An Anatomic Human Body
for Motion Capture

Ronan BOULIC, Pascal FUA, Lorna HERDA, Marius SILAGHI,
Jean-Sébastien MONZANI, Luciana NEDEL and Daniel THALMANN
Computer Graphics Laboratory, EPFL CH-1015 Lausanne, Switzerland
Tel: +41.21.6935246; Fax: +41.21.6935328 ;
Email:{boulic,fua,herda,msilaghi,jmonzani,nedel,thalmann}@lig.di.epfl.ch

Abstract. Motion capture is the process of capturing the live motion of a person in order to animate a virtual character. Until now the approach providing the best results was marker recognition from multiple camera views (in the infrared range). However, it still requires significant human intervention in the post-processing phase due to weaknesses in the marker tracking. In the framework of the MOCA project we propose a motion capture methodology based on a 3D anatomic human model. This model encompasses a precise description of the skeleton mobility associated with an approximated envelope. Its first objective is to ensure the best matching between the real performer and the articulated structure supporting the animation description. The second objective is to allow the prediction of the 3D location and the visibility of markers.

1. Introduction

The recent film "Titanic" has demonstrated the maturity of the virtual actor technology [6]. In the film, the ship was carrying hundreds of digital passengers with a degree of verisimilitude that made them indistinguishable from real actors. Among the many challenges taken over by the production team, the most critical element in the creation of digital humans was the replication of human motion: "no other aspect was as apt to make or break the illusion" as reported in [6]. Optical motion capture has been used to recover the fidelity of the motion of strolling adults, playing children and other lifelike activities. However, though impressive in its ability to replicate movement, the motion capture process is far from perfect. Even with a highly professional system there are many instances where crucial markers are obscured from camera view or when the algorithm confuses the trajectory of one marker with that of another. This requires much work on the part of the animator before the virtual characters are ready for their screen debuts. Paradoxically, the price to pay for getting phenomenally subtle human movement is that it becomes very difficult and very time consuming to alter what has been captured [6]. The issues are only slightly different for game-oriented motion capture. Capturing subtleties is less important because games focus more on big and broad movements. On the other hand, the end user character may considerably differ in shape and proportion from the performer artist (**Figure** 1a,d). This brings in the additional issue of movement distortion.

In the framework of the MOCA project we propose a motion capture methodology based on an anatomic human model (**Figure** 1). This model encompasses a precise description of the skeleton mobility associated with an approximated envelope. It has a double objective: by ensuring a high precision mechanical model for the performer, we can predict accurately the 3D location and the visibility of markers, thus reducing significantly the human

intervention during the conversion process.

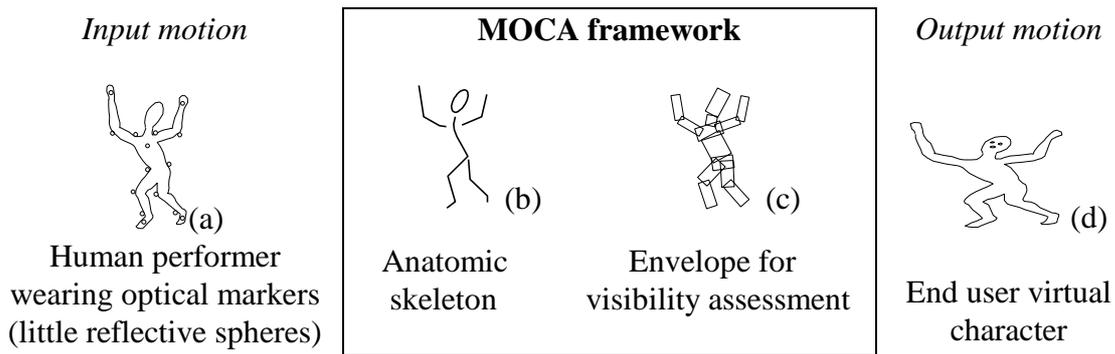| Input motion | MOCA framework | Output motion |

Figure 1: Converting the human performer's motion (a) into the end-user character's motion (d)
with the anatomic human body (skeleton (b) and approximated envelope (c))

The proposed methodology concentrates on body motion capture; face motion capture requires specific devices and approaches (see [12]). One issue addressed in another part of the MOCA project is the simultaneous capture of both the body and the face motions with their respective technologies.

In the remainder of this paper we first review the fields of application of the main motion capture techniques. We then summarise a survey on virtual human modelling and we motivate our choices. The next section gives additional detail on the skeleton topology stressing the original aspects compared to the state of the art. The next sections are focusing on the methodology adopted for improving optical motion capture in the framework of the MOCA project. The paper concludes on the expected business benefits and stresses the important issues of the project.

2. Background in Motion Capture

Motion capture technologies can be grouped into two broad classes [12]:
- *On-line motion capture devices,* where the output of the system can be directly used to pilot in real-time a virtual human body mimicking the performer posture. The main technology is based on magnetic sensors [2,13]. They are mainly used for Virtual Reality and On-line TV shows with synthetic characters [12]. However this technology is too limited in several respects: range of measurement space, noisy data, cumbersome sensors (although they tend to become smaller).
- *Off-line motion capture devices,* where two processing stages are necessary to retrieve the motion of the performer. The technology is based on optical motion capture from multiple camera views (usually in the infrared range). Despite the longer time necessary for visualising the captured motion, it is nevertheless preferred to on-line technology in many cases. It allows to acquire the subtle gestures that are important in high-quality production to convey emotion through motion [6]. It also allows the capture of the large and complex movements that are important in real-time production to maintain a salient visual response to user input. This technology is also used in a clinical context for the assessment of orthopaedic pathologies.

The ESPRIT project MOCA focuses on the off-line technology which will prevail for a significant time before new technologies take over in a production context (the industrial partner Actisystem has a long experience in that field [38]). The MOCA project aims at exploiting a mechanical model of the virtual human that is precise enough to improve the second motion capture processing phase (further referred to as post-processing).

3. Virtual Human Modelling Platform

Numerous companies are now proposing animated virtual human models for the market segments of Biomechanics [32], Ergonomics [15,19,28,34,35], Robotics [30,36], Simulation [26,29], Entertainment [31,33,37] and 3D-Internet [27]. The private research effort is generally trade secret but some of these products directly derive from public institutions with accessible literature [1,3,7,8,9,10,11,14,16,19-25]. Our report [5] lists these products in more details. Among them, one is especially relevant to the animation and motion capture field: *Character Studio Max* a well-known plug-in associated to the *3DS Max* software from *Kinetix* (Autodesk) [31]. The 3DS Max software provides a mechanism of Software Development Toolkit (SDK) allowing independent developers to integrate their original module, the *plug-ins*, in the general architecture of 3DS MAX. Character Studio Max is such a plug-in offering a configurable human model (a `Biped`) that can be animated and deformed. This product derives from earlier researches on legged locomotion by Michael Girard [7,8]. Despite its large acceptance and partly configurable mobility the `Biped` model has two drawbacks:

- The shoulder and the foot regions are too simplified for the needs of the MOCA project. Although most of the end-user characters match such simple skeletons, we want to be able to investigate more complex cases.
- The Character Studio plug-in is oriented towards animator needs. It is not as open as 3DS MAX. As there is no Software Development Toolkit interfacing this module, we are not able to develop new features associated with the `Biped` data structure.

The remainder of our report [5] analyses our participation in four projects integrating virtual human models. **Figure** 2 recalls the major connections between these R&D efforts:

- *ESPRIT HUMANOID* defined our human model for real-time animation of multiple virtual humans.
- *ESPRIT CHARM* focused on the shoulder region.
- *MPEG4-SNHC* [17]: Our contribution was to propose a revised human skeleton model for the Body Animation sub-group.
- *VRML 2.0, H-ANIM* group [18]: a first specification is frozen for the Body skeleton animation and the group moves to the standardisation of facial animation.
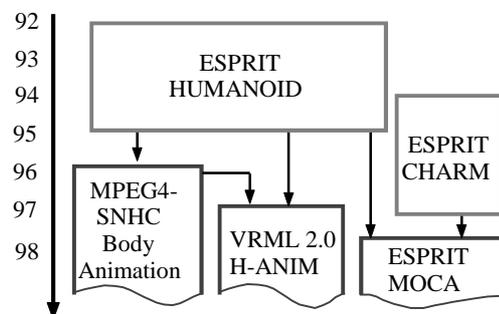

Figure 2 : Timeline of ESPRIT projects and international standardization efforts related to human body mobility modeling

In short, we have retained the 3DS Max software platform to build a plug-in encapsulating the `Body` topology model (adapted from the HUMANOID project). The plug-in is the pivoting element for the anatomic conversion (section 5). Other parts of the algorithmic effort in MOCA are based on in-house libraries implementing the `Body` topology model. This choice is dictated by the need of low-level access to graphic primitives (Open-GL™). We now briefly describe the key characteristics of the `Body` model prior to delineating its use within the proposed motion capture methodology.

4. An Anatomic Human Body Model

We propose a body model allowing the characterisation of a human skeleton regarding all the major mechanical joints. The skeleton mobility (not taking into account the hand

mobility) is modelled with 62 degrees of freedom structured into 31 anatomic joints. One major difference with the structure generally used for virtual characters appears in the shoulder region where we consider the clavicle and the scapula (**Figure 3**). We also dissociate the foot region into four joints, namely the ankle-flexion, the subtalar, the mid-foot and a global toe-flexion. The spine is represented by eight compound vertebrae. We have identified a generic template skeleton by analysing a real three-dimensional skeleton. Each joint rotation centre has been determined based on anatomy concepts and motion range observation. This template can be adjusted on a joint-by-joint basis to match any human skeleton; this is the purpose of the skeleton-fitting stage detailed further on. The complete description of the anatomical skeleton can be found in [4].
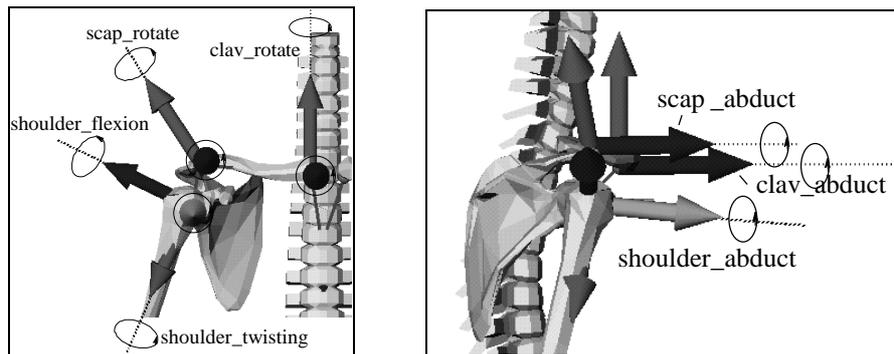

Figure 3: Close-up views of the shoulder region with rotation axis

5. Methodology of Motion Capture

5.1. Overview

The motion capture process consists of four major stages illustrated by **Figure 4**:

- In a typical motion capture session, the performance of standardised "gym" motion is first recorded for the *skeleton fitting* post-processing (**Figure 4**a).
- Then, according to a work plan, multiple takes of each motion are rehearsed until the artistic director is satisfied with the performance of the artist. It is important to remember that only 2D marker images are captured using dedicated hardware (**Figure 4**b, first arrow). So the artistic direction aesthetic evaluation can only focus on the real performer's motion and not on the corresponding virtual character's motion [6]. This is the key problem of motion capture.
- The post-processing phase currently occurs after the motion capture session and can last days or weeks depending on the complexity of the recorded material (**Figure 4**b, two arrows on the right). The first stage is *motion tracking,* to identify the performer's motion expressed in joint angle trajectories.
- The last stage is the *anatomic conversion* producing the virtual character's motion while retaining the emotional subtleties conveyed by the real performer's motion.

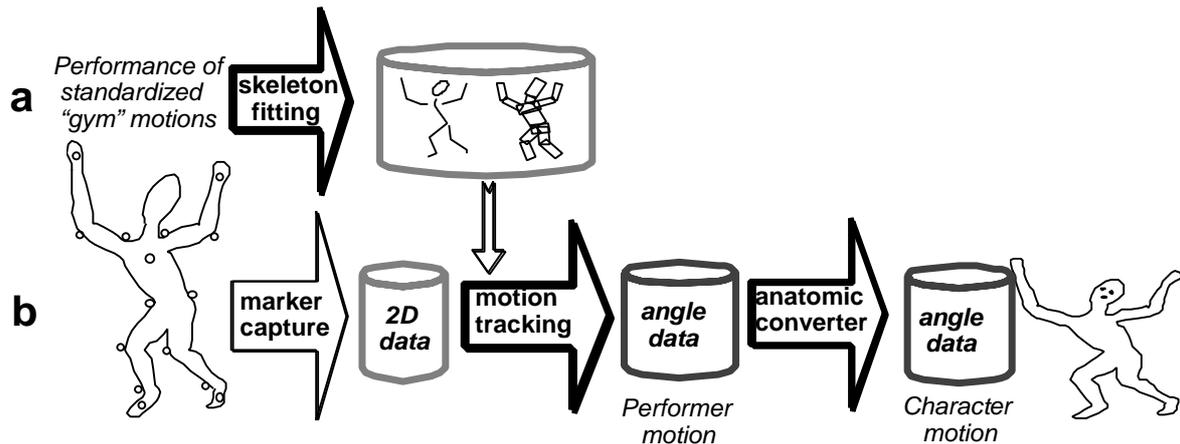The three post-processing stages benefit from the use of the anatomical skeleton as outlined now.

Figure 4: Overview of the motion capture process: initialisation (a) and measurement session (b)

5.2 Skeleton Fitting and Optimal Marker Location

The input to the skeleton fitting procedure are the 3D marker locations computed from simple standardised motions (**Fig.** 5a). The output are the 3D locations of the joint rotation centres (**Fig.** 5b). The difference is clearly visible : the first set forms an exo-skeleton while the one we look for is internal, thus not directly accessible to the measurement technique. The rotation centre identification exploits least square fitting techniques as illustrated on **Fig. 6** on an elbow-type joint (a) and a shoulder-type joint (b). The determination of a pertinent cost function is of decisive importance as the skin-level markers always have an unknown relative motion with respect to the underlying bones. Another outcome of this research phase is to identify some simple motions that optimise the skeleton identification. Besides, this knowledge will also provide guidelines for the optimal placement of markers as a function of the end user requirements (e.g. the considered mobility of the virtual character).
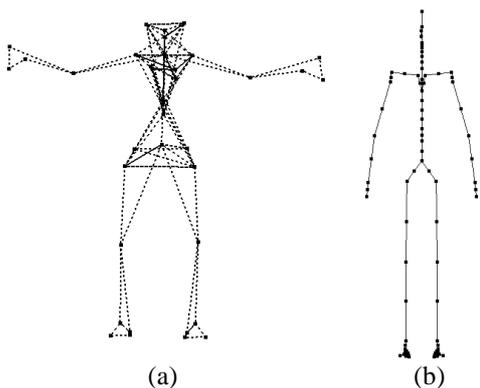

(a)   (b)

Fig. 5: exo-skeleton built from the markers' data (a) and internal skeleton to be adjusted to these data (b)
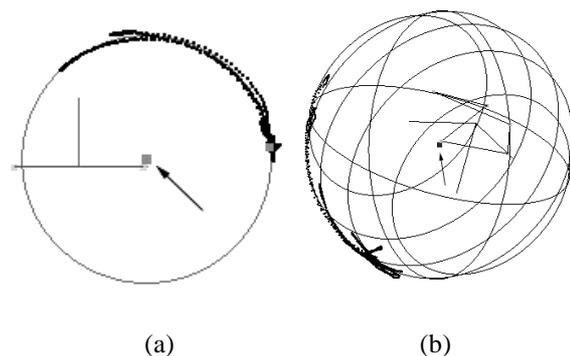

(a)   (b)

Fig. 6: least square fitting of a single joint centre (larger square in the centre) in 2D (a), in 3D (b).

5.3 3D Marker Tracking and Anatomic Motion Identification

Let us review the three levels of motion tracking to understand the benefit of integrating the information of the 3D anatomical skeleton at that stage (**Figure 7**). The input is the 2D marker location expressed in the multiple camera image spaces (from two to seven cameras). The motion tracking algorithm proceeds by using as much as possible the low-

level information (well-known stereo matching from the 2D data, *level one*) prior to switching to higher level tracking when encountering ambiguities. As motion captured in production is often complex, the 2D tracking is frequently supplemented by a 3D tracking (*level two*). At that level, the extrapolated 3D trajectory of the markers helps solving ambiguities by predicting future locations of markers in the camera plane. However, in many cases, occlusions bring the tracking algorithm to a halt, forcing costly manual intervention in order to identify lost markers. A typical example is shown in **Figure 7** where two knee markers should project nearly on the same location on the background camera image. Self-occlusion yields only the right one being visible. In such a context, it often happens that uncertainties induce the tracking algorithm to make the wrong decision about which marker is visible or not. The major interest of the anatomic skeleton and its associated approximated envelope is to allow matching decisions based on marker visibility (*level three*). Such a visibility assessment should significantly reduce false detection cases. Besides, prediction of the 3D marker trajectories (*level two*) can also be slightly improved from the knowledge of natural joint range boundaries.
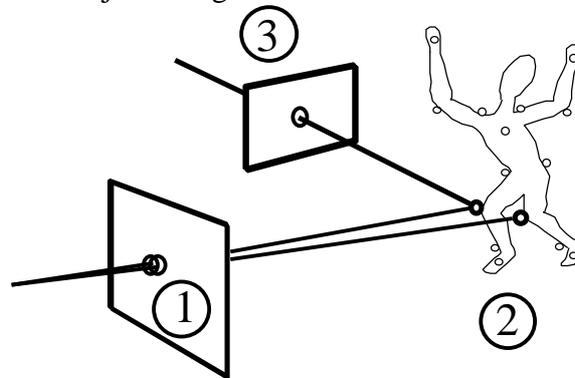

Figure 7: The three levels of marker tracking; in image plane (1), in 3D space (2), and using knowledge of the body motion for visibility assessment (3)

5.4 Conversion from Anatomic Motion to Character Motion

Once the performer's motion has been identified, the goal is to convert it into a motion for the virtual character. The more different the virtual character is in scale and proportion from the real performer, the more artifacts the conversion introduces. For example, self collisions occur between limbs or significant cartesian constraints are lost. We address these problems in two stages, first by working only in joint space. In this space, the end-user can establish the transfer functions between the performer and the character models. Two aspects are considered here: the mobility simplification and the joint motion deformation. Whenever this is not sufficient to recover the desired motion, the second stage is to express cartesian constraints that the character has to enforce all along the motion. The problem that we foresee with this second level is an additional loss in the artistic quality of the motion.

6. Testing the Approach

The MOCA project architecture is built around a central testing phase which consists of three pilots during which the methodology is evaluated in complementary production environments. The end of the project reserves extra time to take into account feedback about issues of robustness, flexibility and efficiency. But before the tools are handed over to the motion capture partner and the end user partners, the methods undergo two specific test stages within the research phase:

- 1) *Optimal marker positioning and the skeleton fitting*: we specify a set of precise motions around a few joints to better estimate the relationship between the underlying skeleton motion and the skin-level marker motion. We have decided to focus on the shoulder region and the twisting rotations of most of the body segments.
- 2) *Motion tracking with the 3D model and visibility assessment*: we need a set of motions highlighting, if possible independently, the three measurement problems: noise (false markers), occlusion, and position uncertainty.

7. Conclusions

The potential benefits induced by the proposed methodology span the motion capture process from the acquisition phase to end-user motion editing. A sharp reduction in the motion capture post-processing time is the first objective of the project. This limitation currently hinders the realisation of complex motion for ambitious projects. The proposed improvements will be evaluated in an industrial context at the settings of Actisystem [38]. The second objective targets the costly motion translation from the measured motion to the character motion. The game market segment is typical of end-user characters that are significantly different in shape and proportions compared to the performer artist. Adjusting the motion obtained from an even perfect motion-capture process is the second deadlock in the widespread use of motion capture. It is currently addressed by asking skillful animators to carry out that tedious and lengthy editing work. Identifying a selection of key editing problems can boost the productivity of animation artists.

The MOCA project is in its first stage but we have already made instrumental decisions about the choice of the human model. The high-resolution skeleton mobility is important to identify fine motion of complex regions such as the shoulder and the spine (mostly for the high-quality production market segment). We also study the optimal marker positioning for the numerous cases where the end user character has a reduced mobility dimension for real-time efficiency purpose (mostly for the game market segment). It is yet too early to conclude on the performances of the approaches proposed for the skeleton fitting, the motion tracking and the anatomic converter. However, we strongly believe that motion capture has an enormous potential for more user-friendly acquisition and manipulation of human motion. The goal of visualising the captured motion on the end-user character during the motion capture session is very ambitious. However, it is the most pertinent goal in terms of global efficiency, both aesthetic and economic, of this technique. When considering the evolution of computing power on PC platforms and the evolution of software integration, we foresee achieving it within the next five years. The MOCA project is an important step in that direction.

References

[1] N. Badler, C. Phillips and B. Webber (1993) "Simulating Humans, Computer Graphics Animation and Control", Chapter 4 "Behavioral Control", Oxford University Press 1993.

[2] N. Badler, M. Hollick, J. Granieri "Real-Time Control of a virtual Human Using Minimal Sensors". Presence 2:1 (1993) 1-5.

[3] R. Boulic & al. "A system for the Parallel Integrated Motion of Multiple Deformable Human Characters with Collision Detection", EUROGRAPHICS' 95 in Computer Graphics Forum 14(3), pp 337-348 , Maastricht, 1995

[4] R. Boulic, L. Nedel, W. Maurel, T. Molet "Anatomic BODY Model, Kinematics", ESPRIT Project 25513 MOCA, D. Thalmann EPFL Project Director, 22 pages, December 1997

[5] R. Boulic "Review of Human Body Modeling Approaches", ESPRIT Project 25513 MOCA, D. Thalmann EPFL Project Director, 30 pages, March 1998

[6] Cinefex 72, Titanic special reprint, 98 pp, P.O. Box 20027, Riverside, CA 92516, USA, December 97

[7] M. Girard, A.A. Maciejewski " Computational Modeling for the Computer Animation of Legged Figures". Computer Graphics 19 (3),1985, pp263-270

[8] M. Girard "Interactive Design of 3D Computer-Animated Legged Animal Motion" , IEEE CGA , 7(6), 1987

[9] J. Hamill, K. M. Knutzen, Biomechanical Basis of Human Movement, Williams & Wilkins 1995, ISBN: 0-683-03863-X

[10] P.L.Y. Lee "Modeling Articulated Figure Motion with Physically- and Physiologically-based Constraints", PhD Dissertation in Mechanical Engineering and Applied Mechanics, University of Pennsylvania, 1993

[11] P.L.Y. Lee, S. Wei, J. Zhao, N.I. Badler "Strenght Guided Motion", Computer Graphics 24 (4), pp253-262, 1990

[12] R. Maiocchi "3D Character Animation Using Motion Capture", in "Interactive Computer Animation", Thalmann&Magnenat-Thalmann (eds.), 1996 ISBN 0-13-518309-X, Prentice-Hall Europe,

[13] T. Molet, R. Boulic, D. Thalmann " Human Motion Capture Driven by Orientation Measurement", to appear in Presence, MIT Press

[14] G. Monheit, N. Badler "A Kinematic Model of the Human Spine and Torso" IEEE CGA 11(2), pp.29-38, 1991

[15] Society of Automotive Engineers Inc. "SAE Cooperative Research Program - Civilian American and European Surface Anthropometry Resource (CAESAR)", Proposed Business Plan, 1996.

[16] Woodson, Tillman, Tillman "Human Factors Design Handbook" 2nd ed, McGraw Hill, ISBN: 0-07-071768-0, 1992

[17] MPEG4-SNHC http://www.es.com/mpeg4-snhc/

[18] WWW VRML 2.0 Human Animation Working Group, Specification for a Standard VRML Humanoid, http://ece.uwaterloo.ca/v-humans, 1998

[19] WWW Lab, Center for Human Modeling & Simulation, http://www.cis.upenn.edu/~hms

[20] WWW Lab, Center for Human Simulation, Colorado University, http://www.uchsc.edu/sm/chs/

[21] WWW Lab, EPFL-LIG, http://ligwww.epfl.ch

[22] WWW Lab, Miralab, http://miralabwww.unige.ch

[23] WWW Lab, Simon Fraser University, http://www.cs.sfu.ca/cs/research/group/GMRL/

[24] WWW Lab, University of Santa Cruz, http://www.cse.ucsc.edu:80/~wilhelms/fauna/index.html

[25] WWW Lab, Wright-Patterson CARD Lab, http://fits-hed.al.wpafb.af.mil/cardlab/home.html

[26] WWW Software, ADAMS-Android, http://www.adams.com/mdi/product/modeling/modeling.html

[27] WWW Software, Blaxxun Avatar Den, http://www.blaxxun.com/vrml/avatarden

[28] WWW Software, Boing Human Modeling, http://www.boing.com/assocproducts/hms/home.htm

[29] WWW Software, Boston Dynamics, http://www.bdi.com/html/interactive_humans.html

[30] WWW Software, DENED-Ergo, http://www.deneb.com

[31] WWW Software, Kinetix-3DS Max-Character Studio, http://www.ktx.com

[32] WWW Software, MusculoGraphics - SIMM, http://www.musculographics.com

[33] WWW Software, Protozoa-ALIVE, http://www.protozoa.com

[34] WWW Software, RAMSIS-TECMATH's 3D, http://www.techmath.com

[35] WWW Software, Safework, http://www.safework.com/Saf-4Dnav/release.html

[36] WWW Software, Technomatix-ROBCAD, http://www.tecnomatix.com/robc1.htm#6

[37] WWW Sofware, The Motion Factory, http://www.motion-factory.com

[38] WWW Motion Capture, ACTISYSTEM, http://www.actisystem.fr