M4-4. Intermediate Advanced MPEG-4 animation system

ViSiCAST European Project

Project Manager: Pr. Françoise PRETEUX

Technical consultant: Eng. Marius PREDA

Unité de Projets ARTEMIS[1]

Institut National des Télécommunications

9 rue Charles Fourier, 91011 EVRY CEDEX

tel : +33 (0) 160764657

fax : +33 (0) 160764381

e-mail : francoise.preteux@int-evry.fr

marius.preda@int-evry.fr

[1] *Advanced Research & TEchniques for Multidimensional Imaging Systems* (www-artemis.int-evry.fr)

1 Problem statement and objectives

The objective of the ViSiCAST Project is to produce adaptable communication tools allowing sign language communication where only speech and text are available at present. These tools will be based on advanced technology for the synthetic generation, transmission, and storage of sign language to be developed by the project.

The issue of face and body animation and coding has been addressed by the Face and Body Animation (FBA) Subgroup of the Synthetic and Natural Hybrid Coding Group (SNHC) of MPEG. Generally, virtual actors can be portrayed visually as 2D icons, cartoons, composited video, 3D shapes or full 3D bodies. The MPEG-4 standard provides tools for the efficient compression of the animation parameters associated with the H-Anim 3D articulated human body model. We emphasise that the MPEG-4 standard aims only at efficiently coding the animation parameters; thus, it does not standardise any specific model, nor the methods for extracting animation parameters. However, tools for gesture analysis or synthesis are of a great interest.

This document reports on the tools developed by INT-ARTEMIS within the ViSiCAST Project for MPEG-4 avatar animation. They comprise two authoring tools (avatar modelling and animation editor and player) and a server-client communication system which are described.

2 MPEG-4 avatar animation tools developed by INT-ARTEMIS

Two authoring tools and a server-client communication system have been developed in order to implement and evaluate the FBA virtual avatar animation framework.

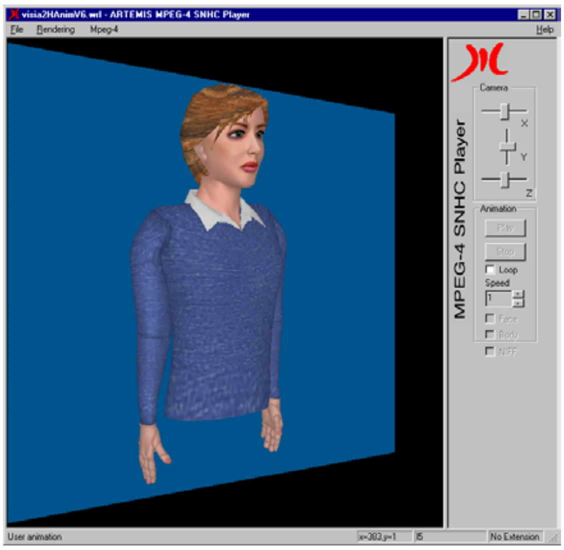
2.1 Body Modelling and Animation Authoring Tools

2.1.1 FBA compliant avatar modelling

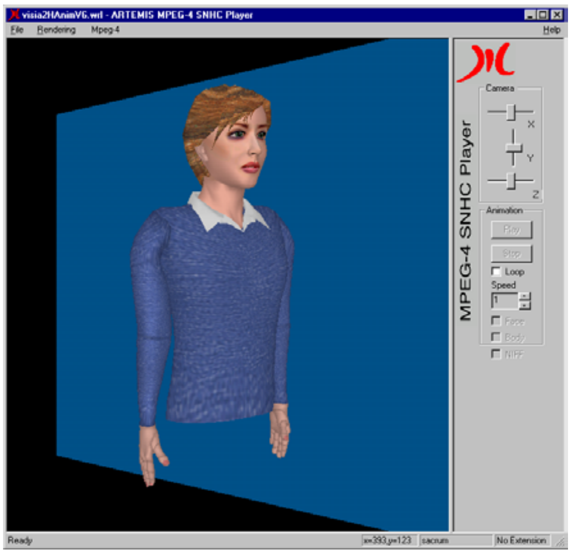
An MPEG-4 FBA compliant virtual avatar is a segmented character. The character definition, strongly based on H-Anim V2.0 specifications, uses scene graph nodes, namely, *Humanoid* node, *Joint* node, *Segment* node, and *Site* node. In addition, the set of human body joints and anatomical segments (names and hierarchical topological graph) is standardised and available. Based on a mesh propagation algorithm, we have designed and implemented a human modelling interface which allows to easily transform a continuous 3D mesh representing a human body into an FBA compliant avatar.

2.1.2 FBA compliant avatar animation

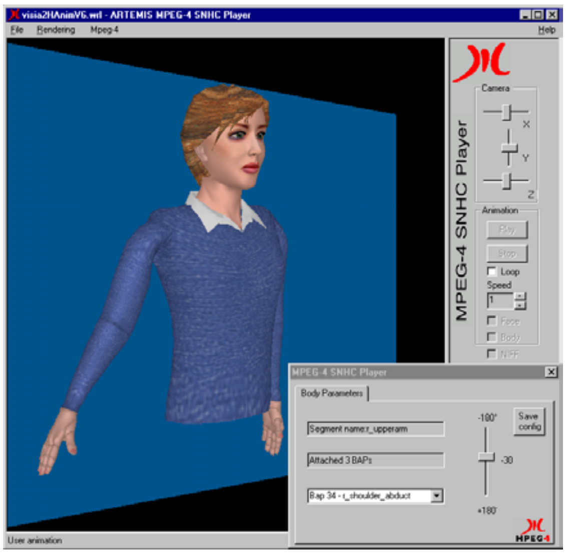
The key concept underlying animation parameters (referred to as BAPs) representation is the orientation of any anatomical segment as the composition of elementary rotations, namely twisting, abduction and flexion. 296 angular joint values are enough to specify any 3D pose of a virtual avatar. Angular values are specified with respect to the local 3D coordinate system of the anatomical segment. The origin of the local coordinate system is defined as the gravity centre of the Joint Contour common to the considered anatomical segment and its parent. The rotation planes are specified (Figure 1) and anatomical segment rotation axes are standardised (Figure 2).



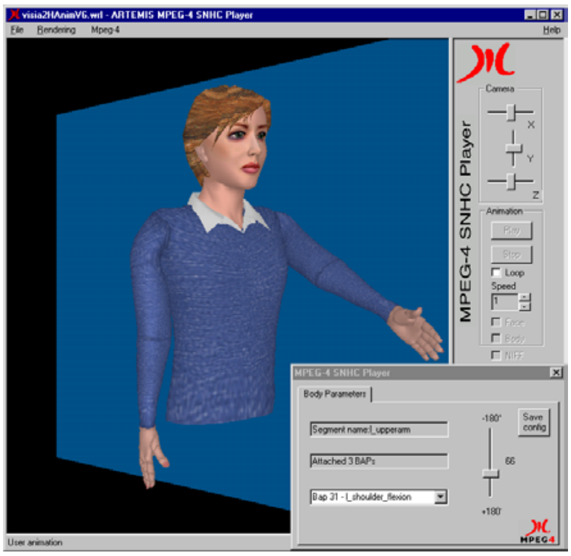
a) Coronal plane (CP)



b) Sagital plane (SP)



c) Left arm motion in coronal plane (BAP #34)



d) Right arm motion in sagital plane (BAP #31)

Figure 1 - Arm rotations with respect to standardised rotation planes.

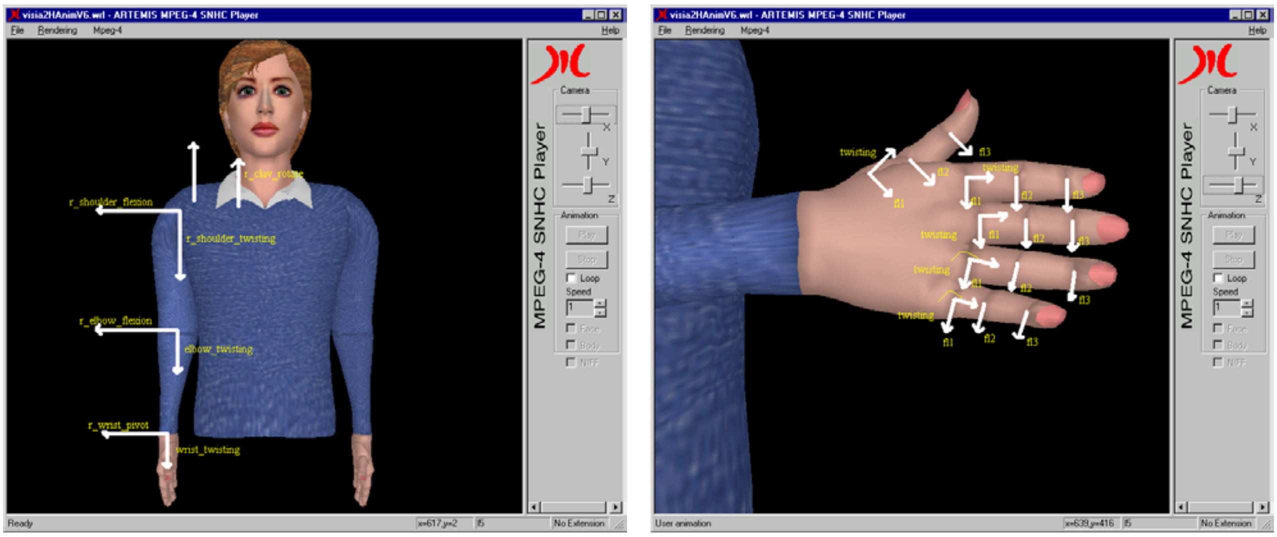


Figure 2 – Standardised rotation axes attached to shoulder, elbow, wrist and fingers.

In order to facilitate BAPs editing, the so-called *ARTEMIS Avatar Animation Interface* (3AI) has been developed by INT-ARTEMIS (Figure 3). A preliminary version has been delivered by ARTEMIS to the consortium members.

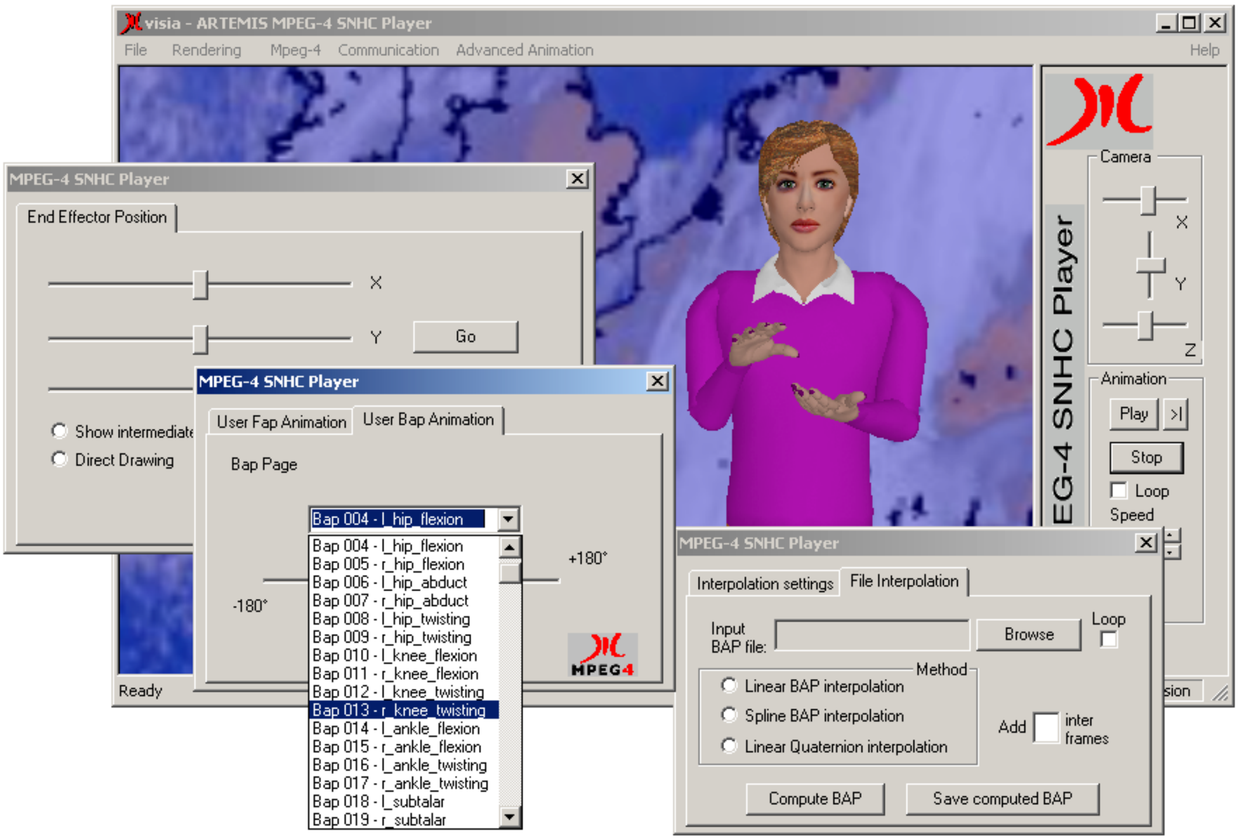


Figure 3 - The ARTEMIS Avatar Animation Interface.

3AI is a C++ user-friendly interface, available on X11/Motif and Windows platforms, which offers the following functionalities: (1) BAP editing, including basic and advanced instantiation techniques as linear, spline and spherical interpolation and inverse kinematics; (2) 3D compositing of objects such as images, video sequences, human body models or anatomical part models (hand, arm), and 3D scenes; (3) calibration of the 3D body model according to the anthropometric characteristics of the actor in the video sequence (size of palm, length of fingers...); (4) interactive extraction of BAPs specifying any gesture posture or corresponding to the posture shown in the video sequence; (5) animation of the virtual human model according to a BAPs file resource.

2.2 FBA compliant avatar compression and transmission

2.2.1 Animation Parameters encoding

On the one hand, the independence between a generic 3D model and the BAP description makes it possible to avoid 3D model transmission during animation. On the other hand, BAP encoding ensures a very low bit rate transmission. Two encoding methods, namely predictive and DCT-based encoding, are included in the MPEG-4 standard.

In predictive encoding, BAPs are quantised and coded by a predictive coding scheme. For each parameter to be coded in frame *n*, the decoded value of this parameter in frame *n-1* is used as a prediction. The prediction error is then encoded by arithmetic coding. This scheme prevents encoding error accumulation. Since BAPs can be assigned with different precision requirements, different quantisation step sizes are applied. They consist of a local (BAP specific) step size and a global one (used for bit-rate control). The quantised values are passed to the adaptive arithmetic encoder. The coding efficiency is increased by providing the encoder with range estimates for each BAP. We have tested the MPEG-4 animation parameters encoding schemes on BAP data corresponding to the ASL alphabet. Related to the motion complexity, we note that all the signs are performed with one hand and the avatar body position and orientation do not change.

The DCT-based coding method splits BAP time sequences into BAP segments made of 16 consecutive BAP frames. Encoding a BAP segment includes three steps achieved for all BAPs: (1) computing the 16 coefficient values by using discrete cosine transform (DCT), (2) quantising and coding the AC coefficients and (3) quantising and differential coding the DC coefficients. The global quantisation step Q for the DC coefficients can be controlled and, the AC coefficients global quantisation step is 1/3 from Q. The continuous component coefficient (DC) of an intra-coded segment is encoded as it is and, for an inter-coded segment, the DC coefficient of the previous segment is used as a prediction of the current DC coefficient. The prediction error and alternative component coefficients (AC), (for both inter- and intra-coded segments), are coded by using Huffman tables.

The implementation of the two encoding schemes is ongoing.

2.2.2 Server-client communication system

In order to transmit the FBA MPEG-4 data over the World Wide Web and MPEG-2 channels, a UDP client-server application has been developed by INT-ARTEMIS.

The functionalities of the server are as follows:

- BAP transmission (the user can choose to send data to a given IP address or to broadcast over the local network; the transmission frame-rate can also be controlled).
- Transmission pause/resume at any time.

In addition, a dedicated UDP client module has been added into the MPEG-4 FBA Player, also developed by INT-ARTEMIS. The control allows to open a listen port and wait until the connection to a push server is established. When a push server sends animation data, the client calls the MPEG-4 FBA decoder and starts the animation process.

3 Conclusion

We have presented in this report the current status of MPEG-4 tools related to virtual human animation developed by ARTEMIS within the ViSiCAST project. Namely, two authoring tools (avatar modelling and animation editor and player) and a server-client communication system have been developed and delivered to the consortium members.

These developments are currently being extended into a complete MPEG-4 virtual character animation system, capable of performing low bitrate animation in a variety of distributed communication frameworks, including the Internet and MPEG-2 channels.

4 Annex: Delivered software

4.1 Face & Body Animation MPEG-4 SNHC compliant player.

The player allows animating H-Anim compliant avatars.

Once an avatar is loaded, three animation methods are supported:

1. directly manipulating the animation parameters via the "Edit MPEG-4 Animation Parameters" dialog.
2. using a local animation parameter file.
3. allowing the reception of broadcaster data by using "Communication" dialog control. This functionality is enabled by the UDP Push Server application provided by INT in the framework of ViSiCAST Work Package One.

The MPEG-4 compliant version of the *Visia* avatar -the reference avatar for the ViSiCAST Project- is preloaded in the software as the default avatar.

4.2 Flexible UDP Server for BAP transmission

The application allows selecting two kinds of data transmission :

1. directly to specified IP address,
2. broadcast into LAN.

The user must specificity the client receiving port.

The "Socket" part of controls allows specifying the size of the socket in bytes. Whenever the "Full socket" option is not checked, the server sends one frame in each socket and the size of the socket will be the size of the frame. If the "Full socket" is checked, the server will add bytes in the socket until the socket size is equal to the value indicated by the user.

The server allows transmitting textual messages or animation files. Also, the transmission frame rate can be modified during transmission. The "Send" button starts the transmission which can be paused or resumed at any time.