Statistical Gesture Models for 3D Motion Capture

Zhenbo LI[1], Patrick HORAIN[1], André-Marie PEZ[2,3], Catherine PELACHAUD[2,3]

[1] Institut TELECOM; TELECOM & Management SudParis,
9 rue Charles Fourier, 91011 Evry Cedex, France

[2] IUT de Montreuil, Université de Paris 8,
140 rue de la Nouvelle France, 93100 Montreuil Cedex, France

[3] Institut TELECOM; TELECOM ParisTech,
46 rue Barrault, 75634 Paris Cedex 13, France

zhenboli@gmail.com, Patrick.Horain@IT-SudParis.eu,
andremarie.pez@telecom-paristech.fr, catherine.pelachaud@telecom-paristech.fr

Keywords: Motion capture, Gesture model, Gaussian Process Dynamical Models

1. Introduction

Monocular video based markerless 3D motion capture has the intrinsic limitation of poor precision and robustness. Many conditions like occlusions, image blur, poor image resolution, depth ambiguities etc. often confuse the tracker.

In this paper, we propose using statistical gesture models as guides to constrain and disambiguate 3D tracking for a set of human motions. A set of example gestures with variations was initially generated using Greta [1] expressive conversational agent. Then gestures are statistically described with Gaussian Process Dynamic Models (GPDM) [2] and used as constraints while tracking specific gestures. The interest of this approach is demonstrated with results.

2. Generation of example gestures

The library of gestures we used for the learning phase of our tracking algorithm is composed of animation sequence of a 3D virtual agent [1]. The gestures of the agent were defined by analyzing a corpus of human dyads that converse together. In this corpus we looked for gestures that appear repetitively. Every person has one's own behaviour style. Some person gesticulates a lot, another does very slow gestures, another ample ones, etc. To encompass such a variety of behaviour expressiveness in our library, we have modulated each gesture of the agent with a set of expressive parameters [1]. Thus for each gesture reproduced from the video corpus, we had a large variety of 'similar' gestures.

3. Gesture modeling and 3D motion capture

Gaussian process dynamic models (GPDM) [2] are a powerful approach for probabilistically modeling high dimensional time related data through dimension reduction. It can learn probability motion models from small training data sets as a mapping between the full-dimensional data space and a low-dimensional latent space

and a dynamical model in the latent space. Fig.1 gives two conversational gestures from our gesture library and their corresponding gesture trajectories in 3D latent space. Each circle on the trajectories corresponds to a body pose.

Those gesture models were used as constraints while tracking specific gestures [3]. Each candidate pose is projected into the latent gesture trajectory, so enforcing the captured motion regularity.

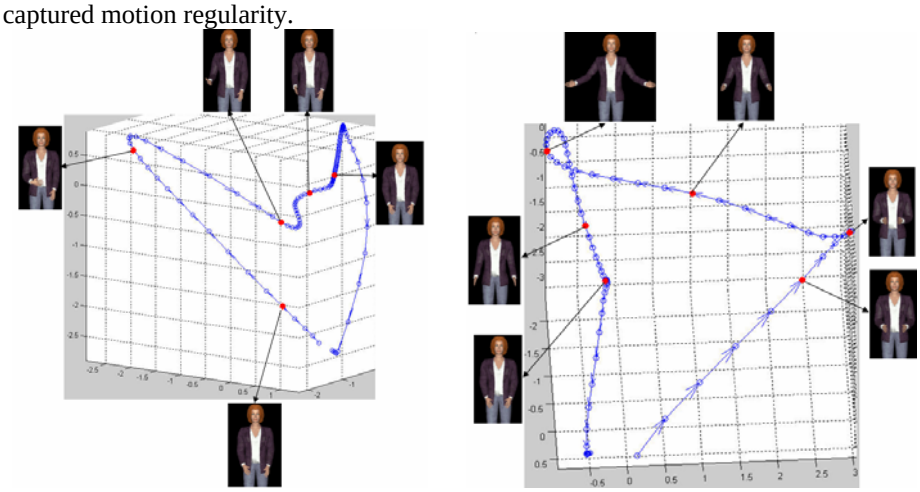

Fig.1 Two conversational gestures and their corresponding gesture models in 3D latent space.

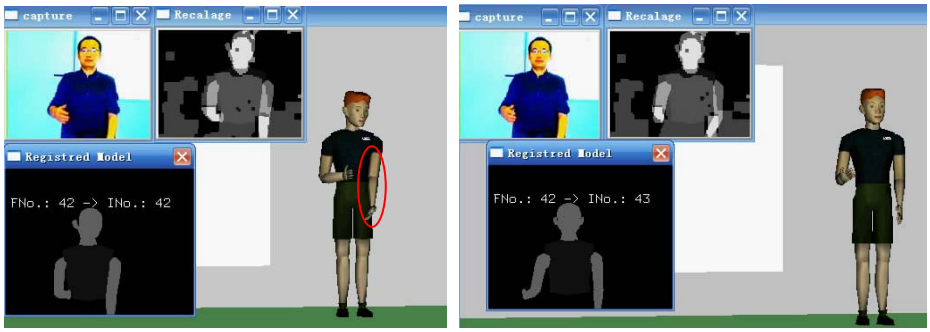

Fig. 2 Gesture models as constraints in monocular video based 3D motion capture. **Left:** captured motion without using gesture models. **Right:** captured motion using gesture models

References

1. E. Bevacqua, M. Mancini, R. Niewiadomski, C. Pelachaud, "An expressive ECA showing complex emotions", *AISB'07 Annual convention, workshop "Language, Speech and Gesture for Expressive Characters"*, Newcastle UK, pp. 208-216, April 2007. http://linc.iut.univ-paris8.fr/greta.
2. J. M. Wang, D. J. Fleet, A. Hertzmann, "Gaussian Process Dynamical Models for Human Motion". In *IEEE Trans. PAMI*. Vol. 30, No. 2. Feb. 2008. pp. 283-298. http://dx.doi.org/10.1109/TPAMI.2007.1167.
3. David Antonio Gómez Jáuregui, Patrick Horain, Fawaz Baroud, " Acquisition 3D des gestes par vision monoscopique en temps réel". In *Proceedings of MajecSTIC 2008*. Marseille, October 2008. http://www-public.it-sudparis.eu/~horain/Publications/MajecSTIC08_Gomez.pdf.