCG 2002, pages 49-54, 2002