

## Doctorat en informatique

*équipe EXPRESSION / IRISA, Université Bretagne Sud, France*

### *Avatar signeur en langue des signes française*

**Sujet :** Avatar signeur : édition, génération, évaluation

**Encadrants :** Sylvie Gibet, Caroline Larboulette, équipe Expression, IRISA

**Contact :** Sylvie Gibet ([sylvie.gibet@univ-ubs.fr](mailto:sylvie.gibet@univ-ubs.fr))

Caroline Larboulette ([caroline.larboulette@univ-ubs.fr](mailto:caroline.larboulette@univ-ubs.fr))

<http://www-expression.irisa.fr/>

**Date :** octobre 2016 – septembre 2019

#### **1. Contexte de l'équipe EXPRESSION/IRISA : le projet *Avatars Signeurs***

Le projet de thèse s'inscrit dans un projet plus large intitulé *Avatars Signeurs* de l'équipe Expression de l'IRISA, projet à vocation interdisciplinaire dédié à la manipulation de contenus numériques utiles aux personnes sourdes et malentendantes ainsi qu'aux personnes entendant s'exprimant en langue des signes française (LSF). Ce projet se situe au croisement de problématiques informatiques qui visent à développer des outils interactifs innovants permettant d'apprendre et de communiquer au moyen de langues gestuelles, et de problématiques sociétales qui visent l'étude de nouvelles technologies dédiées aux usagers sourds (tablettes, Web, etc.) et en particulier d'avatars signeurs.

##### 1.1. Applications visées

Les nouvelles technologies de l'information et de la communication à destination des personnes sourdes, que ce soit le téléphone mobile avec vidéo intégrée, la messagerie MMS, les dispositifs de visioconférence, ou Internet et les réseaux sociaux, sont particulièrement adaptées pour des personnes sourdes puisque les applications utilisées sont essentiellement basées sur une communication visuo-gestuelle. La vidéo ne permettant pas de garantir l'anonymat et imposant des contraintes fortes au niveau du stockage et du transport d'information, la production automatique de messages en LSF et la visualisation au moyen d'avatars 3D semblent constituer une réponse adaptée et permettre à la fois la réduction des informations stockées, l'anonymisation ainsi que la programmabilité des personnages virtuels pour éditer et produire de nouveaux énoncés.

Parmi les domaines applicatifs envisagés nous ciblons plus particulièrement :

- L'éducation en LSF
- L'apprentissage de la LSF
- L'accessibilité à la connaissance pour les personnes pratiquant la LSF (Web)
- La conception de vidéo-book

Cependant, il manque aujourd'hui des outils pour faciliter l'accès aux connaissances aux personnes sourdes, en particulier aux enfants. En effet, l'information disponible aujourd'hui passe principalement par l'écrit ou l'oral. Or, la LSF, langue maternelle des personnes sourdes, exploite le canal visuo-gestuel qui ne permet pas d'appréhender toute forme de connaissance écrite. Ainsi, l'accès à des domaines de connaissance aussi variés que l'histoire, les mathématiques, les langues ou les sciences reste très limité et il n'est à l'heure actuelle pas possible de créer facilement des cours en LSF qui pourraient être utilisés sur des tablettes par exemple. Nous souhaitons dans le cadre de notre projet de recherche *Avatars Signeurs* apporter de tels outils, facilitant ainsi l'accès à la connaissance aux personnes malentendantes.

Ces objectifs ne se limitent pas aux personnes sourdes. De plus en plus d'entendants et malentendants souhaitent apprendre la LSF. Or, il manque de manière cruciale des outils permettant d'apprendre les langues des signes. Il a été montré par ailleurs que les bébés peuvent s'exprimer en langues signées avant de pouvoir parler, cependant aucun outil ou jeu n'est disponible pour eux ou pour leurs parents afin de les aider à apprendre. Notre but final est donc de favoriser l'accessibilité à la connaissance pour tout type d'utilisateur, du plus jeune au plus âgé, sourd ou entendant, en utilisant des technologies avancées. Par extension, nous souhaitons développer des outils de traduction automatique français-LSF à partir de contenus Internet, de façon à procurer aux personnes sourdes la même information que celle que l'on peut lire/entendre sur le Web.

## 1.2. Compétences acquises dans l'équipe Expression de l'IRISA

L'équipe Expression dispose actuellement d'une chaîne complète de synthèse permettant de spécifier une phrase en LSF à partir d'une séquence contextualisée de signes (ou groupes de signes) et de produire automatiquement l'animation de l'avatar signeur en 3D.

Le développement logiciel associé à ce projet a été initié il y a une dizaine d'années à l'UBS. Il a été financé par plusieurs projets de recherche et de valorisation industrielle : projet Robea HuGEx piloté par S. Gibet de 2003 à 2006, projet Région SIGNE piloté par S. Gibet de 2004 à 2008 (thèse d'Alexis Héloir, 2008), projet ANR-SignCom piloté par S. Gibet de 2008-2012, projet investissement d'avenir IA-SIGN3D de 2012 à 2014 (S. Gibet responsable scientifique). Deux projets internationaux ont permis également d'asseoir la crédibilité linguistique du projet. Un projet CFQCU piloté par l'université de Québec à Montréal permettant d'analyser conjointement 3 langues des signes (québécoise, américaine et française), et un projet (financé par le projet ANR SignCom) de collaboration avec l'université de Gallaudet (Washington, USA) avec une thèse de linguistique (Kyle Duarte, 2012)

La chaîne de synthèse concaténative est maintenant opérationnelle. Elle comporte :

- un module d'indexation et de recherche d'information à partir d'une base de données de mouvements duale : données mouvement indexées au niveau du signal et données sémantiques annotées à différents niveaux (plusieurs canaux gestuels) ;
- un module d'édition de mouvement permettant de composer de nouvelles phrases à partir d'éléments de mouvement identifiés sur différents canaux multimodaux (mouvement des mains, configuration des mains, expression faciale, direction du regard) ;
- un module de synthèse de mouvement construit à partir de différents contrôleurs synchronisés et associés à chacun des canaux précédents ;
- un module de rendu 3D permettant de visualiser les mouvements générés par l'avatar.

## 2. Objectifs de la thèse

Le projet de thèse s'appuie sur la plateforme de synthèse concaténative décrite ci-dessus, dédiée à la génération de gestes de la langue des signes française.

Le travail de thèse vise à concevoir et réaliser un ensemble de méthodes et d'outils logiciels permettant de consolider la boucle d'édition-génération / perception nécessaire pour aller vers une application pertinente pour la catégorie d'utilisateurs visés. Il comprend :

- la création de corpus permettant de produire un large panel de phrases, aux vues des applications visées ;
- le développement de méthodes de « motion retrieval » permettant d'extraire des bases de données constituées des meilleurs mouvements pour construire de nouvelles phrases contextualisées ;
- l'amélioration des méthodes de transition, en tenant compte des mécanismes de co-articulation ;
- l'évaluation de la qualité de la LSF produite : animation, rendu 3D des avatars signeurs, compréhension à plusieurs niveaux (signe, phrase, texte) du contenu généré.

A partir de l'édition de phrases simples construites sur un corpus LSF à vocabulaire limité, il s'agit de générer des mouvements en 3D, et de contrôler simultanément les mouvements corporels, manuels, les expressions faciales et la direction du regard. Les nouvelles phrases sont construites à partir de la composition de mouvements extraits d'une base de données annotée. L'objectif est de trouver des représentations appropriées permettant d'une part de construire des phrases intelligibles, suivant des règles sémantiques et grammaticales propres à la LSF. La paramétrisation de l'espace de signation devra permettre d'automatiser la génération de phrases contextualisées (utilisation de verbes directionnels, mécanismes d'identification spatiale utilisés dans les narrations, etc.). D'autre part, afin d'accéder et d'extraire les mouvements existants dans la base de données, il conviendra de développer de nouvelles méthodes permettant d'améliorer l'adaptation à la phrase construite (extraction de sous-séquences optimales). Dans certains cas, lorsque les mouvements extraits sont trop éloignés des précédents ou suivants, il sera nécessaire de synthétiser de nouveaux mouvements suivant le contexte linguistique. Enfin l'exploitation de données capturées devra permettre d'identifier statistiquement les règles de co-articulation et de coordination entre les différentes modalités intervenant dans la formation des signes.

Le travail de thèse devra également s'intéresser aux questions d'évaluation perceptuelle. Différentes expérimentations seront menées pour évaluer la qualité des phrases signées, en considérant à la fois la compréhension et l'acceptation de l'avatar 3D par les personnes sourdes. Le but des études perceptuelles ne consiste pas uniquement à établir et valider la qualité des animations produites, mais également à développer des métriques, des heuristiques et des méthodologies dédiées aux avatars signeurs.

Les questions identifiées dans cette thèse constituent en elles-mêmes une problématique très originale, valorisable en terme d'articles académiques (conférences, journaux, etc.). Malgré l'importance des enjeux économiques liés aux avatars (industrie du jeu, serious games, etc.), il y a très peu de travaux dans le monde ayant développé de manière approfondie ce type d'approche exploitant à la fois des données capturées (séries temporelles multidimensionnelles) et des données langagières à différents niveaux (phonétique, phonologique, syntaxique, sémantique, etc.).

**Lieu** : ce travail sera réalisé à l'IRISA dans l'équipe EXPRESSION (Vannes, France).

**Prérequis** : Le candidat doit avoir un master ou équivalent en informatique. Il(elle) devra avoir des connaissances solides en C++, un excellent niveau d'anglais et si possible des connaissances en Interaction Homme-Machine, en informatique graphique ou en traitement de données. Une connaissance supplémentaire en langue des signes serait fortement appréciée. Dans le cas contraire, une formation en LSF sera incluse dans la formation.

**Financement** : bourse ministérielle française (CDE)

**Salaire mensuel** : autour de 1350 € (brut)

**Pour postuler** : envoyer un CV et une lettre de motivation ainsi qu'une liste de personnes de référence par email à Sylvie Gibet ([sylvie.gibet@univ-ubs.fr](mailto:sylvie.gibet@univ-ubs.fr)) ou Caroline Larboulette ([caroline.larboulette@univ-ubs.fr](mailto:caroline.larboulette@univ-ubs.fr)) dès que possible, et AVANT April 30, 2016.

### Bibliographie

1. Projet Investissement d'Avenir *SIGN3D*, <https://www.youtube.com/watch?v=oRM4nZnokow>, 2015.
2. Projet ANR *SignCom*: <http://www-valoria.univ-ubs.fr/signcom/en/>, 2011.
3. Clément Reverdy, Sylvie Gibet, Caroline Larboulette. Optimal Marker Set for Motion Capture of Dynamical Facial Expressions. ACM Siggraph Conference Motion in Games, MIG 2015, Nov. 2015, Paris, France, 2015.
4. Sylvie Gibet, François Lefebvre-Albaret, Ludovic Hamon, Rémi Brun, and Ahmed Turki. Interactive editing in french sign language dedicated to virtual signers: requirements and challenges. ACM Universal Access in the Information Society, pages 1-15, 2015.
5. François Lefebvre-Albaret, Sylvie Gibet, Ahmed Turki, Ludovic Hamon, Rémi Brun, Overview of the Sign3D Project High-fidelity 3D recording, indexing and editing of French Sign Language content. Sign Language Translation and Avatar Technology, SLTAT 2013, Chicago, October 2013.
6. Ludovic Hamon, Sabah Boustila, Sylvie Gibet. Edition Interactive d'énoncés en langue des signes française dédiée aux avatars signeurs, TALN 2013.
7. Sylvie Gibet, Pierre-François Marteau, Kyle Duarte. Toward a Motor Theory of Sign Language Perception. Lecture Notes in Computer Science (Revised Selected papers), Berlin:Springer, Vol. 7206, 161-172, 2012.
8. Gibet, Sylvie, Nicolas Courty, Kyle Duarte, and Thibaut Le Naour. The SignCom System for Data-Driven Animation of Interactive Virtual Signers: Methodology and Evaluation. In ACM Transactions on Interactive Intelligent Systems TiiS 1(1): 6, 2011.
9. Nicolas Courty et Sylvie Gibet. Why is the Creation of a Virtual Signer Challenging Computer Animation, Motion in Games, LNCS, nov. 2010.
10. K. Duarte, S. Gibet. Heterogeneous Data Sources for Signed Language Analysis and Synthesis, The seventh international conference on Language Resources and Evaluation (LREC 2010), Malta, Mai 2010.
11. K. Duarte, S. Gibet. Corpus Design for Signing Avatars, *Workshop on Representation and Processing of Sign Languages: Corpora and Sign Language Technologies*, Valetta : Malta, 2010.
12. K. Duarte, S. Gibet. Reading Between the Signs: How are Transitions Built in Signed Languages? Thoretical Issues in Sign Language Research (TILSR 2010), Indiana, USA, 2010.
13. C. Awad, N. Courty, K. Duarte, T. LeNaour, S. Gibet. A Combined Semantic and Motion Capture Database for Real-Time Sign Language Synthesis. In Proc. of the 9th Int. Conference on Intelligent Virtual Agent (IVA 2009), LNCS, pp. 432-438, Amsterdam, NL, Sept. 2009.
14. A. Héloir, N. Courty, S. Gibet, et F. Multon. Temporal alignment of communicative gesture sequences. Computer Animation and Virtual Worlds (selected best papers from CASA'06), 17:347-357, juillet 2006.