

# Artus: Animation Réaliste par Tatouage audiovisuel à l'Usage des Sourds

P. BAS<sup>1</sup>, J. LIENARD<sup>1</sup>, J-M CHASSERY<sup>1</sup>, D. BEAUTEmps<sup>2</sup>, G. BAILLY<sup>2</sup>

<sup>1</sup>Laboratoire des Images et des Signaux (LIS) INPG/CNRS,  
BP. 46 38402 Saint Martin d'Hères cedex

<sup>2</sup>Institut de la Communication Parlée (ICP),  
46, avenue Félix Viallet, 38031 GRENOBLE CEDEX 1

Patrick.Bas@lis.inpg.fr,

Joel.Lienard@lis.inpg.fr, Jean-Marc.Chassery@lis.inpg.fr, beauteemps@icp.inpg.fr, bailly@icp.inpg.fr

**Résumé** — Le projet Artus a pour objectif d'insérer dans les émissions télévisées classiques une information imperceptible permettant d'animer à la réception un clone de synthèse. Ce clone 3D communique avec des téléspectateurs sourds grâce à un langage des signes appelé Langage Parlé Complété (LPC). Le LPC repose sur la lecture labiale auquel s'ajoutent les gestes d'une des mains de la personne communicante. Dans le cadre de ce projet, les mouvements du clone sont directement insérés en utilisant des techniques de tatouage. De telles techniques permettent l'insertion d'une information indélébile et imperceptible. Ce projet traite de problèmes scientifiques multiples comme la conversion de texte en LPC, le codage des mouvements du clone, le tatouage haute-capacité ou encore l'animation 3D du clone. Artus est composé de 7 partenaires dont 4 laboratoires universitaires, une chaîne audiovisuelle et 2 partenaires industriels.

## 1 Introduction

Actuellement 3 à 5 millions de Français sont sourds ou souffrent d'un déficit auditif ce qui représente 6 à 8% de la population française. La surdité est un handicap auquel beaucoup de personnes sont ou seront confrontés puisqu'il est souvent lié à l'âge : la population des sourds et malentendants est composée de 11% de personnes de moins de 10 ans, de 27% de personnes âgées entre 10 et 60 ans et de 27% de personnes de plus de 60 ans.

Le projet RNRT Artus propose de développer des applications multimédia permettant d'augmenter l'intelligibilité des émissions télévisées vis à vis des personnes sourdes et mal entendant. Pour ce faire, un clone, composé d'un visage et d'une main, est animé en superposition de documents télévisés classiques. Les membres de ce projet ont choisi d'utiliser le langage parlé complété (LPC) qui est à la fois un langage facile à apprendre pour les personnes non sourdes et à comprendre pour les personnes devenues sourdes. La Figure 1 présente un résumé des différentes tâches du projet Artus. Pour effectuer l'animation du clone, les paramètres de mouvements sont synthétisés à partir du télétexte issu du document télévisé. La transmission de ces différents paramètres au sein du flux audio-vidéo s'effectue par techniques de tatouage haute capacité. Une telle approche permet de pouvoir animer le clone quel que soit le format de diffusion (câble, satellite, hertzien). Un terminal situé chez l'utilisateur est chargé de décoder les informations dissimulées dans l'émission télévisée et d'animer le clone 3D à partir de ces informations.

Ce document a pour objectif de présenter d'une part le contexte scientifique du projet (le langage parlé complété et le tatouage), et d'autre part de souligner les enjeux scientifiques et techniques qui y sont rattachés.

## 2 Présentation du projet

Cette section présente les deux domaines qui sont à l'origine de ce projet, à savoir le Langage Parlé Complété et le tatouage de contenus audiovisuels.

### 2.1 Le Langage Parlé Complété

#### 2.1.1 Présentation générale

Le LPC permet de distinguer les sosies labiaux (groupes de phonèmes qui lors de leur prononciation ont une même forme aux lèvres, également appelés visèmes) et ainsi de lever les ambiguïtés de l'information labiale par la vision. C'est

# Projet ARTUS: Animation Réaliste par Tatouage audiovisuel à l'Usage des Sourds

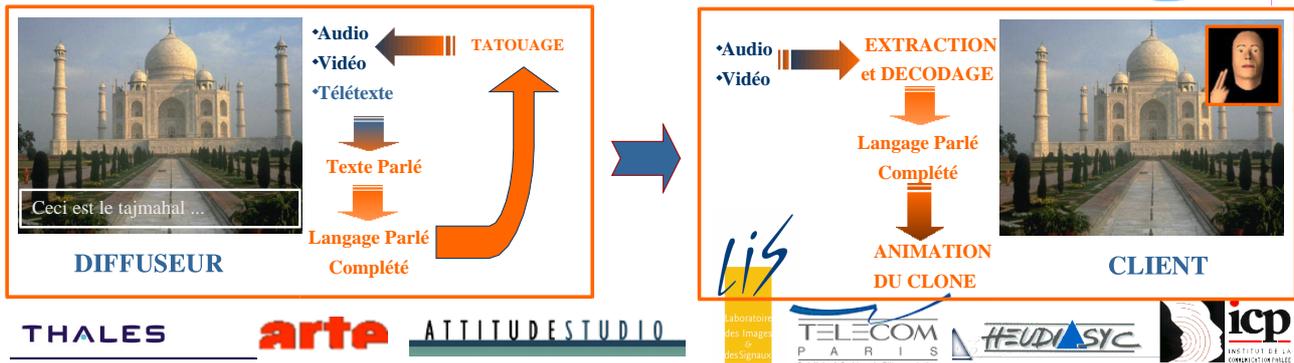


FIG. 1: Présentation générale du projet Artus

un moyen de “rendre les sons visibles”.

Le locuteur exécute, pendant qu’il parle, une succession de gestes de la main, appelés clés. Chaque geste est composé de deux paramètres : la configuration de la main (position des doigts les uns par rapport aux autres) et la position de la main autour de la bouche. La forme adoptée par la main lève les ambiguïtés des consonnes tandis que l’endroit où se place la main lève les ambiguïtés des voyelles (cf Fig. 2).

Les phonèmes qui peuvent être confondus sur les lèvres (par exemple : /p/, /b/, /m/) sont représentés par des clés différentes afin de bien les différencier, tandis que ceux qui ont des images labiales clairement distinctes (par exemple : /p/, /d/, /j/) sont regroupés sous une même clé. Il en est de même pour les voyelles. Ainsi, on peut transmettre toute l’information phonologique avec un nombre limité de clés.

Cependant, il est important de noter que les configurations et les positions de la main ne fournissent pas à elles seules l’information phonologique non ambiguë (elles sont compatibles avec plusieurs phonèmes différents). C’est la combinaison des mouvements des lèvres et des mains qui permet de transmettre visuellement cette information avec la même précision que l’information audio-visuelle pour les entendants.



FIG. 2: Les 5 positions de la main pour coder les voyelles.

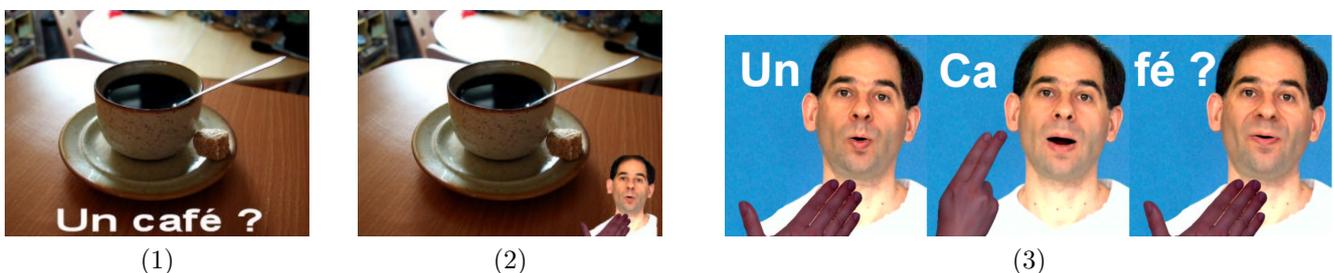


FIG. 3: Emploi du Langage Parlé Complété dans des émissions télévisées: (1): Séquence avec télétexte, (2) Séquence où le télétexte est remplacé par un clone animé, (3) Séquence de gestes en LPC représentant la formulation “un café”

## 2.1.2 Avantages du LPC

L’emploi du Langage Parlé Complété peut être motivé par plusieurs facteurs:

- **l’intégration:** le LPC permet une intégration plus facile des sourds dans la communauté entendant alors que la LSF (Langue des Signes Française) ne permet le dialogue qu’avec les pratiquants. Le LPC est un facteur intégrant.

- **l'apprentissage:** le LPC semble être une langue plus facile à apprendre que la LSF, du moins pour les personnes ayant appris la langue parlée. Le “vocabulaire” du LPC est lié aux traits phonétiques utilisés dans la langue, donc limité à une dizaine de contrastes d'articulations non visibles. En quelques mois, les entendants peuvent facilement apprendre à coder, la pratique ne servant ensuite qu'à accroître la vitesse de codage.
- **la gestuelle:** reposant sur le code phonétique, la gestuelle du LPC est plus simple que celle de la LSF qui encode chaque mot en contexte. En termes technologiques, l'espace exploité par la LSF est tridimensionnel, va de l'estomac à la tête et déborde largement sur les côtés. Le LPC exploite au contraire des gestes “plan” d'une main référencée au visage.
- **le cadrage:** d'un point de vue technique, l'étude et la synthèse des gestes du LPC est donc a priori plus simple que ceux de la LSF, où de toutes façons les gestes du visage et notamment les expressions faciales sont importants. Le LPC permet ainsi un cadrage serré sur le visage et donc l'usage d'une plus petite fenêtre d'insertion sur l'écran que le cadrage plus large sur le haut du personnage imposé par la LSF.

## 2.2 Le tatouage audio et vidéo

Après l'utilisation du LPC, les techniques de tatouage représentent la deuxième notion clef de ce projet. Le tatouage peut se définir comme l'insertion d'une information imperceptible et indélébile au sein d'un contenu. La nature de ce contenu peut être multiple, il peut s'agir par exemple d'images, de sons, de vidéos, de textes ou encore de composants électroniques. Dans le cadre du projet Artus, nous nous attacherons à tatouer des contenus audio-vidéo représentés par les émissions télévisées.

Les techniques de tatouage ont été initialement conçues pour pouvoir protéger les oeuvres numériques en insérant une information renseignant sur l'identité du propriétaire ou de l'acheteur du contenu. Ces techniques cherchant à protéger les droits des auteurs s'avèrent cependant extrêmement complexes à mettre en oeuvre car elles doivent faire face à des attaques directement destinées à effacer le tatouage.

Dans le cadre du projet Artus, l'utilisation du tatouage a été envisagée pour permettre d'étendre un contenu multimédia. Cette extension rend possible l'ajout d'une fonctionnalité initialement non prévue lors de la création du format multimédia. En effet le tatouage permet ici le transport des mouvements du clone dans des émissions télévisées, quelque soit le mode de diffusion (hertzien, numérique, ou par câble) ou l'opérateur de diffusion (CanalSat, TPS, ... ). Dans la littérature, plusieurs autres exemples utilisent le tatouage pour augmenter des contenus, on peut citer l'utilisation du tatouage pour le codage d'images couleurs [3], la création d'images stéréoscopiques étendues [2], la correction d'erreurs pour la transmission de séquences vidéo [4], l'utilisation du tatouage pour faciliter la fouille de données, ou encore l'insertion de code barres invisibles pour servir à des fins publicitaires [1].

## 3 Plusieurs verroux technologiques

Ce projet offre un enjeu applicatif de grande envergure tout en soulevant des problèmes scientifiques extrêmement motivant pour les différents partenaires de ce consortium. L'objectif de cette section est de présenter les tâches et les enjeux scientifiques clés de ce projet.

### 3.1 Transduction du télétexte en Langage Parlé Complété

Dans le cadre d'Artus, nous proposons de générer le LPC directement à partir du télétexte contenu dans la plupart des émissions diffusées. Cette tâche peut être décomposée en deux niveaux. Dans un premier temps, un signal de parole doit être généré à partir du télétexte. Ce signal doit ensuite être analysé pour générer les différents paramètres nécessaires à la compréhension du LPC : la forme des lèvres, la forme et la position de la main et la manière dont s'effectuent les transitions.

### 3.2 Modélisation et codage des mouvements du Langage Parlé Complété

Pour que les paramètres nécessaires à l'animation du clone puissent être transmis par tatouage, il faut que leur représentation soit la plus condensée possible. Il est donc nécessaire de procéder à un codage source de ces paramètres. De par le caractère multimodal de cette information (positions des lèvres, de la main et transitions entre les mouvements), la dépendance entre ces paramètres et l'évolution des paramètres dans le temps doit être analysée puis ensuite modélisée. Ces modèles, linéaires ou non, permettront de réduire le nombre de paramètres à transmettre tout en ne dégradant pas au final l'intelligibilité du clone communicant (cf. Fig. 4). Cette tâche a débuté par l'acquisition de données réelles des mouvements d'une codeuse LPC qui ont été ensuite utilisées afin de concevoir des modèles de mouvements.

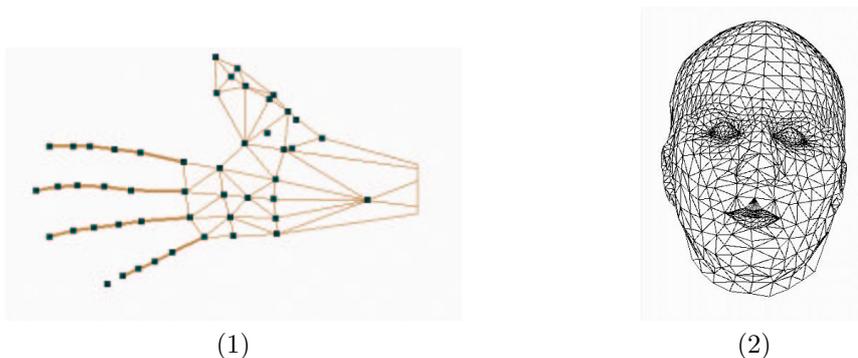


FIG. 4: Maillages utilisés pour analyser et modéliser les mouvements du codeur LCP. (1):Maillage de la main, (2): Maillage du clone.

### 3.3 Développement d’algorithmes de tatouage audio-vidéo

Concernant le tatouage des contenus audiovisuels, les contraintes spécifiques à cette application sont multiples:

- L’information ajoutée doit être aussi importante que possible de manière à ce que l’animation du clone soit aussi réaliste que possible. Il faut donc concevoir des algorithmes de tatouage haute capacité (cf Fig. 5).
- Le décodage de l’information dissimulée doit être possible quelque soit la chaîne de diffusion. Ainsi le message doit être lisible après une compression du signal, une distorsion due à une réception de faible qualité ou encore transcodage du flux audio-vidéo.
- Les techniques de tatouage mises en oeuvre doivent également être de complexité très faible afin de pouvoir être utilisées pour un décodage en temps-réel. La complexité doit être inférieure à la complexité nécessaire au décodage de la source vidéo.
- La nature des contenus doit également être prise en compte, par exemple le mouvement d’objets dans une séquence augmente la visibilité du tatouage: une marque “fixe” ajoutée sur un objet qui se déplace est d’avantage perceptible que si l’objet reste fixe.
- L’information à dissimuler doit être répartie entre deux contenus de natures totalement différentes: le son et la séquence vidéo. Afin d’offrir plus de souplesse dans le fonctionnement du codeur et du décodeur, le débit d’insertion doit également être variable pour chacune des sources.

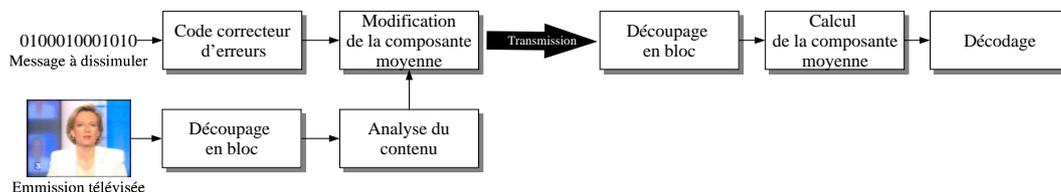


FIG. 5: Principe d’un schéma de tatouage haut-débit vidéo basé sur un codage par quantification indexée.

### 3.4 Animation réaliste du clone

Le terminal client doit être capable de décoder le tatouage en temps réel mais aussi d’effectuer le rendu d’un clone aussi réaliste que possible. Cette dernière condition assure une intelligibilité de bonne qualité. La société Attitude-Studio, partenaire du projet Artus, est spécialisée dans la génération de clones extrêmement réalistes (cf Fig 6). Elle a en charge de créer un moteur d’animation temps-réel spécifique au décodage du clone d’Artus.

## 4 Présentation des partenaires

Le consortium d’Artus est composé de deux partenaires industriels (Thalès et Attitude-Studio), un partenaire audiovisuel public (Arte-France) et 4 partenaires universitaires (LIS/INPG, ICP/INPG, TSI/ENST et HEUDIASYC/UTC). Cette section présente le rôle de chaque partenaire dans ce projet.

### 4.1 Thalès

Thalès Communications -THC- (Gennevilliers) a un très large spectre d’activités dans le domaine des communications couvrant les équipements de radio mobiles, les grands réseaux terrestres, les communications aéronautiques, spatiales



FIG. 6: Exemple d'un clone réaliste conçu par Attitude-Studio.

et navales.

Dans le cadre d'Artus, Thalès, associée à sa spin-off Nextamp, a pour objectif de développer un tatoueur vidéo temps réel et de concevoir le terminal cible permettant le décodage du tatouage.

## 4.2 Attitude Studio

ATTITUDE STUDIO est le premier prestataire spécialisé dans la conception, la fabrication et l'exploitation de personnages virtuels haut de gamme.

Ses missions au sein du projet sont d'aider à la modélisation d'un clone communiquant en LPC ainsi qu'à la conception du moteur 3D permettant l'animation de ce clone.

## 4.3 Arte-France

ARTE France - issue de La SEPT, préfiguration d'une chaîne culturelle à vocation européenne - est chargée de produire, de coproduire et d'acheter des programmes pour ARTE.

Son rôle au sein du projet Artus est d'une part de réaliser le lien entre le monde de l'audiovisuel et celui des personnes malentendantes et, d'autre part de mettre à la disposition du consortium des moyens permettant de tester les techniques développées.

## 4.4 LIS/INPG

Le LIS (Laboratoire des Images et des Signaux) est un laboratoire de l'Institut National Polytechnique de Grenoble (INPG), Unité Mixte de Recherche du CNRS (UMR 5083). Il est spécialisé en traitement du signal et en traitement de l'image.

Le LIS est également chef de file du projet Artus. Outre la coordination du projet, ce laboratoire est impliqué dans le développement de tatoueurs audio et vidéo haute capacité.

## 4.5 ICP/INPG

L'ICP (Institut de la Communication Parlée) est un laboratoire associé au CNRS (UMR n°5009) et à deux universités de Grenoble : l'INPG et l'Université Stendhal. L'ICP développe ses recherches à la fois en Sciences de l'Ingénieur et en Sciences Humaines et étudie la parole dans tous ses états : signaux, langage et cognition.

Le rôle de l'ICP dans ce projet est de fournir des techniques permettant d'acquérir, de synthétiser et modéliser le Langage Parlé Complété.

## 4.6 TSI/ENST

Le laboratoire TSI de l'Ecole Nationale Supérieure des Télécommunications à Paris possède une compétence dans les aspects théoriques des communications numériques, du traitement du signal, de la compression de la parole et des images, ainsi que dans les aspects théoriques du traitement des images.

Dans ce projet, le laboratoire TSI a pour objectif de concevoir et modéliser des algorithmes de tatouage audio et vidéo.

## 4.7 HEUDIASYC/UTC

Le laboratoire Heudiasyc/UTC (UMR CNRS 6599) de l'Université de Technologie de Compiègne est spécialisé dans les domaines de l'analyse de données, de la reconnaissance des formes pour le diagnostic de systèmes et, du traitement du signal et des images.

Ce laboratoire participe à l'étude de la représentation des expressions faciales du codeur LPC ainsi qu'au développement de tatoueurs vidéo.

## 5 Conclusion

Le projet Artus, dont la durée totale est de trois ans, est maintenant à mi-parcours. Les différentes tâches sont en cours de réalisation aussi bien sur la partie concernant le Langage Parlé Complété, que sur la partie tatouage. Ce projet a également permis d'impliquer trois doctorants ainsi que de nombreux étudiants. La deuxième phase de ce projet se focalise actuellement sur l'implémentation et les tests de la chaîne de traitements proposés dans Artus.

## Références

- [1] A. M. Alattar. Smart images using digimarc corporation's watermarking technologie. In *Proc. SPIE*, pages 264–273, January 2000.
- [2] P. Barralon, P. Bas, and J-M. Chassery. Projet de fin d'étude: Création d'images stéréoscopiques étendues par techniques de tatouage. Master's thesis, INPG, 2002.
- [3] P. Campisi, D. Kundur, D. Hatzinakos, and A. Neri. Compressive data-hiding: An unconventional approach for improved color image coding. *Applied Signal Processing*, 2002(3), Feb 2002.
- [4] Robie and Mersereau. Video error correction using steganography. *Applied Signal Processing*, 2002(3), Feb 2002.