

Problèmes posés par la reconnaissance de gestes en Langue des Signes

Bossard Bruno

LIMSI-CNRS - Université Paris XI, Orsay

Bât 508 BP133 91403 Orsay CEDEX

Bruno.Bossard@limsi.fr

date de soutenance prévue : automne 2004

Mots-clefs – Keywords

Reconnaissance de gestes, Langue des Signes
Gesture recognition, Sign Language

Résumé - Abstract

Le but de cet article est d'explicitier certains des problèmes rencontrés lorsque l'on cherche à concevoir un système de reconnaissance de gestes de la Langue des Signes et de proposer des solutions adaptées. Les trois aspects traités ici concernent la simultanéité d'informations véhiculées par les gestes des mains, la synchronisation éventuelle entre les deux mains et le fait que différentes classes de signes peuvent se rencontrer dans une phrase en Langue des Signes.

The aim of this paper is to specify some of the problems raised when one wants to design a gesture recognition system dedicated to Sign Language, and to propose suited solutions. The three topics considered here concern the simultaneity of information conveyed by hand gestures, the possible synchronicity between the two hands, and the different classes of signs that may be encountered in a Sign Language sentence.

1 Introduction

Lorsque la parole ne suffit plus ou n'est pas adaptée à la situation, l'homme utilise le geste comme moyen de communication alternatif ou complémentaire. Il existe différentes communications gestuelles suivant les besoins d'expressivité. Les plus simples sont les gestes de commande qui ne se combinent pas entre eux ou avec un autre canal de communication (par exemple les gestes des plongeurs). Les gestes co-verbaux ou les mimiques sont un peu plus complexes et sont complémentaires vis à vis de la parole. Au plus haut degré de complexité se situent les gestes des Langues des Signes qui se combinent pour former des phrases

Dans cet article, c'est l'interprétation de certains gestes de la Langue des Signes (LS) qui nous intéresse. Le but n'est pas d'être capable d'interpréter n'importe quelle phrase de la LS mais de traiter certains problèmes de synchronisation. En effet, les deux mains sont deux canaux de communication qui peuvent être aussi bien combinés que totalement indépendants. Pour effectuer correctement l'interprétation de phrases gestuelles en LS, on doit alors être capable de déterminer si les deux mains opèrent de manière synchrone, ou si chacune effectue des gestes de manière plus indépendante. A l'heure actuelle il n'y a quasiment aucune étude d'effectuée sur les cas où les mains alternent des étapes de synchronisation avec des étapes où elles sont partiellement ou totalement asynchrones, car la plupart des travaux portent sur d'autres aspects de la LS.

Dans cet article la première partie décrit différents signes que l'on peut rencontrer en LS et leurs caractéristiques. La seconde partie expose les différents problèmes d'interprétation induits par ces signes et le problème de la synchronisation des mains y sera plus particulièrement développé. Enfin la troisième partie expose l'architecture proposée pour traiter ces problèmes de synchronisation.

2 Quelques caractéristiques de la Langue des Signes

La Langue des Signes est le moyen d'expression utilisé par les communautés de sourds ou malentendants pour communiquer entre eux. La Langue des Signes est une véritable langue à part entière (avec un lexique, une syntaxe...) et constitue de ce fait la forme la plus évoluée en communication gestuelle (Cuxac, 2000; McNeil, 1992).

L'expression de phrases en LS ne se réduit pas aux gestes produits par les deux mains, c'est le corps tout entier qui peut être mis à contribution pour exprimer une phrase. On peut distinguer trois principales parties qui interviennent : les mains, la tête (mimique et regard) et le buste. Dans cette étude on se restreint uniquement aux mains, c'est pourquoi les autres parties du corps ne seront pas abordées par la suite.

Chaque phrase est constituée d'une suite de gestes des mains que l'on appelle *signes* et qui sont agencés suivant une syntaxe régie par une logique spatiale et temporelle. C'est principalement l'interprétation des signes et des relations qui s'établissent entre eux qui nous intéressent, et qui vont être décrits avec précision dans les parties qui suivent.

2.1 Composition d'un signe

Chaque geste d'une main peut être décomposé en quatre paramètres qui sont indépendants et peuvent être aussi bien dynamiques qu'invariants durant l'émission du signe.

- La **configuration** correspond à la forme de la main définie par les doigts et la paume (exemples figure 1). Les signes ont souvent des aspect iconiques (Cuxac, 2000). En particulier la configuration est souvent en rapport avec la forme de ce que le signe décrit.
- L' **orientation** de la main est définie par deux axes de la main indiqués dans la figure 1.
- **Le mouvement** correspond à la trajectoire décrite par la main (ligne, arc de cercle...)(voir figure 1).

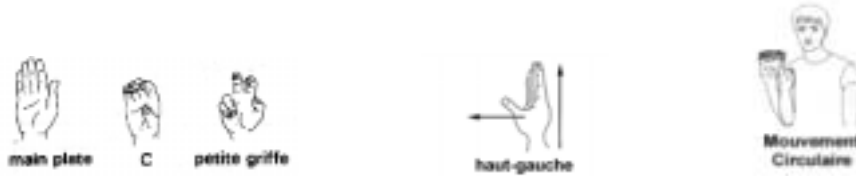


Figure 1: Exemples de configurations, d'orientation et de mouvement.

- L' **emplacement** est la position de la main par rapport au corps. Selon les besoins, il va avoir une granularité plus ou moins fine. Dans la figure 2, l'emplacement du premier signe [GARÇON] est le front. Le second signe indique que le garçon se situe à droite, devant le signeur. En effet, en LS le signeur utilise un espace de narration au sein duquel l'emplacement où il effectue le signe va être utilisé pour indiquer une relation spatiale (ou temporelle) entre entités.



Figure 2: La phrase “Un garçon à droite” constituée des signes [GARÇON] et [ICI]

Chacun de ces paramètres est porteur d'information et contribue au sens d'un signe.

2.2 Interaction entre les deux mains

Un signe peut aussi bien faire intervenir une main que deux mains et les deux mains ont différentes façons d'interagir. Lorsque les deux mains sont impliquées dans un signe, deux cas se présentent.

Dans le premier cas on voit apparaître un rôle pour chaque main. Une main est dite **dominante** et a pour rôle de décrire “l'action”, tandis que l'autre main qui est appelée main **dominée** sert de référence à cette action (Cuxac, 2000). Par exemple, avec le signe [TOMBER DE] (figure 3), on peut voir la main dominante (main droite dans la figure) qui mime l'action de tomber d'un support (référence donnée par la main dominée). En général la main dominante se déplace au cours du geste tandis que la main dominée reste statique.

Dans le deuxième cas, les deux mains sont complètement synchronisées : leurs paramètres sont

identiques ou symétriques. Par exemple, avec le signe [TABLE] (figure 3), les deux mains ont une configuration et une orientation identiques et des mouvements de même trajectoire mais de sens opposés.



Figure 3: Les signes [TABLE] et [TOMBER DE].

Lorsque qu'un signe ne fait intervenir qu'une seule main il y a également deux cas qui se présentent : soit la deuxième main est inactive (comme dans la figure 2), soit elle effectue un autre signe (voir plus loin les trois dernières images dans la figure 5) et dans ce cas on peut parfois retrouver une interaction main dominante-main dominée entre les deux signes (cas de la dernière image dans la figure 5).

On voit donc que les flux de données véhiculés par chaque main peuvent aussi bien être synchrones que complètement asynchrones.

2.3 Différentes classes de signes

En LS on peut distinguer différentes classes de signes, chacune correspondant à un usage particulier. L'ensemble des signes dits "standards" correspond à des mots (nom, verbe, adjectif...) ayant un sens bien établi. On peut leur associer une correspondance relativement précise dans les langues orales. Du fait qu'il y a une correspondance "directe", ce sont les signes qui ont été le plus étudiés en reconnaissance de gestes et il existe des systèmes permettant d'effectuer la reconnaissance de manière fiable sur des vocabulaires de taille limitée. Par exemple dans (Hienz et Bauer, 2000) le taux de reconnaissance atteint un peu plus de 90% pour un vocabulaire d'une centaine de signes.

Il existe deux autres classes de signes qui eux ne permettent aucune correspondance directe avec des mots du langage oral. Il s'agit des *spécificateurs de forme et de taille*, et des *classificateurs*. Ces signes ont des aspects fortement iconiques et font partie de ce que l'on appelle la grande iconicité (Cuxac, 2000). Les spécificateurs permettent de décrire un objet, un animal, une scène. C'est à l'aide de la forme des mains, de leur orientation et de leur mouvement que le signeur décrit une forme et les dimensions d'un objet. Les classificateurs sont similaires aux spécificateurs dans le sens où ils représentent également un objet (ou personne, animal...) et donc la forme de la main est en rapport avec celle de l'objet ou de sa fonction. Mais ils ont un tout autre rôle car ils servent en quelque sorte de pronom. Lorsqu'un objet a été cité dans une phrase à l'aide d'un signe du vocabulaire standard (ou de spécificateurs), un classificateur peut ensuite être utilisé pour représenter cet objet dans le reste de la phrase. On peut avec ce classificateur préciser la position de l'objet ou décrire une trajectoire qu'il a emprunté.

Dans la figure 4, on peut voir une phrase qui regroupe les trois classes de signe. Le premier signe est un spécificateur de forme et de taille, il décrit une caisse en symbolisant ses côtés. Le signe suivant est le signe standard [POMME]. Enfin, la troisième étape montre des signes qui sont des classificateurs symbolisant des pommes que l'on place à différents endroits pour représenter le tas de pomme que contient la caisse.



Figure 4: La phrase “Une caisse pleine de pommes”.

On se doute alors que tous ces aspects (diversité des signes, interaction des mains...) vont soulever des problèmes au niveau de la reconnaissance de ces différents gestes. Cela va être abordé dans la partie qui suit.

3 Problèmes spécifiques à l’interprétation de signes

Dans un système de reconnaissance de forme classique, les données sont généralement traitées en un seul bloc. Dans notre cas, cela voudrait dire que les deux mains seraient considérées comme une seule entité. Nous allons maintenant voir que ce type de système n’est pas adapté aux caractéristiques de la LS.

3.1 Complexité du vocabulaire

Un des principaux problèmes que l’on trouve en reconnaissance de forme est la complexité. En effet, du fait qu’un signe est composé de plusieurs éléments co-occurents, le vocabulaire peut être composé a priori d’un très grand nombre de signes (Vogler et Metaxas, 1999). Pour une main, si on a N configurations possibles, M types de mouvements, P emplacements et Q orientations, on a $N * M * P * Q$ signes possibles¹. Sachant que l’on peut avoir des signes qui combinent les deux mains, on voit tout de suite que créer un système de reconnaissance peut se révéler complexe et nécessiter un long apprentissage.

En réalité, toutes les combinaisons possibles de paramètres ne sont pas utilisées par les signeurs. Car ils ne sont pas combinés au hasard mais en fonction de critères syntaxiques et sémantiques. Toutefois le nombre de signes reste important. Pour remédier à ce problème, il existe une solution qui consiste à traiter séparément les divers composants d’un signe au lieu de traiter le signe dans sa globalité (Liang et Ming, 1996; Vogler et Metaxas, 1999). On peut établir un parallèle avec la reconnaissance de la parole où l’on cherche à reconnaître des phonèmes et non pas directement des mots. Nous allons tâcher de préserver cette approche dans l’architecture que nous allons proposer dans la partie 4.

¹Il n’y a théoriquement pas de limite possible (mis à part les contraintes articulaires) quand au nombre de variantes pour chaque paramètre. Cela est particulièrement valable pour les spécificateurs de forme et de taille dont les paramètres sont liés à ce qu’ils décrivent. Pour les signes standards on recense l’utilisation d’environ 140 configuration, 300 orientations et 50 emplacements (Braffort, 1996).

3.2 Relations entre les deux mains

On a vu précédemment (partie 2.2) que l'on peut rencontrer trois types d'interaction entre les deux mains. Soit les deux mains produisent ensemble un signe, soit une main effectue un signe et l'autre est au repos, soit les deux mains effectuent séparément deux signes de manière plus ou moins indépendante.

Le problème va donc être de savoir distinguer ces différents cas pour ne pas reconnaître un signe à deux mains alors qu'il s'agissait de deux signes distincts effectués au même moment (ou vice-versa). Pour distinguer ces deux cas, on ne peut pas se contenter d'une quelconque similitude des paramètres des deux mains, car on a vu dans la première partie que lorsque les deux mains sont utilisées pour former un signe, elles peuvent être aussi bien complètement synchrones (signe [TABLE] dans la figure 3), qu'être juste corrélées et n'avoir aucune similitude au niveau des paramètres (signe [TOMBER DE] dans la figure 3).

Il existe une autre relation entre les mains, que l'on va devoir détecter. Lorsque les deux mains effectuent séparément des signes et qu'il existe une relation main dominante-main dominée, il existe des *points de synchronisation* pour mettre en rapport les signes. Pendant un instant les deux mains vont être agencées spatialement de manière à véhiculer une certaine information. Par exemple dans la phrase "Le chat est dans la voiture" (figure 5), le signe [VOITURE] est d'abord énoncé (première image) puis remplacé par le classificateur "C" qui représente la voiture en tant que "contenant" et la positionne dans la scène (deuxième image). Ensuite on voit la main gauche qui énonce le signe [CHAT] (troisième image), le remplace par le classificateur "X" (représentant les pattes du chat) et positionne ce dernier relativement au classificateur représentant la voiture (dernière image)(description effectuée dans (Lejeune et al., 2002)).



Figure 5: "Le chat est dans la voiture"

On remarque dans les trois dernières images que la main dominante "maintient" le classificateur "C" alors que la main dominée effectue trois signes successifs. On peut donc penser que les deux mains sont indépendantes. Cependant, il va donc y avoir pendant un certain laps de temps une disposition particulière des mains pour exprimer la relation spatiale "dans" (quatrième image, figure 7). Il va donc falloir être capable de détecter cet instant pour interpréter cette relation. Aucune solution n'a été proposée à ce jour pour résoudre ce problème.

3.3 Interprétation des signes non standards

Un troisième problème encore plus difficile à résoudre concerne l'interprétation des signes non standards. A ce jour, seule une catégorie de ces signes (les verbes directionnels) a été abordée dans le cadre d'un système de reconnaissance (Braffort, 1998; Sagawa et Takeuchi, 1999). Parmi les catégories qui n'ont pas été abordées en reconnaissance de la LS, on trouve les classificateurs et les spécificateurs. On a vu dans la partie 2.3 que ces derniers n'ont pas de sens propre, c'est la phrase et son contexte qui permettent de le déterminer.

Les classificateurs se rapportent souvent à un signe standard qui a été signé précédemment mais ce n'est pas toujours le cas. On peut cependant déterminer pour chaque signe la liste des classificateurs potentiels. Les spécificateurs posent un problème beaucoup plus complexe du fait qu'ils sont propres à la description et peuvent varier énormément. Contrairement aux langues orales qui possèdent un vocabulaire délimité, en LS on peut très bien "inventer" un signe de toutes pièces. Par définition, les paramètres d'un spécificateur sont associés à ce qu'ils décrivent et donc à chaque nouvelle description peut correspondre de nouveaux signes. De plus, la manière de faire une description peut dépendre du contexte social et culturel du signeur (Cuxac et al., 1998)

Il va falloir chercher le sens de ces signes dans le contexte de la phrase et en analysant les formes qu'ils décrivent. Ceci est un problème très complexe et il n'y a actuellement aucune solution qui y remédie. Il faut à notre niveau que l'architecture proposée permette une interaction bidirectionnelle entre le processus de reconnaissance de forme et le processus d'interprétation prenant en compte les aspect syntaxiques et sémantiques.

Pour pouvoir traiter des phrases de la LS les plus riches possibles il faut donc que le système de reconnaissance prenne en compte les différents problèmes soulevés ici.

4 Réalisation d'un système de reconnaissance

4.1 Architecture proposée

Ce qui suit va exposer différentes propositions visant à résoudre le problème de la complexité et celui des relations entre mains, mais aussi à prévoir l'interaction entre les niveaux de reconnaissance de forme et d'interprétation. L'architecture proposée vise également la généricité pour permettre l'utilisation de divers modèles de reconnaissance.

- Proposition 1 : reconnaissance des paramètres

La reconnaissance d'un signe va être décomposée en quatre parties correspondant aux différents paramètres. Cela afin de réduire la complexité (comme vu en 3.1), mais aussi pour être capable de modéliser n'importe quel type de signe. En effet pour les signes comme les classificateurs et les spécificateurs, on va être capable de reconnaître leurs différents paramètres et à partir de ceux-ci faire une interprétation de ces signes à des étapes ultérieures.

La figure 6 nous montre ces quatre modules qui fonctionnent en parallèle (pour les signes à deux mains on aura évidemment huit modules). Le cinquième module (en pointillés sur la figure) qui reçoit les résultats de la reconnaissance des paramètres va devoir effectuer diverses tâches : il va tout d'abord construire des quadruplets à partir des symboles reçus depuis les quatre modules, il va ensuite chercher à associer chaque quadruplet à un signe (cela peut être effectué soit en cherchant une correspondance dans une base de données de quadruplets, soit par des méthodes stochastiques), enfin il va transmettre le signe reconnu avec un score de reconnaissance basé sur celui des paramètres et la qualité de la correspondance signe-quadruplet. Dans le cas où aucune correspondance n'est trouvée (on est probablement face à un classificateur ou un spécificateur), c'est le quadruplet qui est transmis (avec un score basé sur celui des paramètres) pour une analyse ultérieure en relation avec les niveaux d'interprétation. En plus du score de reconnaissance associé à l'élément retourné par le module, la durée de cet élément est fournie (à quel instant il a commencé et à quel instant il a fini).

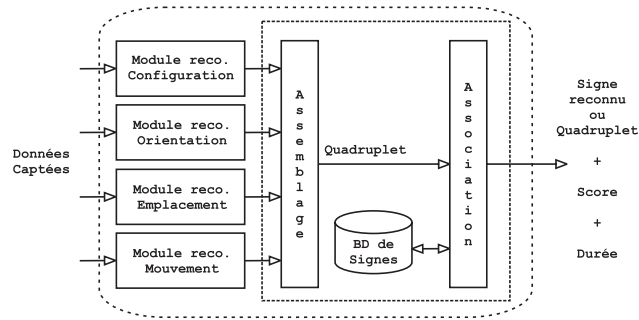


Figure 6: Module de reconnaissance des gestes d'une main

La composition même des modules de reconnaissance des paramètres peut-être variable suivant la complexité du vocabulaire que l'on veut reconnaître. On peut aussi bien utiliser un simple système de classification statistique (Rubine, 1991), que des modèles stochastiques tel que les chaînes de Markov cachées (Vogler et Metaxas, 1999).

- Proposition 2 : reconnaissance à la fois synchrone et asynchrone

Si l'on considère la problématique liée à l'utilisation d'une ou deux mains, on peut identifier trois types de signes possibles : ceux effectués avec les deux mains, ceux effectués par la main droite et ceux effectués par la main gauche. Pour interpréter correctement un signe il va donc falloir déterminer à laquelle de ces trois catégories il est associé. Pour cela, la solution proposée est la suivante : pour chaque geste émis, trois modules correspondant aux trois types possibles vont essayer de le reconnaître (figure 7). Le module qui parvient à obtenir le "meilleur" score de reconnaissance détermine alors le type du signe. Chacun des modules a une architecture

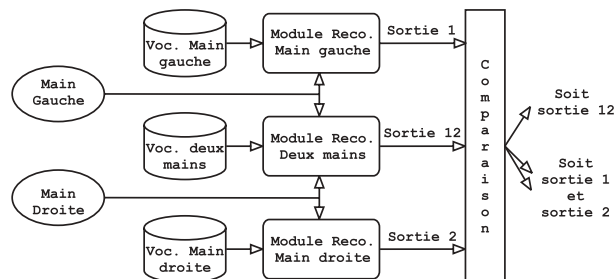


Figure 7: Les trois modules permettant de reconnaître les trois types de gestes

correspondant à celle du module de reconnaissance décrit précédemment dans la proposition 1. Les trois modules vont travailler en parallèle à partir des données émises par la main droite et la main gauche. Mais chacun sera associé à un vocabulaire propre, c'est à dire que chacun aura fait un apprentissage avec des signes que n'auront pas appris les deux autres modules.

Il va falloir ensuite choisir un des éléments reconnus par les trois modules. C'est un module de comparaison qui va se charger de cela en utilisant une fonction d'évaluation qui peut être plus ou moins sophistiquée. Au niveau le plus simple correspond une fonction qui fait une évaluation "naïve" : le choix d'un signe sera fait en fonction de son score de reconnaissance et sans tenir compte du contexte. Pour des niveaux plus évolués, il faudra concevoir une fonction qui se base sur les signes précédemment reconnus dans la phrase (en se basant sur des règles de syntaxe et de sémantique, on peut émettre des hypothèses sur le signe à reconnaître).

- Proposition 3 : établissement des relations entre signes

Suivant le résultat obtenu à l'étape de comparaison (proposition 2), on va éventuellement rechercher d'éventuels relations entre les signes : ainsi si l'on a reconnu deux signes à une main, il va falloir chercher s'il n'y a pas une relation spatiale qui est exprimée par la disposition des mains. On va donc devoir rechercher parmi les signes reconnus s'il n'y a pas un point de synchronisation entre eux.

Pour détecter cet instant bien particulier, la solution proposée est la suivante. Une première étape va consister à synchroniser les signes de la main gauche avec ceux de la main droite, car deux signes n'ont pas forcément la même durée et ne sont donc pas reconnus au même instant par les modules de la main droite et de la main gauche. Cette synchronisation va s'effectuer en assemblant les signes main gauche main droite par couple selon l'instant auquel ils ont débuté et l'instant auquel ils ont fini. Par exemple si la main gauche effectue un signe A pendant que la main droite effectue les signes B,C et D, on va construire les couples suivants : (A,B),(A,C),(A,D).

Les étapes suivantes consisteront à utiliser ces couples de signes durant les phases de traitement syntaxique et sémantique pour détecter les éventuelles relations entre mains. La décision de l'existence d'une relation se fera en fonction de la disposition des mains mais aussi des informations déduites sur les signes durant les analyses syntaxique et sémantique. Ces traitements ne seront pas abordés dans cet article.

4.2 Réalisation et expérimentation

Pour tester cette architecture, le choix s'est porté sur des phrases mêlant des classificateurs et des signes standards, et décrivant des positions d'objets (ou animaux) par rapport à d'autres : par exemple "Le chat est dans la voiture" (figure 5). La phrase est composée successivement du signe standard [VOITURE], du classificateur "C" (qui dure jusqu'à la fin de la phrase), du signe standard [CHAT] et enfin du classificateur "X". L'avantage de ce type de phrase est qu'il permet de tester les différents cas qui peuvent soulever des problèmes : signe à une ou deux mains, synchronisation variable, classificateurs ...

Afin de faciliter la phase de capture des gestes, l'acquisition de ces phrases se fait à partir de gants numériques et de capteurs de position 3D. Les données sont récupérées de manière synchrone à partir des quatre périphériques, cela afin que les modules de reconnaissance puissent travailler correctement.

La réalisation de l'architecture est actuellement en cours. A ce jour l'implémentation des propositions 1 et 2 est achevée. Pour l'instant les modules de reconnaissance sont basés sur un système de classification de type statistique (Rubine, 1991) pour pouvoir tester la comparaison des sorties des trois modules. Mais l'objectif à moyen terme est de tester une approche stochastique qui est beaucoup plus adaptée aux phrases gestuelles car elle permet de modéliser la coarticulation entre deux gestes enchaînés. L'objectif à court terme est de tester différentes méthodes de comparaison avant de progresser plus en avant dans la réalisation de l'architecture.

5 Conclusion

Le fait que les mains sont deux moyens de communications indépendants permet en LS d'exprimer diverses relations entre les signes. On ne peut ignorer l'information que ces relations apportent

car avec certains signes comme les classificateurs, c'est tout le sens de la phrase qui en découle. D'autre part, du fait que les mains peuvent être indépendantes, on rencontre aussi bien des signes à une main que des signes à deux mains, il faut être alors capable de distinguer ces différents cas.

C'est ce à quoi tente de donner réponse l'architecture qui a été présentée ici. D'autre part elle n'ignore pas les autres problèmes existants, à savoir la complexité et la difficulté à interpréter certains signes. Elle utilise certaines solutions proposées pour gérer la complexité du vocabulaire et laisse la possibilité d'interpréter les signes non reconnus en propageant les paramètres du signe vers les phases d'analyse syntaxique et sémantique. Actuellement la première partie de l'architecture a été réalisée et commence à être testée. La réalisation des parties concernant l'interprétation des relations lors de l'analyse syntaxique et sémantique ne sera effectuée qu'après avoir obtenu un module de comparaison suffisamment fiable.

Du fait de certains points communs entre certains gestes de la LS et d'autres moyens de communication gestuelle, il est envisagé d'étendre le travail effectué à d'autres domaines. Par exemple, on trouve de fortes analogies entre la description de scènes utilisant des classificateurs et des spécificateurs et l'interaction gestuelle en Réalité Virtuelle. D'autre part le travail qui sera effectué sur la détection de points de synchronisation entre les mains pourra être adapté à d'autres domaines où l'on a des synchronisations ponctuelles entre des flux d'information.

Références

- Braffort A. (1996), *Reconnaissance et compréhension de gestes, application à la langue des signes*, thèse, Université Paris XI - Orsay.
- Braffort A., (1998), *Traitement automatique des langues des signes : interprétation des informations spatiales*, dans *Revue d'Interaction Homme-Machine*, Vol. 1, Num. 1, 65-87.
- Cuxac C. (2000), *La langue des Signes Française (LSF), les voies de l'iconicité*, Faits de Langues 15-16, Editions OPHRYS.
- Cuxac C., Fusellier-Souza I., Sallandre M.A., (1998), *Iconicité des Langues des Signes et Catégorisation*, dans *SEMIOTIQUES*, Num. 16, 181.
- Hienz H., Bauer B. (2000), Video-Based Continuous Sign Language Recognition Using Statistical Methods, Actes de *IEEE International Conference on Automatic Face and Gesture Recognition*, 440-445.
- Lejeune F., Braffort A., Desclés J.P. (2002), Study on Semantic Representations of French Sign Language Sentences, Actes de *Gesture Workshop'01*, Springer, LNAI 2298 (à paraître), 440-445.
- Liang R.H., Ming O., (1996), A Sign Language Recognition System Using Hidden Markov Model and Context Sensitive Search, Actes de *ACM Symposium on Virtual Reality Software and Technology'96*, 59-66.
- McNeil D. (1992), *Hand and Mind : What gestures reveal about thought*, University of Chicago Press.
- Rubine D. (1991), *The Automatic Recognition of Gestures*, Ph-D thesis, Carnegie-Mellon.
- Sagawa H., Takeuchi M. (1999), A Method for Analyzing Spatial Relationships Between Words in Sign Language Recognition, Actes de *Gesture Workshop'99*, Springer, LNAI 1739, 197-210.
- Vogler C., Metaxas D.N. (1999), Toward Scalability in ASL Recognition: Breaking Down Signs into Phonemes, Actes de *GEsture Workshop'99*, Springer, LNAI 1739, 211-224.
- Vogler C., Metaxas D.N. (1999), Parallel Hiden Markov Models for American Sign Language Recognition, Actes de *IEEE International Conference on Computer Vision*, 116-122.