BIOMECHANICS AND VIRTUAL REALITY

(BIOMECHANIKA A VIRTUÁLNÍ REALITA)

Vladimír Štěpán[*], **Stanislav Otáhal**[*]

[*]*Department of Anatomy and Biomechanics,FPES, Charles University, Prague, Czech Republic*

Abstract:
This article takes as its purpose to briefly introduce to biomechanics community the state-of-the-art of modelling and animating humans in computer produced 3D environment known as Virtual Reality (VR). Virtual Reality is also an interdisciplinary study that brings together various sciences including biomechanics. We explain basics of human modelling and animation, then we point out several interesting animation tasks that share a problem of reduced input data and we continue by suggesting the possible solution based on the knowledge of real human joint anatomy and kinematics. Moreover we remind that VR can be generally useful visualization tool even for biomedical sciences.

KEYWORDS: biomechanics, virtual reality, human animation, inverse kinematics

1 Introduction

The term *virtual reality* (VR) usually refers to completely or partially nonexistent environment created using the computer system. More narrow definition speaks of the 3D environment created in real time, which means the time that will not be perceived as a delay by the user. Thus the most common and widespread example of VR system would be real-time computer games because of their considerable emphasis on realism both visual and physical. Obviously in such a 3D environment a large variety of real world phenomena is being modelled and that inevitably includes humans.

Moreover the term *virtual reality* also refers to the interdisciplinary science that studies the problems and solutions of creating highly realistic virtual environments. Since a great deal of attention is being given to modelling of humans, the biomedical sciences contribute to virtual reality research as well. With growing focus on simulating the human motion, also the part of biomechanics that studies human motion is becoming important.

Besides being a tremendous source of entertainment, VR is a powerful visualization tool. As such it has been entering new fields of application demonstrating various analyzed phenomena, many of that concern human motion. The artificial intelligence research has been dealing with systems capable of recognition and interpretation of human activities. These can be used for educational purposes and so they need visualization tool. The two opposite requirements meet at this point, the tendency to abstract typical for artificial intelligence tasks and the need for instantiation necessary for understandable visualization. This can be considered an entirely new task for VR and human animation as it is required to provide clear visual effect with very limited input information.

This and the ever growing demand for realism of animated virtual human models produces increasing interest in modelling approaches based on deeper anatomical knowledge. Also other research areas face problems of insufficient input information. The information available could be complemented possibly to sufficient amount with realistically behaving human model.

Structure of the Document

In this document, let us briefly describe the wide topic of virtual human modelling and animation in section 2 with closer focus on principles and some typical problems and solutions of creating (2.1) and animating (2.2) the model. Then, in section 3 we will point out what impact can biomedical research still have on human modelling and animation related tasks. And finally in section 4 we propose the possible contribution in opposite direction which is quite simple, yet possibly important.

2 Virtual Humanoid

With the development of virtual environments also the techniques to model and display humans in these environments are evolving. Computer models of humans differ in many an aspect according to their intended usage. On one end of the spectrum there are the simple projections of user to multiuser virtual environments with the sole purpose of providing the observable countenance, usually not realistic from the functional point of view. On the opposite side there exist highly realistic and sophisticated models both visually and functionally, serving the purpose of various simulations.

A model, no matter its realism, that serves to visualize the user in virtual environment is called *avatar*. This is a word, originally Sanscrit, to denote the materialization of a god. In VR, the term avatar often refers to any virtual character including autonomous agents, although correctly only user driven characters should be called that name. Basically, an avatar is a communication tool for VR visitors and it doe not necessarily need to be humanoid in shape.

The humanoid shaped models are refered to as virtual humans or *virtual humanoids* since they do not always depict really human beings. The use of these models is not limited to the area of VR as it was defined in the introduction.

The largest application area would be the entertainment. We have already mentioned the games and everyone can think of motion picture industry although that is not the VR application as it trades the real-time feature for greater realism. Besides that VR is the medium addressed by modern artists to create performances and installations that often use human modelling and animation techniques.

The visualization power of virtual environments is often utilized in industrial design, testing the usability without really building the prototype. The Boeing Human Modelling System based on JACK[7] software has been used at testing the accessibility of airplane's internal spaces by the ground personnel and also at the design of International Space Station

Cultural heritage is rapidly evolving area of VR. The archeological sites and pieces of ancient architecture are being reconstructed in virtual environment. Moreover they are inhabited with virtual human population to make the illusion of ancient life complete. Great deal of research on autonomous agents has been done here. Also some interesting project has been carried out such as augmented reality enhancement of ancient Pompeii (LIFEPLUS).

Thus we get to the educational capabilities of VR. Even games can be used that way as some military schools use them to teach combat tactics. Except for that, purely educational applications have appeared. We have experience with sport techniques demonstration [11], where the freedom of viewpoint can greatly enhance the explanatory effect. We also used animated virtual human in a drum-set tutorial systém [7]. Even projects that aim at different areas of research ask for VR presentation modules to visualize analyzed phenomena [12]. If the project focuses on human activities, the human modelling and animation techniques are necessary tool for visualization.

2.1 Human Modelling

We have introduced briefly the world of human modelling and animation. Let us now see in more detail how such a human model is constructed. Modelling a human being in VR is quite a complex task. If we generalize a little, we can say that there are several distinct areas of research.

- *Body and motions*. The base of model for animation from computer science point of view is the hierarchical structure of geometric transformations – joints. These are interconnected by links (segments) – bones. The simplest models implement only this hierarchical structure and the body segments are visualized in a simplistic way. This concerns mostly the applications for motion data processing that do not have to solve the surface deformation. Human perception can evaluate the realism of motion even if it is shown on wireframe structure or just a selection of moving points.
- *Soft tissue deformation*. Human body surface is difficult to model for its ability to deform in motion and at interactions with the environment. Visually realistic model must be able to cope with this problem. We will mention the specific problems and solution in separate paragraph.
- *Facial animation*. Human face is capable of a large variety of motions that cannot be modelled using the mechanism of geometric transformation hierarchy. The face is usually modelled by a single polygonal mesh where the relative positions of points or groups of points can be changed. Facial modelling and animation is relatively independent research area with a very low impact on joint modelling so we will not go into detail. It is often closely connected to speech synthesizer research.
- *Complements*.. Real humans are not just the bodies, they usually grow hair, wear clothes and other attached objects and these must be modelled too. This area does not connect to anatomical human modelling, but is as important for final model realism as the hair and clothes are important for real humans.

- *Behavior*. We list this only for completion. This is a top of the pyramid at the development of autonomous agents and it belongs to the artificial intelligence rather than anywhere else.

We can interpret this division as the levels of realism, but up to certain measure it also present the functional layers of the model itself. All the model is dependent on geometric transformation hierarchy (skeleton). The soft tissue deformation are inferred from the state of the hierarchy and determine the surface that in turn influences the appearance of complements.

2.1.1 Model Structure

Typically the model is composed of two dependent layers, *skeleton* and *skin*.

Skeleton at this context denotes a kind of compensatory anthropomorphic structure describing the hierarchy of transformations necessary to animate the model. At certain measure these transformation correspond to real joints. The graph describing the dependencies among these transformations is a tree. It is minimal graph called skeleton in graph theory terminology. Transformations – *joints* – are the nodes of this graph, the edges correspond to the links between the joints that are called *bones*. It can seem that the term skeleton would be meaningful also in anatomical point of view, but it is not the case. Many models contain in skeleton layer large number of joints and bones that have no anatomical meaning and serve solely the purpose of visual effect.

Joint is defined as a geometric transformation, particularly rotation around the firm center. Thu the skeleton layer of virtual humanoid is a hierarchy of rotations.

Skin is the layer that includes the part of the model that is really displayed. The implementation is usually based on polygonal mesh that is being rendered. Process of creating this mesh differs according to the complexity of the model. There may be one for whole model or for each rendered segment.

The model could possibly consist merely of the skin layer. Any animation would then have to contain positions of all the skin vertices for each animation frame (per-vertex animation), which is ineffective. Therefore other layers, mainly the skeleton, are introduced to simplify the animation. The model specifies the dependency on particular joint for each skin vertex and the animation can contain only skeleton configurations (skeletal animation).

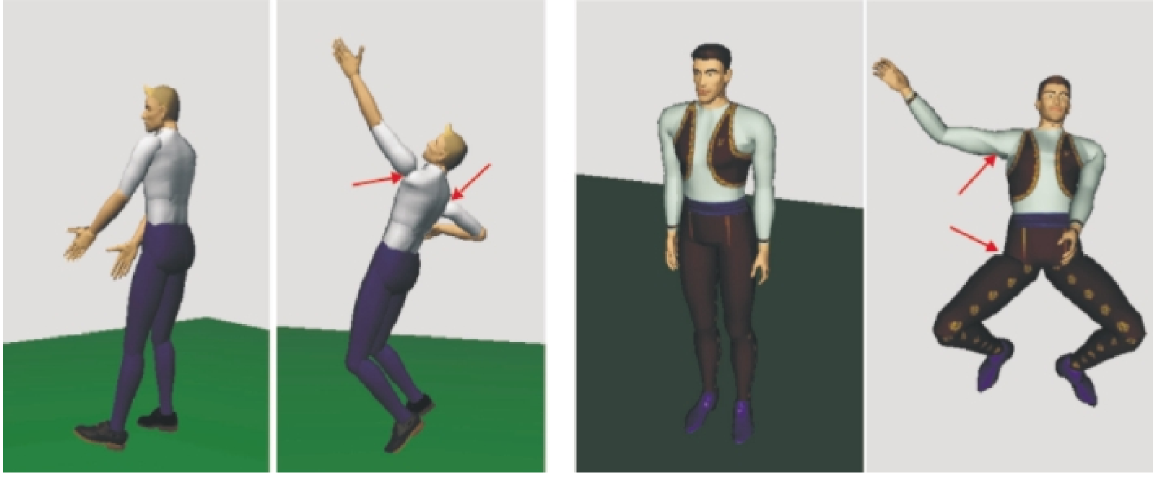

Figure 1. Left: Example of compact segment humanoid.Right: Seamless shape implementation. In both cases note the artifacts indicated by arrows

2.1.2 Compact Segments

The simplest skin implementation is by compact segments. Each displayed segment of the body consists of separate mesh. all the vertices of this mesh are dependent on the same transformation. Thus the relative positions of the vertices do not change over the animation. The computational complexity depends only on the number of vertices. More points per mesh can be used for greater realism. It is not possible to implement the surface deformations and the undesirable artifacts of dislocation occur frequently. See Figure 1 (left).

2.1.3 Seamless Shape

Models of higher computational complexity and visual quality use the skin layer implemented as a single deformable mesh. The basic idea is that vertices of single seamless mesh are assigned to various transformations of the skeleton. Suppose the vertex with reference position *v* is assigned to a joint *j*. Resulting position *v'* is computed

$$v' = A_0 F_j v \tag{1}$$

Where $\mathbf{A_0}$ is global transformation of root joint and $\mathbf{F_j}$ is the transformation of *j*-th joint.

All the methods that concern with relationship between skin and skeleton are known as *skinning*. This basic approach eliminates the discontinuous dislocation artifact, but produce another ones of its own – pinches around the joints. Various adaptations of a technique known as *vertex blending* are employed to eliminate the artifacts. The idea of vertex blending is assigning a vertex to more joints in the hierarchy and resulting position is obtained by convex combination of transformations by those joints.

Suppose the vertex with reference position *v* is assigned to a joints $j_0, ..., j_n$. Resulting position *v'* is computed

$$v' = A_0 \sum_{i=0}^{n} w_i F(j_i) v \tag{2}$$

where w_i are weights that do not break the condition that sum of all of them must be equal to 1.

We can interpret the weights w_i as a measure of influence joint j_i on the vertex *v*. They are set either manually or they are computed using some rule such as a distance from the bone. The vertex is usually assigned to no more than 4 joints. Many vertices are assigned to only one, as the vertex blending has some effect around the joints.

The biggest drawback of vertex blending is the deformation artifact known as *candy-wrapper artifact* that occur at bigger axial rotations. Adaptations of the method to avoid this problem are subject to research[6].

Vertex blending is not the only technique of skinning, but it is fully sufficient to create the idea of the problems and solutions. Other methods, such as[5] are basically similar, as different vertices of the mesh are transformed with different transformations. Character on the Figure 1 (right) was created using the last cited method.

For more detail on human body modelling see for instance the book [1].

2.2 Animation Techniques

Computer animation is the process of building the dynamic scene resulting in a series of images. The model of scene is built and rendered in discreet time moments, like in classical animation. The advantage of computer processing is that most of the animation frames can be created automatically using various simulation algorithms.

Speaking of computer animation, we mean all the forms of motion ranging widely from translocating the object, change of its shape, to the changes of other visual parameters, such as color. In this text we will be restricted to animation of virtual humans, that has its specifics on the higher levels, although the low level theory is common to animation in general.

2.2.1 Keyframing and Animation Curves

The idea of classical *keyframing* has been widespread in computer animation. The animator creates the important key-frames while the program computes all the rest. The key values can concern just about anything, positions, angles, colors, textures, transparency, etc. In human animation it is the value of joint rotation.

The task of keyframing in computer graphics is to find the interpolation curve connecting the key-frames. Displayed animation data are samples of the curve parameterized by time sampled with the viewing frequency.

The animation curves also be represented by a vector of key values $\mathbf{S} = (s_1, \dots, s_n)$ and vector o key times $\mathbf{T} = (t_1, \dots, t_n)$.

In case of articulated structure of human body, there is usually more animated values -- more joints – and therefore more separately interpolated curves is necessary.

2.2.2 Articulated Structure and State Space

Human body is a good example of what computer animation calls *articulated structure*. Such objects consist of sequence of rigid bodies that are connected by rotating joints. In robotics where similar terms are used also prismatic joints are considered, but those have no meaning in human animation.

Articulated structure can be expressed by tree graph with one root and at least one free end. This statement decreases the universality, but we can afford that when we are restricted to human animation.

Let us introduce the term *kinematic chain* as a sequence of joints, subset of the whole articulated structure. For the needs of the animation, kinematic chain is usually fixed within the coordinate system on one end called *base*. The other, loose, end is called *end effector* (EE). Contrary to the name, end effector is not a muscle or any other object with capabilities to actively produce motion. These terms are illustrated by Figure 2.

To determine the position of rigid body in space, we need the coordinate system and total of six numbers, three for spatial coordinates of position, the other three for rotations about coordinate axes. These numbers represent *degrees of freedom* (DOF) and provide unique description of arbitrary system. The number of DOF tells us how many state parameters can be changed. If we add a body to the system, the number of DOR rises. If we bind bodies with constrains, the number of DOF decreases.

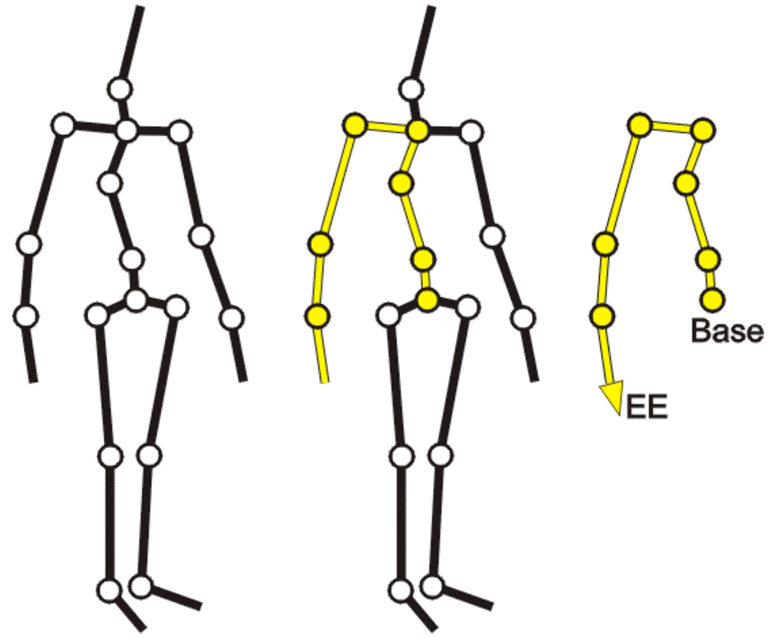

Figure 2. Articulated structure and a kinematic chain. End-effector (EE) is hand, kinematic chain for animation in this case includes joints between hand and root of the whole structure (Base).

The *state space* of the articulated structure comprises of all states it can be at. Current state can be described with *state vector* whose length equals the number of DOF. The number of DOF is the dimension of state space. We will designate the state vector:

$$\Theta = (\theta_0, \theta_1, \ldots, \theta_n) \tag{3}$$

where θ_i are the parameters of the structure. In case of human animation, the state vector is a vector of rotations in joints of the articulated structure.

2.2.3 Forward and Inverse Kinematics

To create the effect of animation it is necessary to define state vector values, which means the joint rotations and position of the base, for each animation frame. There are two approaches to do that in computer animation – *forward kinematics* and more commonly *inverse kinematics*.

From the previous it is clear that state vector defines the position of EE (**X**). Formally written as

$$X = f(\Theta) \tag{4}$$

This is the principle of *forward kinematics*. We set the joint values in order to place EE to desired location. This has recently been used mainly when the sequence of joint values had already been known.

To interactively create the animation this is not very suitable and the *inverse kinematics* (IK) approach has been used. The inverse kinematics problem is literally an inversion of the forward kinematics problem. The goal is to find the joint values based on known EE location. Written formally:

$$\Theta = f^{-1}(X) \tag{5}$$

It is obvious that this problem does not have unique solution. For some EE location the **X** the inverse function f^{1}, does not have to exist, the solution can be out of state space of the kinematic chain.

For other EE locations there can be infinite number of solutions. Number of the acceptable solutions can be reduced by introducing constrains to limit joint range of motion, thus shrinking the state space. Other constrains can act as a priority for some solutions - the parameter traditionally called "*stiffness*" by the animators.

The problem of IK lies in the fact that the function *f* is generally non-linear. The analytical solution of its inversion is practically impossible for a little more complicated hierarchies (DOF > 6). Commonly used approach it the linearization of the problem, computing the solution in the vicinity of current position. The generalization of one variable difference called the *Jacobian* matrix is being used for this purpose.

In the equation (4) the dimension of **X** is *m* (usually *m* = 3 in 3D space) and dimension of **Θ** is *n*. Jakobian is *m x n* matrix, that maps changes **ΔΘ** to changes **ΔX**.

$$\Delta\mathrm{X} = J(\Theta)\Delta\Theta \tag{6}$$

Note that Jacobian just as the function *f* depends on current state of hierarchy which is reflected by notation **J(Θ)** instead of **J**.

The above equation is linearized and theoretically invertible. In fact, the matrix **J(Θ)** is typically rectangular and therefore the inverse does not exist. Other techniques must be used instead, such as rectangular matrix pseudoinversion. We can even successfully use transposed matrix $\mathbf{J}^{\mathbf{T}}(\mathbf{\Theta})$ instead of inversion $\mathbf{J}^{-1}(\mathbf{\Theta})$.

The jacobian methods are iterative. Until the EE does not reach the destination, we repeat solving the equation:

$$\Delta\Theta = J(\Theta)^{-1}\Delta\mathrm{X} \tag{7}$$

We use small step in the direction to the destination as **ΔX**, with computed angle changes we update the state vector and compute new Jacobian and its inversion.

There also exist iterative method that does not work with Jacobian matrix. It is called *CCD* (*Cyclic Coordinate Descent*) and is based on traversing the kinematic chain and setting the joint values to minimize the EE error. The method is very intuitive although not that well theoretically described. The computational time for a single iteration is low compared to Jacobian based methods. Typically it needs more iterations to converge, although many an application can successfully trade the precision for speed and so this method has been widely used. Moreover it is quite robust to singularities that can complicate the pseudoinversion computation. For more detail, we can recommend [4] or an appropriate chapter in [1].

3 Contribution of Biomechanics

Virtual humans often make use of anthropometric measurements to determine the dimensions of body and its parts and locations of joints and other important locations. All of this is useful for visual realism of non-moving model.

For animation purposes there is yet another set of parameters that limits the state space of articulated structure of human body. The H-Anim standard[16] concludes the experience from previous applications and introduces the parameters of *motion range* (upper and lower rotation limit) and priority (*stiffness*) of the joint. These parameters are usually set in order to produce realistic output and the measure of their universality is not obvious.

Animation tasks have been appearing lately, that need to somehow complement missing input information. A human body model behaving realistically from the anatomical and biomechanical point of view might be able to reduce this lack of input information up to the point of satisfactorily realistic animation result.

This model would be an articulated structure whose state space would be sufficiently similar to "state space" of real human body. Moreover it should be equipped with a mechanism that would always select among the possible solutions the one that would be visually the most realistic.

Creating such a model is therefore a task that can be converted to creating a model of joint with suitably defined limiting parameters of *motion range* and *priority*. This task requires a research of human joints motion ranges and constraining relationships among them oriented at possibilities to generalize the observations.

3.1 Limited Input Data

The most obvious of animation tasks that could profit from a human model with realistic motion parameters is the display of insufficiently described motion. Suppose the example of visualization of animated scene for the system that does not focus on human motion description.

The goal of IST-2001-32184 ActIPret project[15] was to create the methodology of understand, learn and recognize human activities (i.e. manipulation tasks). In another words, the system was supposed to observe the activity and create generalized concept for future recognition. The application of this system was to be the education and training, therefore it was to be equipped with the module for results visualization in 3D virtual environment [12].

This VR presentation module had to animate virtual character although all the data available on any level of the system was quite insufficient for realistic animation. The satisfactory effect was at the end obtained with significant help of "tuning" the joint constrain parameters. This method has no universality at all and would inevitably fail if we changed the analyzed scenario.

Example of the ActIPret project serves well to identify the task of realistically modelled joint constrains to complement missing information in certain type of animation task.

3.2 Motion Parametrization

The most realistic animations have been created by recording real motion using Motion Capture technology. The problem of using such motions appears when we need to use them in slightly different situation. Gait motion is usually created by repeating the recorded step. What if some of the steps need to be longer or higher? This example clearly shows that using more recorded samples is not the solution as there can always be yet another step length.

The parametrization of the animation could be a solution. Most of the methods[10][13][2] approaches this problem with some sort of interpolation among several instances of the same class of motion. One of them can be taken as "average" while the others represent particular parameters (long step, high step). Interpolation along the curve connecting the "average" and "long" step will yield steps of varying length. Equation (8) demonstrates the interpolation of animation *C* from source animations *A* and *B* using the parameter *s*:

$$C = (1 - s)A + sB \tag{8}$$

In most cases, especially when it lies between the recorded sources, $(0 < s < 1)$, the result is realistic too. Sometimes, mostly, but not only, in the extrapolation cases when $s > 1$ or $s < 0$, it is possible that resulting caricature of motion would not be feasible for the human. Therefore it is useful to test each interpolated motion and change it to remain within the state space of animated structure if necessary.

The work[11] suggests the use of VR at tennis training. It can be extended to other precise sport techniques. The data input was not discussed deeply in the paper. One of the ways how to provide the motion data to be analyzed, would utilize the very idea described above. The motion of the trainee would be modelled by the means of interpolation among the entries of atabase of recorded typical examples. We would not analyze the real recording of the trainees motion. On the other hand, we would avoid disturbing the trainee with measurements that can not be completely non-invasive.

Such an application would be even more sensitive to the realism of interpolated motion. Thus we identify another important use of joint model with realistic motion constrains.

3.3 Computer Vision

Data input in the sport training application example in previous section could be solved by some new measuring method, both non-invasive and with low hardware requirements. Reconstruction of 3D motion using a single common camera recording (monocular 3D reconstruction), however it sounds like a science-fiction, is the topic addressed by part of computer vision community[8][14].

Monocular recording gives us only 2D information on the objects projected to the camera projection plane. The missing depth information must be complemented somehow. It might be possible to search for posture of the virtual human model that would match the recorded posture in the camera projection plane. If we could measure and ensure the feasibility of such pose, we might be able to find the 3D reconstruction we search for.

Basically this scenario also uses the realistic model would be used to verify the computed body posture. Only this time the computed posture would be the result of much more ambitious task.

4 VR for Biomechanics

Let us say again, VR is a strong visualization tool. Whenever we want to visualize human in motion we would be using some of the techniques mentioned before. Moreover, the VR uses many simulation techniques to gain the desirable visual effect. This can be utilized for further benefit. In our case we discuss the human body model with realistic motion parameters. If we can create one, we can use it to gain further information on the motion we might be interested in.

This VR model would be made to produce realistic visual effect by simulating the physical possibilities of real anatomy. To record the real motion, the visual clues are often used. This is the principle of optical motion capture. By measuring the motion of surface points, we infer the inner conditions of virtual joints of the model, thus obtaining the animation data to be visualized. If these virtual joints are modelled with respect to the real anatomy, the animation we get as a result can give us feasible information on behavior of the real joints.

We have already mentioned some practical applications (e.g. sport techniques training assistance) that can benefit from the research on realistic joint model. We can also believe that many a theoretical research application can use this model as well.

5 Acknowledgements

This work has been supported by the project GAČR 106/03/0958 and GAUK C/114/2005.

References

[1] EBERLY D. H. (2001). *3D Game Engine Design: A Practical Aproach to Real-time Computer Graphics*. Morgan Kaufmann publishers.

[2] LIM I. S. and THALMANN D. (2001). ‘A Vector-Space Representation of Motion Data for Example Based Motion Synthesis’ in *Deformable Avatars*, Kluwer Academic Publishers. pp 169 – 179.

[3] WATT A. and WATT M. (1992). *Advanced Animation and Rendering Techniques*. ACM Press.

[4] WELMAN CH. (1993). ‘Kinematics and geometric constraints for articulated figure manipulation’. Master’s thesis, Simon Fraser University.

[5] BABSKI C. and THALMANN D. (1999). ‘A seamless shape for HANIM compliant bodies’. Proc. of VRML 99, ACM Press. pp 21 – 28.

[6] KAVAN L. and ŽÁRA J. (2005) ‘Spherical blend skinning: A real-time deformation of articulated models’. Proc. of 2005 ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games. ACM Press, April 2005.

[7] LINDA M., ŠTĚPÁN V. and SPORKA A. (2005) ‘The Drum Set Tutorial System by Means of Inverse Kinematics’. Proc. of CEMVRC ’05, Eurographics Digital Library. June 2005. pp 75-80.

[8] LOY G., ERIKSSON M., SULLIVAN J., and CARLSSON S. (2004) ‘Monocular3D Reconstruction of Human Motion in Long Action Sequences’. In Proc. of ECCV 2004, Springer-Verlag, 2004. vol. 4, pp 442 – 455.

[9] PHILLIPS C. B. and BADLER N. I. (1988) ‘Jack: A Toolkit for Manipulating Articulated Figures’. In Proc. of the 1st annual ACM SIGGRAPH symposium on User Interface Software, pp 221 – 229.

[10] SEMANČÍK J., PELIKÁN J. and ŽÁRA J. (2004). ‘Synthesis of Terrain/Adapted Walk from Example Steps’. In Proc. of VIIP ’04.

[11] ŠTĚPÁN V. and ŽÁRA J. (2002) ‘Teaching Tennis in Virtual Environment’. In Proc. of SCCG’02, ACM. April 2002. pp 49–54.

[12] ŠTĚPÁN V., ŽÁRA J. and HLAVÁČ V. (2005) ‘Presenting Generalized Human Activities in Virtual Environment’. In Proc. of SCCG ’05, ACM. May 2005. pp 49–56.

[13] UNUMA M., ANJYO K. and TAKEUCHI R. (1995). ‘Fourier Principles for Emotion based Human Figure Animation’. In Proc. of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, ACM Press. pp 91 – 96.

[14] WILHELMS J., VAN GELDER A., ATKINSON-DERMAN L. and LUO A. (2000). ‘Human Motion from Active Contours’. In Proc. of Workshop on Human Motion (HUMO’00).

[15] ACTIPRET project homepage. http://actipret.infa.tuwien.ac.at.

[16] H-Anim (Human Animation Working Group). http://www.hanim.org.