ViSiCAST Technical Meeting

INT - 17th November 2000

Participants :

UEA	Nunio Dionisio - Ralph Elliott - John Glauert – Richard Kennaway – Silko Kruse
TV	Sanja Rankov
UH	Thomas Hanke – Constanze Schmaling
IRT	Werner Brückner
INT	Marius Preda - Françoise Prêteux – Nicolas Rougon

Meeting agenda :

1. Project status on animation
2. MPEG-4 SNHC avatar and player
3. Motion data formats: technical overview
4. SiGML: technical overview
5. Future plans

1. Project status on animation

INT summarises current project status on animation (*see PowerPoint slides*)

- TV developments consist of:
  - ✓ generation of avatar models (NewDan, Visia),
  - ✓ high-quality motion capture system (delivering MaskVR data),
  - ✓ implementation of avatar players (MaskVR ALL player, IHost (DirectX), ActiveX Control plug in + specialised broadcast-intended version (OpenGL)).

  The so-called NIF (New Interface Format) motion format, consisting of compressed bone sets, has been devised by Farzad Pezeshkapour to facilitate data exchange between UEA and TV.

- INT developments follow a standardisation path and consist of:
  - ✓ generation of MPEG-4 compliant avatar from TV models (MPEG-4 NewDan and Visia),
  - ✓ implementation of a MaskVR (EXR) to MPEG-4 animation parameters translator,
  - ✓ implementation of a MPEG-4 SNHC avatar player (OpenGL).

- UEA and UH developments follow a standardisation path and consist of the definition of the SiGML description language on the basis of the Hamnosys notation system.

2. MPEG-4 SNHC avatar and player

INT demonstrates version 1.1 of its MPEG-4 SNHC avatar player (version 1.0 has been presented at the Munich general workpackage meeting). This version features include :

- real time animation of MPEG-4 avatar from BAPs and FAPs data (demonstrated on MPEG-4 Visia driven by home made BAP / FAP data generated by manual editing from a video signing sequence provided by TV. Motion files);
- animation speed tuning via BAP / FAP subsampling / oversampling (demonstrated with a subsampling factor = 2);
- manual BAP / FAP editing for each MPEG-4 segment;
- virtual lighting and camera;

- VRML background scene inclusion.

In its current version, the player can render at 21 fps in full speed and 26 fps in double speed (using a 633 Mhz PC equipped with an OpenGL graphic board) (video including HDTV is 25 fps which can be regarded as acceptable).

UH found that animation rates above 15 fps are OK for most signs; 12 fps to 15 fps shows degradation; below 10 fps is becoming unacceptable. UH distributes video at 12.5 fps.

Information is provided on relationships between VRML, HAnim and SNHC :

- SHNC has incorporated HAnim, added new functionalities and modified some features. Joint efforts are currently made by SHNC and HAnim to restore compatibility (hence, HAnim will become a true subset of SNHC).
- VRML is only partially included in SNHC. HAnim is actually a subset of the intersection between SNHC and VRML. The player implements a VRML scene reader and can load a VRML avatar model. However, animation from BAP/FAP data is then not possible since it requires preliminary decomposition into MPEG-4 segments.

Richard Kenneway has developed a NIF (bones) to VRML translator. For interpolating motion, data are dropped to 3 significant figures, then frame data for a particular bone where the motion is less than a millimetre are dropped. 15s of motion (94 frames) becomes a 220k VRML file (including avatar code) from an original 916k bone file (78 bones) (which reduces to 610k at 3 significant figures).

Richard Kenneway reports that a couple of bones in the head seem to be in the odd directions.

Action: UEA will give feedback to TV to investigate anomalous bone data and correct software bugs.

3. Motion data formats: technical overview

Sanja Rankov presents a technical overview of motion data formats:

- The motion capture system delivers MaskVR ALL data. These data are fed into the various TV players. ALL files consist of 3 streams (for hands, body and face) and comprise raw motion data together with separate calibration information (provided as pointers to external calibration files). Streams are connected through a high level combination file. Raw data and calibration information can be accessed through the ActiveX control player.
- NIF files are compressed bone files. A NIF frame consist of 87 bones with 13 floats:
  - ✓ Length;
  - ✓ X axis (x,y,z coords – local)
  - ✓ Y axis (x,y,z coords – local)
  - ✓ Z axis (x,y,z coords – local)
  - ✓ Origin (x,y,z coords - global)

Françoise Prêteux proposes to place some standard data on a central server.

Sanja Rankov says that broadcast data can be shared: ALL files and Video are available.

Action: UEA will provides up to a GB of server space.

It is agreed that motion capture file format and animation file format echo to specific tasks and therefore must distinguished.

Action [required by TV]:

- ✓ NIF files will be the basis for the standard motion data archive
- ✓ NIF files need scaling factors for spatial and temporal units
- ✓ NIF files need to represent blend points

It is suggested to try also to get simultaneously:

- ✓ Video of a capture session
- ✓ Motion data (raw or cleaned)
- ✓ Screen dump (raw or cleaned)

Action [required by TV / + 1 week]:

- ✓ Written NIF file documentation
- ✓ 5 cleaned NIF files (approximately 15 seconds each) + capture session video

Action [required by TV / + 3 weeks]:

- ✓ 5 extra cleaned NIF files + capture session video

Action [TV + UEA + INT]:

- ✓ Investigate ways of providing avatar screen dump as digital video

4. **SiGML: technical overview** *(PPT slides to be distributed)*

Ralph Elliott presents an overview of SiGML for the core of Hamnosys along the lines of the ASSETS'00 presentation.

SiGML main concepts (for manual gestures) consist of :

- Hand configuration:
  - ✓ Hand shape (hundreds of them)
  - ✓ Hand orientation (finger base, palm). Orientation space is quantized (26-connectivity).
- Location in signing space:
  - ✓ Positions on hand and body
  - ✓ Modifiers indicates *position* on the left-centre-right spectrum and *contact* distance (touching, close, normal, far)
  - ✓ Positions on (non dominant) arm and hand
- Motions:
  - ✓ Absolute i.e. *targeted:* give new hand position and/or configuration
  - ✓ Relative i.e. *directed:* motion direction from initial configuration, implicit target
  - ✓ Composition: sequence, repetition, concurrency
  - ✓ Modality: speed, acceleration, tension

SiGML is mainly semantic and is actually not very precise in a geometrical sense (a certain amount of variations are permitted).

Richard Kenneway demonstrates its work on SiGML driven avatar animation in the HAnim framework. Contact detection is implemented but should be refined.

John Glauert says that further developments of Richard's work will be based on TV avatars.

Action [required by UEA (Richard Kenneway) / + 1 week]:

- ✓ Documentation on current version of SiGML. To form the starting point for revision by WP5.

5. Future plans

It is important for all WPs concerned with animation to be rapidly delivered a first specification of SiGML:

- ✓ Short term: SiGML animation is required. Route is not important.
- ✓ Long term: ViSiCAST expects SiGML to be the interface to all avatars. It is desirable to achieve *independence* of avatar definitions.

WP4 will interact with WP5 and provide feedback on SiGML semantics adequacy and completeness for animation purpose (*see PowerPoint slides*).