Capturing and Visualization of Human Body Motion

Martin Fědor *
fedor@dcse.fee.vutbr.cz

Martin Dobšík **
dobsik@dcse.fee.vutbr.cz

Abstract This paper gives a short review of methods used for acquisition and definition of human body motion using an optical motion capturing and layered approach for skin surface modelling. An introduction of simple and low-cost capturing and modelling system is presented.

Key Words: Forward kinematics, inverse kinematics, skeleton model, articulated structure, motion capturing, deformation model, physically-based simulation, layered construction

1 Introduction

As the computer animation comes to be the most powerful tool to simulate and demonstrate the reality and the virtual reality seems not to be only a dream anymore, it happened that the visualization of human body motion became an objective of many research laboratories. Human being design in the virtual reality is very exciting task as well as a big challange. Since a designed representations and simulations of the real world could not be called complete with no human and without their realistic behaviour, many graphics programmers implemented methods and tools for human body animations. In order to gain a realistic look and feel of the human body movement, graphics algorithms and their implementations offer both the sufficient level of realism and simplification of the animation design.

The task of realistic visualization of human body in computer graphics could be divided into two successive parts: The first one dealing with a rigid part of human body – the skeleton and its animation, the second one is the realistic rendering of non-rigid parts including muscles, skin and body surface in general.

This work would like to discuss a short preview of both project parts and depict possible solutions as present in many papers by world research teams. Moreover we would like to present a possible way how to create a low-cost but sufficiently usable system of human body visualization and its motion capturing.

2 Human Skeleton Animation

Many research teams introduced a system for animating human bodies for several purposes: as virtual actors within a fully synthetic movies, for virtual reality engines, internet conferencing (poses recognition), entertainment industry, medicine and many others. All those system have one thing in common: They use a simplified abstract of a human body kinetic system - the

* Department of Computer Science and Engineering, Božetěchova 2, CZ-612 66 Brno
** Department of Computer Science and Engineering, Božetěchova 2, CZ-612 66 Brno

simplified skeleton consisting of set of rigid links (bones) connected by joints into the articulated structure. As this structure is very similar to the well-known robot manipulator schemes, animators imported mathematic methods previously widely used in robotics.

The manipulation with an articulated structure could be achieved in two manners: using forward or inverse kinematics (IK). Except the mathematic formulas they use, these two methods differ in a difficulty in which the animator can enliven synthetic legged figure. Both methods are very well described in many publications and articles dealing with computer graphics and animation. Nevertheless let me introduce a brief description of these methods for their importance in computer animation of legged figures.

2.1 Forward and Inverse Kinematics for Articulated Structure Animation

When the forward kinematics is used, the motion of all joints is specified by the animator. The motion of the last joint of structure is determined indirectly as the accumulation of all transformations that lead to that end effector:

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

where Θ is joints state-vector (usually angles of rotation) and X is computed end-effector position.

In order to animate human body defined by an articulated structure the animator must perform all necessary actions leading to desired end effector position in a scene. He must rearrange all angles of joints antecedent to the last one.

IK is sometimes called 'goal directed motion', the animator defines a position of the end-effector only. IK solves for the position and the orientation of all joints in the link hierarchy that lead to the end-effector.

$$\Theta = f^{-1}(X) \tag{2}$$

It is apparent that IK is more computational expensive and complicated than the forward kinematics but it allows the animator to easily place each limb of legged figure in a desired position. Appropriate join angle values are computed automatically with respect to defined joint constraints (maximal and minimal angles between each two links in the structure).

As stated above, the evaluation of end-effector position using forward kinematics is very simple task. All we need is just to multiply all of the transformation matrices assigned to each joint. These matrices describe the transition and orientation of every link with respect to previous one (see [1],[8]).

There are many publications about computing state vector of articulated structure from given end-effector position. Many of them were printed in conference proceeding, transactions and journals about robotics and automation. Early works involved in IK were strictly dedicated to robotic application ([19],[20]), but achieved results were very soon applied in computer graphics as well. The basic idea over IK lies in finding the inverse function mapping the end-effector position into a space of state vectors and consequently finding the result state vector. As the degree of freedom (DOF) of the articulated structure is much bigger than degree of vector X, the system is said to be redundant and very hard to solve.

The methods solving the IK could be divided into a few categories:

- *Algebraic* (see [20], [8] for details)
  This is a fast method of IK computation, the limitation is that found algebraic equations are valid only for a specified structure.

- *Geometric* ([9])
  As opposite to the algebraic method, a solution is derived using the geometry of the manipulator. The possible limitation is that the solution only applies to a specified geometry.
- *Iterative* ([19], [1], [8])
  This method solves IK iteratively, unfortunately it is reasonably slower than previous and converges to only one solution. Iterative method is well documented in literature and is very suitable for arbitrary structures. Jacobian matrix is very intensively used during the iteration process – it maps the end-effector velocity in cartesian space to velocity of joints in state space (see equation 3):
  $$\dot{\Theta} = J^{-1}(\dot{X}) \tag{3}$$
  Since the articulated structure of human body is redundant, the Jacobian matrix is usually singular and non-square. Instead of simple matrix inversion a pseudo-inversion must be applied ([17]). Detailed description of Jacobian construction and iteration is in [1] and [8], further reading about ill-conditioned and singular systems could be found in [12].

2.2 Human Skeleton Model

A design of the human skeleton model is very important part of realistic animation system. The model's degree of freedom affects program difficulty in solving IK, so that it is recommended to simplify the model keeping as much functionality and flexibility as possible.

Presented model with only 6 DOF for arms and legs is sufficiently simple to reduce computation time, although it may bound the animator and consequently reduce authenticity of figure motion. In [18] proposed body model consists of 31 anatomic joints and the skeleton mobility is modelled with 62 DOF. Such model is very computational expensive, not mentioning very hard design of algorithms handling such amount of joints in reasonable time.

Using Denavit-Hartenberg notation ([1]) we can describe the kinematics of each human model link (bone) relative to its neighbours by attaching an own coordinate system to each link. In order to define the transformation matrix it is necessary to define at least 4 parameters for every (rigid) link coordinate system: rotation angles and the distance of current link according to the previous. For computation reasons we further need to limit each other link's rotation angles to reflect the physiological limitations of the human limbs. Defining such constraints helps us to find a solution of redundant system of IK.

3 Capturing the Motion

Previously mentioned techniques of animation using forward or IK can help the animator to manually design an arbitrary human body motion, but the level of realism strongly depends on the animator artistic skills as well as his patience. Recently many professional 3D modeling tools imported the inverse or forward kinematics and improved their ability in articulated structures design. Even though more and more sophisticated tools were included in 3D programs such as physically based modelling or automatic generation of motion transitions, still the biggest part of animation design lies on the animator.

Motion capturing (further referred as *MoCa*) is the method giving truly realistic motion of almost any legged figure. It involves measuring an object's position and orientation in a space and recording that information in a computer-usable form. Objects of interest could be any element in a scene including camera and light position, facial expression or a position of any object featured during an capturing session (ball, rocket – simply anything what the actor is handling).

Mostly used *MoCa* systems are based on magnetic or optical principles:

- *Magnetic systems* use a set of sensors attached to an actor body and the source of magnetic field. A typical magnetic motion capture system has one or more electronic units into which the source and sensors are connected. It is generally fast enough for real time animation although not so accurate.
- *Optical systems* are based on high contrast video imaging of retro-reflective markers which are attached to the object whose motion is being recorded. The captured animation sequence of the actor's motion is post-processed and an extraction of marker's position in a scene is performed. As this process is usually very time-consuming it is not recommended for real-time systems, nevertheless its advantage is in high accuracy of motion estimation.

4 Human Body Visualization

Modeling and visualization of human body is undoubtedly very complex task. The first problem of it - natural-looking motion of articulated skeleton, was briefly discussed in previous sections. The second problem is the deformation of the skin surface shape according to the skeleton pose change. If we intend to model human body for visualisation in computer animation the only thing we need to do, except for skeleton motion specification, is to find the change of the surroundig skin surface as a response to this motion. It is the inverse task comparing to common state of matter. Most of the systems for computer animation (research and commertional), current and developed in the past as well, use this approach. These systems differ mainly in the method used for definition of relation between the skeleten motion and resulting skin deformation. Using the criteria of this relation we will classify the deformation models into these cathegories: geometrical deformation models, simulation based deformation models and layered deformation models. Let me review the main deformation techniques and models developed in the past few years as it is very important to the development of our work.

4.1 Geometrical Deformation Models

It was very important for soft object modelling like human body, to have tools powerful enough for "sculpturing" complex shapes. First such a tool developed Alan Baar [1] using a set of geometrical operators based on traditional linear transformations. Sederberg and Perry [1] introduced the concept of Free-Form Deformations (FFD) that provides global and local deformation using Bèzier hyperpathes with Bernstain polynomials. Cocquillart and Jancène extended these FFDs for more general deformations [2], and developed FFDs for interactive animation techniques [3]. More recently Moccozet and Thalmann [15] proposed the Dirichlet Free-Form Deformations (DFFD). There are many other deformation techniques which were usually developed for special purposes, e.g.: [11], [6] and others.

We can achive ease of control and low computational cost for interactive modeling of synthetic actors using these techniques. However clear geometrical modeling tends to represent the bodies using geometrical surfaces or uniform solids. It has little relation to the physical reality and therefore the resulting model is not realistic enough.

4.2 Simulation-based Deformation Models

Lack of realism accompanying geometrical modeling techniques forced a creation of the new computer graphics research area which is focused on object motion and shape deformation modeling through applied physics.

Physically-based deformable models are beased on elastic and viscous properties of human or animal body components (bones, flesh, fat, skin, . . .). This approach is used in projets by Terzopoulos and Witkin [23], Metaxas and Terzopoulos [14] and many others.

There are two main drawbacks related to physical models: complexity and computational cost. Threfore, good constraints should be used to make them useful. Moreover, because they represent physical media as a large number of discreet elements, they can be very CPU-intensive. Despite of these problems the results of physically-based techniques are much better then results of modeling with geometrical information only.

4.3 Layered Deformation Models

Working with layers is very well known common method for animators using traditional animation techniques. They sketch stick figure first, which represents simplified skeleton, then draw rounded forms to represent the flesh, and finally complete the animation frame with the outline representing skin. The same sort of approach is used in some other types of traditional animation, like clay animation.

So called ”layered approach” for constructing and simulating artifical characters is also derived from this concept. Magnenat-Thalmann and Thalmann used two-layered approach in their film ”Randez-vous à Motréal” [13] – the digitized polygonal skin surface was deformed according to the skeleton pose change using a set of so called joint-dependent local deformation operators (JLD) ([10]).

Turner and Thalmann in [24] use the name ”Layered elastic model” to describe layered model with some added physically based elastic components. The combination of physical simulation with layered approach improves realism of animation results rapidly.

Chadwick, Haumann and Parent presented some sophisticated results of their animation system (Critter) in [5]. They used four layered approach, but only one of these layers were physically-based. Another example of succesful layered model is ”Elastic surface layer model” implementated in LEMAN system as presented in [24]. Gourret et al used finite element method to describe a human hand as a volume element mesh surrounding bones in [4].

Recent research projects still use layered principles, morover, they tend to create human body model which is as close to reality as possible. Nedel and Thalmann used anatomically-based approach [16] and Chen and Zeltzer biomechanically-based muscle model [7].

5 Capturing and Visualization System Design

We propose following complex system consisting of two independent parts: the *MoCa* as a source of animation data and the *Visualization* as a post-processing tool for rendering, simulation and further editing of acquired motion.

Optical *MoCa* sub-system is devided into a few successive steps:

- Image sequences from calibrated camera are processed to extract and identify markers. The acquisition and markers extraction steps are demonstrated on fig.1 – the actor has set of optical markers attached to his body, grabbed video images are thresholded in order to extract markers positions in a simple frame.

- The 2D markers positions are subsequently passed to 3D position estimation block which computes the actor's limb positions in a scene according to the camera parameters. The camera intrinsic and extrinsic parameters and physiological constraints of the human body are heavily used within this step.
- 3D positions are assigned to a skeleton model. Since more then one optical marker is attached to a single human joint, this step is responsible for propriety multiple markers–joint assignment.
- Verification of physiological constraints is performed by using IK. The results affect the identification and prediction process as well as the 3D position estimation.
- When the captured motion is completely assigned to a skeleton model, the abstract motion specification is derived and passed to the *Visualization* sub-system.

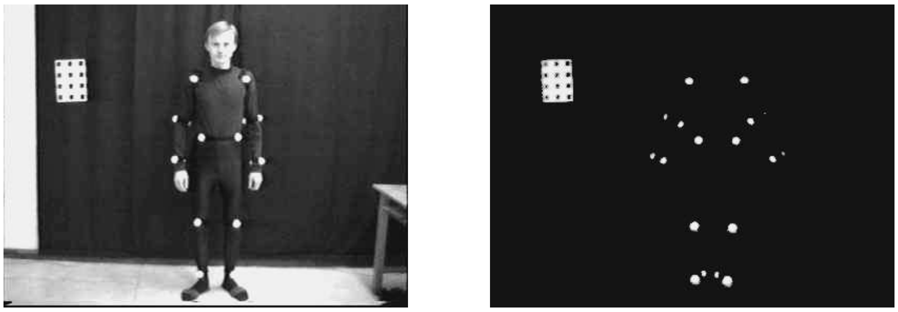

Figure 1: MoCa session using single camera view

The prediction phase is necessary within this sub-sytem because of frequent occlusions and the disappears of markers from the camera view. In order to minimize this unwanted behaviour of optical *MoCa* systems it is proposed to cover the active area by at least four cameras. Unfortunately even this solution does not fully eliminated these optical limitations, so that the prediction system tries to estimate most probable position of markers in a scene according to previously computed linear and angular velocities and defined skeleton constraints using IK.

The Visualisation subsystem we propose (as presented on fig.2) is better to describe by enumeration of its functions. It provides:

- simple user interface useful for: human body model construction, interactive texture mapping, motion specification using forward and IK, modifications and corrections of captured motion
- simulation module for experiments with physical muscle behavior,
- collision detection module,
- and rendering module providing displaying functions for interactive model construction and motion specification and also for simple final animation visualization.

We intend to use priciples of elastic layered construction for the simulation module in our project. The first phase will implement simple layered model based on principles briefly introduced in section 4. Concurrently we want to develop better physical simulation model of muscle behaviour.

As we are still working on the skin surface deformation model, the visualisation subsytem should offer general functionalities which allow to incorporate reasonably complex deformation model we will develop in the future. It will be also useful for experimenting with other new

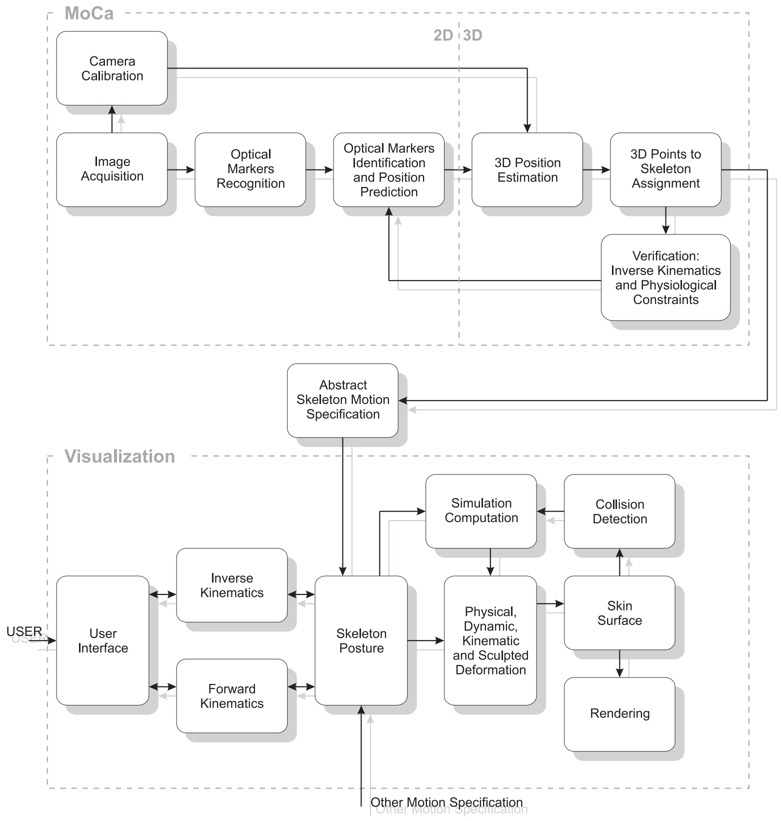

Figure 2: Motion capturing and human body visualization system

developments in this field. No matter which deformation and simulation model is used for skin shape visualization, the system is ready to be always able to visualize and adjust the results of *MoCa* subsytem at least as a simple stick skeleton model.

6 Conclusion

We presented the complex animation system comprises both the animation data acquisition as well as their visualization. Our approach is based on the optical motion capturing using low-cost camera equipment (it is under consrtuction with only one camera available at this time).

We want to compensate the lack of high-level optical systems by applying a prediction phase which should remove the problems with marker-marker and body-marker occlusions. Mathematics methods of IK and definition of physiological constraints of human skeleton model are strongly used within this computational phase of *MoCa* sub-system. Concurrently they are used in the *Visualization* for post-editing as well.

Discussed visualization sub-system will use the resulting data from the previous capturing phase enabling further manual improvements of animation sequence by the animator. It should provide functionality to remove remaining inaccurately predicted or completely lost positions of optical markers attached to the actor's body. Moreover simulated muscles tension works as a feedback affecting the skeleton motion.

Both parts of our system are solved independently as the output from the *MoCa* stage could be simulated by temporally created animation data sets. That way the visualization does not depend on a source of provided data although we assume it will affect the *MoCa* during the work progress.

References

1. M. Watt A. Watt. Advanced animation and rendering techniques. Addison-Wesley, 1992.
2. S. Coquillart. Extended free-form deformation: A sculpturing tool for 3d geometric modeling. In *Computer Graphics (SIGGRAPH '90 Proceeding)*, 1990.
3. S. Coquillart and P. Jancene. An interactive animation technique. In *SIGGRAPH'91*, pages 23–26, 1991.
4. P. Gourret, J., N. Magnenant Thalmann, and D. Thalmann. Simulation of object and human skin deformations in a grasping task. In *SIGGRAPH '89*, pages 21–30, 1989.
5. J.E. Chadwick, D. R. Haumann, and R. E. Parent. Layered construction for deformable animated characters. In *Computer Graphics (SIGGRAPH '89 Proceedings)*, pages 243–252, 1989.
6. Y. Chang and A. P. Rockwood. A generalized de casteljau approach to 3d free-form deformation. In *Proceedings of SIGGRAPH '94, Computer Graphics*, pages 257–260, 1994.
7. T. Chen, D. and D. Zeltzer. Pump it up: Computer animation of a biomechanically based model of muscle using the finite element method. In *Computer Graphics (SIGGRAPH '92 Proceedings)*, pages 89–98, 1992.
8. K. W. Chin. Closed-form and generalized inverse kinematic solutions for animating the human articulated structure. Bachelor project, Curtin University of Technology, 1996.
9. N. I. Badler J. U. Korein. Techniques for generating the goal-directed motion of articulated structures. IEEE CG&A, November, 1982.
10. L. Laperriere, N. Magnenant Thalmann, and D. Magnenant Thalmann. Joint-dependent local deformations for hand animation and object grasping. In *Graphics Interface'88, Edmonton*, 1988.
11. R. MacCracken. Free-form deformations with lattices of arbitrary topology. Technical report, Computer Graphics Reserach Laboratory, University of California, 1996.
12. A. A. Maciejewski. Dealing with ill-conditioned equations of motion for articulated figures. IEEE CG&A, May, 1990.
13. N. Magnenant-Thalmann and D. Thalmann. The direction of synthetic actors in the film *randez-vous à montréal*. *IEEE Computer Graphics & Applicaions*, pages 41–51, November 1988.
14. D. Metaxas and D. Terzopoulos. Dynamic deformation of solid primitives with constraints. In *Computer Graphics (SIGGRAPH '92 Proceedings)*, pages 309–312, 1992.
15. L Moccozet. *Hand Modeling and Animation for Virtual Humans*. PhD thesis, University of Geneva, 1996.
16. P. Nedel, L. and D. Thalmann. Modeling and defomration of the human body using anatomically-based apprach. In *Computer Animation'98*, pages 34–40, 1998.
17. M. H. Wright P. E. Gill, W. Murray. Numerical linear algebra and optimization, volume 1. Addison-Wesley, 1992.
18. L. Herda R. Boulic, P. Fua. An anatomic human body for motion capture. EMMSEC'98, Bordeaux, 1998.
19. G. Mayer R. P. Paul, B. Shimano. Differential kinematic control equations for simple manipulators. IEEE Transactions on System, Men and Cybernetics, No.6, June, 1981.
20. G. Mayer R. P. Paul, B. Shimano. Kinematic control equations for simple manipulators. IEEE Transactions on System, Men and Cybernetics, No.6, June, 1981.
21. J. Zulauf S. Dyer, J. Martin. Motion capture white paper. Silicon Graphics Technical Report, December, 1995.
22. R. Boulic D. Thalmann T. Molet, Z. Huang. An animation interface designed for motion capture. Computer Animation '97 Conference, IEEE Press, pp. 77-85, 1997.
23. D. Terzopoulos and A. Witkin. Physically based models with rigid and deformable components. *IEEE Computer Graphics & Applicaions*, pages 41–51, November 1988.
24. R. Turner and D. Thalmann. The elastic surface layer model for animated character construction. In *Computer Graphics International*, 1993.