

## **Système d'annotation et de segmentation de gestes de communication capturés**

Alexis Héloir, Sylvie Gibet, Nicolas Courty, Mickaël Raynaud

Samsara Valoria – UBS  
BP 573, 56017 Vannes Cedex  
prenom.nom@univ-ubs.fr

**Mots-clés :** Annotation, Synthèse du Geste, Capture de Mouvements, Segmentation, Indexation

**Keywords :** Annotation, Gesture Synthesis, Motion Capture, Segmentation, Indexation

**Résumé** Nous proposons un outil permettant l'annotation de gestes préalablement acquis par des méthodes récentes de capture du mouvement. La segmentation et l'annotation peuvent être réalisées selon différents niveaux de représentation du geste. Cet outil est conçu pour être utilisé aussi bien par des signeurs non informaticiens que par des spécialistes de l'animation par ordinateur. Dans le cadre de notre projet, cet outil a pour vocation la création d'une base de données de représentations numériques de gestes capturés capable d'alimenter des méthodes originales de synthèse et de spécification de gestes de communication. Nous pensons qu'il peut également s'avérer utile dans le contexte de l'analyse linguistique des langues signées.

**Abstract** We present an annotation tool handling motion capture gestures captured via recent techniques. Segmentation and annotation can be carried out according various representation levels. This tool is meant to be used by computer graphics experts as well as signing individuals with no computer oriented prerequisites. The main goal of this tool, according to our objectives, is to set up a coherent data base of signals. Data provided should support our investigation on designing original methods for specification and synthesis of expressive gestures. We also believe that such a tool can prove to be useful in the context of linguistic analysis of signed languages.

### **1 Introduction**

Les méthodes d'analyse et de segmentation du geste permettent d'indexer des bases de données de mouvement en introduisant des éléments structurels contribuant à enrichir les modèles de synthèse et de spécification du mouvement existants. Cette étape d'analyse nécessite des outils appropriés. Plusieurs projets ont mis au point ou sont en train de réaliser des logiciels d'annotation et de transcription de séquences gestuelles. Ces outils ont pour principal flux de données des séquences vidéo pouvant être enrichies de signaux tels que la direction du regard (Wittenburg, 2002) ou les configurations manuelles (Koizumi, 2003).

Bien que certains logiciels proposent des traitements évolués tels que la reconstruction de postures par analyse d'images (Dalle, 2001), à notre connaissance, aucune des solutions existantes n'offre la possibilité de travailler directement sur des représentations numériques du mouvement telles qu'elles sont obtenues par l'intermédiaire des outils de capture récents. Ce résumé décrit le logiciel de segmentation et d'annotation de mouvements capturés que nous sommes en train de développer en présentant dans un premier temps les signaux pris en compte et l'intérêt de tels signaux dans le cadre de nos recherches et dans le cadre de la linguistique. Nous décrivons ensuite la structure et le format des annotations. Nous concluons par les objectifs et les perspectives de notre travail.

## 2 Description des signaux acquis

Nous disposons d'un système d'acquisition composé d'un système optique Vicon réalisant la capture des mouvements corporels et de l'expression faciale synchronisés avec une paire de gants de données. Ce système permet de capturer l'évolution de la majorité des articulateurs entrant en jeu dans la langue des signes (corps, visage et mains). En sortie du système d'acquisition, nous obtenons des vecteurs de positions (angulaires ou cartésiennes) caractérisant l'évolution de chaque marqueur de mouvement au cours du temps (figure 1). Ces données sont structurées selon un squelette articulaire modélisant le corps et les mains de l'acteur, ainsi qu'un treillis relatif aux marqueurs disposés sur son visage.

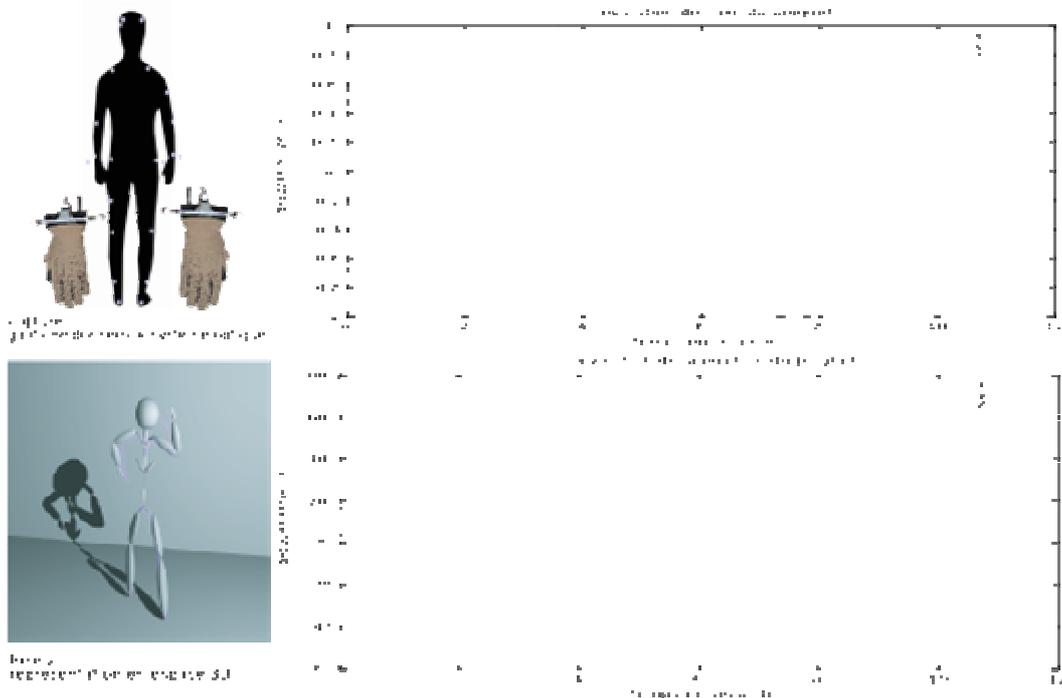


Figure 1 : signaux numériques capturés

L'outil d'annotation est en cours de développement. Un précédent prototype nous a permis d'évaluer sa faisabilité et de valider certains choix d'interface. Outre les fonctionnalités propres à la plupart des logiciels d'annotation existants, notre outil propose une représentation animée du mouvement dans un espace tridimensionnel permettant à l'utilisateur de choisir son point de vue. Cette représentation est accompagnée d'une vidéo de l'acteur enregistrée pendant la séquence de capture et synchronisée avec la visualisation dans l'espace

tridimensionnel. Enfin, notre outil permet l'affichage des trajectoires des signaux numériques éventuellement inclus dans les pistes d'annotation sous la forme de courbes intégrées aux lignes de temps des pistes d'annotation.

L'utilisation de la représentations numérique du geste offre la possibilité d'effectuer des traitements automatiques pour la segmentation (Barbic, 2004), facilitant ainsi le processus de segmentation manuelle et d'annotation.

Ce type de représentation permet également la création d'une base de données permettant des recherches multi-critères (spatialisation du geste, dynamique du geste, analyse en composantes principales, etc.).

### 3 Structure et format des annotations

La structure des fichiers annotés obéit à un format de données basé XML et est conçue pour répondre aux préoccupations de différents utilisateurs, signeurs, linguistes ou informaticiens. Le format d'annotation est composé de quatre sections (Figure 2).

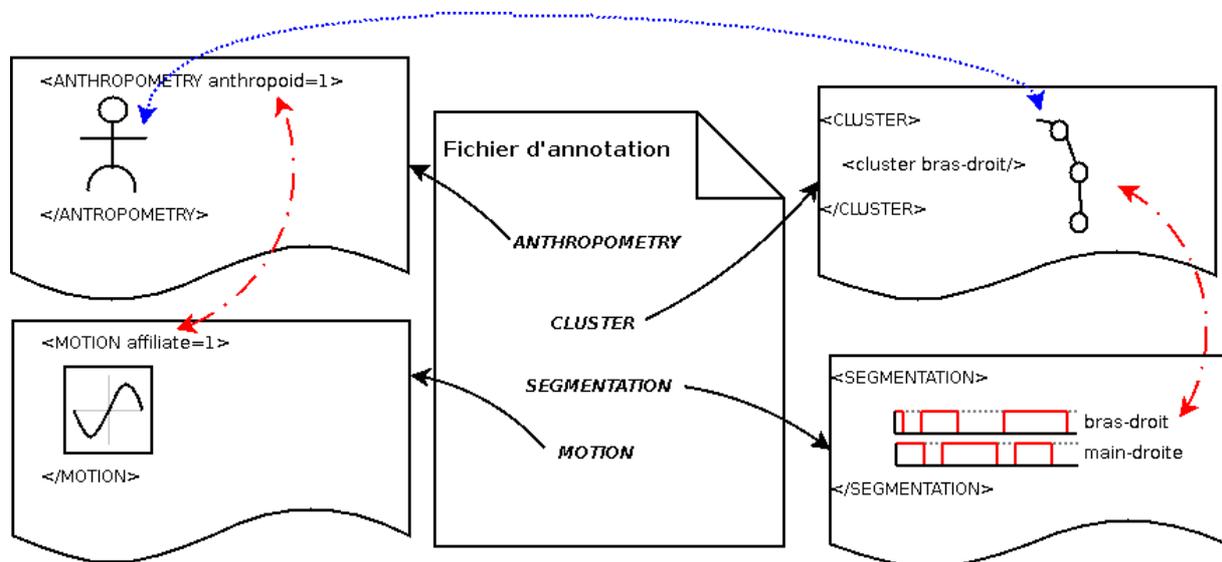


Figure 2 : Structure des annotations

La section ANTHROPOMETRY décrit de manière hiérarchique la morphologie du squelette articulaire de l'acteur (corps et mains pour le moment).

A cette anthropométrie, sont affiliés des signaux capturés et définis dans la seconde partie intitulée MOTION.

La section CLUSTER permet de regrouper en sous-structures hiérarchiques significantes (bras gauche, main droite, etc) des éléments de la hiérarchie décrite dans la première section. Ces clusters constituent le support à l'annotation.

Enfin, la dernière partie, intitulée SEGMENTATION contient les pistes d'annotation. Ces pistes sont spécialisées selon l'annotation que l'on désire effectuer. Une segmentation de type signal permet de décomposer les mouvements élémentaires en phases caractéristiques ; une segmentation de plus haut niveau permet d'étiqueter les gestes suivant leurs paramètres formationnels (configurations manuelles et digitales, primitives de mouvement, emplacement et orientation). De nouveaux types peuvent être définis pour permettre l'extension des possibilités de l'annotation et l'adaptation à d'autres catégories de mouvements.

## 4 Perspectives

Un outil de segmentation et d'annotation du geste de communication a été brièvement présenté. Il se base principalement sur des représentations numériques du mouvement obtenues par des appareils de capture récents. L'accent a été mis sur l'ergonomie du logiciel ainsi que sur la souplesse des structures d'annotation utilisées. Les données annotées et segmentées obtenues nous permettront d'étendre les langages de spécification du geste (Lebourque, 99). Nous nous appuierons également sur ces données pour enrichir les méthodes de synthèse déjà utilisées (Gibet, 2001) et explorer les méthodes récentes de synthèse par graphes de transition (Kovar, 2002). Bien que ce logiciel soit dédié à l'annotation de données pour la synthèse, nous pensons qu'il peut s'avérer utile pour l'analyse linguistique du geste, notamment pour étudier les relations de spatialisation propres à la langue des signes (Lejeune 2001).

## Remerciements

Travail réalisé dans le cadre du projet Signe, financé par la région Bretagne (Réf. B/1042/2004/SIGNE) et du projet RobEA HuGEx, financé par le département STIC du CNRS. La capture des données présentée dans cet article a été rendue possible grâce au LPBEM de Rennes 2, au LESP de Toulon et du Var et au LINC de Paris 8.

## Références

- WITTENBURG P., LEVINSON ST., KITA S., BRUGMAN H.(2002), Multimodal Annotations in Gesture and Sign Language Studies, *Int. Conf. on Language Resources and Evaluation*.
- KOIZUMI A., SAGAWA H., TAKEUCHI M.(2003), An Annotated Japanese Sign Language Corpus, *Actes de Meeting of the Association for Computational Linguistics (ACL)*.
- DALLE P., LENSEIGNE F., HUDELLOT C.(2001), Apport d'un système d'analyse d'images à l'étude de la Langue des Signes, *Actes de Journées d'études Recherches Sur Les Langues Des Signes*.
- BARBIC J., SAFONOVA A., PAN JY., FALOUTSOS C., HODGINS J., POLLARD N.(2004), Segmenting Motion Capture Data into Distinct Behaviors, *Actes de conf. on Graphics interface*, 185–194.
- LEBOURQUE T., GIBET S.(1999), High level specification and control of communication gestures : the GESSYCA System, *Actes de Computer Animation*, 24-36.
- GIBET S., LEBOURQUE T., MARTEAU PF.(2001), High level Specification and Animation of Communicative Gestures, *journal of Visual Languages and Computing*, Vol. 12, pp.657-687.
- KOVAR L., GLEICHER M., PIGHIN F.(2002), Motion graphs, *Actes de conf. on Computer graphics and interactive techniques*, 473-482.
- LEJEUNE F., BRAFFORT A., DESCLÉS JP.(2001), Study on Semantic Representations of French Sign Language Sentences. *int. Gesture Workshop*, 197-201.