WP1Television and Broadcast Transmission

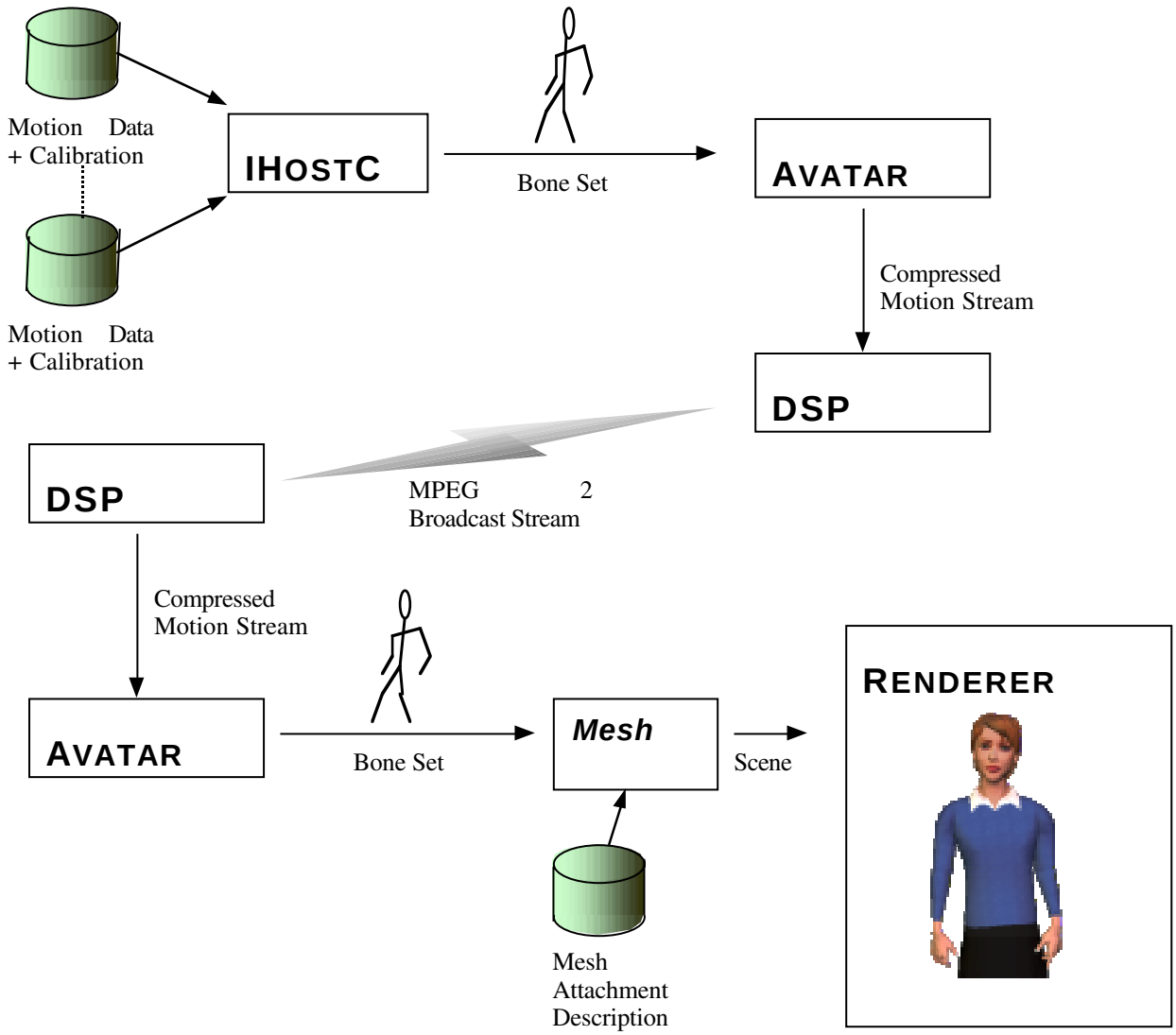
Progress against the Workplan

For the early deliverable in this WP (Direct TV Transmission of Virtual Human Signing), it was agreed by the consortium to build a system based on the Televirtual Mask-VR animation system. The adaptation of the core animation system necessary to permit this is described in the WP4 report.

In the Broadcast system, the transmitter essentially uses the BoneSet as derived from IHOSTCOM (see WP4) to generate a compressed representation of the bones for each frame. The receiver reproduces the BoneSet from this compressed format and uses it to animate an avatar.

Televirtual, conducting the work within WP4, has also co-operated in early trials of this system, designed to mimic a Broadcast environment. This involved setting up a signing server PC – representing the TV studio or transmission end, and a signing client – representing the consumer and the STB (Set Top Box) likely to be used to decode the signing in the home. This system has been used to test the feasibility of a one-way data feed and it's robustness in the face of a TX interruption.

The following diagram illustrates were the Broadcast transmission layer has been inserted in the Mask-VR pipeline.



Within the framework of Work Package 1, the ARTEMIS Project Unit (APU) at the Institut National des Télécommunications (INT), has released version 1 of a MPEG-4 video codec in June 2000. Preliminary results, including evaluation of coding performance and perceptual quality with respect to MPEG-2 compression, have been presented during on the 3rd General Work Package Meeting in Holland (28-30 June 2000). Developments have focused on coding efficiency adaptation with respect to motion amplitude. Specifically, two distinct codecs have been

released, respectively adapted to encoding sequences with small and large motion amplitudes. In each case, the superiority of MPEG-4 technologies over MPEG-2 compression in term of bit-rate has been established: MPEG-4 compression of a MPEG-2 stream with high video quality yields bit-rate reduction with a factor varying from 2 to 9. Depending on scene features (in particular, motion activity and texture complexity), the resulting bit-rate for a PAL sequence with standard frame size and rate varies from 500kbps to 2.5Mbps.

APU is currently working on version 2 which will extend version 1 by integrating video object-based coding functionalities, including (i) separate (spatial and motion) encoding of individual moving video objects throughout the sequence, and (ii) coding adaptivity via an automatic partitioning mechanism into frame groups with globally high or low motion activity to be processed by the corresponding codec. Preliminary results obtained with version 2 indicate that a bit-rate reduction could be achieved while maintaining a constant video quality.

Moreover, for small frame dimensions, MPEG-4 compression proves to yield bit-rates compatible with MPEG-2 Supplementary Packets (SP) size requirements. Hence, the idea of encoding avatar animation as a ***video*** MPEG-4 stream and to embed it into MPEG-2 SP. This leads to propose two additional solutions for TV broadcasting which maintain full compliance with the MPEG-2 standard, allowing therefore to take advantage of existing hardware/software equipment :

Broadcast video as a MPEG-2 stream, and embed avatar animation, encoded as a ***video*** MPEG-4 stream, as MPEG-2 SP.

Convert the original MPEG-2 video ***and*** avatar animation into two ***video*** MPEG-4 streams, and embed them into a single MPEG-2 stream as two MPEG-2 SP.

In both schemes, MPEG-2 is used as a transport mechanism that takes care of all system requirements (packetizing, multiplexing, synchronisation), avatar animation being combined with video data at the decoding stage. This requires only the insertion of a video MPEG-4 coder/decoder prior/after the MPEG-2 multiplexer/demultiplexer. Solution 2 is clearly superior to Solution 1. The main limitation of these schemes relies on the assumption of video encoding of the animated avatar, which considerably restrains scene composition possibilities at the decoder side.

WP1 is therefore examining a range of possible approaches to broadcasting realistic Virtual Humans, seeking both to exploit emerging standards, while retaining the best possible quality.