3D modeling for multimedia applications

Patrick HORAIN

Institut National des Télécommunications
Signal & Image Processing Department
9, rue Charles Fourier, 91011 Evry Cedex, France
Patrick.Horain@int-evry.fr
http://www-sim.int-evry.fr

1. MPEG-4 context

1.1 The MPEG-4 standard

Image coding standards usually focus on storing and transmitting signals efficiently, in a frame based approach. The coming MPEG-4 standard [1] adds a new challenge, namely bringing new services to image communication users. It moves towards representing a scene as a composition of audiovisual objects rather than "just" a collection of pixels.

MPEG-4 objects can be captured naturally by one (or more) camera(s), or be synthesized by a computer. They can be 2D or 3D. Their composition is called Synthetic / Natural Hybrid Coding (SNHC).

The MPEG-4 SNHC standard is being extended to support human face and body representation. Face and body animation parameters (FAPs and BAPs) have been defined.

1.2 MPEG-4 core experiment on hand animation

INT / SIM is involved with EPFL / LIG in validating the BAPs exchange format [2].

Hand animation sequences in the MPEG-4 format have been produced independently by each partner, and re-played successfully by the other (figure 1) [3, 4].

2. 3D object registration

INT / SIM contributes to a France Telecom / CNET project on scene representation from image sequences. Other partners are INRIA / SYNTIM and INRIA / TEMIS.

This project aims at restituting a 3D representation of a dynamic scene from an uncalibrated video image sequence. The scene is assumed to be composed of objects for which a generic deformable model is available, surrounded by an unknown environment. The unknown fixed areas can be reconstructed using projective techniques. Known objects are detected and identified by an indexation based method applied on the moving areas [5].

Figure 1 : MPEG-4 compliant interface for gesture animation [3].

In this project, INT / SIM is in charge 3D objects registration in image sequences. Knowing the 3D object model defined as a mesh or as a set of segments with adjacency properties, the 3D / 2D matching is achieved by minimizing the distance between the projected segments of the 3D model and the image edges [5, 6, 7].

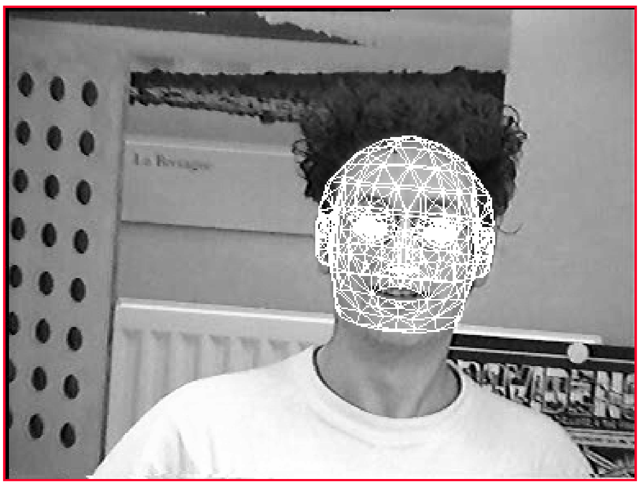

Figure 2 : 3D object registration on images [5].

3. Hand modeling and tracking

Gesture analysis can be of interest for such applications as low bit rate gesture transmission, human-machine interaction by gesture for object designation or manipulation in space, and sign language recognition.

Our work aims at analyzing gestures from a single image sequence and without any marker on the body of the operator. To process 3D gestures possibly with auto-occlusions, we take advantage of a generic hand model that is matched against image data in an analysis-synthesis procedure.

The generic hand model consists of articulated primitives, either meshes or simple geometrical volumes. Its dimensions are adjusted to the user's morphology from an image of his hand open and perpendicular to the camera axis, with respect to morphological constraints. The hand silhouette is extracted from images by thresholding the skin color, and fingers are separated by morphological filtering [8].

The model is animated with biomechanical constraints from this initial pose to reproduce movements observed in a video sequence. Joint angles and hand pose are adjusted to match the projected model contours with image edges. A Kalman filter is used to regularize the tracking process [9].

So far, this procedure has been experimented with translation and abduction / adduction image sequences. It will be extended to other finger movements.

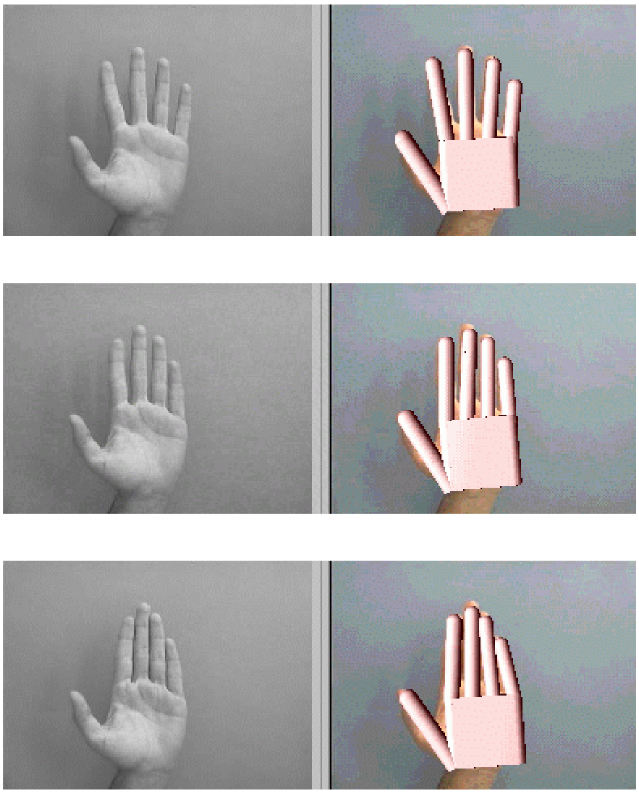

Figure 3 : Hand registration and tracking on some images of the abduction / adduction sequence (28 frames) [9].

4. References

[1] R. Koenen, F. Pereira, L. Chiariglione, "MPEG-4 : Context and Objectives", *Signal Processing : Image Communication* (9) 4, Special issue on MPEG-4, 1997, pp. 295-304 ; http://www.cselt.stet.it/ufv/leonardo/icjfiles/mpeg-4_si .

[2] F. Prêteux, P. Horain, H. Ouhaddi, M. Preda, *Report on Core Experiment 3 on Hand Baps interpretation*, ISO/IEC JTC1/WG1 MPEG97/M3332, Tokyo, Japan, March 1998 ; http://www-sim.int-evry.fr/Publications.

[3] F. Prêteux, G. Mozelle, M. Preda, *INT-SIM : MPEG-4 SNHC Core Experiment "Hand Animation"* ; http://www-sim.int-evry.fr/Research/MPEG_Core.html.

[4] T. C. Capin, and J. Esmerado, *EPFL-LIG's MPEG-4 Body Animation Page* ; http://ligwww.epfl.ch/mpeg4 .

[5] L. Oisel, F. Fleuret, P. Horain, L. Morin, J-M. Vezien, F. Prêteux, A. Gagalowicz, C. Labit, P. Leray, "Analyse de séquences non calibrées pour la reconstruction 3D de scène", *Actes du 11ème Congrès AFCET / RFIA (RFIA'98)*, Clermont-Ferrand, France, vol. 1, janvier 1998, pp. 189-198 ; http://www-sim.int-evry.fr/Publications.

[6] M. Bourges-Sévenier, P. Horain, F. Prêteux, P. Leray, "Recalage d'un modèle générique sur une séquence d'images 2D", *Actes des 3èmes Journées d'Etudes et d'Echanges "COmpression et REprésentation des Signaux Audiovisuels" (CORESA 97)*, CNET / France Telecom, Issy-les-Moulineaux, France, 26-27 mars 1997, pp. 163-171 ; http://www-sim.int-evry.fr/Publications.

[7] F. Prêteux, S. Curila, M. Malciu, "3D object registration in Image Sequences", *SPIE Proceedings Vol. 3304 "Nonlinear Image Processing IX"*, Editors : E. R. Dougherty, J. T. Astola, San Jose, CA, USA, January 1998, paper 3304-20, pp.186-196.

[8] H. Ouhaddi, P. Horain, "Conception et ajustement d'un modèle 3D articulé de la main", *Actes des 6èmes Journées de Travail du GT Réalité Virtuelle*, France Telecom / CNET , Issy-les-Moulineaux, France, 12-13 mars 1998, pp. 86-90 ; http://www-sim.int-evry.fr/Publications.

[9] H. Ouhaddi, P. Horain, K. Mikolajczyk, "Modélisation et suivi de la main", *Actes des 4èmes Journées d'Etudes et d'Echanges "COmpression et REprésentation des Signaux Audiovisuels" (CORESA 98)*, France Telecom / CNET, Lannion, France, 9-10 juin 1998 ; http://www-sim.int-evry.fr/Publications.