# Captured Motion Data Processing for Real Time Synthesis of Sign Language

Alexis Heloir[1], Sylvie Gibet[1], Franck Multon[2], and Nicolas Courty[1]

[1]Université de Bretagne Sud, Laboratoire Valoria, Bâtiment Yves Coppens, BP 573, 56017 Vannes Cedex, France
[2]Université de Rennes 2, Laboratoire de Physiologie et de Biomécanique de l'Exercice Musculaire, Av. Charles Tillon CS 24414, 35044 Rennes, France
[1]firstname.surname@univ-ubs.fr, [2]franck.multon@uhb.fr

**Abstract.** This study proposes a roadmap for the creation and specification of a virtual humanoid capable of performing expressive gestures in real time. We present a gesture motion data acquisition protocol capable of handling the main articulators involved in human expressive gesture (whole body, fingers and face). The focus is then shifted to the postprocessing of captured data leading to a motion database complying with our motion specification language and capable of feeding data driven animation techniques.

**Issues.** Embodying a virtual humanoid with expressive gestures raises many problems such as computation-cost efficiency, realism and level of expressiveness, or high level specification of expressive gesture [1]. Here, we focus on the acquisition of motion capture data from the main articulators involved in communicative gesture (whole body, face mimics and finger motion). We then show how acquired data are postprocessed in order to build a database compatible with high level gesture specification and capable of feeding real time data-driven motion synthesis techniques. A recent automatic segmentation algorithm based on Principal Component Analysis (PCA) is then evaluated.

**Motion acquisition protocol.** The motion data acquisition protocol is designed to capture the whole range of articulators involved in order to produce human communicative gestures. This protocol relies on two complementary techniques, as shown in figure 1. The first technique aims at capturing facial and body motions and relies on a set of reflective markers placed on standardized anatomical landmarks and a network of 12 *Vicon-MX*[1] infrared cameras located all around the subject. The aim of the second technique is to capture finger motion thanks to pair of *Cybergloves*[2] measuring finger abduction and flexion. This technique is well-suited to finger motion, as it is robust by finger occlusions that may appear too often during the signing performance. The two sets of data acquired are post processed, synchronized and merged offline.

[1] `http://www.vicon.com/products/viconmx.html`
[2] `http://www.immersion.com/3d/products/cyberglove.php`

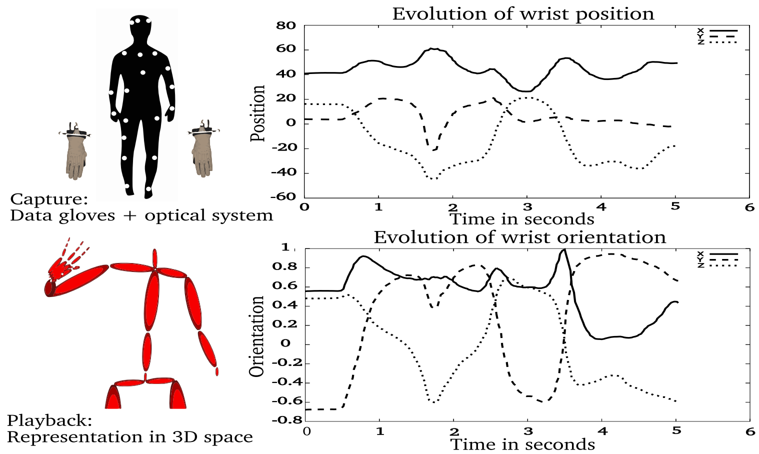


**Fig. 1.** motion acquisition protocol

The resulting dataset describes the evolution along $n$ frames of a skeleton hierarchy composed of $k$ joints. For each joint $i$, a set of $l_i$ degrees of freedom is defined, $1 \leq l_i \leq 3$. The size of a posture vector $v$ can then be easily computed.

$$size(v) = \sum_{i=0}^{k} l_i$$

Table 1 sums up the composition of a posture vector according to our representation model.

| segment | coord type | number of joints | DOF per joint | size of segment subvector |
|---|---|---|---|---|
| body | angular | 18 | $1 \leq l \leq 3$ | 54 |
| hand | angular | 18 | $1 \leq l \leq 3$ | 25 |
| total | angular | 36 | — | 79 |

**Table 1.** detail of a posture vector

**Processing motion data.** One of the early steps of data processing consists of segmenting motion into relevant chunks. Extracted chunks must be short enough to guarantee the synthesis of a wide range of new motions conveying sufficient meaning to comply with high level task oriented specification language [2].
Even though it has been shown that low level motion segmentation can be achieved in a straightforward manner [7][4], Hodgins and al. recently showed that higher level motion segmentation could be efficiently achieved thanks to the principal component analysis (PCA) approach. According to the results they

presented [6], PCA segmentation method applied on simple motions representing typical human activities, such as walking, running, sitting, standing idle, etc. achieved very good results: up to 80% precision call for simple body movements. This algorithm is based on the assumption that the intrinsic dimensionality of a motion sequence containing a single behavior should be smaller than the intrinsic dimensionality of a motion sequence containing multiple behaviors. Thus, from one motion sequence to another, the reconstruction error of the frames projected onto the optimal hyperplane of dimension $r$ increases rapidly, for a fixed $r$. Figure 2 illustrates the motion transition detection between two hand configurations.

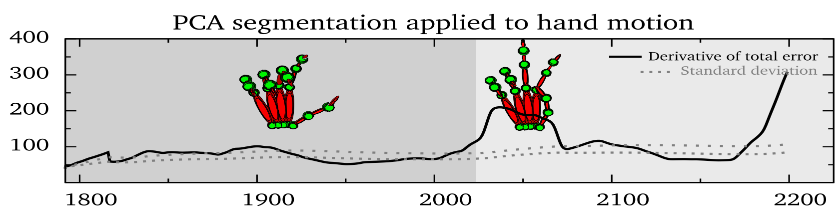


**Fig. 2.** automatic segmentation using PCA. Cut is performed when derivative of error reaches three standard deviation from the average.

**Evaluating the PCA approach to hand motion segmentation.** We apply the PCA-based segmentation algorithm allocated to a sequence representing a non signer subject finger spelling French dactylologic alphabet [3]. The sequence is 7200 frames long with 120 frames per second. To carry out PCA, decomposition is thus performed on a $7200 \times 25$ matrix extracted from the total motion data and representing the right hand motion. According to our experiments, the ratio $Er$ which indicates how much information is retained by projecting the frames onto the optimal r-dimensional hyperplane reaches acceptable range [6] when $r \leq 3$ for all the 20 segments we extracted manually from the alphabet spelling sequence. Further experiments led us to set the window parameter $k$ originally fixed at 2 seconds to 1.3 seconds, considering that hand motion is much faster than body motion.

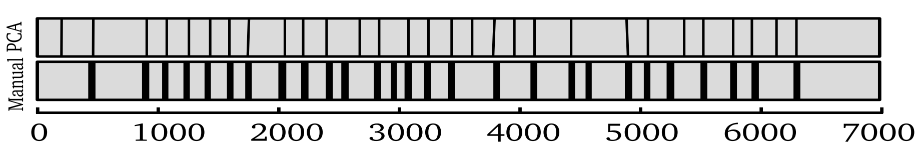


**Fig. 3.** segmentation results

**Results.** In parallel to automated PCA based hand motion segmentation, a human observer manually segmented the finger spelling sequence by identifying probable zones of motion transition. Figure 3 compares how the two methods designated motion separation zones. The human observer identified 27 zones while the PCA based motion segmentation algorithm identified 29 zones. Among those, 22 zones were overlapping.

**Conclusion.** We have presented a motion acquisition framework designed to manage several articulators involved in communicative gesture and in sign language performance. We then rely on the data provided by this framework to evaluate a recent automatic motion segmentation technique based on principal component analysis of hand motion. This method proves to be capable of solving high level segmentation required by our needs. In the near future, we wish to extend this technique to the whole of upper body motion. In parallel, we would like to provide a better evaluation framework based on data acquired and annotated by French sign language specialists. Such a framework will provide us with the grounds required to perform reliable motion analysis and performance comparisons.

**Acknowledgements.** This work is part of the RobEA HuGEx project and the Signe projet, respectively funded by the department STIC of CNRS and the regional council of Brittany (Ref. B/1042/2004/SIGNE). Motion capture presented in this document has been carried out in collaboration with the following research institutes: the LPBEM based at the Universit de Rennes 2, the LESP based at the Universit de Toulon et du Var and the LINC based at the Universit de Paris 8.

## References

1. Gibet S., Courty N., Multon F., Pelachaud C. and Gorce P.: HUGEX: virtual humanoids embodied with expressive communicative gestures. Proc. of Journées Bilan ROBEA (2005), 233–238, Montpelier, France.
2. Gibet S., Lebourque T., Marteau P. F.: High level Specification and Animation of Communicative Gestures. Journal of Visual Languages and Computing, **12** (2001), 657–687.
3. Moody B., La langue des signes - Tome 1 : Histoire et grammaire. International Visual Theatre (I.V.T.). Éditions Ellipses.
4. Kovar L., Gleicher M. and Pighin F.: Motion Graphs. Proc. of Int. Conf. on Computer graphics and Interactive Techniques (2002), 473–482, San Ãntonio, USA.
5. Keogh F., Palpanas T., Zordan V., Gunopulos D. and Cardle M.: Indexing Large Human-Motion Databases. Proc. of Int. Conf. on Very Large Data Bases (2004), 780–791, Toronto, Canada.
6. Barbič J., Safonova A.,Pan JY, Faloutsos C, Hodgins J.C. and Pollard N.S.: Segmenting Motion Capture Data into Distinct Behaviors. Proc. of Graphics Interface (2004), 185–194, London, Ontario, Canada.
7. A. Fod, M. J. Mataric, and O. C. Jenkins.: Automated derivation of primitives for movement classification. Autonomous Robots, **12** (2002) 39–54, .