



**HAL**  
open science

# Modélisation du squelette pour la génération réaliste de postures de la langue des signes française

Maxime Delorme

► **To cite this version:**

Maxime Delorme. Modélisation du squelette pour la génération réaliste de postures de la langue des signes française. Autre [cs.OH]. Université Paris Sud - Paris XI, 2011. Français. NNT : 2011PA112312 . tel-00672085

**HAL Id: tel-00672085**

**<https://theses.hal.science/tel-00672085>**

Submitted on 20 Feb 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***Thèse***

*pour l'obtention du titre de*

Docteur de l'Université Paris-Sud 11  
Spécialité Informatique

Modélisation du squelette pour la génération  
réaliste de postures de la langue des signes  
française.

*Maxime Delorme*

ED-Informatique de Paris-Sud

*LIMSI/CNRS*

07 décembre 2011

**Jury :**

Annelies Braffort – Directrice de recherche au CNRS – Directrice de thèse  
Patrice Dalle – Professeur à l'Université Paul Sabatier – Examineur  
Michael Filhol – Chargé de recherche au CNRS – Encadrant scientifique  
Jean Claude Martin – Professeur à l'Université Paris-Sud – Examineur  
Catherine Pelachaud – Directrice de recherche au CNRS – Rapporteur  
Daniel Thalmann – Directeur de recherche à l'EPFL – Rapporteur

**Thèse financée par l'Université Paris-Sud**

## **Résumé**

Les avancées récentes en matière d'animation ont permis le déploiement de personnages virtuels à des fins diverses et variées. Les signeurs virtuels sont des personnages en trois dimension s'exprimant en langue des signes et permettant la diffusion de messages aux personnes sourdes et malentendantes signantes de manière anonyme et modulaire. Cependant, la génération d'animations pour ces personnages dépend de la description lexicale des signes, modèle linguistique dépendant du système de génération. Les signes décrits par ces modèles sont généralement des réalisations parfaites et géométriques menant à des mouvements robotiques et peu naturels de la part du signeur. Cette thèse s'intéresse à l'addition d'informations anatomiques au squelette de contrôle du personnage virtuel de manière à le faire signer de manière plus humaine et réaliste. Ces informations supplémentaires sont regroupées sous l'appellation de "modèle anatomique" et sont divisées en cinq contributions principales : une nouvelle représentation informatique du squelette, une étude anthropométrique sur la main, l'unification de dépendances articulaires, un nouveau modèle de complexe carpo-métacarpien permettant l'opposition aisée du pouce et enfin un modèle calculant le confort d'une posture. Ces apports sont intégrés à une plateforme de génération au moyen de techniques adaptées aux contraintes imposées par les modèles linguistiques. Les travaux sont conclus par une évaluation du système ainsi qu'une réflexion sur les travaux futurs pouvant être élaborés à partir de cette thèse.

## **Mots-clés**

Informatique, Animation, Génération automatique, Cinématique inverse, Modèle anatomique, Langue des signes

## **Abstract**

Recent progresses in animation have allowed the use of virtual character to many extents. Virtual signers (or signing avatars) are three-dimension characters expressing themselves in sign language. These characters allow the the broadcasting of audio information to deaf and hearing-impaired signers in an anonymous and modular fashion. However, automatic generation of animation for such characters strongly rely on the lexical description of signs. Signs described through these models are usually perfect and geometric performances leading to robotic and unrealistic movements. These thesis focuses on adding information to the control skeleton of the signer to help him perform signs in a more human and realistic way. Such information are grouped under the name of "anatomic model" and are divided in five main contributions : a new computer-based description of the skeleton, an anthropometric study of the hand, the merging of articulatory dependencies, a new model of the carpo-metacarpal complex allowing an easier opposition of the thumb and finally a model computing posture comfort. These contributions are then implemented in a generation system through linguistic constraints-adapted techniques. The study ends with an evaluation of the system and a presentation of future prospects.

## **Keywords**

Computer science, Animation, Automatic generation, Inverse kinematics, Anatomic modelling, Sign Language

# Table des matières

Table des matières .....	1
Introduction.....	7
Chapitre I : Contexte et élaboration de la problématique .....	9
1. Langue des signes.....	10
1.1. Surdit� en France.....	10
1.2. Langue des Signes Franaise .....	12
1.3. Affichage de LSF dans les lieux publics.....	13
1.4. Aspects linguistiques de la LSF .....	14
2. G�n�ration de langue des signes .....	17
2.1. Histoire de l’animation .....	17
2.2. Principes de base de l’animation.....	19
2.3. Animation de personnages virtuels.....	21
2.4. Approches pour la g�n�ration d’animations.....	22
2.5. Mod�les linguistiques pour la g�n�ration.....	31
3. Mod�les anatomiques pour personnages virtuels.....	34
3.1. Qu’est ce qu’un mod�le anatomique ?.....	34
3.2. Anthropom�trie.....	35
3.3. Le probl�me des limitations angulaires .....	35
3.4. Solutions multiples .....	39
3.5. Gestion des collisions .....	40
3.6. Dynamique du mouvement.....	41
3.7. Expressions faciales .....	42
3.8. R�sum� et �laboration du cahier des charges .....	44
Chapitre II : Proposition de mod�le pour la repr�sentation du corps humain.....	46
1. D�composition des probl�matiques et propositions pr�liminaires de solutions .....	47
1.1. Mesures anthropom�triques et repr�sentation du squelette.....	47
1.2. Articulations, limites et naturel des postures .....	49
2. Repr�sentation du squelette.....	52
2.1. Blocs fonctionnels.....	52
2.2. Sites corporels .....	59
2.3. D�finition et nomenclature du squelette utilis� .....	61
3. Mesures anthropom�triques des bras et de la main .....	64
3.1. De la n�cessit� de mesures anthropom�triques.....	64
3.2. M�thodologie de conception des mesures.....	65

3.3. Résultats de l'étude .....	66
3.4. Conclusion de l'étude anthropométrique .....	67
4. Dépendances articulatoires.....	68
4.1. Dépendances simples .....	68
4.2. Dépendances dynamiques .....	69
4.3. Dépendances distributives .....	70
4.4. Résultats de l'utilisation de dépendances.....	71
5. Complexe carpo-métacarpien .....	72
5.1. Main humaine vs. main virtuelle .....	72
5.2. Chaîne articulaire du pouce .....	72
5.3. Modélisation du complexe carpo-métacarpien .....	73
5.4. Création d'un axe de rotation pour l'opposition.....	73
5.5. Limitations articulatoires.....	74
6. Modèle de confort.....	76
6.1. Définitions de l'effort et du confort .....	76
6.2. Elaboration d'un protocole .....	77
6.3. Corpus d'étude .....	79
6.4. Extraction du modèle .....	80
7. Conclusion du chapitre .....	87
Chapitre III : Intégration au système de génération automatique .....	88
1. Approches de cinématique inverse .....	89
1.1. Exemple .....	89
1.2. Méthodes basées sur la Jacobienne.....	91
1.3. Descente de Coordonnées Cycliques (CCD) .....	92
1.4. Exploration par Monte-Carlo.....	92
1.5. Descente de gradient .....	92
1.6. Autres méthodes et méthodes hybrides.....	93
2. Problématiques liées à la génération automatique de langue des signes.....	94
2.1. Placements relatifs et absolus.....	94
2.2. Dépendance de contraintes .....	95
2.3. Résolution multiple .....	95
3. Elaboration d'une méthode de cinématique inverse adaptée .....	96
3.1. Définition des cibles .....	96
3.2. Phase exploratoire.....	98
3.3. Optimisation locale des solutions .....	100
3.4. Sélection du résultat par le modèle de confort.....	101

3.5. Limites de l'approche .....	102
4. Ingénierie des modèles .....	104
4.1. Bibliothèque de squelette .....	104
4.2. Bibliothèque de séquences .....	104
4.3. Système de génération.....	104
4.4. Gestionnaire et connexion à l'environnement du LIMSI.....	105
5. Conclusion du chapitre.....	106
Chapitre IV : Résultats et discussions.....	107
1. Modèle de pouce opposable.....	108
1.1. Résultats.....	108
1.2. Apports.....	111
1.3. Perspectives.....	112
2. Système de génération et modèle de confort.....	114
2.1. Etude préliminaire .....	114
2.2. Protocole .....	114
2.1. Pertinence du modèle de confort .....	117
2.2. Résultats de génération.....	118
2.3. Résultats du confort .....	119
2.4. Apports .....	123
2.5. Perspectives.....	124
3. Conclusion du chapitre.....	127
Chapitre V : Perspectives générales et conclusion.....	129
1. Perspectives.....	133
1.1. Extension du modèle aux expressions faciales .....	133
1.2. Animation du signeur .....	133
1.3. Gestion des collisions .....	134
1.4. Evolutions du modèle en rapport avec la peau.....	135
1.5. Mouvement ambiants .....	135
1.6. Application à la reconnaissance de langue des signes .....	136
1.7. Applications à d'autres domaines .....	136
2. Résumé et conclusion.....	130
Bibliographie .....	140
Annexe 1 : Terminologie .....	144
1. Terminologie biologique.....	144
1.1. Repérage en anatomie .....	144
1.2. Nomenclature du squelette .....	146

1.3. Mouvements du corps humain .....	148
2. Terminologie d'animation .....	149
2.1. Notions sur le squelette .....	149
2.2. Quaternions.....	150
2.3. Cinématique inverse.....	151
Annexe 2 : Définition du squelette et dépendances.....	154
1. Normes des blocs terminaux du squelette.....	154
2. Dépendances articulaires connues .....	155
Annexe 3 : Paramètres du système de génération .....	157
Annexe 4 : Comparatif des systèmes de génération.....	159

## Introduction

Depuis les années 90, les images de synthèse et leur incrustation au sein de vidéos sont monnaie courante. Les avancées de la technologie, liées aux progrès de la recherche dans le domaine du graphisme, ont permis la création de personnages virtuels animés par ordinateur. Particulièrement présents dans l'industrie cinématographique, on leur trouve d'autres applications, notamment dans le champ de l'accessibilité à l'information pour les personnes sourdes.

L'accès à l'information pour les personnes sourdes peut être grandement simplifié par l'utilisation de ces personnages virtuels dans un contexte de traduction de l'information en langue des signes. Cependant, l'information est rarement disponible sous forme monolithique et complète, il est souvent nécessaire de morceler le discours en petites unités gestuelles que l'on va chercher à faire réaliser au personnage virtuel. De nombreuses méthodes d'animation et de contrôle existent pour donner vie à ces personnages virtuels et potentielle leur permettre de pratiquer la langue des signes. Cependant, celles qui répondent au mieux à cette problématique donnent généralement des résultats robotiques et peu naturels voire parfois impossible à imiter par un humain.

Ce manque de naturel peut limiter la compréhension des signes et nécessite d'être corrigé. De nombreux travaux de recherche s'intéressent à la génération de mouvements naturels en développant de nouvelles méthodes d'animation, mais peu s'intéressent à améliorer la compréhensibilité des signes. C'est pourquoi nous choisissons dans cette thèse de nous centrer sur le processus en amont de l'animation : la définition d'une structure et d'une hiérarchie permettant de contrôler le personnage virtuel tout en incluant des informations supplémentaires visant à pallier ce manque de naturel.

Il existe à l'heure actuelle un certain nombre de standards de représentation du personnage virtuel mais le niveau de détail du modèle est généralement fonction de l'application. Les modèles plus complexes sont trouvés dans le domaine de la médecine et sont d'un niveau de détail trop élevé pour être utilisables en animation. Par ailleurs, les modèles courants en animation sont généralement dénués de connaissances biologiques et anatomiques, menant la génération aux situations peu naturelles.

Nous nous posons ainsi la question suivante : de quelle manière est-il possible d'accroître le réalisme des postures et des animations du personnage virtuel, de manière à améliorer la

compréhensibilité ? Nous proposons d'y répondre dans cette thèse par l'intégration d'un modèle biologique/anatomique pour la représentation du personnage virtuel. Nous l'appliquons au domaine de la génération automatique de postures de langue des signes. L'approche que nous proposons est pluridisciplinaire puisque centrée sur un travail en informatique incluant des éléments de biologie, de statistique et de linguistique.

Le premier chapitre s'attache à décrire le contexte de l'étude, présente la langue des signes ainsi la génération d'animation pour ce domaine. A partir d'un état de l'art des techniques d'animations existantes, il pose la problématique générale de l'étude et définit le cahier de charges de cette thèse. Le second chapitre présente le modèle théorique de squelette ainsi que les différents apports de recherche de la thèse, notamment en termes de représentation abstraite de l'anatomie. Le troisième chapitre s'intéresse à l'implémentation et les liens créés entre le modèle théorique du chapitre II et le processus de génération des postures/animations. Il comporte un bref état de l'art des techniques de génération de posture par cinématique inverse, ainsi qu'une nouvelle méthode élaborée à partir des contraintes du contexte. Le quatrième chapitre présente les protocoles d'expérimentation ainsi que les résultats du système de génération de postures. Enfin, le dernier chapitre ouvre les travaux de cette thèse à différentes perspectives, futurs travaux et valorisations puis conclut l'étude.

## **Chapitre I : Contexte et élaboration de la problématique**

L'objectif de ce chapitre est de présenter le contexte des travaux décrits dans cette thèse. La première section présente le contexte socio-linguistique de l'étude : la surdité de manière générale mais aussi la langue des signes et les moyens mis en place au niveau national pour l'accès des sourds à l'information. La seconde section fait l'état de l'art du contexte scientifique et présente la génération d'animation ainsi que les problématiques spécifiques à notre domaine d'étude. Enfin, la troisième et dernière section du chapitre liste les modèles existants pour la représentation de personnages virtuels et plus particulièrement les modèles anatomiques permettant de décrire le squelette humain.

## 1. Langue des signes

Cette section pr sente le contexte g n ral de l' tude   savoir le domaine de la langue des signes. Nous y verrons dans un premier temps (1.1) une d finition des diff rents niveaux de surdit  ainsi que les causes du fort taux d'ill trisme chez les sourds. La seconde section (1.2) pr sente le moyen de communication consid r  comme le plus naturel pour les sourds : la langue des signes. La troisi me section (1.3) donne un exemple d'application des nouvelles lois sur l'accessibilit    tous les types de handicaps : l'affichage de langue des signes fran aise sur des  crans dans les lieux publics. Enfin, nous revenons dans une derni re partie (1.4) sur quelques aspects linguistiques li s   la langue des signes fran aise.

### 1.1. Surdit  en France

M dicale­ment parlant, la surdit  est un  tat pathologique d signant une perte de l'audition chez une personne. Celle-ci peut  tre li e   de multiples facteurs et survenir   diff rents stades de la vie. L'affection peut  tre de s v rit  variable et handicaper plus ou moins la personne concern e. A partir d'une perte de 30 d cibels, la pose d'un appareil d'amplification est possible, permettant de combler partiellement le d ficit auditif. La conversation et de mani re g n rale l'apprentissage de la langue vocale devient difficile   partir de pertes de 60 d cibels. Pass  le cap des 80 d cibels de perte, la communication audio devient quasi-impossible et l'apprentissage de la langue vocale pour les enfants se complique. Enfin, au-del  de 90 d cibels de perte, on qualifie la personne de « sourde profonde ». Il est n cessaire d'employer d'autres moyens pour contourner le d ficit d'audition. En 1998, Dominique Gillot, d put  du Val d'Oise,  crit un rapport sur le Droit des Sourds (Gillot 1998) o  elle comptabilise en France pr s de 3 millions de personnes sourdes, dont 100 000 sourds profonds de naissance, ainsi qu'entre 60 000 et 80 000 sourds pratiquant la langue des signes. Ce m me rapport indique que pr s de 80% des sourds fran ais sont illettr s.

La cause principale de cet illettrisme est due   l'insistance du corps enseignant   n'utiliser que le fran ais comme langue d'enseignement et de donner une (trop) grande importance   l'apprentissage de l'oral (L'Huillier 2004). Un exemple particuli rement grave en est la d cision prise au Congr s de Milan en 1878 d'interdire l'apprentissage de la langue des signes en France. Cons quence de cette interdiction, les sourds se retrouvent isol s, noy s dans des masses d'entendants ne leur permettant ni une communication ais e, ni l'int gration et l' panouissement social dont ils ont besoin. De nombreux pays dont la

France déclinent alors très rapidement en termes d'éducation des sourds. Il faudra attendre 2005 en France avant que ce bannissement de la langue des signes soit annulé par un ensemble de lois sur l'accessibilité. Désormais, on peut distinguer deux approches sur les méthodes à mettre en place pour l'éducation des enfants sourds. Ces deux approches reposent avant tout sur la langue d'apprentissage à l'école.

La première approche concerne l'apprentissage uniquement basé sur la langue française : l'oralisme. Cet apprentissage se propose en particulier de compenser les défaillances de réception du canal audio par la lecture labiale. L'émission d'information passe toujours par le canal phonatoire. Les sourds, par le biais d'orthophonistes, apprennent à vocaliser des sons et donc à parler. Enfin la dernière composante de cet apprentissage concerne l'acquisition du français écrit. Il a été montré à plusieurs reprises (Marschark 2001) que l'oralisme n'est pas une technique probante. Premièrement car l'apprentissage du français pour un entendant passe par l'écoute avant de permettre l'acquisition de l'écrit, ce qui ne peut être transposé aux sourds. Deuxièmement car il est désormais acquis (Alegria 2007) que la lecture labiale est particulièrement ambiguë dans certains contextes et ne permet pas un déchiffrement suffisamment efficace des informations.

La seconde approche conjugue l'apprentissage du français écrit avec celui de la langue des signes. Ce bilinguisme permet l'acquisition pour les sourds d'une langue première (la langue des signes) tout en en apprenant une seconde, la langue écrite (le français). Par ailleurs, l'acquisition rapide de la langue des signes permet aux enfants de recevoir des cours en langue des signes ne focalisant pas uniquement sur l'apprentissage de la langue mais aussi sur les connaissances que les enfants sont supposés acquérir durant leur scolarité.

Cette approche reste encore peu développée en France même si elle progresse. Il aura fallu attendre le 11 février 2005 pour que la loi *pour l'égalité des droits et des chances, la participation et la citoyenneté des personnes handicapées* stipule pleinement dans le code de l'éducation que la langue des signes est une langue naturelle. En outre, ce décret reconnaît officiellement la langue des signes française comme une langue à part entière.

Nous nous focalisons, dans la suite de ce chapitre, sur la langue des signes que nous considérons comme étant le moyen le plus naturel de communiquer pour une personne sourde. La prochaine section présente plus en détail les langues des signes et plus particulièrement la langue des signes française : son implication sociale ainsi que

quelques rudiments de grammaire qui nous permettront de mieux cerner les problématiques de cette thèse.

## 1.2. Langue des Signes Française

La ou plutôt les langues des signes (LS) sont des langues dites visuo-gestuelles. Elles emploient le canal visuel pour la réception d'information et le canal gestuel pour l'émission. Le canal gestuel n'emploie pas que les mains. La construction de signes en LS implique l'utilisation de l'intégralité du corps, des mains au buste en passant par chaque élément du visage (yeux, sourcils, bouche, etc.).

Les LS sont des langues utilisées au quotidien par des personnes sourdes mais aussi malentendantes et entendants. Bien qu'il soit couramment pensé qu'il n'existe qu'une seule langue des signes, universelle, il s'agit bien de plusieurs langues naturelles et différentes ayant chacune une localité géographique ainsi qu'une vie et une évolution qui leur est propre. On parle de Langue des Signes Française (LSF) pour la LS principalement pratiquée et institutionnalisée sur le territoire Français. Par ailleurs, il faut noter l'existence de langues des signes internationales (LSI). Ces dernières sont issues de pratiques naturelles entre sourds, mais le fait qu'elles ne soient pas rattachées à un territoire particulier contraint leur utilisation à un cadre strictement international. Les LS sont qualifiées de langues, et pas de langages, car elles remplissent les mêmes fonctions que les langues vocales. Elles ne sont en aucun cas un codage de l'écrit comme le braille peut l'être pour l'alphabet latin. Pourtant, longtemps considérées comme telles, il aura fallu attendre les travaux de Stokoe (Stokoe 1960) pour qu'elles soient observées scientifiquement comme des langues à part entière.

La loi du 11 février 2005 invite à la mise en place d'actions en faveur de la diffusion de la LSF. Nous pouvons citer entre autres la traduction de certains programmes télévisés tels que les informations ou le sous-titrage des programmes aux grandes heures d'écoute. Plus généralement, la loi contraint les lieux publics à procurer un accès aux informations diffusées en leur sein quelque soit le handicap. Ce qui sous-entend pour les personnes sourdes un accès en langue des signes aux informations audio diffusées dans les lieux public. Cette perspective entraîne depuis 2005 un intérêt considérable en termes de recherches sur la langue des signes, et plus particulièrement sur les méthodes de génération d'énoncés. Il n'est pas raisonnable d'envisager qu'un ou plusieurs interprètes soient embauchés à plein temps pour traduire en LSF et au cas par cas toutes les informations audio diffusées dans des

lieux public de haute affluence telles que les gares ou les aéroports. En effet, les informations diffusées dans ces contextes sont générales et s'appliquent à un grand nombre de voyageurs. Elles ne peuvent pas être traitées au cas par cas pour chaque personne sourde.

La production d'énoncés en langue des signes devient dès lors un défi qu'il faut résoudre. Le cadre induit deux problématiques sous-jacentes : Premièrement, comment diffuser à grande échelle des informations en langue des signes ? Deuxièmement, comment répondre aux attentes des personnes sourdes en termes d'information personnalisée (e.g. à un guichet).

La suite de ce chapitre se concentrera sur la première problématique.

### 1.3. Affichage de LSF dans les lieux publics

La solution intuitive à ce problème serait de disposer d'écrans dans les lieux publics et de diffuser des annonces en langue des signes en même temps que les versions audio. Ceci soulève néanmoins d'autres questions : de quelle forme doivent être les annonces ? Doivent-elles être jouées une ou plusieurs fois ? Comment prévenir les personnes sourdes qu'un message va être diffusé et qu'il faut qu'ils soient attentifs aux écrans ?

Sur la question du format deux solutions s'offrent à nous. La première, intuitive, serait de diffuser des vidéos d'interprètes traduisant le message audio. Elle a néanmoins deux inconvénients. Le premier concerne l'anonymisation. (Aznar et al. 2005) décrit ce problème en disant que « l'anonymat permet au récepteur du message de se concentrer sur le message et non le locuteur ». Il s'agit d'un point important : une personne ne peut être amenée à représenter l'organisme pour lequel elle signe le message. En découle d'autres problèmes (comme la responsabilité du dit locuteur vis-à-vis des clients), qui poussent à penser que l'utilisation de vidéos d'interprètes ou de locuteurs sourds n'est pas une option réellement viable.

Le second inconvénient, technique, concerne la modularité des vidéos. La constitution des vidéos nécessaires à l'interprétation des différents énoncés diffusés peut effectivement être coûteuse. Mais ces films, *in fine*, ne seront pas modifiables. Dans bien des cas ce n'est pas optimal. Par exemple lorsque l'on doit afficher des énoncés à structure fixe et parties variables, que l'on peut trouver par exemple dans les transports en commun : « Le train numéro ### arrivera avec ### minutes de retard ». Il semble particulièrement déraisonnable (et impossible en pratique) de constituer des vidéos couvrant toutes les possibilités de

phrases ce type. On pourrait alors imaginer des segments d'énoncés concaténés<sup>1</sup> les uns à la suite des autres. Les vidéos ainsi générées seraient coupées au début de chaque segment d'énoncé ce qui produirait un rendu particulièrement artificiel avec des ruptures d'images. Le message final n'en serait que plus difficile à comprendre.

La seconde solution consiste à générer des énoncés en langue des signes via un personnage virtuel appelé *signeur virtuel*. Le processus de génération étant totalement informatisé, il permet d'outrepasser les limitations induites par l'utilisation de vidéos d'interprètes. Cette solution présente à son tour un certain nombre d'inconvénients, comme par exemple la difficulté technique à reproduire le naturel du mouvement humain ou encore l'articulation des différents signes entre eux pour rendre l'énoncé réaliste et compréhensible. Ces obstacles constituent des défis scientifiques pouvant être résolus grâce aux efforts de la recherche et de l'ingénierie en linguistique informatique ainsi qu'en graphisme et en animation.

Le choix d'un tel système nécessite une stratégie en amont de la diffusion afin de produire les animations. La production peut se faire via trois modalités que nous allons développer dans la section 2. Avant de considérer ces aspects informatiques, nous allons expliciter quelques éléments de grammaire de langue des signes qui permettront de mieux cerner les problématiques.

#### 1.4. Aspects linguistiques de la LSF

L'objectif de cette section est de présenter quelques éléments du fonctionnement de la LSF permettant de mieux cerner les difficultés liées à l'animation d'un signeur virtuel. Il ne s'agit pas de présenter une liste exhaustive de propriétés linguistiques mises en jeu, mais plutôt d'apporter quelques éléments à la connaissance du lecteur en termes de définitions et de contexte.

La LSF, comme nous l'avons évoqué auparavant, implique l'utilisation des articulateurs de la moitié supérieure du corps : les mains, les bras, le torse, la tête, les yeux et le visage. La conjonction de tous les articulateurs crée un flot d'information en simultané que l'interlocuteur va capter et interpréter. La composante manuelle est importante puisqu'elle est source d'une grande part de l'information émise. Le visage l'est aussi puisque les expressions faciales et le regard ont un rôle déterminant dans la désambiguïsation de certains signes, dans la qualification (gros, petit, etc.) des énoncés ainsi que dans les modes

---

<sup>1</sup> C'est-à-dire mis bout à bout.

de discours (questions, négations), etc. La zone devant le signeur où les signes sont réalisés est appelée *espace de signation* (cf. Figure 1).

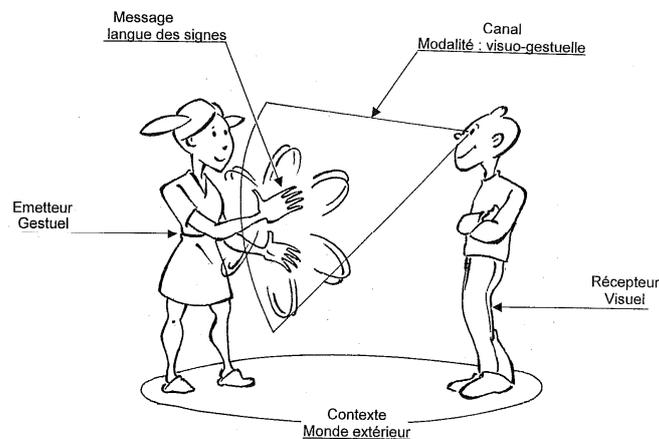


Figure 1 L'espace de signation désigne la zone devant le locuteur où les signes sont réalisés

L'une des particularités de la LSF est sa capacité à représenter des scènes et des situations de manière iconique (Cuxac 2000). Certains signes lexicaux se basent sur cette iconicité pour se réaliser comme par exemple le signe [IMMEUBLE] (cf. Figure 2).

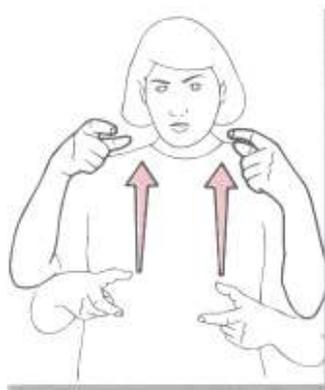


Figure 2 Signe [IMMEUBLE] en LSF.

Cependant, l'iconicité ne se limite pas aux signes standards et prend partie intégrante dans des structures narratives plus spécifiques de la LSF. Par exemple, il est possible d'effectuer une mise en abîme lorsque le locuteur raconte un dialogue entre deux personnes, en incarnant chacun des personnages tour à tour. Pour ce faire, le locuteur va tourner son buste de manière à jouer un personnage, puis se tourner de l'autre côté pour jouer le second.



Figure 3 Prise de position par un locuteur en langue des signes.

De même il est courant en LSF de décrire des scènes en plaçant des éléments narratifs dans l'espace de signation. Ces éléments peuvent être rappelés ultérieurement de manière simple en se contentant de désigner leur position dans l'espace de signation, à l'aide du regard ou d'un geste de pointage.

De cette très courte description du fonctionnement de la LSF, il faut retenir trois caractéristiques fondamentales :

- *Simultanéité* des informations émises
- *Iconicité* des signes et des énoncés
- *Spatialisation* des éléments du discours en les plaçant dans *l'espace de signation*.

La section suivante s'attache à décrire les différentes méthodes permettant d'employer ces connaissances pour générer des animations pour signeur virtuel.

## 2. Génération de langue des signes

L'animation d'un personnage virtuel, et à plus forte raison pour la langue des signes, est une discipline requérant deux types d'informations : un personnage à animer et une séquence de postures indiquant comment le personnage doit bouger dans le temps. Cette partie décrit dans un premier temps l'historique (2.1) et les principes fondamentaux de l'animation informatique (2.2). Nous présentons dans un second temps les modèles (2.3) et les techniques (2.4) utilisés pour la création d'animation de signeurs virtuel. Enfin, nous nous focalisons d'avantage sur la problématique de la langue des signes, en présentant les modèles linguistiques utilisés pour la génération d'animations en LSF (2.5).

### 2.1. Histoire de l'animation

L'animation en trois dimensions par informatique prend ses racines au début des années 1970. A cette époque, les séquences animées sont très courtes et en fil de fer. Les capacités et l'intérêt particulièrement limité de ces courtes séquences en font un objet de dérision appuyé par la réalisation de films médiocres. Cependant, la sortie de deux films employant les techniques d'animation par ordinateur en 1991 renversent cette tendance et encouragent l'industrie cinématographique à investir plus dans la recherche et le développement de telles techniques.

Le premier de ces films est *La Belle et la Bête*, publié par les studios Walt Disney. Il permet une révolution dans la production de films d'animation en étant le premier succès du système CAPS (Computer Animation and Production System) développé en partenariat avec Pixar. Ce système est le premier à fournir aux animateurs des outils informatiques pour dessiner, encre et animer entièrement leurs scènes.

Le second film est *Terminator II : Le jugement dernier* de John Cameron. Il reste dans les annales pour ses effets spéciaux révolutionnaires, animant un personnage « liquide » de manière réaliste et convaincante, mélangeant le jeu d'un véritable acteur avec des images de synthèse.

A partir de ce moment, l'utilisation de l'informatique pour la génération d'effets spéciaux et d'animations en trois dimensions prend un véritable essor. En 1995, Pixar et Disney produisent *Toy Story* le premier film entièrement généré informatiquement. Les années 2000 marquent un tournant pour les effets spéciaux au cinéma. Les animations générées par ordinateur deviennent alors les principales sources d'effets spéciaux reléguant leurs

prédécesseurs (stop-motion<sup>2</sup> et animatroniques) au rang d'antiquités. La technologie permet à cette époque de créer des doubles virtuels d'acteurs et de les intégrer dans les films de manière réaliste. Il s'agit des premiers personnages virtuels photoréalistes.

---

<sup>2</sup> Le stop-motion est une technique consistant en la prise de photos successives de maquettes, donnant l'illusion du mouvement. Ce sont ces techniques qui étaient par exemple utilisées dans Godzilla ou King-Kong.

Les années 2000 représentent aussi la consécration de la 3D dans les jeux vidéo. L'aspect narratif et immersif de ces derniers ainsi que leur grand succès commercial permet un développement accru des personnages virtuels et des techniques d'animation. Désormais, les jeux vidéo représentent un banc d'essai pour différentes techniques d'animation et de rendu, créant un réel pont entre les disciplines de recherche et l'industrie du spectacle. La considérable augmentation des puissances de calcul des ordinateurs et des stations de travail permettent désormais l'intégration d'images animées de plus en plus détaillées et réalistes.

Les techniques permettant d'animer automatiquement, ou d'aider les graphistes à animer des personnages virtuels sont désignées dans la suite de ce texte par les termes *génération* et *synthèse* d'animation. Malgré des origines sémantiques légèrement différentes, ces deux termes sont utilisés pour indiquer le processus permettant de créer une animation finale pour un personnage virtuel.

L'étude de moyens permettant de synthétiser des animations a été largement détaillée au cours des trois dernières décennies. La section suivante présente plus en détails les principes fondamentaux d'animation en informatique.

## 2.2. Principes de base de l'animation

L'animation d'une image sur un écran, qu'il s'agisse d'un film ou d'une vidéo quelconque, est un enchaînement à un rythme rapide d'images statiques. Au cinéma par exemple, les films sont projetés à une vitesse de 24 images par seconde. En Europe, le système PAL/SECAM correspond à 25 images par seconde alors qu'aux Etats-Unis et au Japon, la norme NTSC fixe la projection à 30 images par seconde.

Animer une vidéo informatiquement revient donc à créer une succession d'images statiques (*frames*) qui vont être enchaînées à une vitesse prédéfinie. Il en est de même pour la création d'animations avec un signeur virtuel. Le signeur sera traditionnellement placé dans une *scène* en trois dimensions. La scène est un espace virtuel disposant d'*objets* pouvant être des représentations matérielles (des personnages, des décors) ou virtuelles ayant une incidence sur la scène comme des sources de lumière ou une *caméra*. La caméra est une figuration abstraite représentant la position des yeux de l'observateur de la scène. Le procédé consistant à transformer une scène en image vue depuis la caméra est appelé *rendu*.

L'animation d'une vidéo consiste donc à rendre successivement des images de la scène en faisant varier très légèrement la position, ou la configuration des objets. Il n'est cependant pas raisonnable de vouloir faire varier chaque paramètre à chaque image. Pour plusieurs heures d'animation à 24 images par seconde, le travail serait particulièrement fastidieux. Il existe plusieurs façons d'accélérer le traitement. Par exemple, il est possible de définir des *images clés* (ou *keyframes*) pour l'animation. Ces images clés sont des instants particuliers où l'on définit la position et la configuration des objets de la scène. Puis à l'aide d'éditeurs de courbes, on indique au programme d'animation une interpolation pour effectuer automatiquement la transition entre les différentes images clés (cf. Figure 4).



Figure 4 Deux exemples d'images clé du squelette. Le squelette est représenté par des os en formes d'octaèdres.

Considérant les objets virtuels de la scène comme immuables (pas de changement de lumière ni de position de la caméra), les seuls paramètres susceptibles d'évoluer lors de la génération d'une animation en langue des signes sont ceux du signeur virtuel.

La section suivante détaille le modèle de représentation du signeur virtuel en informatique ainsi que ses paramètres d'animation.

### 2.3. Animation de personnages virtuels

L'animation a défini au cours des années ses propres standards en termes de modèles de représentation. Majoritaires<sup>3</sup> dans le domaine, les modèles à base de peau/squelette sont les plus simples à utiliser et conviennent à tous les types de figures animées, incluant les signeurs virtuels. Ils consistent en deux systèmes fonctionnant en parallèle : un *squelette* permettant le contrôle des articulations principales du personnage, et une *peau* permettant une visualisation du personnage.

L'utilisation d'un tel modèle permet de distinguer clairement les deux caractéristiques principales du personnage virtuel : son visuel, et son interface de contrôle. Alors que les recherches sur la peau du personnage ont plutôt tendance à s'intéresser aux techniques de rendu et notamment de photoréalisme (Courgeon, Jacquemin, et al. 2008), les travaux sur le squelette sont en lien avec le contrôle de l'animation. Dans la suite de cet exposé, nous nous concentrerons sur le squelette du signeur.

L'objectif d'un modèle de squelette est de représenter l'objet à animer sous la forme d'un ensemble d'os articulés entre eux. Il est important de noter que la notion d'os et la mécanique interne du squelette du signeur virtuel sont très loin de coller au fonctionnement de leurs homologues biologiques. Il existe des modèles de squelette simulant précisément l'anatomie du corps humain, mais ces derniers sont généralement utilisés dans le domaine médical plutôt que dans celui de l'animation (Chao 2003). Les deux domaines ne considèrent pas les mêmes problématiques, la médecine a tendance à se concentrer sur la précision et le réalisme des modèles pour des applications telles que la simulation d'opérations ou l'apprentissage. Cette précision n'est pas nécessaire pour l'animation et peut même être une source de temps de calcul inutile.

Les os sont modélisés au sein du squelette comme étant des objets rectilignes (cf. Figure 4) connectés entre eux par des articulations. Chaque articulation permet une rotation selon un (e.g. une phalange), deux (e.g. le poignet) ou trois (e.g. l'épaule) axes ou *degrés de liberté*. Ces rotations permettent de positionner des articulations dans l'espace. Malheureusement, la représentation du squelette ne permet pas un positionnement direct d'un articulateur dans l'espace. Elle nécessite d'affecter à chaque articulation le bon jeu d'angles. Le processus permettant de calculer automatiquement les angles des articulations

---

<sup>3</sup> A notre connaissance, l'utilisation efficace d'autres modèles implique une réelle spécificité de l'application. Par exemple pour la médecine.

du squelette afin de positionner un effecteur dans l'espace est appelé *cinématique inverse*. La représentation précise du squelette, ainsi que les processus de cinématique directe et inverse sont présentés plus en détail dans la suite de cette thèse (chapitres II et III). La constitution d'une animation pour un personnage virtuel revient donc à affecter à chaque image une *configuration* du squelette. C'est-à-dire un jeu d'angles qui seront affectés aux articulations du personnage virtuel afin de le positionner dans une posture particulière.

Lorsque l'on considère la création d'animations dédiées à la langue des signes, le réalisme des animations constituées est particulièrement important. Afin de générer des vidéos réellement utilisables en situation, il est nécessaire de créer un mouvement à la fois fluide et réaliste afin de produire une animation rendant l'énoncé compréhensible par les sourds. Pour cela, l'animation d'un personnage virtuel nécessite l'apport d'informations supplémentaires. La prochaine section s'attache à décrire puis comparer trois méthodes de génération d'animations pouvant être utilisées pour la synthèse de langue des signes.

## 2.4. Approches pour la génération d'animations

Il existe diverses méthodes permettant la génération d'animations pour signeurs virtuels. On peut citer les techniques requérant l'intervention d'un animateur professionnel. Celles-ci nécessitent une aide et sont dites *manuelles* selon le degré d'intervention de l'animateur dans le processus de génération et sur la finalisation de l'animation. Il existe deux méthodes principales de génération manuelle : la *rotoscopie* et la *capture de mouvement*, présentées respectivement en sections 2.4.1 et 2.4.2. En dehors de ces deux techniques, il existe aussi des méthodes de génération *automatique* d'animations pour signeurs virtuels. Celles-ci seront détaillées dans la section 2.4.3. Enfin, la section 2.4.4 établit un comparatif des différentes méthodes et synthétise l'approche.

### 2.4.1. Rotoscopie

La rotoscopie est une technique d'animation brevetée en 1915 par les frères Fleischer. A l'époque, technique de dessin à base de photographie, le rotoscope est une table à dessin sur laquelle sont projetées des images (cf. Figure 5). L'artiste peut ainsi dessiner par-dessus la projection comme s'il s'agissait d'un calque. Le principe est par la suite utilisé pour projeter les images successives d'une pellicule vidéo et permettre la création de dessins animés.

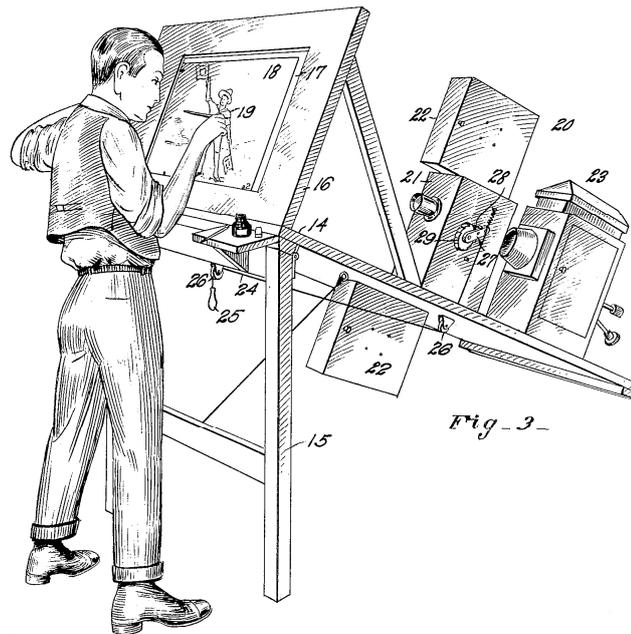


Figure 5 Un rotoscope. Image extrait de Wikipedia.

Cette méthode est désormais utilisée en infographie pour créer des animations réalistes de personnages à partir de vidéos d'acteurs. L'application de ce concept à la génération d'animations pour signeurs virtuels permet la création de mouvements particulièrement réalistes du point de vue de leur dynamique (Braffort et al. 2011). Bien évidemment, les techniques informatiques permettent désormais de s'abstraire du dispositif du rotoscope. Des prises de vues d'un *signeur source* sont réalisées au moyen de caméras placées face à lui et sur un des côtés (gauche s'il est gaucher, droit s'il est droitier). Ces prises de vues sont ensuite synchronisées et affichées sur des plans dans un logiciel de modélisation et d'animation en 3D (tels que Maya, 3DSMax ou encore Motion Builder). L'animateur place le signeur virtuel au centre de la scène, et calque les mouvements du personnage de façon à recouvrir ceux du film (cf. Figure 8). Pour cela, il choisit les images clés significatives et calque la posture du personnage (de son squelette) sur celle du signeur source. L'animateur doit aussi calibrer la dynamique du mouvement de façon à ce que les transitions entre les images clés se fassent de la manière la plus naturelle et réaliste possible. Les logiciels d'animation fournissent les outils nécessaires à l'animateur pour qu'il puisse s'assurer de la qualité de l'animation finale.

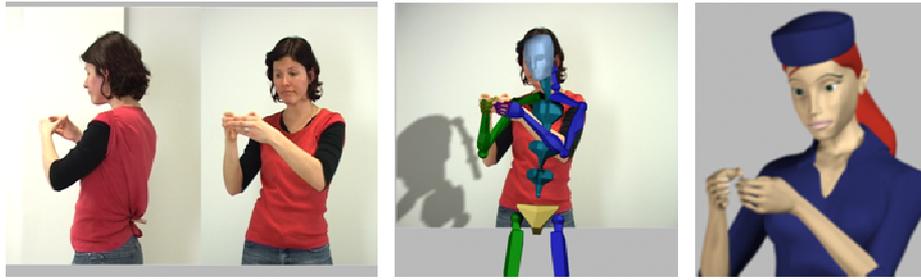


Figure 6 Exemple d'animation créée par rotoscopie.

Malgré des résultats d'une qualité remarquable, la rotoscopie n'a malheureusement que peu de succès pour la génération d'animations en langue des signes. Les causes sont multiples. Premièrement, la technique calque l'animation du signeur sur celle d'une véritable personne. Bien que le signeur soit généralement un personnage virtuel dont le physique ne ressemble que peu à celui du locuteur de la vidéo source, la dynamique et la façon de signer peut être reconnue et attribuée au locuteur sourd filmé. La rotoscopie ne permet donc pas de pallier intégralement les problèmes d'anonymisation.

Comme nous venons de le mentionner, la physionomie du signeur est généralement différente de celle de la personne dont on va copier les mouvements. Cela demande une adaptation des animations à la morphologie du signeur. Les signeurs virtuels, pour répondre au critère d'anonymisation, sont choisis de manière à être « neutres » physiquement, par exemple d'une taille et d'une corpulence moyenne. L'animation doit être adaptée de manière à prendre en compte les variations physionomiques du signeur source.

Cette adaptation ainsi que le choix des images clés à définir dans l'animation nécessite une connaissance approfondie de l'animation, du mouvement et de la dynamique, ainsi qu'un réel talent artistique (l'utilisation de la rotoscopie pour l'animation de signeurs virtuels nécessite non seulement de l'expérience et du talent, mais aussi une connaissance approfondie de la langue des signes. Très peu de personnes possèdent cette double expertise). Il est important de noter aussi que la génération d'une animation par rotoscopie est un processus coûteux en temps. Pour ces raisons, l'animation par rotoscopie n'est que peu utilisée à l'heure actuelle dans les domaines de recherche tels que la génération d'énoncés à parties variables. Elle reste tout de même très prisée dans certains domaines de production de vidéos de langue des signes, telles que la narration d'histoires<sup>4</sup>, ou les applications éducatives.

<sup>4</sup> Par exemple, « A forest » proposé par l'entreprise VCom3D.

### 2.4.2. Capture de mouvement

La capture de mouvement est une autre technique de génération d'animation manuelle. A l'origine créée pour des applications médicales (Furniss 1999), la capture de mouvement a vu son plein potentiel exploité dans l'acquisition de mouvements pour la création d'animations. Désormais très largement répandu au cinéma et dans les jeux vidéo, elle permet l'acquisition de volumes de données importants en un temps plus raisonnable que la rotoscopie.

Le terme « capture de mouvement » désigne toutes les techniques d'acquisition du mouvement, et à plus forte raison du mouvement humain. Elle permet, de manière générale, de recopier automatiquement le mouvement d'un acteur sur celui d'un personnage virtuel. Usuellement, ces techniques utilisent des caméras pour suivre le mouvement de marqueurs dans une scène en trois dimensions. On dispose les marqueurs sur le corps de l'acteur de façon à permettre la reconstitution informatique de son squelette. Ils sont donc placés sur ses articulations et sur différents points facilement identifiables du squelette.

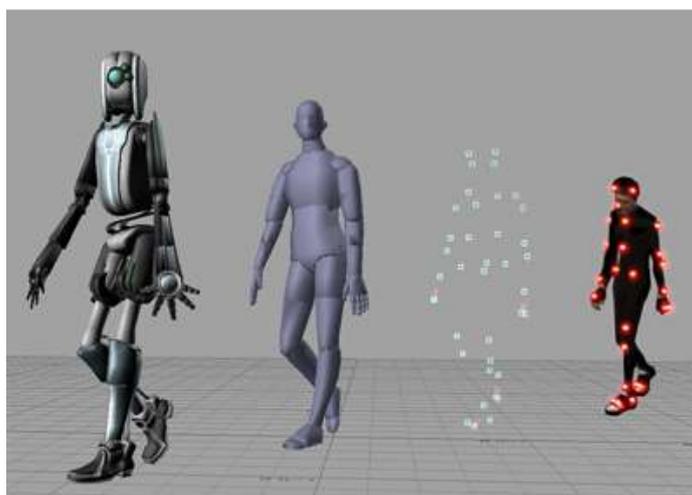


Figure 7 Exemple de capture de mouvement<sup>5</sup>.

Pour cela, l'acteur est placé au centre d'une scène constituée de multiples caméras. Ces dernières vont suivre des marqueurs disposés sur le corps de l'acteur (cf. Figure 7) et reconstituer informatiquement leur position en trois dimensions. Les marqueurs sont disposés judicieusement de façon à reproduire la position des articulations du personnage (épaules, coude, poignet, colonne vertébrale, etc.). Ainsi l'ordinateur enregistre des séquences de postures du squelette reconstitué de l'acteur. L'animation ainsi constituée est ensuite transférée au personnage virtuel.

---

<sup>5</sup> Image extraite de l'article Capture de Mouvement de l'encyclopédie Wikipedia.

La capture de mouvement permet la génération d'animations particulièrement réalistes puisqu'elles sont calquées sur les mouvements d'un véritable acteur. En revanche, ces technologies sont généralement très coûteuses pour obtenir des résultats corrects (nous verrons par la suite pourquoi). On retrouve de la génération de langue des signes par capture de mouvement au sein de plusieurs équipes et projets de recherche tels que SignCom (Duarte et al. 2010), ATLAS (Lombardo et al. 2010) ou encore pour la génération automatique d'ASL<sup>6</sup> (Lu et al. 2009). Cependant, son utilisation se limite généralement à des domaines très restreints au lexique parfaitement contrôlé (Cox et al. 2002)(Duarte et al. 2011). Se placer dans une approche générale de la langue des signes (capacité à générer n'importe quel signe du lexique de la LSF) nécessite d'avoir stocké une animation de chaque signe existant. Par ailleurs, cette technique est loin d'être modulaire. Par exemple, elle ne permet pas directement de relocaliser un signe dans l'espace, phénomène particulièrement courant en LSF. Une telle transformation nécessite l'utilisation de techniques additionnelles particulièrement coûteuses et difficiles à mettre en place (Gleicher 1998).

Notons aussi parmi les inconvénients de la capture de mouvement la nécessité de retravailler chaque animation générée sur des articulateurs fins comme les mains. Nous avons déjà insisté sur l'importance de la composante manuelle dans la langue des signes. Une génération de piètre qualité sur les mains entrave grandement la compréhension des signes générés. Les animations générées par capture de mouvement nécessitent de ce fait un long travail de calibration (Lu et al. 2009) ainsi que d'être retravaillées en post-production afin d'être nettoyées des erreurs liées aux capteurs. Ces erreurs varient en fonction de la qualité du matériel utilisé. Le bruit qui en découle doit être nettoyé au moyen d'algorithmes et/ou manuellement. Un tel nettoyage est coûteux en temps ce qui contrebalance la rapidité de l'acquisition de données.

Notons aussi que l'acquisition de données par ce moyen ne peut être faite par n'importe qui. La mise en place d'un système de capture de mouvement nécessite de disposer d'un studio et du matériel de capture. L'acquisition de nouvelles données nécessite obligatoirement la présence d'un signeur natif comme acteur, ainsi qu'une compatibilité des données enregistrées (définition du squelette, placement précis des marqueurs, etc.).

Malgré la forte popularité de cette technique, ses inconvénients nous poussent à penser qu'elle n'est pas la méthode la plus adaptée à la génération de langue des signes

---

<sup>6</sup> ASL : « *American Sign Language* », la langue des signes américaine.

### 2.4.3. Génération automatique

La troisième et dernière catégorie de génération que nous présentons est automatique. Elle regroupe tous les systèmes permettant l'animation du signeur virtuel par des moyens algorithmiques. Il est important de noter que les techniques de rotoscopie et de capture de mouvement sont populaires dans le domaine de l'animation en général. Ce n'est pas le cas des méthodes de génération automatique qui sont employées quasi-exclusivement dans le domaine de génération de langue des signes et de geste co-verbaux.

Ces méthodes nécessitent en entrée des modèles représentant l'information à générer. Cette représentation peut être de deux natures différentes : linguistique et anatomiques. Les modèles *linguistiques* permettent de fournir des informations sur le contenu à générer. Il s'agit principalement de la description formelle des signes ainsi que de leur organisation au sein d'un énoncé. Ces modèles seront présentés plus en détail dans la section 2.5. Les modèles *anatomiques*, quant à eux, ajoutent une information n'ayant pas sa place au sein des modèles linguistiques. Etant au cœur de cette thèse, ces descriptions seront approfondies dans la section 3.

La qualité de l'animation générée peut être résumée à deux facteurs : le niveau de détail des modèles utilisés et la précision des algorithmes de génération. Ces derniers sont en effet responsables du calcul des postures du signeur virtuel. Ils doivent notamment prendre en compte la génération de poses statiques permettant la constitution des images clés de l'animation, ainsi que la dynamique du mouvement permettant l'enchaînement des images de manière cohérente et compréhensible par la personne réceptrice du message. A l'heure actuelle, les différents modules de génération automatique d'animation existants donnent des résultats encore très robotiques (R. Elliott et al. 2008) et peu naturels (Fotineas et al. 2008). La génération automatique permet une adaptation instantanée au contexte car elle n'est pas basée sur une réalisation spécifique comme la capture de mouvement ou la rotoscopie. En revanche, une telle méthode de génération nécessite beaucoup d'informations complémentaires (modèles linguistiques, modèles anatomiques, gestion de l'information non-manuelle). Ces modèles complémentaires sont directement responsables de la qualité du signe et leur qualité trop simpliste à l'heure actuelle font de la génération automatique une des méthodes de génération les moins réalistes, mais les plus ouvertes aux développements futurs.

Cette thèse propose de contribuer à l'enrichissement des méthodes de génération automatique par l'apport d'un modèle anatomique, détaillé dans le prochain chapitre.

#### 2.4.4. Comparatif des méthodes de génération

Le choix d'une méthode de génération d'animation se doit d'être adapté à l'usage que l'on en fait. Chacune des méthodes présentées dans les sections précédentes dispose d'avantages et d'inconvénients qui leur sont propres et il est nécessaire de s'interroger sur les objectifs finaux du système avant de décider quelle méthode choisir. Ces avantages et inconvénients sont résumés dans le Tableau 1. Six critères principaux sont utilisés pour évaluer ces méthodes de génération :

- Le *réalisme* quantifie la qualité finale de l'animation, en termes de naturel et de compréhensibilité des signes. C'est un aspect primordial de la génération de langue des signes puisque l'objectif, *in fine*, est de faire passer un message.
- La *dépendance au matériel* indique si le système nécessite la mise en place d'un équipement particulier ou s'il peut être utilisé sur n'importe quelle configuration. Idéalement, un système disposant du moins de dépendances possibles sera utilisables dans un plus grand nombre de situations.
- *Pré/Post-Traitement* indique si les données ont besoin d'être préparées ou retravaillées avant d'être utilisables.
- La *modularité* est la capacité du système à adapter un signe particulier à son contexte. Il n'est pas rare que la LSF nécessite de réaliser des signes à des emplacements précis de l'espace de signation. Le signe reste globalement le même, mais sa réalisation se fait à un endroit différent de cet espace.
- Le *temps de génération* exprime la vitesse à laquelle les données sont acquises et rendues utilisables. Ce critère va de pair avec la modularité. Si un système est à la fois modulaire et rapide, alors l'adaptation des énoncés et la construction de phrases complexes peut être fait rapidement pour subvenir aisément aux besoins de l'utilisation.
- Enfin, les *dépendances annexes* indiquent si le système de génération nécessite d'autres apports de données ou se suffit à lui-même.

Un comparatif des méthodes utilisées par les différentes équipes générant de la langue des signes peut être trouvé en Annexe 4. Il recense par ailleurs les équipes utilisant des

configurations prédéfinies pour la main, et les équipes utilisant des modèles anatomiques comme celui que nous définissons dans cette thèse.

	Avantages	Inconvénients
<b>Rotoscopie</b>	<ul style="list-style-type: none"> <li>• Réalisme : <i>fort</i></li> <li>• Dépendance au matériel : <i>nulle</i></li> <li>• Pré/Post-traitement : <i>aucun</i></li> </ul>	<ul style="list-style-type: none"> <li>• Dépendances annexes : <i>graphiste, bases de données de signes pour la génération d'énoncés</i></li> <li>• Temps de génération : <i>long</i></li> <li>• Modularité : <i>faible</i></li> </ul>
<b>Capture de mouvements</b>	<ul style="list-style-type: none"> <li>• Réalisme : <i>fort</i></li> </ul>	<ul style="list-style-type: none"> <li>• Temps de génération : <i>long</i></li> <li>• Dépendance au matériel : <i>importante</i></li> <li>• Pré/Post-traitement : <i>importants</i></li> <li>• Modularité : <i>faible</i></li> <li>• Dépendances annexes : <i>bases de données de signes pour la génération d'énoncés</i></li> </ul>
<b>Génération automatique</b>	<ul style="list-style-type: none"> <li>• Temps de génération : <i>faible</i></li> <li>• Modularité : <i>importante</i></li> <li>• Dépendance au matériel : <i>nulle</i></li> </ul>	<ul style="list-style-type: none"> <li>• Réalisme : <i>faible</i></li> <li>• Dépendances annexes : <i>modèles linguistiques et anatomiques</i></li> </ul>

Tableau 1 : Comparatif des systèmes de génération de langue des signes

A l'analyse de ces critères il apparaît clair que la modularité pointe le champ d'application du système. Les systèmes non adaptables se verront restreints à un environnement où le lexique et les structures des énoncés sont prédéfinis. On peut fournir comme exemple le système ATLAS (Lombardo et al. 2011) dédié à la génération de bulletins météo ou encore le système Octopus développé au LIMSI-CNRS (Braffort et al. 2011) et intégré à une application grand public de la SNCF pour la diffusion de messages en LSF dans les gares. Par ailleurs, les systèmes modulaires permettront la génération d'énoncés de nature plus variée mais seront contraints par manque de réalisme à des applications de recherche.

L'équipe Modélisation et Traitement Automatique de la Langue des Signes (M&TALS) du LIMSI-CNRS s'intéresse à la modélisation informatique du fonctionnement de la LSF et aux questions de génération de langue des signes. Deux systèmes distincts y sont développés : Octopus, cité précédemment, qui permet la concaténation et le rejeu d'animations générées par la rotoscopie; et GeneALS un système de synthèse automatique d'animations. Cette thèse s'inscrit dans les travaux menés sur le système GeneALS en réponse aux critères de modularité et de temps de génération. Ces deux critères permettent d'assurer la création d'un lexique généré de manière automatique, rapide et flexible.

Comme nous l'avons vu dans la section 2.4.3, la qualité de la génération, et donc son réalisme, dépend fortement des modèles utilisés en entrée. La section suivante présente de manière très succincte quelques modèles linguistiques utilisés en génération automatique et explique notre choix.

## 2.5. Modèles linguistiques pour la génération

La génération automatique s'appuie de prime abord sur des modèles linguistiques pour générer les signes et les énoncés. Ces modèles représentent l'essence linguistique des signes et de permettre en théorie la génération d'une animation nécessaire à la compréhension du signe ou de l'énoncé. De tels modèles dédiés à la génération sont encore rares du fait que ce thème de recherche est récent et que peu d'équipes travaillent sur ce sujet. De plus, la qualité des signes générés n'est, à l'heure actuelle, pas suffisante pour envisager sérieusement la synthèse d'énoncés complets, nécessitant par ailleurs la mise en place de systèmes pour les articuler. Nous n'aborderons donc dans cette section que les modèles de description des signes.

La majorité de ces modèles s'inspire de modèles linguistiques tels que ceux de Stokoe (Stokoe 1960) ou Liddell et Johnson (Liddell et al. 1989) qui sont des modèles paramétrique. Chaque signe est décrit par un tuple d'attributs prenant leur valeur dans une liste fermée. Il s'agit de représenter les signes sous la forme d'un ensemble de paramètres pour chaque main : position et orientation dans l'espace, configuration de la main et enfin mouvement. La représentation informatique la plus aboutie qui s'inspire de ces modèles s'appuie sur HamNoSys(Prillwitz 1989), un système graphique dérivé du modèle de Stokoe. Il est utilisé pour la génération de langue des signes sous une forme enrichie (incluant non seulement l'information linguistique mais aussi des éléments co-verbaux et ambiants) et redéfinie en XML : *SigML* (J.R.W. Glauert et al. 2011). Un exemple illustré et détaillé peut être trouvé en Annexe 5.

Les limitations de l'approche paramétrique sont répertoriées par Filhol dans (Filhol et al. 2006). Il présente par ailleurs une nouvelle approche, non paramétrique, ayant abouti à *Zebedee* (Filhol 2008), un modèle basé sur une approche géométrique et temporelle des signes. L'idée fondamentale de ce modèle est de représenter un signe comme une succession de postures fixes (dites *postures clé*) enchaînées par des transitions. Cette approche s'inspire de celle de Liddell et Johnson pour l'aspect temporel en ajoutant une dimension supplémentaire : chaque posture est décrite en termes de contraintes géométriques motivées linguistiquement (cf. Figure 8). Les différents articulateurs (ne se limitant pas uniquement aux mains) sont placés et orientés dans l'espace. Un exemple détaillé de description en *Zebedee* est donné en annexe 5.

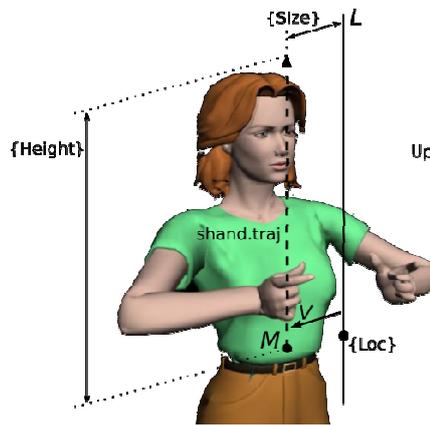


Figure 8 Exemple de contraintes géométriques définies par Zebedee. Signe [IMMEUBLE] en LSF.

La génération d'un signe consiste à résoudre un ensemble de contraintes géométriques nécessaires et suffisantes. Elle évite la sur-spécification de signes en n'imposant pas systématiquement de contraintes à l'intégralité du corps, contrairement à l'approche paramétrique. De plus, elle permet d'exprimer des contraintes sur d'autres articulateurs que les mains. Enfin, la génération des mains est faite à la volée, permettant la définition de configurations manuelles changeantes et modulaires, là où les modèles paramétriques ne savent que réutiliser des configurations de la main prédéfinie.

Concrètement, on voit une posture-clé comme un ensemble de règles à appliquer sur le squelette du signeur. Chaque règle contraint un ou plusieurs degrés de libertés du squelette. Cependant, la condition « nécessaire et suffisante » de la définition permet à une partie du squelette de ne pas être contrainte. Cette partie du squelette est donc totalement libre et n'est plus soumise qu'à des contraintes anatomiques telles que les limitations des articulations du squelette. Cette indépendance n'est pas assurée en SigML qui autorise par exemple la définition de mouvements ambiants<sup>7</sup> du signeur.

Zebedee permet par ailleurs une grande modularité dans la génération de signes en proposant un certain niveau d'abstraction dans la définition des signes. Là où SigML ne décrit qu'une instance de signe, Zebedee permet la création de *dépendances contextuelles*. Ces dépendances permettent d'indiquer quelles parties du signe sont variables et de les modifier lors de la génération. Elles permettent une adaptation du signe à son contexte, et autorisent l'utilisation d'une seule description générique pour différentes réalisations d'un même signe, alors qu'il faudrait autant de descriptions que de réalisations du signe en HamNoSys. Par exemple, on peut abstraire la position d'un signe dans l'espace afin de

<sup>7</sup> On appelle mouvements ambiants les actions effectuées n'ayant aucune incidence supposée sur le discours. Par exemple le balancement des épaules, la respiration ou certains clignements d'yeux.

pouvoir le réaliser sous sa forme standard (devant le signeur) ou le placer en contexte (quelque part dans l'espace de signation). C'est ce modèle qui est utilisé dans la plateforme de génération GeneALS.

Il est important de noter que la nature géométrique des descriptions mène naturellement à des réalisations géométriques parfaites. Un être humain ne bouge pas en termes de vecteurs ou d'arcs de cercle : la mécanique du corps humain rend le mouvement beaucoup plus complexe que de simples descriptions géométriques. De ce fait, les animations générées apparaissent robotiques à l'utilisateur. L'apport de modèles anatomiques – par exemple, gérant la dynamique du mouvement – est aussi d'assouplir l'animation pour la rendre plus réaliste et plus naturelle.

### 3. Modèles anatomiques pour personnages virtuels

Cette partie présente l'utilisation de modèles anatomiques pour la génération de postures et d'animation. Elle détaille la problématique générale de cette thèse, établit un cahier des charges des zones d'ombres de la recherche et liste les problèmes que nous avons choisis de traiter dans nos travaux.

#### 3.1. Qu'est ce qu'un modèle anatomique ?

Les modèles anatomiques sont des représentations informatiques du fonctionnement du personnage virtuel. Ils ne sont nullement réservés au domaine de l'animation et sont, par exemple, particulièrement prisés dans des champs d'application tels que la médecine ou la biomécanique. Ces modèles regroupent toute forme de contrôle du squelette liée de près ou de loin à un fonctionnement biologique.

Il existe un grand nombre de représentations valides de tout ou partie du corps humain. Chaque modèle se place à un niveau de granularité qui lui est propre : les modèles appliqués à la médecine ou la biomécanique sont généralement bien plus fournis que les modèles plus simples pour l'animation ou le jeu vidéo. Il n'existe à l'heure actuelle aucun standard de représentation du squelette en informatique, et ce malgré quelques tentatives plus ou moins réussies telles qu'H-Anim<sup>8</sup> dans le domaine de l'animation sur Internet. De même, l'absence de convention déterminant le type de modèle à utiliser en fonction de l'application pousse généralement les développeurs et les chercheurs à établir leurs propres modèles.

Lorsque l'on s'intéresse à l'animation, les modèles sont généralement simplistes, se contentant de l'information hiérarchique des os. Cet état de fait est principalement dû à leur utilisation : la grande majorité des modèles de représentations utilisés en animation sont destinés au cinéma ou au jeu vidéo. Ils ne sont pas véritablement sujets à un contrôle automatique comme nous souhaitons le faire avec la génération automatique de langue des signes. De ce fait, très peu d'information est requise en supplément de la hiérarchie car la quasi-totalité du squelette est contrôlée par l'animateur. Il convient de tempérer cette affirmation en indiquant que certains modèles utilisés par des animateurs incluent des informations additionnelles afin d'aider l'animateur à contrôler plus rapidement et plus simplement le personnage virtuel (McDonald et al. 2002).

---

<sup>8</sup> <http://www.h-anim.org>

En considérant les modèles anatomiques comme une source d'information complémentaire aux modèles linguistiques, on peut s'interroger sur l'apport que l'on souhaite avoir de leur part : quelles informations vont nous permettre de combler le manque de réalisme de la génération automatique. Les prochaines sections établissent une liste non exhaustive des apports possibles de modèles anatomiques à l'animation de manière générale et à la génération automatique de langue des signes plus spécifiquement.

### 3.2. Anthropométrie

L'anthropométrie est définie dans le *Trésor de la Langue Française*<sup>9</sup> comme les « Mesures pratiquées sur l'homme en vue de déterminer quantitativement ses caractères morphologiques, soit individuels, soit de groupes, et suivant le sexe, l'âge, le lieu, l'époque, etc. ». La constitution de bases de données d'anthropométrie permet la mise à disposition de larges quantités d'informations mesurées sur des humains. En pratique, elles permettent la calibration des os du squelette du signeur virtuel à une taille standard pour un humain. De nombreuses sources d'information sont disponibles pour la taille moyenne du buste, ainsi que pour les longueurs des membres. La suite de cette thèse se base sur les mesures publiées dans le corpus anthropométrique de la NASA<sup>10</sup> et sur des mesures propres faites pour les valeurs manquantes comme la longueur moyenne des doigts de la main (cf. section 3 du chapitre II). Les bases de données d'anthropométrie mesurent dans une certaine limite les *amplitudes de mouvement* ou ROM (*Ranges of motion*) des articulations du corps. Les amplitudes de mouvement sont représentées informatiquement par un couple de valeurs : le minimum et le maximum de l'angle. Nous nous basons sur ces mesures pour les plus simples telles que l'amplitude de rotation sur la colonne vertébrale ou sur le coude.

### 3.3. Le problème des limitations angulaires

Le manque de réalisme et de naturel de la génération automatique est issu comme nous l'avons dit précédemment d'une approche géométrique de la synthèse. Ajoutons à cela que cette synthèse n'est pas pilotée par un enregistrement réel à l'origine (comme le serait la MoCap ou la rotoscopie). Il est donc tout à fait possible au signeur virtuel de réaliser des mouvements qu'un humain ne pourrait pas faire. C'est-à-dire qu'en l'absence d'informations anatomiques, la hiérarchie articulée des os du squelette ne suit aucune règle. Le signeur virtuel peut donc se « casser une articulation » pour atteindre une posture demandée (cf. Figure 9). L'apport de modèles anatomiques permet d'estomper voir de supprimer ces effets

---

<sup>9</sup> <http://atilf.atilf.fr/tlf.htm>

<sup>10</sup> <http://msis.jsc.nasa.gov/sections/section03.htm>

en créant des *limites articulaires*. La nature de ces limitations varie selon la représentation utilisée et le niveau de granularité souhaité. Elles peuvent être statiques : un simple ROM qui n'évoluera pas dans le temps ; ou dynamiques : les valeurs des ROM de l'articulation sont interdépendantes et évoluent en fonction de la valeur des angles des autres degrés de liberté.

Les ROM suffisent à limiter le mouvement des articulations dont le comportement est le plus aisément modélisable telles que le coude ou les articulations de la colonne vertébrale. Cependant les complexes articulaires tels que ceux de l'épaule ou de la main requièrent une modélisation plus fine si l'on veut obtenir des résultats de génération plus réalistes et plus précis.

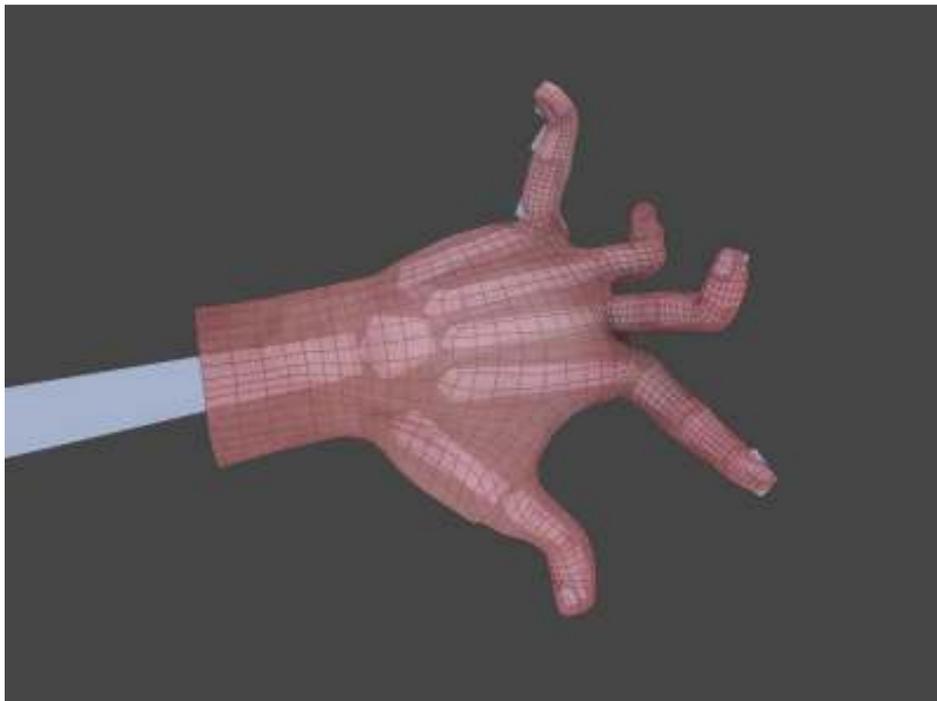


Figure 9 Posture impossible prise par la main du squelette.

### ***3.3.1. Limitations articulaires de la main***

La modélisation des mains bute généralement sur deux problèmes principaux : la modélisation de l'opposabilité du pouce et celle de l'interdépendance des doigts.

Le premier, lié à la complexité des articulations du complexe carpo-métacarpien<sup>11</sup> n'est pas modélisable en termes d'articulations simples à trois degrés de liberté. Les résultats d'une modélisation simple (Losson 2000) n'apportent que peu de satisfaction quant à la crédibilité des configurations générées. (Chang et al. 2008) et (Hollister et al. 1992) donnent des pistes de recherche sur la modélisation de l'axe de rotation permettant l'opposabilité du

<sup>11</sup> Complexe osseux permettant l'opposition du pouce

pouce. Cependant, ces méthodes développées pour la création de modèles anatomiques et parfois informatiques de la main mènent généralement à des modèles dont la complexité coûte particulièrement cher en temps de calcul. Leur utilisation dans le domaine de la simulation médicale ne semble néanmoins pas justifier une réelle pertinence en termes d'animation (comme indiqué en section 2.3). Une modélisation trop complexe amène les algorithmes de génération à résoudre des problèmes de complexité très élevée pour une plus-value moindre. (Van Nierop et al. 2007) présente un modèle de main naturelle dont les résultats semblent prometteurs pour l'animation. Malheureusement, le système, se bornant à des configurations simples de la main, ne permet pas la génération de configurations complexes telles que la lettre [T] (cf. Figure 10). En marge, (McDonald et al. 2001) présente une approche différente pour la définition interactive de configurations de la main à partir d'un modèle paramétrique du pouce. Cette approche modélise la main par une surcouche au dessus du niveau des os, contrôlant la configuration de la main grâce à un ensemble de paramètres tels que la « flexion de l'index ». Elle nécessite l'intervention d'un humain, mais ce modèle peut être envisagé pour une automatisation ultérieure permettant une utilisation au sein de méthodes de génération automatique.



Figure 10 Lettre [T] de l'alphabet dactylogologique.

L'interdépendance des doigts est un problème désormais connu et bien couvert par la littérature. Il est dit que les ROM de certains degrés de liberté des doigts covarient. Considérons par exemple la flexion de l'index : la flexion d'une phalange proximale entraîne nécessairement une limitation de l'extension des phalanges intermédiaires et distales. La très forte interdépendance des doigts rend la modélisation fine de la main difficile de par les nombreuses forces en action et la complexité des systèmes dynamiques à résoudre. Tout comme l'opposabilité du pouce, les modèles les plus complexes tels que présentés par (Tsang et al. 2005) proposent des solutions efficaces mais coûteuses. (Losson 2000) et (Lee et al. 1995) présentent quant à eux des systèmes d'équations simples permettant une limitation dynamique des ROM de la main.

### ***3.3.2. Limitations articulaires de l'épaule***

La modélisation de l'épaule pose les mêmes problèmes que ceux de la main. L'épaule est le siège de mouvements complexes qu'une simple combinaison de six rotations (trois pour la clavicule, trois pour l'épaule) ne suffit pas à modéliser. Du fait de sa géométrie et de son fonctionnement, les modèles biomécaniques de l'épaule se contentent généralement de décrire certains mécanismes spécifiques tels que l'articulation gléno-humérale (Wuelker et al. 1995), la réponse des articulations de l'épaule aux charges (Niemi et al. 1996) ou encore la force des articulations lors de mouvements sur le plan scapulaire (Makhsous et al. 1999).

Certains modèles cinématiques développés pour l'animation permettent la simulation simplifiée de l'épaule. (Klopkar et al. 2007) propose un modèle cinématique à six degrés de liberté simulant le comportement général de l'articulation scapulo-humérale afin d'évaluer la portée du bras humain. (Badler et al. 1999) dans son livre « *Simulating Humans* » consacre une section entière à la modélisation de l'épaule humaine et aux dépendances angulaires entre la clavicule et l'épaule. Il n'existe pas à notre connaissance de modèles d'animation permettant une simulation précise de tous les mouvements de l'épaule. Cette absence est globalement due à la très forte interdépendance des degrés de liberté de l'épaule. La modélisation du complexe de l'épaule par un système standard à six degrés de liberté nécessite une adaptation dynamique des ROM. Les limitations en jeu n'ont pas été formalisées à l'heure actuelle et ne sont donc pas intégrables à un modèle de l'épaule, contraignant la synthèse automatique de langue des signes à reposer sur des modèles à la complexité trop élevée ou trop simplifiés.

### ***3.3.3. Limitations articulaires du coude et de la colonne vertébrale***

Les limitations d'articulations « simples » (par comparaison avec les articulations de la main et de l'épaule) sont décrites exhaustivement dans la littérature. Bien qu'encore une fois, il y ait une très forte variabilité entre les individus, des valeurs moyennes ont été extrapolées de bases de données d'anthropométrie comme indiqué en section 3.2. Pour ces articulations plus triviales de simples ROM suffisent à présenter un comportement suffisamment précis pour générer des animations en langue des signes.

Toutes ces limitations laissent tout de même une marge de manœuvre aux algorithmes de génération leur permettant d'exploiter toutes les postures possibles dans ces intervalles. Il convient de créer un lien encore plus fort entre les algorithmes de génération et les

modèles anatomiques de façon à ce que l'ensemble des « postures possibles » se transforme en ensemble des « postures probables »<sup>12</sup>.

### 3.4. Solutions multiples

La résolution d'un jeu de contraintes fournies par le modèle linguistique peut mener à trois cas : aucune solution possible (par exemple lorsqu'on demande au signeur d'atteindre avec son bras un point hors de portée), une solution unique, ou une infinité de solutions. Ce dernier cas est très courant lors de l'utilisation de techniques de cinématique inverse (cf. Figure 11).

La prise d'une posture en particulier sur cet ensemble infini est *a priori* arbitraire et il n'existe pas dans la littérature de moyen précis de quantifier la qualité d'une solution par rapport à une autre. (McDonald et al. 2002) utilise cette propriété à son avantage pour la définition de contacts entre le doigt du signeur et un point du visage. Le doigt est placé sur la zone de contact et un animateur peut calibrer à sa guise la position du poignet. Cette méthode requiert cependant l'intervention d'un humain dans la génération, ce dont nous souhaitons nous affranchir.

---

<sup>12</sup> C'est-à-dire en un ensemble de postures crédibles du corps humain.

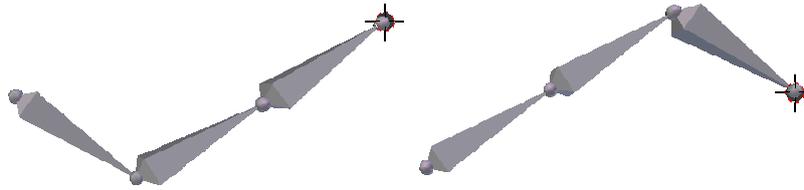


Figure 11 La cible (croix) peut être atteinte par une infinité de configurations. En voici deux exemples.

### 3.5. Gestion des collisions

Le principe d'animation en image clés implique la définition de postures statiques et une interpolation entre elles. Du fait de l'utilisation d'un système automatique à cent pour cent, il n'est pas rare que l'interpolation soit sujette à artéfacts, le plus courant d'entre eux étant la collision de parties du corps du signeur virtuel. Rappelons que le personnage virtuel, en l'absence de modèle anatomique, n'est pas sujet aux « lois de la physique ». Il dispose donc de toute la liberté pour traverser ses membres, ses mains ou n'importe quelle partie de son corps lors d'un mouvement.

Bien évidemment, ce genre de comportements est à proscrire si l'on considère une génération naturelle et réaliste. Cependant, la langue des signes impose parfois des contacts entre les différents articulateurs (cf. Figure 12). La détection et l'évitement de collisions entre les parties du corps du signeur doit donc être maîtrisée de manière à autoriser les contacts mais à empêcher les mouvements peu plausibles.

La résolution des collisions est un élément très important en génération automatique de langue des signes et à plus forte raison en animation. Cependant, il n'existe pas à l'heure actuelle de technique permettant de résoudre ce problème. La principale cause de cet échec vient de la séparation en deux entités du personnage virtuel. Le processus de génération automatique ne s'intéresse qu'au squelette, alors que les collisions ont lieu au niveau de la peau du personnage. La prise en compte de la détection de collision nécessite donc un calcul intégrant le squelette ET la peau, ce qui est particulièrement coûteux en temps de calcul. Cette problématique est spécifique à la génération d'animation et n'apparaît pas dans les études plus classiques sur l'animation.



Figure 12 Signe [PARDON] en LSF.

Tout comme les limitations articulatoires, la détection et la gestion des collisions peuvent se faire à plusieurs niveaux de granularité. Il est évident que certaines collisions sont plus simples à détecter que d'autres. Par exemple, la collision entre le bras du signeur et son torse peut être évitée simplement (du fait de la simplicité d'un volume tel que le torse), alors que la collision entre plusieurs doigts lors de la prise de configuration d'une main peut s'avérer bien plus épineuse. Il est admis que la gestion de collisions complète à partir du squelette du signeur s'avère trop coûteuse en temps si elle doit être calculée pour chaque image de l'animation. Les techniques de détection de collision sont donc limitées soit aux simples os du squelette ou à une représentation simplifiée de la peau par volumes primitifs tels que décrits par (Kavan et al. 2005). Bien que la détection de collision ne soit pas en soi une composante des modèles anatomiques, la représentation schématique de la peau du signeur par volumes l'est. Cet ensemble de volumes ne peut pas être calculé automatiquement à partir du squelette et nécessite donc un prétraitement utilisant la peau finale du signeur. La plus-value d'un tel système est la réduction de la peau du squelette à des volumes pour lesquels la détection de collision est simple et rapide.

En marge de la détection de collision, il faut aussi permettre au système la planification de l'évitement pour les mouvements les plus complexes. Cette problématique n'est pas liée à la modélisation anatomique du signeur et ne sera donc pas traitée dans la suite de l'exposé.

### 3.6. Dynamique du mouvement

Les limitations de la génération automatique ne s'arrêtent pas aux simples postures statiques. La dynamique du mouvement est primordiale pour la reconnaissance et la compréhension des signes. La construction de signes de manière géométrique contraint généralement la dynamique du mouvement créant des trajectoires trop parfaites. De plus,

en l'absence d'informations complémentaires, le système ne sait pas *a priori* comment gérer les accélérations du signeur. (Filhol 2008) dans son mémoire de thèse insiste sur l'existence de différents types de mouvements. La description linguistique *Zebedee* présentée en section 2.5 permet l'affectation d'une dynamique simple aux mouvements de transitions. Le mouvement est soit à vitesse constante, soit en accélération, soit en décélération. Cependant, la dynamique du mouvement est aussi une composante liée aux articulations et à leur fonctionnement. Les modèles fins décrits en section 3.3 présentent parfois des indications biomécaniques quand à la dynamique de mouvement. Néanmoins, ces modèles très fins posent le même problème que pour les limitations articulaires : ils sont trop coûteux en temps de calcul pour être utilisables par un système de génération automatique d'animation. On trouve une solution intéressante à ce problème dans l'analyse automatique de mouvements. (Lefebvre-Albaret 2010) présente dans son mémoire de thèse plusieurs analyses intéressantes de dynamique de mouvements élaborées par des méthodes de reconnaissance automatique. L'utilisation de telles courbes pour simuler l'accélération et la décélération du mouvement permettraient à la synthèse de signes de gagner en crédibilité et en réalisme. Ces travaux, très récents, n'ont pas été incorporés dans l'étude présentée dans cette thèse mais constituent des perspectives intéressantes.

### 3.7. Expressions faciales

La compréhension des signes ne nécessite pas seulement la génération d'une animation de bonne qualité des membres supérieurs. Le visage, le regard et les expressions faciales représentent une composante primordiale de la langue des signes. Ces expressions jouent de multiples rôles dans les énoncés en langue des signes : désambiguïsation de signes, modes de discours (interrogation, négation, conditionnel, etc.). La prise en compte de ces expressions faciales est nécessaire pour la génération d'animation de qualité. Une étude à notamment montré l'importance linguistique des sourcils et des clignement d'yeux en langue des signes (Chételat-Pelé 2010).

Cependant, tout comme la problématique liée aux collisions, la génération d'expressions faciales ne se fait pas au niveau du squelette mais de la peau du signeur. La littérature concernant la modification du maillage du visage présente généralement le problème selon deux approches : les systèmes à base de codage d'expressions faciales, tels que FACS ou MPEG4 (Bailly et al., 2009) et les systèmes à base de morphing d'expressions prédéfinies. La nature des expressions faciales est loin de l'approche géométrique que nous utilisons *a priori*

pour les descriptions manuelles. Il n'est pas possible de définir une expression faciale en termes géométriques sans faire appel à des artefacts arbitraires. Il est donc nécessaire de prédéfinir des expressions faciales et de les combiner pour créer une expression finale comme le font (Braffort et al. 2011) ou (Glauert et al. 2011).

Anatomiquement, les expressions faciales sont contrôlées par la multitude de muscles composant le visage. L'interdépendance de ces muscles et leur action sur la peau rend la simulation précise d'expressions faciales quasi-impossible et il n'existe pas, à notre connaissance, de modèle précis du visage. Pour la génération automatique, il est possible de s'inspirer des modèles utilisés en animation de manière générale, notamment dans le domaine des agents conversationnels animés expressifs tels que (Courgeon, Martin, et al. 2008). Il est aussi possible de considérer chaque composante de l'expression faciale (sourcils, clignement des paupières, forme de la bouche, placement de la langue) comme autant de pistes d'expression parallèles (cf. Figure 13). Il est donc possible d'animer chacune de ces composantes séparément. (Schnepp et al. 2010) utilisent cette technologie pour composer les expressions faciales de leur signeur, en la combinant à la labialisation<sup>13</sup> qui accompagne parfois les signe.

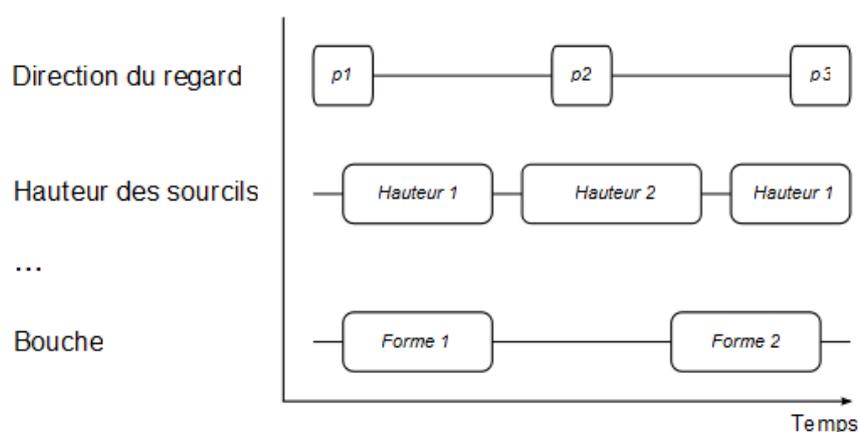


Figure 13 Exemple de pistes d'expression parallèles, animant les composantes non-manuelles du signeur.

Par ailleurs, la formation d'expressions faciales entraîne très souvent la formation de rides sur le visage. Les plus évidentes sont celles sur le front liées aux haussements des sourcils. (Courgeon et al. 2008) modélisent ces rides en créant de réelles déformations sur la peau du personnage virtuel, alors que (Wolfe et al. 2011) utilisent des textures en lieu et place des rides. Cette dernière solution nous semble plus efficace pour le calcul rapide d'un

<sup>13</sup> « Mouvement d'arrondissement des lèvres qui intervient comme articulation secondaire dans la réalisation des phonèmes dits labialisés ». Définition du *Trésor de la Langue Française* : <http://atilf.atilf.fr/tlf.htm>

résultat mais la première nous semble plus réaliste. Dans ce dernier cas, le modèle anatomique permet la définition précise des points amovibles du visage.

### **3.8. Résumé et élaboration du cahier des charges**

Nous avons vu dans cette section un ensemble de problèmes découlant de la génération automatique d'animations. Ces problèmes sont liés soit à la nature intrinsèque de la génération automatique, soit aux modèles linguistiques utilisés comme entrées au système. Nous nous proposons dans cette thèse d'élaborer un modèle anatomique et une méthode de génération permettant de répondre aux problématiques suivantes :

- Définition aisée de mesures sur le corps, suivant les résultats d'une base d'anthropométrie (cf. Chapitre II, section 3)
- Définition de limites articulaires statiques, dynamiques et permettant la sélection d'une posture parmi une infinité de solutions (cf. Chapitre II, sections 4 et 5)
- Amélioration de la structure de représentation afin de prendre en compte plus de détails (cf. Chapitre II, sections 2 et 5)



## Chapitre II : Proposition de modèle pour la représentation du corps humain

Comme nous l'avons vu dans le chapitre précédent, la synthèse de la langue des signes soulève des problèmes spécifiques. L'utilisation de processus automatiques de génération permet une flexibilité accrue de la synthèse mais nécessite un guidage au moyen de modèles linguistiques et anatomiques. Le modèle anatomique détaillé ci-après cherche à répondre aux problématiques énoncées en section 3.8 du chapitre précédent et constitue les contributions que propose cette thèse au domaine de la génération automatique d'animations en langue des signes.

Dans une première section les problématiques et les solutions proposées sont récapitulées. Les sections suivantes s'attachent chacune à la résolution d'un problème : la section 2 présente une nouvelle méthode de représentation du squelette; la section 3 détaille une étude anthropométrique sur l'élaboration d'une main humaine moyenne; en section 4, nous présentons un ensemble d'outils mathématiques pour contraindre le squelette à des limites physiologiques empêchant la définition de postures impossibles; la section 5 présente une nouvelle modélisation du pouce permettant une opposition compatible avec la plupart des systèmes de génération automatique; enfin, en section 6, nous présentons un modèle permettant la simulation de confort de posture, permettant la discrimination entre plusieurs solutions valables pour un même problème.

## 1. Décomposition des problématiques et propositions préliminaires de solutions

Les quatre problématiques énoncées en 3.8 du chapitre I ne sont pas toutes du même ordre de difficulté. Ajoutons que chacun de ces points peut dépendre du niveau de détail du modèle linguistique utilisé. Il convient donc d'analyser en détail chacun des problèmes pour les redécouper en tâches aisées à résoudre.

### 1.1. Mesures anthropométriques et représentation du squelette

Dans l'optique de créer un signeur virtuel crédible, nous choisissons de représenter le signeur virtuel comme un personnage à l'aspect moyen, sans caractéristique physique particulière. Le processus de conception du signeur se passe en deux temps : d'abord le squelette est élaboré ; ensuite le graphiste conçoit une peau pouvant être contrôlée par ce squelette. Bien que les questions d'ordre esthétique (dont fait partie celle des proportions du corps humain) tiennent plus du graphisme que de la recherche, il est nécessaire d'établir un modèle permettant la modification aisée de la structure et de la forme du squelette du signeur afin de l'adapter au personnage qui sera créé par le graphiste.

Par ailleurs notons que les systèmes de contrôle de l'animation faisant intervenir le processus de cinématique inverse mentionné en section 2.4 du chapitre I vont permettre de placer des éléments du corps dans l'espace. Connaître uniquement la structure du squelette du signeur ne permet pas un contrôle efficace. Pour illustrer cet aspect, prenons l'exemple d'un signe comme [MALADE] (cf. Figure 14).



Figure 14 Signe [MALADE] en LSF

Ce signe bi-manuel nécessite deux contacts entre des zones du corps :

- Le premier concerne la configuration de la main droite : l'index et le pouce se touchent.
- Le second implique le contact entre les deux mains : la main droite va placer la configuration citée précédemment sur le dos de la main gauche.

Ces deux contacts illustrent deux problèmes principaux : D'abord, le contact entre les deux doigts se fait au niveau de la peau et non du squelette : ce ne sont pas les os qui se touchent mais bien les deux surfaces. Deuxièmement, le contact sur la main gauche se fait à un endroit où aucun os n'est repéré *a priori* : il ne s'agit ni du point de naissance ou de terminaison d'une articulation. Or, nous avons défini précédemment que le squelette stockait la hiérarchie des articulations et rien de plus. Il n'est donc *a priori* pas possible de construire la posture nécessaire au signe [MALADE].

Nous proposons de régler ce problème en ajoutant une liste de *sites corporels* comme présentés dans (Filhol 2008) et qui correspondent à des sites utilisés pour la LSF<sup>14</sup>. Ces sites représentent des points sur la peau, repérés grâce aux articulations du squelette. Ils permettent ainsi la définition aisée de zones permettant d'entrer en contact. De ce fait, un faible nombre d'os du squelette peut décrire un grand nombre de sites corporels permettant le placement et le contact efficace de différents éléments du corps. La définition des mesures anthropométriques du squelette ainsi que sa nomenclature précise sont détaillées en sections 2 et 3.

<sup>14</sup> Cette liste à été constituée après un étude ayant porté sur plus de 2000 signes.

## 1.2. Articulations, limites et naturel des postures

Cette section se penche plus en avant sur les problèmes soulevés dans les sections 3.3 et 3.4 du premier chapitre : les problèmes de limitations angulaires et la question des solutions multiples. Dans un premier temps (section 1.2.1), il est nécessaire de traiter le problème des dépendances articulaires. Dans un second temps (section 1.2.2), nous détaillons la question de la modélisation de certaines articulations afin d'éliminer au maximum les postures qu'un humain ne pourrait pas réaliser du fait de contraintes anatomiques. Enfin, en section 1.2.3, nous nous intéresserons au problème de l'infinité de solutions données par les méthodes de cinématique inverse et proposerons une façon de choisir une posture parmi une infinité.

### 1.2.1. Dépendances articulaires

Nous avons vu précédemment l'existence de limitations articulaires. Nous nommons dépendances articulaires une catégorie d'objets mathématiques permettant de définir le comportement d'un degré de liberté sur l'une des articulations. La variété des articulations du corps humain, ainsi que leurs différents degrés de complexité nous pousse à considérer les dépendances articulaires comme une famille d'outils permettant de régir la rotation des articulations. En l'état, les limitations des angles à de simples amplitudes de mouvement ne contraignent le mouvement des articulations qu'à de simples sections de cercle pour un degré de liberté, et de sections de sphères pour deux ou plus degrés de liberté. Ces sections de sphère ne suffisent pas à décrire la complexité de certaines articulations, notamment lors de la gestion de l'interdépendance des doigts. Il est donc nécessaire d'ajouter, par le biais du modèle anatomique, les informations permettant une simulation plus précise de ces mouvements complexes.

Nous proposons l'intégration de trois types de dépendances articulaires :

- Les *dépendances simples* expriment la valeur d'un degré de liberté en fonction d'un simple intervalle nommé *amplitude de mouvement*. La valeur de l'angle du degré de liberté est donc limitée à cet ensemble.
- Les *ROM<sup>15</sup> dynamiques* expriment une variation des amplitudes de mouvement en fonction de la valeur d'un autre degré de liberté. C'est le cas de l'abduction/adduction des phalanges. Ces amplitudes varient en fonction de la

---

<sup>15</sup> Ranges Of Motion (Amplitudes de mouvement).

flexion des doigts. Lorsqu'ils sont étendus, il est possible d'écarter les doigts de la main, alors que c'est impossible lorsqu'ils sont repliés.

- Les *distributions* sont des relations impliquant le mouvement de plusieurs degrés de liberté pour créer un mouvement général. Par exemple, le fait de lever le bras entraîne une rotation générale du complexe de l'épaule. Ce complexe distribue la rotation selon une relation précise entre la clavicule et l'épaule.

Les deux premiers types de dépendances permettent de répondre au problème des limitations simplifiées du mouvement. Elles créent une limitation forte empêchant le squelette de prendre des postures impossibles. Le dernier type, quant à lui, permet une distribution réaliste des angles lors de la formation de complexes articulaires dont on ne peut se permettre la simulation précise. Les dépendances de type distribution nécessitent la mise en place d'une hiérarchie particulière du squelette permettant de spécifier quels groupes d'articulations sont liés. Cette hiérarchie est décrite en détail en section 2.1. Les dépendances simples, dynamiques et distributives sont présentées respectivement en sections 4.1, 4.2, 4.3.

Il existe cependant une exception pour laquelle aucune de ces trois types de dépendances ne convient pour modéliser correctement le comportement : le pouce.

### ***1.2.2. Modélisation du complexe carpo-metacarpien***

Le pouce est usuellement modélisé comme les autres doigts par une chaîne cinématique simple. Cette première approche se révèle insuffisante. Nous nous sommes inspirés des travaux de (Hollister et al. 1992) cités précédemment. Ces travaux proposent vers l'élaboration d'un modèle de pouce permettant la liberté nécessaire à un système de génération et suffisamment de contraintes pour garantir l'intégrité des configurations manuelles générées.

Cette modélisation du pouce, comme nous le verrons dans la section 5 inclut deux nouveaux degrés de liberté. Généralement modélisé avec trois degrés de liberté, ce nouveau complexe permet ainsi de supprimer un degré de liberté et de contraindre suffisamment le mouvement de façon à obtenir des configurations de la main crédibles.

Tout comme pour les dépendances articulaires, la limitation de certains degrés de liberté n'empêche pas la multiplicité des solutions à un même jeu de contraintes. Il est donc

impératif d'établir un moyen de choisir entre les différentes solutions possibles afin de maximiser le réalisme et le naturel des animations générées.

### 1.2.3. Multiplicité des solutions

La multiplicité des solutions est un problème récurrent et malheureusement très mal couvert de la cinématique inverse. Du fait du nombre élevé de degrés de liberté du squelette et du nombre (généralement restreint) de contraintes qu'on essaie de lui appliquer, les problèmes posés sont généralement sous-spécifiés ce qui mène à une infinité de solutions (cf. Figure 15).

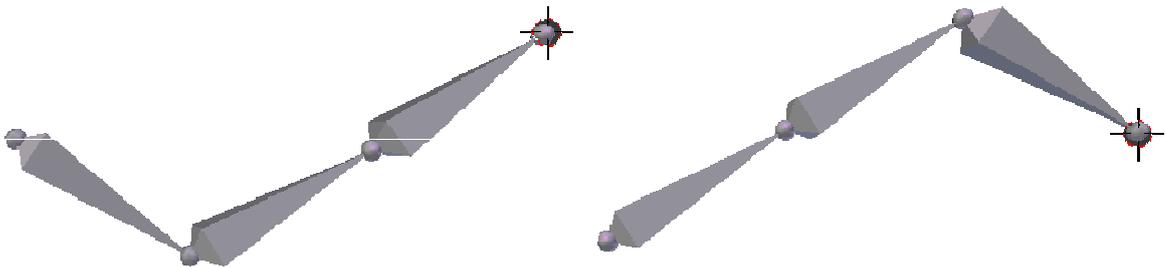


Figure 15 Deux solutions possibles pour un même problème de cinématique inverse.

La section 6 présente le problème de multiplicité des solutions en introduisant au sein du modèle une définition de *confort de posture*. Cette notion de confort nous permet de quantifier à quel point une solution particulière est naturelle ou non. Elle permet donc de discriminer entre plusieurs solutions et de faire un choix. Tout comme pour la formation du squelette, l'utilisation de ce modèle nécessite une intégration particulière au système de génération. Cette intégration, présentée dans le chapitre III, consiste en une approche différente pour la cinématique inverse.

Commençons par détailler une nouvelle représentation du squelette permettant un contrôle plus efficace.

## 2. Représentation du squelette

Le squelette est un élément primordial de la structure du signeur virtuel. Il permet le contrôle de sa posture et plus généralement de ses mouvements. En amont de la recherche de nouvelles techniques d'animation, il nous semble prioritaire de considérer les approches standards pour la représentation du squelette. Ces approches doivent être questionnées et corrigées dans le cas où elles seraient insuffisantes. Cette section décrit dans une première partie une nouvelle approche de représentation du squelette dite en *blocs fonctionnels*. Dans un second temps nous définissons la notion de *sites corporels* en tant qu'éléments additionnels à cette représentation pour le placement précis d'éléments du corps lors de la résolution de postures. Enfin, en troisième partie, nous présentons le squelette précis que nous utilisons dans la suite de cette thèse.

### 2.1. Blocs fonctionnels

#### 2.1.1. Approche et problématique des blocs fonctionnels

La représentation usuelle d'un squelette articulé se fait au moyen d'arbres. Les normes tentant de définir les squelettes telles que H-Anim<sup>16</sup> ou encore le bipède de 3DSmax utilisent généralement le bassin du personnage virtuel comme racine de l'arbre et décrivent chaque articulation comme un nœud de cet arbre (cf. Figure 16). L'orientation et la position d'un os dépendent donc de l'état de ses parents, notion intuitive lorsqu'il s'agit par exemple de parler de la colonne vertébrale. L'ensemble des vertèbres crée une chaîne articulaire dont l'état d'un os dépend de tous ceux qui le précèdent.

---

<sup>16</sup> Norme ISO/IEC FCD 19774:200x disponible sur [http://h-anim.org/Specifications/H-Anim200x/ISO IEC FCD 19774/](http://h-anim.org/Specifications/H-Anim200x/ISO_IEC_FCD_19774/)

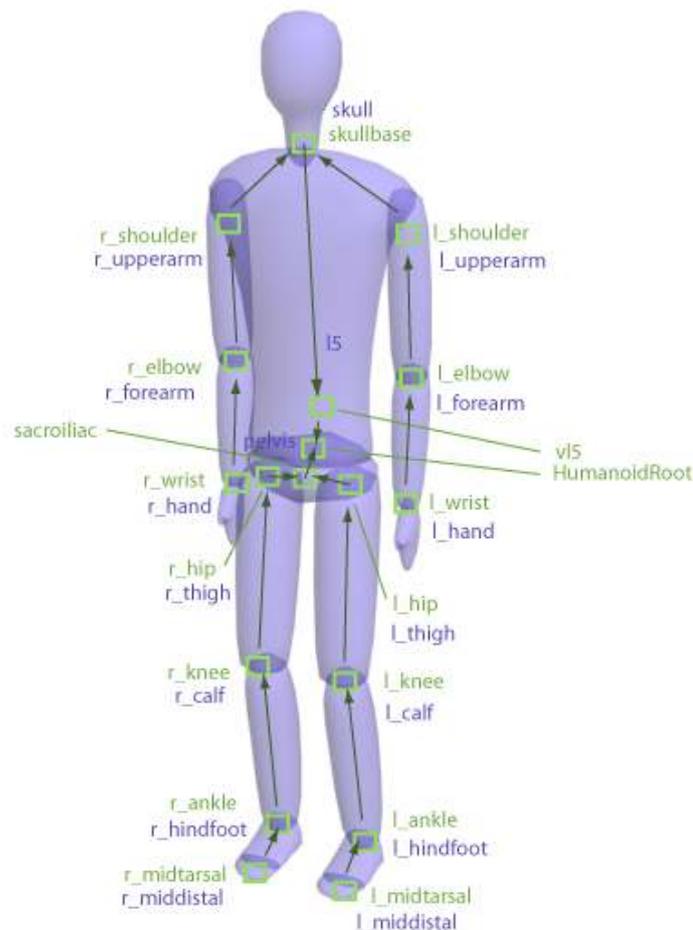


Figure 16 Exemple de représentation du squelette en arbre. Les flèches indiquent la parenté (source = fils, destination = père) (Image © L.Boissieux, INRIA)

Cette forme de représentation a pour avantage de fournir un moyen très simple de décrire la hiérarchie du corps humain. Elle est généralement suffisante pour une modélisation basique pilotée par les données telle que l'animation par capture de mouvement. Cependant les avancées dans la génération automatique ont rendu cette représentation obsolète, notamment lors de la modélisation de phénomènes d'interdépendance entre des os.

Considérons le cas de la main. Il semble intuitif de modéliser *a priori* cet ensemble fonctionnel par une hiérarchie en arbre telle que celle décrite sur la Figure 17. Cependant, les os de la main sont tous très interdépendants du fait de la configuration complexe des os et des tendons du carpe. Il est quasiment impossible de ne faire bouger qu'un seul doigt de sa main sans affecter les autres. Ainsi les os de la main dépendent non seulement des os les précédant dans la chaîne articulaire, mais aussi des autres os de la main (plus précisément les os voisins). La représentation de la main se fait donc via un graphe tel que Lee l'a représenté (Lee et al. 1995), où chaque nœud de la main est connecté de manière

transversale à ses voisins. Cependant cette représentation nous semble confuse au niveau structurel. Nous choisissons de garder à l'esprit cette idée d'interdépendance tout en se penchant sur une approche plus précise et plus cohérente.

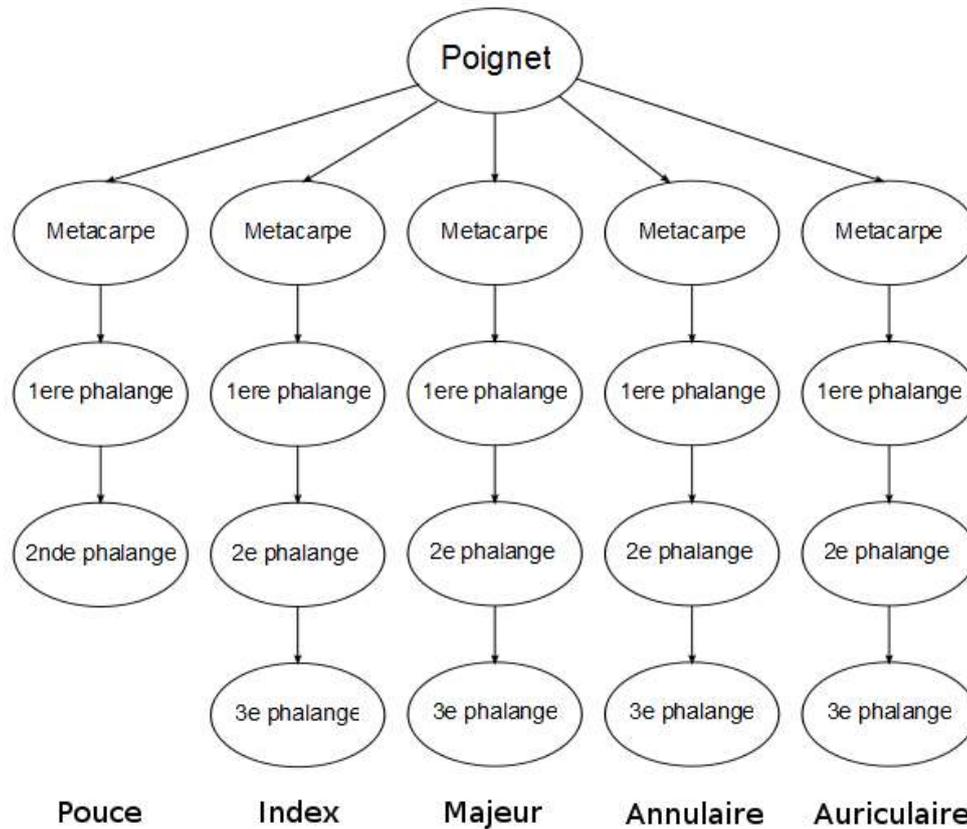


Figure 17 Représentation standard en arbre des os de la main.

Le raisonnement au niveau des os-mêmes peut être abstrait à un niveau supérieur. La modélisation précise du corps os par os sous entend une compréhension intrinsèque de leur fonctionnement. Cependant, comme nous l'avons expliqué précédemment, la modélisation du squelette en animation ne tient absolument pas compte des muscles, des tendons et de tout l'attirail moteur dont le corps humain dispose. Modéliser les os du carpe n'a donc aucun sens sans réelle compréhension de leur fonctionnement. Il est par ailleurs difficile de placer toutes les articulations du corps humain au même niveau. Le modèle linguistique fournissant en entrée les contraintes de posture se base sur un placement et une orientation consciente des os du corps. Illustrons le propos : il est tout à fait possible pour un humain de « plier son coude » mais il paraît difficilement réalisable de « plier son os pisciforme » consciemment. Passé un certain niveau de détail, les os ne réagissent plus de manière indépendante mais en complexes dont le fonctionnement, au niveau du système de génération, se rapproche d'une boîte noire. L'exemple du carpe est flagrant : on aura tendance à vouloir plier les « doigts »

qui sont en soi des groupements d'os plus que de bouger des os indépendamment. Le fonctionnement intrinsèque de ces os ne devrait pas être géré par le système de génération, mais par des relations au niveau-même du squelette.

### 2.1.2. Définition des blocs fonctionnels

Nous proposons la modélisation de ce phénomène au moyen d'une hiérarchie de *blocs fonctionnels* imbriqués. Les blocs représentent soit des os<sup>17</sup> (e.g. le premier métacarpe), ou des groupes d'os (e.g. la colonne vertébrale, le bras droit, etc.). Dans le cas où le bloc représente un groupement d'os, il contient autant de sous-blocs que d'os sous-jacents. Cette relation crée un ordre de parenté similaire à celui utilisé dans le vocabulaire des arbres en informatique. Le bloc contenant est appelé *père* alors que les blocs contenus sont appelés *fils*. En conséquence, il existe des blocs *terminaux* sans enfants. Ces blocs représentent les os effectifs du squelette : ceux que l'on trouve dans la représentation traditionnelle en graphe/arbre du squelette.

Le squelette est donc découpé selon différents niveaux d'abstraction, des blocs les plus génériques aux blocs les plus spécifiques. On organise le squelette en quatre niveaux (minimum) d'abstraction. Le Tableau 2 détaille les différents niveaux d'abstraction et donne quelques exemples de blocs appartenant à ces niveaux. Notons la distinction entre les niveaux « unités fonctionnelles conscientes » et « unités fonctionnelles inconscientes » qui répondent précisément à la problématique formulée en 2.1.1.

Niveau d'abstraction	Appellation	Exemples de blocs
0	Squelette	
1	Grands groupes fonctionnels	<ul style="list-style-type: none"> <li>• Bras gauche</li> <li>• Main droite</li> <li>• Colonne vertébrale</li> </ul>
2	Unités fonctionnelles conscientes	<ul style="list-style-type: none"> <li>• Complexe de l'épaule droite</li> <li>• Coude gauche</li> <li>• Index de la main gauche</li> </ul>
>2	Unités fonctionnelles inconscientes	<ul style="list-style-type: none"> <li>• Clavicule gauche</li> <li>• 2<sup>e</sup> phalange de l'index gauche</li> <li>• Os scaphoïde droit</li> </ul>

Tableau 2 Les différents niveaux d'abstraction de blocs fonctionnels

<sup>17</sup> Lorsque l'on se réfère à une articulation, on parle aussi implicitement de l'os auquel elle donne naissance, et vice versa.

La notion de parenté ne retranscrit pas la hiérarchisation des os comme dans le modèle traditionnel. Elle n'est en aucun cas synonyme de lien entre deux articulations. Ce lien est représenté par des connexions au sein d'un même bloc parent. Ces connexions sont orientées et définissent une nouvelle relation d'ordre : *successeur/antécédent*. Par soucis de clarté, on qualifie les relations de parenté de relations *verticales*, et celles de succession de relations *horizontales*. Les relations horizontales restent confinées au bloc parent dans lesquelles elles se trouvent. Chaque contenu de bloc est sujet à des contraintes sur leur placement venant du père dans lequel ils se trouvent. Les relations entre les différents types de blocs sont résumées sur la Figure 18.

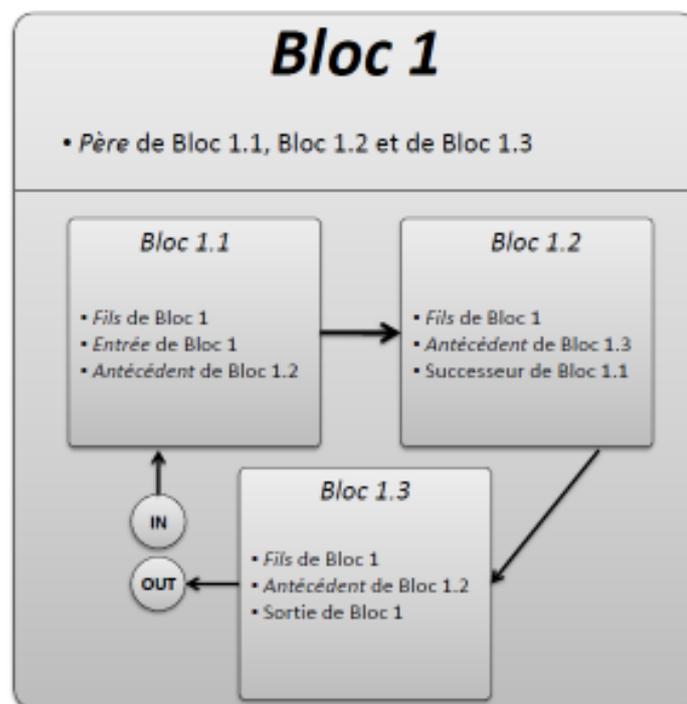
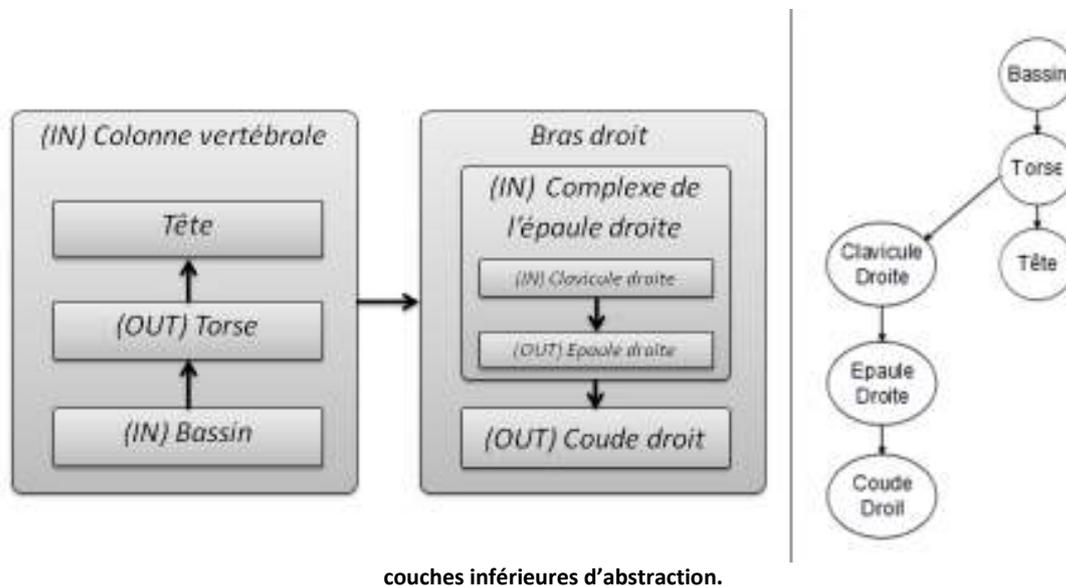


Figure 18 Représentation en blocs hiérarchiques et exemples courants illustrant la nomenclature du système.

Bien que la connexion de blocs non terminaux puisse paraître étrange, il faut reconnaître que la liaison entre le torse et l'épaule droite se fait à deux niveaux d'abstraction. Il y a d'abord la connexion physique entre les os (dans la représentation virtuelle, on dira que le thorax est connecté à la clavicule) et la connexion au niveau des grands groupes fonctionnels (la colonne vertébrale est reliée au bras). Le modèle en blocs fonctionnels permet ces deux représentations mais favorise la représentation abstraite. Concrètement, le lien donné en exemple se fait par une connexion entre le bloc « colonne vertébrale » et le bloc « bras droit ». Le lien n'explique cependant pas la liaison physique, qui est faite grâce à deux types de marqueurs : chaque bloc non-terminal dispose nécessairement d'un sous-bloc

qualifié d'*entrée* et d'un sous-bloc qualifié de *sortie* (ou les deux). Le lien physique est ensuite retrouvé par récursivité : la connexion s'effectue entre le bloc terminal en sortie du premier bloc et le bloc terminal en entrée du second bloc (cf. Figure 19).

Figure 19 Gauche : Représentation d'une partie du squelette ; Droite : Equivalent en représentation standard des



La représentation du squelette selon ces blocs fonctionnels permet un contrôle plus précis des mouvements, notamment grâce à l'inclusion des niveaux d'abstraction, autorisant la représentation de l'appartenance de plusieurs blocs à un même groupe fonctionnel. Cette relation d'appartenance est exploitée dans la section 4 consacrée aux dépendances articulaires.

### 2.1.3. Algorithmique des blocs fonctionnels

L'utilisation des blocs fonctionnels permet une représentation du squelette du personnage virtuel plus précise mais nécessite cependant la mise en œuvre d'algorithmes particuliers pour son parcours. Avant de détailler l'algorithme de parcours, nous proposons une structure générique des blocs (i.e. l'élément Bloc en pseudo-code) qui nous servira de base pour la suite :

Nom de l'attribut	Type de l'attribut	Commentaire
bloc_entrée	Référence de Bloc	Référence directe au sous-bloc qui correspond à l'entrée. NUL si le bloc est terminal.
bloc_sortie	Référence de Bloc	Référence directe au sous-bloc qui correspond à la sortie. NUL si le bloc est terminal.
nb_connexions	Entier	Le nombre de connexions sortantes de ce bloc.
connexions	Vecteur de références de Bloc	Un vecteur regroupant des références aux blocs connectés au bloc courant.
nb_enfants	Entier	Le nombre de sous-blocs contenus dans ce bloc.
enfants	Vecteur de références de Bloc	Un vecteur regroupant des références aux sous-blocs contenus dans ce bloc.
parent	Référence de Bloc	Référence direct au père du bloc courant. Si le bloc courant est la racine, alors parent est NUL.
[... contenu ...]	...	Le contenu sémantique du bloc : les données propres du squelette telles que la longueur des os, les repères de transformation, etc.

Tableau 3 : Structure Bloc, détail des attributs utilisés pour les algorithmes suivants

Le parcours d'une hiérarchie en blocs sous-entend souvent un traitement des blocs terminaux distinct des autres. L'Algorithme 1 présente le parcours des blocs en profondeur, traitant en priorité les nœuds terminaux et leurs connexions puis remontant aux nœuds englobants à la fin.

```

Procédure Parcours(courant : Bloc(E))
Variables
  id_connexion : Entier
Debut

  /* Si le bloc courant n'est pas terminal on descend d'un niveau */
  Si (courant.bloc_entrée != NUL)
  Alors
    Parcours(courant.bloc_entrée)
  Sinon
    [... Actions du parcours sur les blocs terminaux ...]
  Fin

  [... Actions du parcours sur tous les blocs ...]

  /* Pour chaque bloc connecté, on effectue le traitement */
  Pour id_connexion de 0 à courant.nb_connexions-1 faire
    Parcours(courant.connexion[id_connexion])
  Fin
Fin

```

Algorithme 1 : Parcours de la hiérarchie en bloc en profondeur.

Bien entendu, il est possible d'insérer des instructions selon l'ordre de parcours que nous souhaitons (infixe/postfixe/préfixe). Cet algorithme de parcours peut être par exemple utilisé pour convertir automatiquement la hiérarchie en blocs en une représentation traditionnelle en arbre du squelette. Ce procédé demande la création d'un nœud pour

chaque bloc terminal et le lien du nouveau nœud créé à l'ascendant du bloc courant comme le montre l'Algorithme 2.

```

Procédure Construire_Arbre(courant : Bloc(E), resultat : Arbre(E/S),
                             lookup : Tableau associatif<Bloc, Noeud>(E/S))
Variable
    nouveau_noeud : Noeud
Debut

    /* Si le bloc courant est non terminal, on descend d'un niveau */
    Si (courant.bloc_entree != NUL)
    Alors Construire_Arbre(courant.bloc_entree, resultat, lookup)
    Sinon
        nouveau_noeud <- resultat.ajouter_noeud(courant)

        /* Construction des connexions du nœud à son prédécesseur */
        Si (courant.antecedent = NUL)
        Alors
            resultat.definir_racine(nouveau_noeud)
        Sinon
            resultat.ajouter_arete(nouveau_noeud, lookup[courant.antecedent])
        Fin

        lookup[courant] = nouveau_noeud
    Fin

    /* Application du même traitement pour tous les successeurs */
    Pour id_connexion de 0 à courant.nb_connexions-1 faire
        Construire_Arbre(courant.connexion[id_connexion], resultat, lookup)
Fin
    
```

Algorithme 2 : Algorithme de transformation d'une hiérarchie en blocs en arbre de représentation classique.

### 2.1.4. Contenu des blocs fonctionnels

Jusqu'à présent, nous nous sommes attachés à décrire l'organisation hiérarchique des blocs fonctionnels. En dehors de cet aspect purement organisationnel, chaque bloc terminal contient des informations qui lui sont propres, représentées par le terme « contenu » dans le Tableau 3. Ces informations sont communes aux représentations classiques des os du squelette virtuel. La réelle différence de contenu se situe au niveau des dépendances articulaires présentées en section 4 et des distributions de confort abordées en section 6.

En marge de ce contenu spécifique aux blocs, on définit la notion de sites corporels que nous allons maintenant détailler.

## 2.2. Sites corporels

Le placement et l'orientation des os grâce aux méthodes de cinématique ne permettent pas *a priori* la précision nécessaire pour la génération de langue des signes. Par exemple, beaucoup de signes nécessitent un contact entre diverses parties du corps comme mentionné en section 1.1. Ce contact, expressément indiqué dans la représentation linguistique, ne peut se faire directement au niveau des os. Lorsque deux doigts se touchent comme dans le signe [OUI] (cf. Figure 20) le contact a lieu au niveau de la peau et non des

os. De même, le signe [DEMI] demande le contact entre les index des deux mains à un endroit non représenté par l'origine ou la terminaison d'un os.

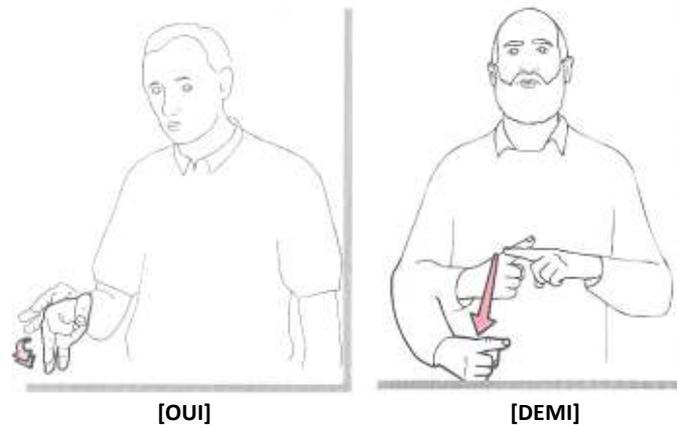


Figure 20 Deux exemples de contacts. Images extraites du dictionnaire IVT.

Les sites corporels (Filhol 2008) sont des points remarquables du corps souvent utilisés pour le placement d'éléments lors de la résolution de postures. Les sites sont donc des localisations virtuelles et ponctuelles sur la peau du signeur virtuel. On les représente en l'état comme une simple translation par rapport à l'origine d'un bloc terminal. Idéalement, les sites corporels devraient être liés précisément au maillage de la peau du signeur. Cependant le besoin d'indépendance entre le squelette et la peau, ainsi que la nécessité d'avoir des modèles rapidement calculables nous pousse à conserver une représentation simple.

La Figure 21 donne un exemple des différents sites corporels identifiés sur une main du squelette. On voit par exemple que chaque doigt est composé de trois phalanges (donc trois blocs terminaux) et de multiples sites présentés sur chaque face, au niveau des articulations, sur le bout des doigts, etc. Ces sites permettent une multiplication des possibilités de placement sans multiplication des os/blocs.

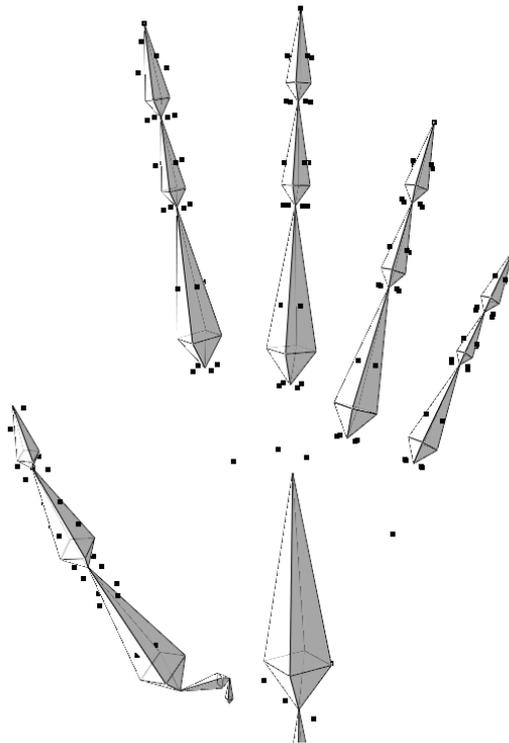


Figure 21 Les différents sites corporels identifiés sur la main du squelette

### 2.3. Définition et nomenclature du squelette utilisé

La définition du squelette dépend bien évidemment de l'emploi que l'on compte en faire. Le squelette que nous utilisons dans la suite de cette thèse est dédié à la synthèse de la langue des signes française. Ce choix n'influe pas sur la longueur des membres et des os, mais sur leur nombre (par exemple sur la précision des éléments de la main) et sur le nombre et la précision des sites corporels. La définition complète et exhaustive du squelette est donnée en annexe 2. Cette section s'attache à détailler les blocs fonctionnels ainsi que les sites utilisés pour la suite de cette étude.

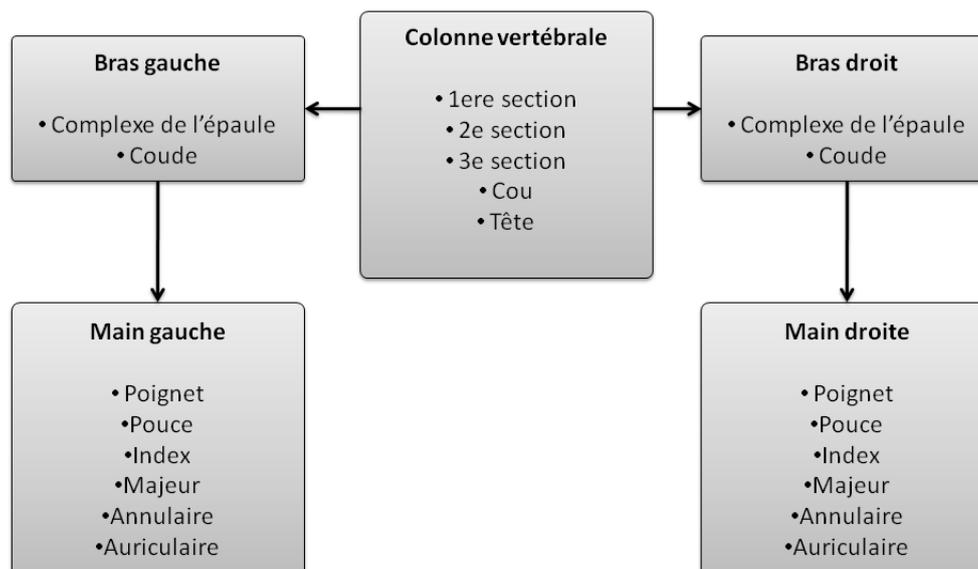


Figure 22 Grands groupes fonctionnels de la représentation en blocs du squelette.

La racine logique de la hiérarchie est le niveau d'abstraction 0 : le squelette. La Figure 22 représente les grands groupes fonctionnels du squelette. Ils sont au nombre de cinq :

- Un groupe fonctionnel pour la *colonne vertébrale* incluant la tête. L'articulation du dos est faite au moyen de trois os simulant la double courbure de la colonne vertébrale. La tête est gérée au moyen de 2 articulations : la base du cou et la base du crâne.
- Deux groupes fonctionnels pour les bras. Chacun dispose d'une articulation inconsciente (le groupe scapulo-huméral appelé ici « complexe de l'épaule » contenant deux articulations conscientes : la clavicule et l'épaule), ainsi que le coude. Les bras sont reliés à la colonne vertébrale grâce à l'articulation entre le « thorax » et la clavicule. A noter que le modèle ne cherche pas à simuler précisément l'interaction entre l'ulna et le radius. L'avant-bras est modélisé par un seul et unique os.
- Enfin, deux groupes fonctionnels pour les mains.

Les mains étant les sections du squelette les plus complexes, elles sont détaillées plus précisément sur la Figure 23. Chaque main dispose d'une articulation au niveau du poignet permettant la jonction avec l'avant-bras. Les doigts sont tous modélisés par un groupe de trois os : phalange proximale, intermédiaire et distale. Les métacarpes des quatre doigts

sont considérés comme un seul et même os pour simplifier le nombre de degrés de liberté en action dans la main.

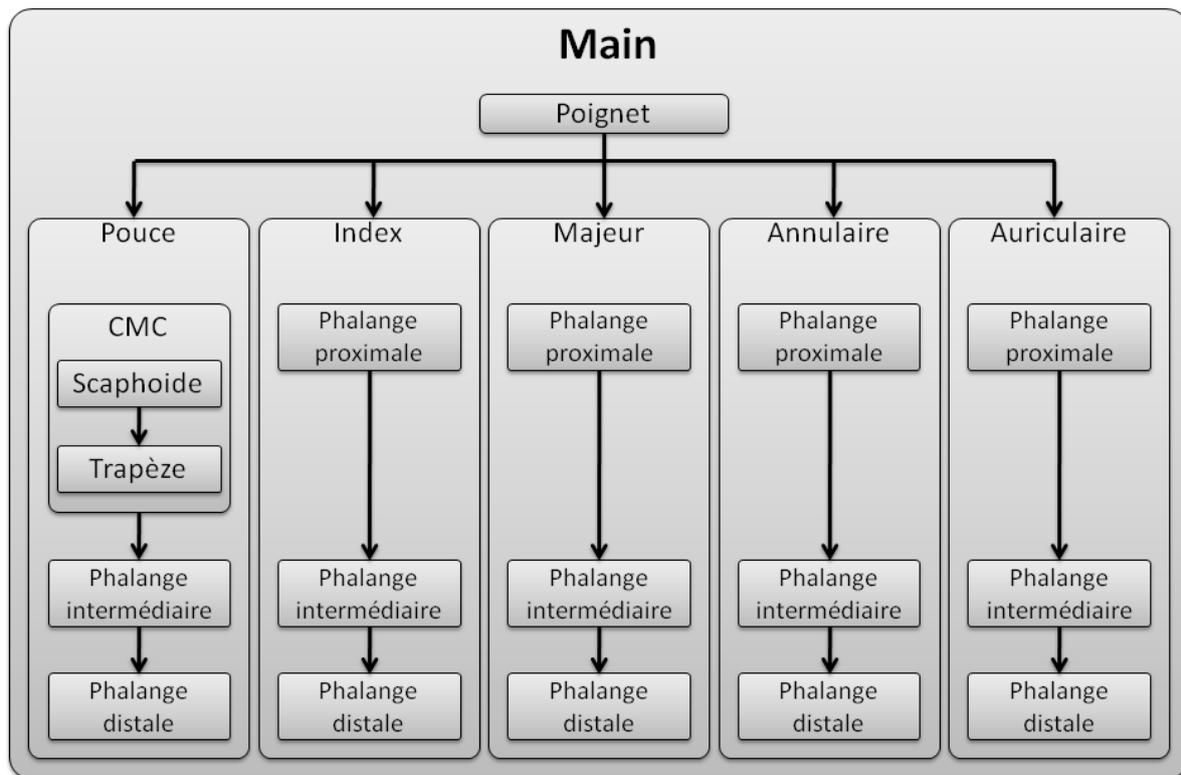


Figure 23 Organisation en blocs de la main

Les sites corporels utilisés sont détaillés dans l'annexe conjointement avec le squelette. Comme mentionné à la section 2.1.4, cette organisation ne donne aucune information sur les valeurs effectives de la configuration basique utilisée ou sur la longueur des os. Nous allons maintenant détailler ce dernier point.

### 3. Mesures anthropométriques des bras et de la main

L'utilisation d'un squelette bien formé est la première étape de la construction d'un signeur virtuel. Nous adoptons le parti pris d'avoir un personnage neutre visuellement.

#### 3.1. De la nécessité de mesures anthropométriques

La question que traite cette section est la suivante : « *est-il possible d'obtenir des mesures permettant l'élaboration d'un personnage moyen ?* ». Par moyen, nous entendons représentant une moyenne de la taille des membres d'un humain. Bien évidemment, il n'est pas ici question de la peau du personnage virtuel. Ce point est entièrement destiné aux graphistes qui devront rendre le signeur neutre sur certains points et distinctifs sur certains autres (comme l'appartenance à l'organisme pour lequel les messages sont transmis dans le cas d'applications réelles pour le grand public).

Les mesures anthropométriques sont utilisées dans des domaines autres que l'animation. La conception de vêtements et de tenues nécessite la prise de mesures et l'élaboration de statistiques permettant de déterminer quelles quantités seront, en moyenne, les plus utilisées. L'anthropométrie est de même particulièrement utilisée en médecine. Malheureusement, les études de la NASA<sup>18</sup> ou encore du projet DINBelg<sup>19</sup> proposent généralement l'élaboration de mesures sur la peau plutôt que sur les os. Comment rendre compte alors de valeurs moyennes permettant de rendre naturel notre personnage virtuel et de le piloter de manière plus efficace ?

Les études déjà effectuées donnent des mesures assez précises pour la colonne vertébrale, le complexe de l'épaule et les bras. Le réel problème se pose généralement pour les mains dont les mesures ne permettent pas une évaluation précise des tailles respectives des phalanges. Si la définition des normes du squelette peut être laissée aux bon jugement des graphistes responsables de la peau il reste tout de même intéressant d'étudier les normes des éléments de la main de manière à proposer un ensemble de mesures cohérent avec les données anthropométriques recueillies jusqu'ici. L'absence de mesures dans ce domaine est principalement due à la haute variabilité inter-individu et inter-os. La variabilité de la taille des bras ou des avant-bras englobe généralement la mesure de la taille des phalanges et il est nécessaire de présenter les données sous un autre angle. Les sections

---

<sup>18</sup> <http://msis.jsc.nasa.gov/sections/section03.htm>

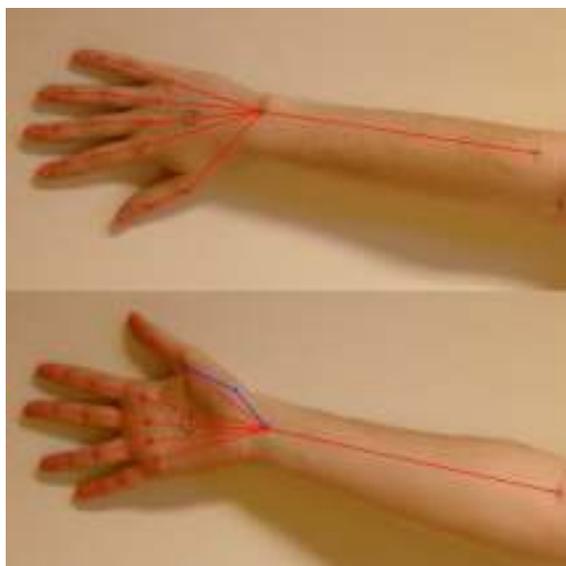
<sup>19</sup> <http://www.dinbelg.be/anthropometrie.htm>

suivantes présentent une étude anthropométrique de la longueur des os constituant la main, relativement à la longueur de l'avant bras du personnage.

### 3.2. Méthodologie de conception des mesures

La méthodologie que nous proposons permet d'établir une longueur normalisée des doigts par rapport à la taille de l'avant bras. La méthode se veut applicable aisément et non invasive. Elle ne donne qu'une information approchée, il n'est pas question de mesurer précisément la longueur des os, mais de faire un rapport pertinent entre différents points d'insertion articulaires. Cette méthode a nécessité la participation d'un panel de sujets.

**Figure 24** Points de contrôle disposés sur l'avant bras pour les mesures anthropométriques. En bleu, mesures pour l'élaboration du modèle de pouce présenté en section 5.3.



On commence par marquer 42 points sur le bras du sujet (cf. Figure 24):

- 20 de chaque côté de la paume :
  - un sur chaque extrémité des doigts (total : 5),
  - un sur chaque articulation phalangienne (total : 14)<sup>20</sup>,
  - un sur le milieu du poignet, entre l'extrémité du cubitus et celle du radius,
- un sur l'insertion du coude côté intérieur,
- un sur l'insertion du radius à l'extérieur du coude

La distance entre les points d'articulations est ensuite reportée comme un rapport de la longueur de l'avant bras. On détermine cette longueur comme étant la distance de l'insertion du radius à l'extérieur au poignet. On extrait ensuite de la population la moyenne

---

<sup>20</sup> On compte douze points pour l'index, le majeur, l'annulaire et l'auriculaire auxquels on ajoute deux points pour le pouce qui ne dispose pas de phalange intermédiaire.

et l'écart-type de chaque longueur qui nous servira de base pour la définition des cotes du squelette.

### 3.3. Résultats de l'étude

La méthode a été appliquée sur 24 sujets divisés équitablement entre les deux sexes. Les résultats sont indiqués dans le tableau suivant :

	<b>Métacarpe</b>	<b>Phalange Proximale</b>	<b>Phalange Intermédiaire</b>	<b>Phalange Distale</b>
<b>Pouce</b>	0.336(0.037)	0.116(0.013)	N/A	0.114(0.009)
<b>Index</b>	0.385(0.045)	0.163(0.018)	0.085(0.010)	0.084(0.010)
<b>Majeur</b>	0.359(0.044)	0.187(0.021)	0.103(0.010)	0.089(0.012)
<b>Annulaire</b>	0.333(0.050)	0.171(0.021)	0.095(0.008)	0.089(0.010)
<b>Auriculaire</b>	0.323(0.041)	0.132(0.017)	0.068(0.008)	0.080(0.008)

Tableau 4 : Moyenne (et écart-type) des longueurs relatives à l'avant-bras des phalanges de la main

Ces données ont été employées pour créer une main moyenne dont les cotes varient en fonction de la taille de l'avant bras. La Figure 25 donne un exemple de squelette d'avant bras construit sur cette base.

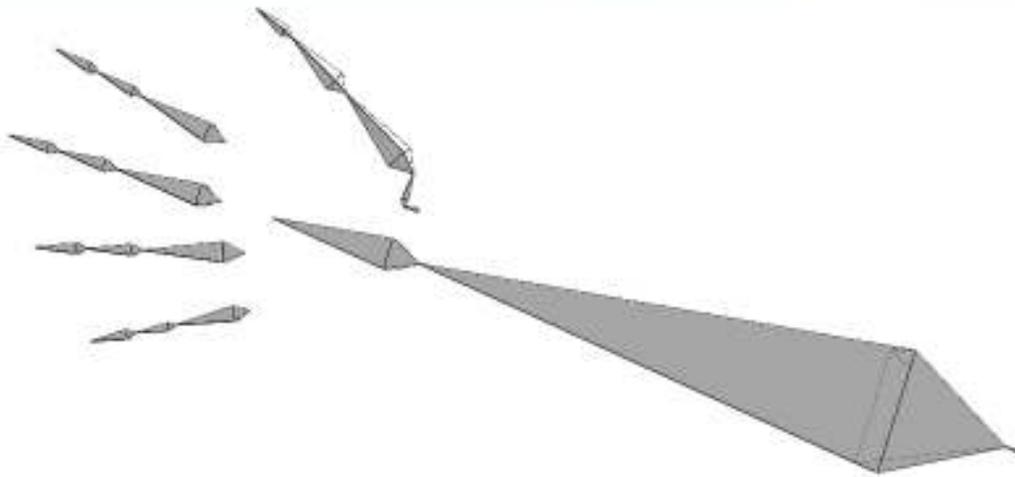
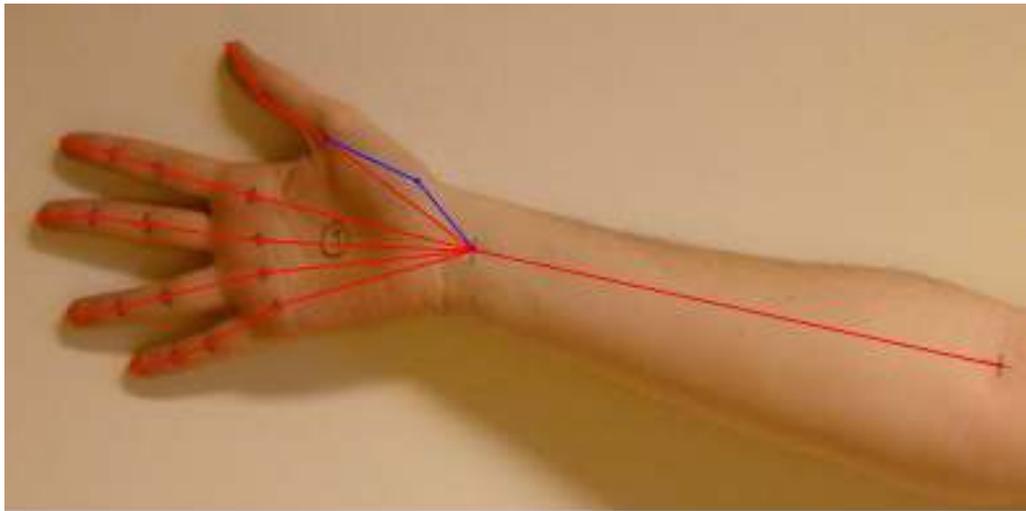


Figure 25 Exemple d'avant-bras réalisé en accord avec les normes établies par l'étude.

### 3.4. Conclusion de l'étude anthropométrique

L'étude anthropométrique nous permet d'accroître le réalisme du signeur en lui donnant des proportions humaines. Cependant, la définition de proportions réalistes n'accroît que de peu le naturel du signeur si celui-ci n'est pas piloté correctement par le système de cinématique inverse. Nous avons vu dans les deux précédentes sections comment définir un squelette, lui donner une hiérarchie cohérente et des proportions réalistes. Nous allons maintenant nous intéresser au problème des dépendances articulaires et la limitation du squelette à un domaine de mouvements physiquement possibles.

## 4. Dépendances articulaires

Le système de cinématique inverse utilisé génère les postures de manière mathématique. Que la méthode de résolution des postures soit analytique ou numérique<sup>21</sup>, le système ne dispose d'aucune « conscience » interne du squelette. Ces méthodes de résolution ne tiennent compte que des limitations imposées par les degrés de libertés du squelette : le nombre d'axe de rotation pour chaque bloc.

Il est impossible de simuler de manière réaliste la rotation d'une articulation du squelette humain en se contentant de limiter son nombre de degrés de libertés. De ce fait, la cinématique inverse est fortement sujette à la création de postures peu réalistes du squelette telle que celle illustrée précédemment pour la main.

Il est nécessaire d'employer d'autres méthodes pour que la physiologie du squelette soit prise en compte de manière effective et que les mouvements de rotation appliqués aux articulations soient limités de façon à reproduire des postures réalistes du corps humain. En anatomie, la limitation des rotations des articulations ne dépend pas simplement de la forme des os, mais aussi des jeux de tendons, de muscles et de ligaments permettant le mouvement. Ces interactions sont trop lourdes à simuler précisément et nécessitent des approximations que nous proposons de modéliser.

Cette section présente la gestion des limitations du corps au moyen de dépendances articulaires. Comme indiqué en section 1.2.1, il existe trois types de dépendances articulaires que nous allons voir dans les sections suivantes : simples (4.1), dynamiques (4.2) et distributives (4.3).

### 4.1. Dépendances simples

Les dépendances simples sont des dépendances s'appliquant de manière statique à un degré de liberté. Elles permettent la limitation de la rotation de ce degré de liberté à un intervalle appelé *amplitude de mouvement*. Cet intervalle est fixé dans la description même du squelette et, malgré qu'il puisse théoriquement être sujet à variations inter-individu, il est défini comme immuable quelque soit le signeur. Le Tableau 5 regroupe les différents degrés de libertés sujets à dépendance simple dans notre définition du squelette.

---

<sup>21</sup> Une méthode de résolution analytique implique la recherche d'une solution exacte au problème de cinématique inverse en employant les outils trigonométriques. Les méthodes numériques utilisent des heuristiques permettant la découverte de solutions approximatives répondant au problème.

Les valeurs de ces limitations sont extraites des différentes sources de mesures anthropométriques citées en section 3.1. Le respect de ces contraintes par le système est simple : lors de la modification de la valeur de l'angle du degré de liberté, on vérifie que la nouvelle valeur est bien dans l'intervalle imposé. Si ce n'est pas le cas, l'angle est placé sur la borne la plus proche.

Bloc	Degré de liberté	Angle minimum	Angle maximum
Coude	X	0	150
	Y	-6	160
Poignet	X	-17	120
	Y	-40	0
	Z	-45	120
MCP des doigts <sup>22</sup>	Z	0	90

Tableau 5 : Amplitudes de mouvement (en degrés) pour les articulations non dépendantes du squelette.

Ces limitations sont particulièrement simples à appliquer mais ne représentent qu'une minorité de cas. La plupart des articulations sont sujettes à interdépendance, qu'il s'agisse de dépendances au niveau des amplitudes de mouvement, ou d'une distribution du mouvement d'une articulation consciente sur plusieurs articulations inconscientes.

#### 4.2. Dépendances dynamiques

Les dépendances dynamiques correspondent au cas où l'amplitude de mouvement d'un degré de liberté d'un bloc est modulée par les degrés de libertés d'un autre bloc. Par exemple, plier la première articulation inter-phalangienne de n'importe quel doigt contraint aussi la seconde. Ces dépendances sont régies par des lois mathématiques simples de la forme  $(min_i; max_i) = f(\alpha_j)$  où  $i$  et  $j$  sont des degrés de liberté,  $\alpha_j$  l'angle attribué au degré de liberté  $j$  et tel que le bloc contenant  $j$  ne se situe pas dans le sous-graphe de connexions partant de  $i$ .

La modélisation simplifiée de comportements réalistes justifie l'élaboration de telles formules approximatives. Il n'est encore une fois pas question de simuler de la manière la plus précise qui soit le fonctionnement des articulations du corps humain mais bien de proposer des heuristiques permettant au signeur virtuel de paraître humain. Les dépendances dynamiques permettent de modéliser les relations d'interdépendances de la main telles que celles figurant dans (Losson 2000).

<sup>22</sup> Cette ligne concerne les articulations MétaCarpoPhalangiennes de tous les doigts à l'exception de celles des deux pouces, qui sont gérés différemment. Les détails de cette exception sont exposés en section 5.

L'évaluation des dépendances dynamiques est moins directe que celle des dépendances simples. Il est nécessaire, à la modification d'un angle, de vérifier tous les autres degrés de liberté dépendant de celui-ci, de recalculer les angles en conséquence et d'appliquer ce traitement en cascade. Ce dernier point peut être épineux dans le cas de dépendances circulaires (comme pour un mouvement d'adduction des doigts).

L'intégralité des dépendances dynamiques appliquées à notre modèle de squelette est détaillée en annexe 2. Celles-ci permettent de couvrir, avec l'aide des dépendances simples, la majorité des cas. Il reste néanmoins une catégorie de dépendances importantes justifiant les appellations d'articulations « conscientes » et « inconscientes » : les dépendances distributives.

### 4.3. Dépendances distributives

Ce dernier type de dépendance articulaire distribue le mouvement d'un niveau d'abstraction aux niveaux directement inférieurs. L'objectif de ce type de dépendance est de modéliser de manière plus précise le phénomène d'articulation consciente et inconsciente. La distribution s'applique depuis un bloc conscient (e.g. le complexe de l'épaule) sur plusieurs blocs inconscients sous-jacents (e.g. la clavicule et l'épaule). Les dépendances distributives sont donc des systèmes d'équations où les angles des degrés de liberté sous-jacents sont exprimés en fonction de l'angle du degré de liberté conscient.

L'expression de ces distributions n'est pas mutuellement exclusive avec les précédents types de dépendance car elles ne concernent pas les amplitudes de mouvement. Les dépendances distributives ne s'appliquent que sur la valeur des angles de degrés de liberté concernés. Il est donc possible qu'un degré de liberté soit à la fois sujet à une dépendance distributive et une dépendance dynamique. En conséquence, il est possible que les deux types de dépendances entrent en conflit, auquel cas la priorité est toujours laissée au plus contraignant : la dépendance dynamique.

Badler dans son livre (Badler et al. 1999) donne un exemple de dépendance distributive, celle de l'épaule. D'autres dépendances sont élaborées à partir de données empiriques telles que celles de la colonne vertébrale. Les lois régissant la colonne vertébrale sont connues du domaine médical mais du fait de notre modélisation simplifiée de cette dernière, il est nécessaire de procéder à une adaptation. Ces deux distributions sont indiquées en annexe 2. Par ailleurs, l'interdépendance de la flexion des doigts peut être représentée par une

dépendance distributive sur les phalanges. Cette relation, étudiée en (Losson 2000), est aussi décrite en annexe.

L'utilisation de méthodes de cinématique inverse ne permet pas le contrôle à différents niveaux d'abstraction. Les dépendances distributives, au premier abord, ne sont destinées qu'à un contrôle manuel. Cependant l'expression de ces dépendances est généralement simple et permet l'élaboration de fonctions réciproques permettant le passage de la valeur d'un bloc inconscient au bloc conscient supérieur, et donc à nouveau la distribution sur tous les degrés de liberté sous-jacents. L'intégration de ces dépendances au système nécessite aussi l'utilisation d'algorithmes particuliers. Ces derniers sont donnés dans le chapitre III.

#### **4.4. Résultats de l'utilisation de dépendances**

Les dépendances articulaires permettent une gestion statique ou dynamique des limitations physiologiques du corps humain. Il est cependant nécessaire de calibrer effectivement ces limitations et de bien concevoir que la modélisation simplifiée du corps humain nécessite un paramétrage adéquat. Une limitation insuffisante conduit généralement à des postures non réalistes du corps, alors qu'une limitation excessive empêche (par le surplus de contraintes) le système de cinématique inverse de réaliser les postures qu'on lui demande.

Les dépendances sur les mains permettent un contrôle des doigts plus réaliste du fait des liens créés entre les phalanges. De même le mouvement distribué sur l'épaule rend l'élévation du bras plus réaliste qu'un contrôle séparé des articulations de la clavicule et de l'épaule. De manière générale, ces dépendances mènent à une modélisation plus fine et plus réaliste, alliée à un contrôle plus précis du système de cinématique inverse.

Cependant, malgré les efforts d'une limitation plus contraignante, l'articulation du pouce (notamment lors de mouvements d'opposition) n'a pas un rendu réaliste, ni un contrôle aisé du fait de sa modélisation trop simpliste.

## 5. Complexe carpo-métacarpien

Comme discuté en section 1.2.2 il est nécessaire de modéliser le pouce de manière plus réaliste afin de contraindre les méthodes de cinématique inverse à un espace de solutions plus viable qu'un système trop libre. Pour ce faire, nous modélisons le complexe carpo-métacarpien de manière à approximer l'opposabilité du pouce.

### 5.1. Main humaine vs. main virtuelle

La main humaine est un élément du corps dont la complexité interroge toujours la communauté scientifique. Plusieurs modèles de main virtuelles existent pour l'animation mais ils présentent généralement des limites pour simuler correctement le comportement du pouce. Comme nous l'avons vu auparavant (section 0 du premier chapitre), ce problème est principalement dû aux modèles employés dans l'animation, se basant sur la réalisation d'un acteur et ne posant que rarement le problème de la « posture impossible ». Afin d'établir un modèle répondant aux besoins du système de génération il est nécessaire de revenir sur le fonctionnement de la main en anatomie puis d'établir un modèle permettant de simuler son comportement de manière satisfaisante.

### 5.2. Chaîne articulaire du pouce

La chaîne articulaire qui nous intéresse part du poignet et remonte le long du pouce. Le premier élément faisant le lien entre le poignet et les phalanges du pouce est appelé complexe carpo-métacarpien (CMC<sup>23</sup>). Il est composé en réalité de deux os bien distincts : le scaphoïde et le trapèze. L'articulation de ces deux os crée une jointure de type selle (cf. Figure 26) permettant la rotation du pouce selon deux axes. Les deux rotations ne suivent pas le même centre, ce qui a pour conséquence de placer l'articulation du métacarpe sur une trajectoire circulaire.



Figure 26 Articulation en selle, autorisant la rotation selon deux degrés de liberté.

---

<sup>23</sup> Carpo-Metacarpal Complex

La suite de la chaîne articulaire est triviale, il s'agit des trois articulations des phalanges (MCP, proximale et distale). La littérature modélise généralement ces trois articulations sous la forme de jointures de type charnières, ce qui est une modélisation satisfaisante pour notre domaine. Notre objectif est de modéliser le CMC de manière à permettre une rotation effective du pouce en opposition.

### 5.3. Modélisation du complexe carpo-métacarpien

L'utilisation du terme anatomique *complexe carpo-métacarpien* nous suggère tout d'abord qu'il est intéressant (voire nécessaire) de ne considérer ces deux os (le scaphoïde et le trapèze) comme une seule et même entité consciente. La rotation de ce complexe en intégralité permet l'opposition du pouce et il ne serait pas judicieux de dissocier les deux os à un niveau d'abstraction élevé. Le squelette intègre donc un bloc articulaire représentant le complexe et contenant les deux os sous-jacents, sous la forme de blocs inconscients. Etant donné que la jointure formée ne comporte que deux degrés de liberté, la distribution des rotations au sein du complexe s'effectue de manière triviale. Cependant, les deux os intégrés au CMC ne sont pas des copies anatomiques du scaphoïde ou du trapèze. L'objectif est de simuler la trajectoire circulaire de l'articulation entre métacarpe et première phalange observée en utilisant deux degrés de liberté. Les deux os sont donc articulés de manière à former un cône. Le premier (le scaphoïde) est autorisé à tourner selon les axes directeur (X) et normal (Z). Le second (le trapèze) est complètement fixé de manière à créer une translation permettant la localisation efficace de l'articulation de la phalange proximale. Ce complexe crée un cône de mouvement permettant l'opposition du pouce et sa contrainte sur une trajectoire circulaire.

### 5.4. Création d'un axe de rotation pour l'opposition

Une telle modélisation du CMC permet la création d'un axe de révolution non aligné avec les phalanges du pouce (cf. Figure 27). Cet axe garantit l'opposabilité du pouce en plaçant l'articulation métacarpo-phalangienne (MCP) sur une trajectoire circulaire. L'existence anatomique de cet axe n'est pas confirmée. Le complexe carpo-métacarpien ne peut être résumé à un simple axe de rotation : l'opposition du pouce est en fait le résultat de l'action d'un jeu de tendons, de muscles et d'os créant un mouvement complexe. Cependant, nous posons comme hypothèse que *la simulation de ce mouvement complexe par un axe de rotation virtuel est une approximation satisfaisante au problème du pouce et donne de bien meilleurs résultats que la simulation traditionnelle*. Il faut cependant garder à l'esprit que ce

que nous appelons ici trapèze et scaphoïde ne sont pas des os simulant précisément le comportement de leurs contreparties anatomiques.



Figure 27 Axe de rotation créé pour simuler l'opposition du pouce par l'action conjointe du scaphoïde et du trapèze.

L'orientation de l'axe formé est fonction de la configuration du CMC, et donc de la taille des deux os ainsi que de l'angle  $\alpha$  formé entre le poignet et le CMC. L'axe principal de la rotation étant l'axe directeur du scaphoïde, il est nécessaire de trouver une valeur de  $\alpha$  permettant une rotation réaliste et crédible. Du fait du décalage entre le modèle et le réel fonctionnement de l'anatomie humaine, déterminer l'orientation de cet axe se fait de manière empirique car aucune donnée biologique ne peut venir appuyer le choix de cet angle. Placer l'axe de manière à former un angle de 9 degrés avec le vecteur directeur de l'avant bras sur le plan coronal fournit des résultats satisfaisants sur le placement du pouce en opposition.

### 5.5. Limitations articulaires

Le calibrage de l'axe de rotation passe aussi par l'étude de ses dépendances articulaires comme détaillées en section 4. Le CMC étant le premier élément de la chaîne articulaire, la présence de dépendances articulaires complexes (comme des ROM dynamiques) ne serait liée qu'au positionnement du poignet. Cette dépendance, peu présente dans la littérature à notre connaissance, est considérée comme inexistante : l'opposition du pouce ne dépend pas de la configuration du poignet. Les limitations articulaires du CMC ne sont donc considérées que comme dépendances simples : l'amplitude de mouvement est en moyenne de 42 degrés (Austin 2005). Les limitations articulaires de l'axe de flexion/extension du CMC et de chaque phalange quant à elles ne

changent pas par rapport aux valeurs données traditionnelles de la littérature : 20 degrés pour le scaphoïde et 70 degrés pour le premier métacarpe.

Le réalisme apporté par ce modèle de pouce ne peut être assuré pour les autres chaînes articulaires en raison du peu de contraintes sur le corps humain. La majorité des problèmes de cinématique inverse restent sous-spécifiés, et mènent à une infinité de solution parmi lesquelles il faut choisir. Ce problème est traité dans la section suivante.

## 6. Modèle de confort

Le modèle de confort est sans doute la contribution majeure de cette thèse. Il s'agit de construire, à partir de données basées sur une réalisation effective, une méthode de sélection d'une posture parmi un ensemble. Ce modèle permet de résoudre le problème de sous-spécification en sélectionnant la posture la plus naturelle parmi un ensemble fini de solutions.

### 6.1. Définitions de l'effort et du confort

En préface de cette section, définissons les termes *effort* et *confort* :

Le Trésor de la Langue Française (TLF<sup>24</sup>) définit l'effort comme étant la « *mise en œuvre de toutes les capacités d'un être vivant pour vaincre une résistance ou surmonter une difficulté* ». Lorsque l'on parle d'effort en anatomie et en biomécanique, il s'agit généralement de mesurer les forces déployées par le corps humain dans certaines situations très particulières (e.g. les forces déployées par les muscles du bras lorsque l'on porte une charge). Dans le cadre de la synthèse automatique de la langue des signes, le problème de multiplicité des solutions peut être résolu théoriquement en cherchant à minimiser l'effort produit par le signeur. La littérature en biomécanique tend toujours vers un certain consensus indiquant que le corps humain (et la nature biologique de manière générale) minimise la dépense d'énergie et, en ce qui nous concerne, l'effort.

Nous définissons le *confort* (dans le contexte de cette thèse) comme étant une mesure inverse de l'effort<sup>25</sup> déployé par le signeur dans une posture particulière. En d'autres termes, une posture du squelette est dite confortable à partir du moment où elle cherche à minimiser l'effort déployé par le signeur. Enfin nous posons l'hypothèse suivante : suivant un ensemble de contraintes imposées au corps humain, un signeur aura toujours tendance à prendre la posture la plus confortable pour répondre à ces contraintes.

Posons-nous la question suivante : « Quels critères nous permettent de différencier une posture naturelle d'une autre ? ». Considérons un signe simple tel que [TABLE] en LSF. La réalisation de ce signe, telle que représentée dans le dictionnaire de l'IVT<sup>26</sup>, est donnée en Figure 28. Si nous analysons la construction du signe, nous pouvons le résumer à la séquence suivante :

---

<sup>24</sup> <http://atilf.atilf.fr>

<sup>25</sup> Comprendre que le confort est inversement proportionnel à l'effort

<sup>26</sup> IVT : International Visual Theatre est un organisme français promouvant la LSF, notamment par la constitution de dictionnaires français/LSF.

- Mettre les deux mains à plat l'une à côté de l'autre, devant le signeur.
- Ecarter progressivement les mains sur le plan transverse

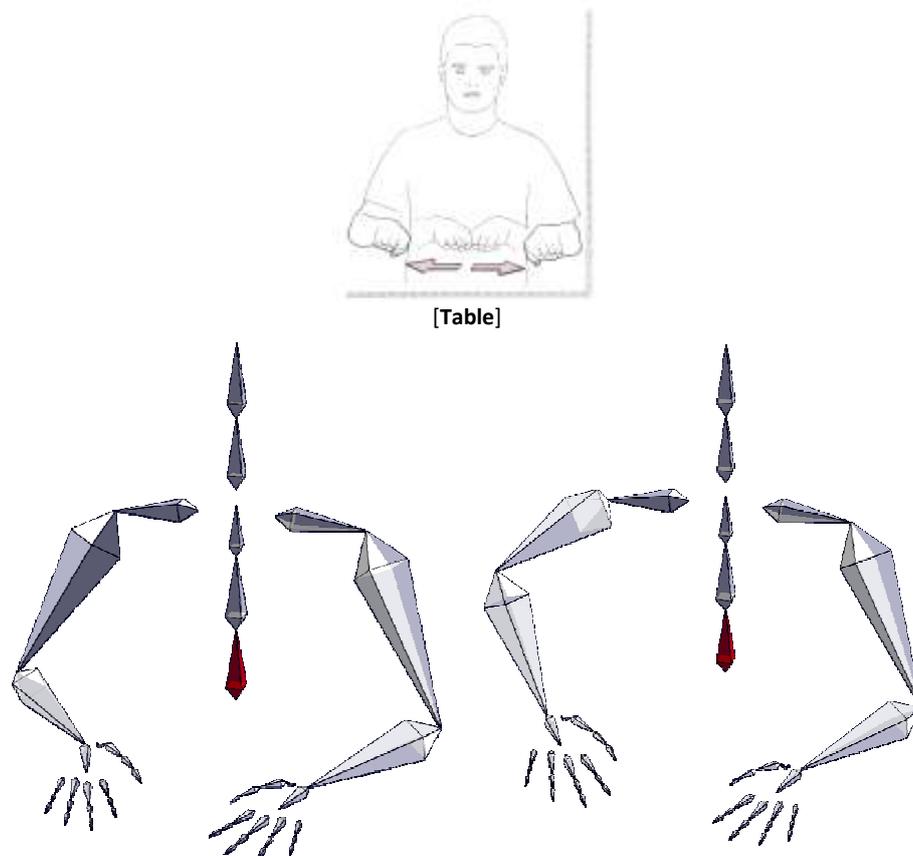


Figure 28 Haut : Le signe [TABLE] en LSF, extrait du dictionnaire de l'IVT. Bas : Deux solutions possibles de postures lors de la transition. A droite, la solution semble moins confortable du fait du coude levé.

La première étape de la construction de ce signe est un exemple typique de problème sous-contraint. Contraindre les mains revient à contraindre les poignets du signeur. Cependant, rien n'est dit au sujet des épaules et des coudes, ce qui nous permet de considérer plusieurs solutions à un même problème comme figuré sur la Figure 28.

Ces deux solutions sont linguistiquement valables. Pour autant, il nous paraît naturel que la solution de gauche soit la plus correcte des deux. En effet, les coudes repliés sur la figure de gauche nous semblent instinctivement répondre à cette question de confort : lever les coudes pour construire ce signe nécessiterait une dépense d'énergie inutile.

## 6.2. Elaboration d'un protocole

Maintenant que le confort est défini, il s'agit de modéliser cette notion de minimisation de l'énergie de manière à pouvoir l'employer dans un système de génération automatique de langue des signes. Il existe de nombreux modèles en biomécanique et en médecine permettant la simulation de l'effort musculaire et donc, dans une certaine mesure, du

confort. Il faut cependant noter qu'une fois de plus, ces modèles sont extrêmement lourds et ne peuvent pas être utilisés tels quels pour l'animation.

Il est donc nécessaire de trouver d'autres modèles ou d'adapter les modèles existants. (Neff et al. 2006) donne un exemple d'adaptation d'un modèle biophysique (effet de la gravité sur la main) pour des applications en temps réel. Il procède par étapes de pré-calcul créant ainsi de grands tableaux de données réutilisables par la suite en temps réel. Cependant, ce modèle ne tient compte que d'un seul effet parmi une multitude. De plus, si nous tentions d'appliquer un tel système à la modélisation du confort, nous serions confrontés à un volume de données trop important. En effet, le modèle présenté en référence ci-dessus ne concerne que les doigts de la main et donne des informations sur des chaînes cinématiques relativement courtes. Notre modèle de confort doit être capable de fonctionner sur les épaules, les coudes et sur la colonne vertébrale.

Poussons le questionnement plus loin et interrogeons-nous sur la portée du modèle de confort. Nous venons de voir qu'il serait bon qu'il s'applique aux coudes, épaules et dos du signeur (l'exemple [TABLE] en est un exemple). Mais est-il nécessaire de disposer d'un tel modèle pour les mains ? La définition de configurations manuelles revient généralement à la résolution de problèmes à très fortes contraintes. Ces problèmes ne laissent pas ou peu de liberté aux articulations, ne nécessitant donc pas de correction de confort. Les seules articulations généralement libres sont celles du CMC qui ne peut être modélisé en termes anatomiques puisqu'il ne fonctionne pas comme son analogue biologique (cf. section 5.3). Ainsi, dans notre étude, domaine de définition du modèle de confort ne concerne dans notre contexte d'étude que la colonne vertébrale, les épaules et les coudes.

Comme mentionné précédemment, il n'est pas pertinent d'utiliser un modèle précis de biomécanique pour gérer le confort des articulations en raison du grand nombre de degrés de liberté du système. Le modèle, au lieu d'être dynamique, est donc statique, à partir de données précalculées. Nous avons choisi de construire un modèle du confort basé sur une approche statistique dirigée par les données. Le modèle est construit à partir de données de capture de mouvement de laquelle nous extrayons la fréquence de chaque posture pour chaque degré de liberté.

### 6.3. Corpus d'étude

Les données utilisées pour cette étude sont tirées de la base de données de capture de mouvement de l'université de Carnegie Mellon (CMU)<sup>27</sup>. Ce corpus est constitué d'un total de 4 268 904 frames d'animation. Le squelette enregistré dans le corpus est une version partiellement différente du notre. Les différences principales sont au nombre de deux : la présence d'une section inférieure (ne nous servant pas), et l'absence de doigts sur le squelette du corpus. Cependant, comme mentionné précédemment, les mains ne sont pas incluses dans le modèle de confort du fait d'un nombre de contraintes généralement élevé.

Le corpus fut à l'origine constitué sur la thématique du mouvement au sens général du terme. Il regroupe des animations d'interaction (jeu, discussion), de sport, de marche, etc. Il n'est donc pas centré sur la langue des signes. Pour autant, le grand volume de données nous pousse à proposer l'hypothèse suivante : *Les données sont en quantité suffisantes pour recouvrir la palette de mouvements proposés en langue des signes*, et ce, considérant aussi le domaine de définition que nous nous sommes données en 6.1. Cette hypothèse très forte crée une perspective importante à ce travail : la nécessité dans le futur de confronter ces données avec un corpus de capture de mouvement dédié à la langue des signes lorsqu'un tel corpus sera disponible.

Le choix de ce corpus par rapport à un corpus d'étude de langue des signes répond à trois critères :

- Tout d'abord, peu de corpus de langue des signes ont été constitués pour le domaine de l'animation, et aucun d'entre eux n'est disponible librement à l'heure actuelle.
- Par ailleurs, les corpus existants et centrés sur la langue des signes disposent de peu d'heures d'enregistrement, rendant un modèle statistique moins robuste et moins fiable.
- La constitution d'un nouveau corpus répondant spécifiquement à nos besoins aurait nécessité une expertise, un temps et des moyens matériels dont nous ne disposons pas.

---

<sup>27</sup> CMU pour *Carnegie Mellon University*. Le corpus est disponible à l'adresse suivante : <http://mocap.cs.cmu.edu/>

- Enfin, le corpus que nous utilisons est un corpus utilisé par la communauté scientifique depuis des années. Il est non seulement de grand volume et de bonne facture mais aussi nettoyé et disponible librement.

Enfin, il est important de noter que le corpus de CMU présente des enregistrements effectués par plusieurs acteurs. En tout, le corpus compte 144 sujets différents. Cette information est capitale pour la constitution d'un modèle statistique, car la diversité des acteurs permet d'écartier les biais liés à la personnalité d'un seul acteur. Le modèle extrait ne reproduira pas le comportement d'un seul acteur, mais de tous, comme une moyenne à la manière des mesures anthropométriques moyennes détaillées en section 3.

#### 6.4. Extraction du modèle

Nous sélectionnons d'abord les parties du corpus correspondant au domaine de définition soit un total de vingt-sept degrés de liberté. Chaque articulation est représentée par trois degrés de liberté. Le modèle considère les articulations suivantes :

- Les trois segments de la colonne vertébrale (neuf degrés de liberté)
- Complexe de l'épaule<sup>28</sup> gauche/épaule droite (douze degrés de liberté)
- Coude gauche/Coude droit (six degrés de liberté)

##### 6.4.1. Méthodologie

On extrait pour chacun de ces degrés de liberté l'histogramme des occurrences des angles. La résolution de chaque histogramme (bien qu'accessoire pour l'instant) est d'un dixième de degré. On estime ainsi visuellement le confort pour chaque degré de liberté. De ces histogrammes sont aussi extraits les quantiles afin de fournir une représentation numérique grossière.

Sous cette forme, les histogrammes ne sont que des représentations permettant à l'utilisateur de visualiser la distribution des courbes de confort. Ces données sont inexploitable en l'état. Nous cherchons donc à extraire de ces histogrammes une densité de probabilité permettant de donner un score pour chaque degré de liberté. Afin d'extraire ces densités de probabilités, nous utilisons la méthode dite d'estimation par noyaux (Parzen 1962). Cette méthode est non-paramétrique et ne nécessite donc pas que les données suivent une loi normale. Elle est considérée comme une méthode permettant de généraliser, d'abstraire, un histogramme en densité de probabilité.

---

<sup>28</sup> Le complexe contenant deux articulations : la clavicule et l'épaule.

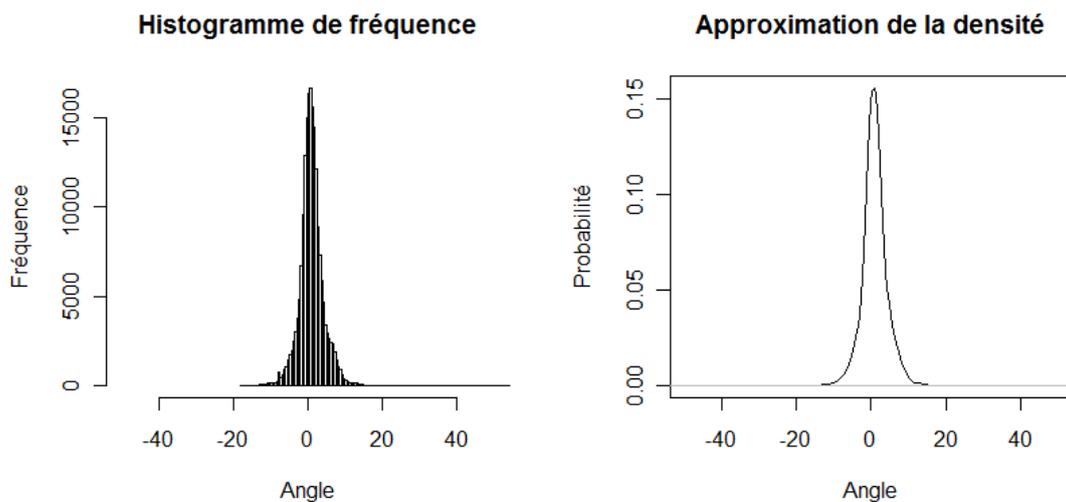


Figure 29 Approximation d'un histogramme de fréquence en densité de probabilité au moyen de la méthode des noyaux.

La méthode des noyaux estime la densité de probabilité en fonction d'un échantillon issu d'une variable aléatoire. On définit ainsi le confort de chaque degré de liberté, comme étant autant de variables aléatoires à estimer par la méthode d'extraction par noyaux. La Figure 29 donne un exemple de transformation d'histogramme en densité de probabilité au moyen de la méthode des noyaux. Par ailleurs, le Tableau 6 résume les noms donnés à chaque variable aléatoire.

Nom de la variable aléatoire	Articulation	Degré de liberté
$CMF_x^{belly}$	Premier segment de colonne vertébrale	Rotation
$CMF_y^{belly}$		Adduction/Abduction
$CMF_z^{belly}$		Flexion/Extension
$CMF_x^{waist}$	Second segment de colonne vertébrale	Rotation
$CMF_y^{waist}$		Adduction/Abduction
$CMF_z^{waist}$		Flexion/Extension
$CMF_x^{thorax}$	Troisième segment de colonne vertébrale	Rotation
$CMF_y^{thorax}$		Adduction/Abduction
$CMF_z^{thorax}$		Flexion/Extension
$CMF_y^{left\_clavicle}$	Clavicule gauche	Élévation/Abaissement
$CMF_z^{left\_clavicle}$		Flexion/Extension
$CMF_x^{left\_shoulder}$	Epaule gauche	Rotation
$CMF_y^{left\_shoulder}$		Élévation/Abaissement

$CMF_z^{\text{left\_shoulder}}$		Antépulsion/Rétropulsion
$CMF_x^{\text{left\_elbow}}$	Coude gauche	Pronation/Supination
$CMF_z^{\text{left\_elbow}}$		Flexion/Extension
$CMF_y^{\text{right\_clavicle}}$	Clavicule droite	Elévation/Abaissement
$CMF_z^{\text{right\_clavicle}}$		Antépulsion/Rétropulsion
$CMF_x^{\text{right\_shoulder}}$	Epaule droite	Rotation
$CMF_y^{\text{right\_shoulder}}$		Elévation/Abaissement
$CMF_z^{\text{right\_shoulder}}$		Antépulsion/Rétropulsion
$CMF_x^{\text{right\_elbow}}$	Coude droit	Pronation/Supination
$CMF_z^{\text{right\_elbow}}$		Antépulsion/Rétropulsion

Tableau 6 : Liste des variables aléatoires du modèle de confort en fonction des degrés de liberté du système

Les sections suivantes détaillent les histogrammes et les courbes obtenues.

#### 6.4.2. Calibration des données

Le squelette du signeur virtuel n'est pas exactement le même que celui des enregistrements de capture de mouvement. Notamment, la position au repos est très différente, ne respectant pas les contraintes d'axes que nous nous sommes fixés pour l'utilisation conjointe avec Zebedee<sup>29</sup>. Avant de commencer l'intégration des données il est nécessaire de transformer chaque angle de capture de mouvement en angle que nous utilisons sur le squelette du signeur virtuel.

#### 6.4.3. Confort de la colonne vertébrale

Les trois segments de la colonne vertébrale sont traités indépendamment et de manière à être calqués sur le squelette final. Les données n'ont pas besoin d'être prétraitées. L'analyse et l'extraction des courbes de confort sont donc directes. Les figures 29 à 31 détaillent les densités résultantes respectives du premier, second et troisième segment de la colonne vertébrale.

<sup>29</sup> Rappel : Zebedee est le modèle de description linguistique des signes que nous utilisons. Il est présenté en section 2.5 du premier chapitre.

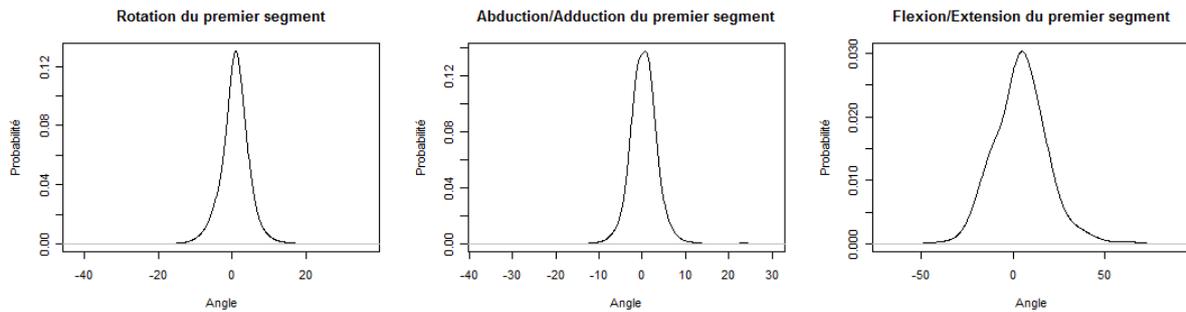


Figure 30 Densités de probabilité du premier segment de colonne vertébrale

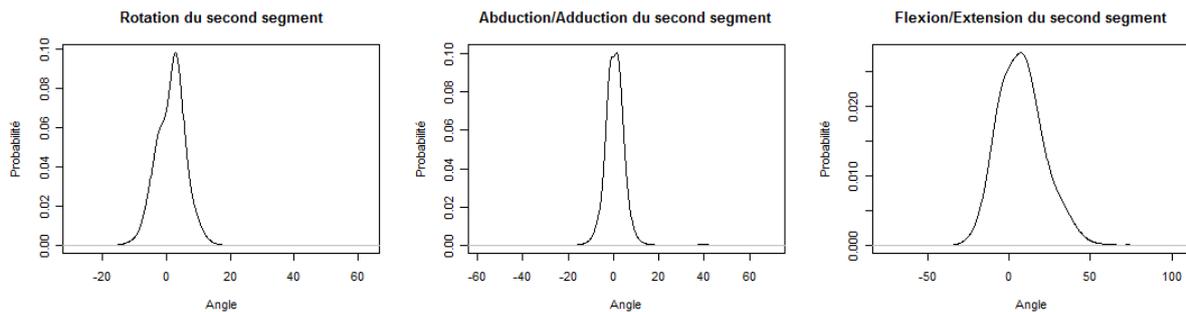


Figure 31 Densités de probabilité du second segment de colonne vertébrale

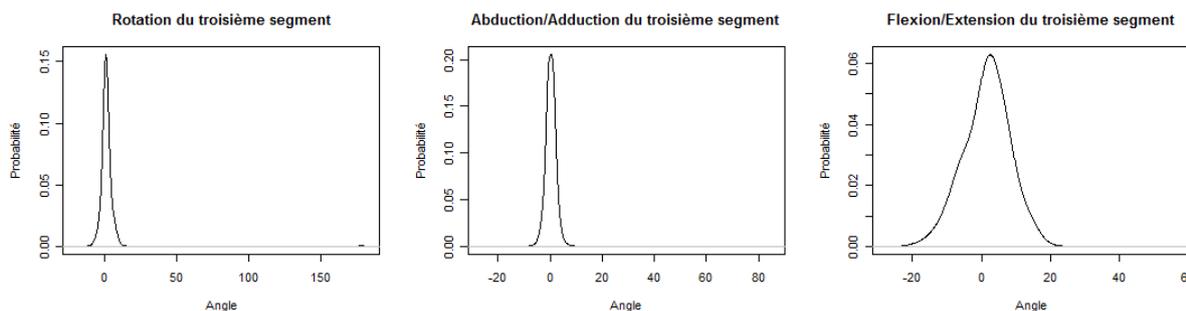


Figure 32 Densités de probabilité du troisième segment de colonne vertébrale.

On peut remarquer à l'analyse de ces données que la rotation d'axe directeur s'effectue pour les trois segments sur un intervalle de 40 degrés centrés autour de la position de repos. L'abduction/adduction semble plus prononcée sur le second segment (40 degrés d'amplitude) de colonne que sur les deux distaux (20 degrés d'amplitude). On explique cette variation du fait de la double courbure de la colonne nécessitant, pour une adduction ou une abduction forte une pliure dont le point d'inflexion est au milieu, la colonne vertébrale prenant alors une forme sigmoïde. Enfin, la flexion et l'extension ne se font pas de manière symétrique comme les deux types de rotations précédentes. A partir de la position de repos, il est plus fréquent pour un humain de fléchir la colonne que de l'étendre, donnant lieu à une courbe plus marquée d'un côté que de l'autre.

#### 6.4.4. Confort du coude

Le coude est un élément du corps auquel nous attribuons deux degrés de liberté : la flexion de l'avant bras et la rotation du poignet. La rotation du poignet est due à l'inversion de l'ulna et du radius, les deux os de l'avant bras. L'effet de cette inversion est particulièrement visible sur la peau, et bien que la simulation de ce phénomène par une rotation axiale au niveau du coude donne des résultats acceptables au niveau du squelette du personnage virtuel, il faut considérer avec attention les effets de cette rotation sur la peau du signeur. Il est aussi important de remarquer que la rotation axiale du poignet, s'effectuant au niveau du coude, est latéralisée. Elle n'est pas égale sur les deux coudes, ce qui n'est pas le cas de la flexion.

L'extraction des données par la méthode des noyaux nous donne les deux courbes suivantes :

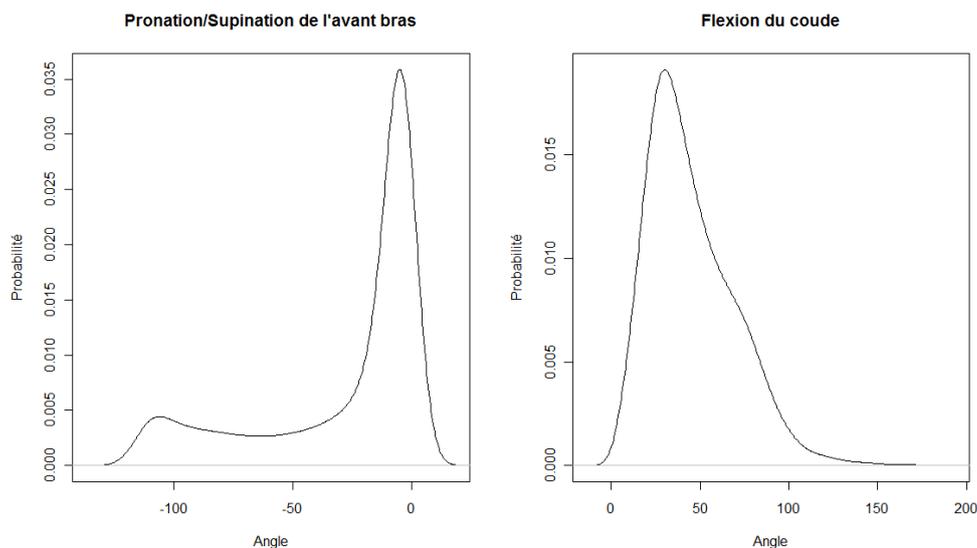


Figure 33 Densités de probabilité du coude

L'amplitude du mouvement de la flexion du coude est particulièrement large du fait de l'utilisation intensive de ce type de mouvement. De même on peut remarquer une forte probabilité que le poignet soit en position de repos, et une probabilité rehaussée en position inversée. Il est plus rare que le coude soit dans une position intermédiaire.

#### 6.4.5. Confort du complexe de l'épaule

Le complexe de l'épaule regroupe deux articulations : la clavicule et l'épaule pour un total de cinq degrés de liberté. Le complexe de l'épaule nécessite un travail important. Le placement des capteurs sur la clavicule des acteurs ne permet pas de capter la rotation de

cette dernière lors de mouvements d'antépulsion/rétropulsion<sup>30</sup> du bras. Toute la transformation se retrouve donc dans le bras et il est nécessaire de faire un prétraitement des données pour retrouver la valeur des angles des deux degrés de liberté concernés. Ce traitement correspond exactement à la distribution inverse que nous avons évoquée en section 4.3 de ce chapitre.

L'analyse par la méthode des noyaux donne les courbes suivantes pour la clavicule (cf. Figure 34).

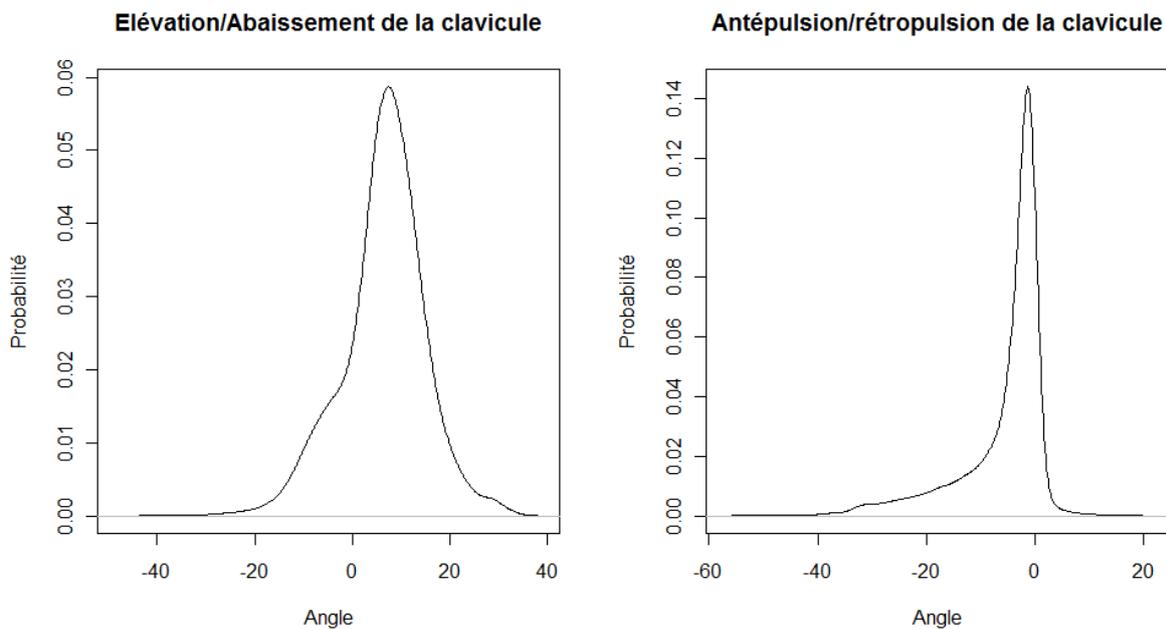


Figure 34 Densités de probabilité de la clavicule

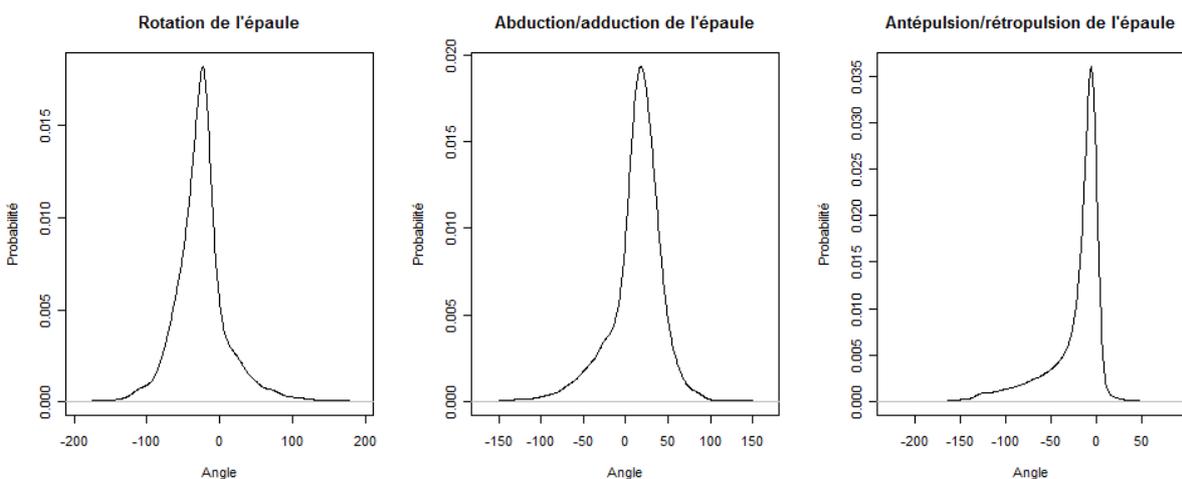


Figure 35 Densités de probabilité de l'épaule

<sup>30</sup> Mouvements d'élévation du bras le long du plan sagittal.

Le mouvement d'abduction/adduction de la clavicule dispose d'une amplitude d'environ 80 degrés centrés autour du point de repos. L'amplitude du même mouvement sur l'épaule est bien plus large du fait d'une mobilité accrue de l'articulation de l'épaule. En dépit de cette forte mobilité, l'amplitude effective se limite à une amplitude de 50 degrés pour l'épaule et de 30 degrés pour la clavicule. Le mouvement d'antépulsion/rétropulsion est bien plus limité, pouvant aller jusqu'à 180 degrés dans le meilleur des cas mais est principalement limité à 50 degrés. Il est important de noter que le mouvement d'antépulsion/rétropulsion au niveau de la clavicule est aussi compensé par le mouvement d'abduction/adduction. Le dernier degré de liberté, uniquement sur l'épaule est un mouvement de rotation axiale. Il s'étend sur environ 180 degrés de manière effective.

## 7. Conclusion du chapitre

Dans ce chapitre nous avons élaboré un modèle anatomique du squelette nous permettant une simulation à la fois plus précise et plus détaillée à un moindre coût en temps de calcul. L'élaboration de structures du squelette plus complexes (section 2, 3 et 4) permet un contrôle plus fin de la modification des degrés de liberté du squelette, de leur limitation et des phénomènes de distribution au sein de plusieurs articulations inconscientes.

La construction d'un nouveau modèle de pouce permet de même une simulation plus fine de son opposition, phénomène particulièrement présent lors de la génération de postures issues de la langue des signes française où la configuration des mains joue un rôle primordial dans la compréhensibilité des signes et où les dites configurations peuvent atteindre un degré de complexité dépassant largement celui de la gestuelle co-verbale.

L'élaboration d'un modèle théorique ne peut cependant être séparée de son implémentation et de sa réalisation effective. L'emploi d'une nouvelle hiérarchie et de données supplémentaires pour le contrôle du squelette nécessite une remise en question du processus de génération, tâche que nous allons présenter dans le chapitre suivant.

## Chapitre III : Cinématique inverse et intégration au système de génération automatique

Les détails du modèle présentés dans le chapitre précédent nécessitent une intégration particulière, d'une part du fait du langage de description linguistique que nous utilisons mais aussi du fait du modèle lui-même. Nous évoquons depuis le début de ce mémoire les méthodes de cinématique inverse ainsi que le problème de sous-spécification lié aux méthodes numériques. Ce chapitre s'attache à proposer une méthode de cinématique inverse permettant la gestion du modèle anatomique et plus particulièrement une intégration menant à une gestion facilitée des courbes de confort présentées en section 6.4 du chapitre précédent.

La première section présente quelques méthodes de cinématique inverse courantes. Nous évoquerons dans un second temps les problèmes généraux liés à ces méthodes et l'approche que nous souhaitons avoir pour la résolution de postures. La section 3 présente la méthode élaborée pour répondre à ces problèmes spécifiques. Enfin, en quatrième section nous présentons les différents programmes et bibliothèques développées dans le cadre de ces travaux, en accord avec les modèles du chapitre II et la méthode de génération détaillée dans ce chapitre.

## 1. Approches de cinématique inverse

Les méthodes de cinématique inverse, comme définies dans l'annexe 1, sont l'ensemble de méthodes permettant de trouver les valeurs des angles d'une chaîne cinématique afin de placer cette chaîne dans une certaine posture, répondant à un certain nombre de contraintes. Cette section fait un bref état de l'art des techniques courantes utilisées, basées sur des méthodes d'optimisation, où chaque contrainte est exprimée comme un problème dont il faut minimiser une quantification. Il s'agit de minimiser la distance entre les effecteurs et leur cible dans le cas de placements et de minimiser l'angle entre les os et l'alignement qu'on souhaite leur donner dans le cas d'orientations.

### 1.1. Exemple

Prenons un exemple simple illustrant le propos des méthodes d'optimisation. On cherche à placer le poignet droit dans l'espace à un point  $T$  de coordonnées  $(x_T, y_T, z_T)$ . On désigne la position courante du poignet par le point  $E$  (pour effecteur) de coordonnées  $(x_E, y_E, z_E)$ . Pendant toute la durée de la résolution,  $T$  restera constant alors que  $E$  variera. Sachant que l'on souhaite au final que  $E = T$  on définit la fonction  $f(E) = \|\overrightarrow{TE}\|$  qu'il s'agira de minimiser. Le placement de  $E$  dépend des valeurs d'angles de la chaîne cinématique que l'on va traiter. Pour cet exemple, elle ne sera composée que de deux degrés de liberté dont les valeurs sont représentées par le vecteur  $\theta = (\theta_1, \theta_2)$ . Puisque  $E$  dépend des valeurs de  $\theta$ , on peut définir la fonction  $g(\theta) = E$ , et donc par transitivité :  $f(g(\theta)) = \|\overrightarrow{TE}\|$ .

Le critère de minimisation dépend d'un espace d'autant de dimensions que de valeurs contenues dans  $\theta$ . Dans notre cas, l'espace de recherche des solutions est en dimension 2. Une valeur de  $f$  est assignée pour chaque point de cet espace. Pour notre exemple nous représentons l'espace de recherche sur la Figure 36.

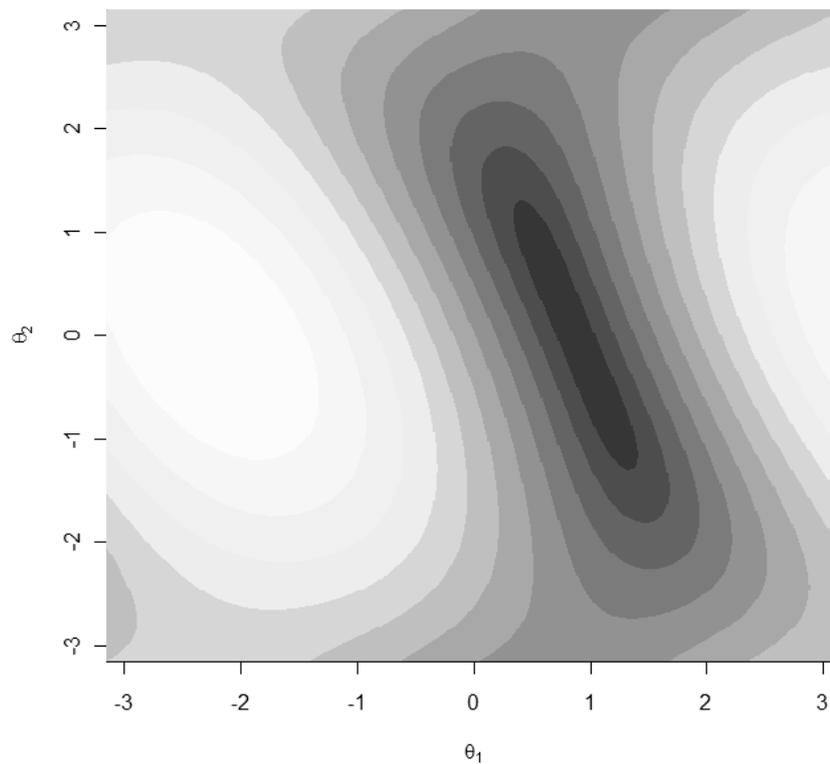


Figure 36 Exemple d'espace de recherche en deux dimensions. Les niveaux de gris correspondant au résultat de  $f$  en fonction des deux angles représentés sur les axes. En clair le maximum, en foncé le minimum à atteindre.

L'objectif de l'optimisation est de parcourir cet espace à la recherche d'un ou plusieurs minima. On appelle dans ce cadre *minimum global* le minimum de l'espace en entier. Ce point est la meilleure solution au problème de cinématique inverse, mais reste généralement difficile à trouver du fait de la taille de l'espace de recherche. Comme le montre la figure, l'espace des solutions forme un paysage dans lequel il est possible de trouver des vallées représentant des *minima locaux*, c'est-à-dire des solutions qui s'approchent du minimum sans l'atteindre. Les méthodes d'optimisation les plus simples, du fait de leur nature heuristique, s'arrêtent à la découverte d'un minimum local, donnant une solution approchée mais non nécessairement la meilleure.

L'exemple couvre une portion restreinte du domaine puisqu'il faut souvent faire face à plusieurs contraintes simultanément. Il s'agit alors de trouver une fonction permettant de quantifier la qualité d'une solution sur plusieurs objectifs (par exemple, une somme pondérée de chaque objectif séparé) ou utiliser des méthodes d'optimisation multicritères.

La suite de cette section présente des méthodes usuelles de résolution de cinématique inverse. Cette section n'a pas vocation à être exhaustive mais à présenter quelques

techniques afin de pouvoir définir une problématique adéquate en fonction des besoins du système de génération.

## 1.2. Méthodes basées sur la Jacobienne

Les méthodes les plus courantes en termes de cinématique inverse font intervenir la matrice Jacobienne du système. La matrice Jacobienne représente les dérivées premières du système pour chaque degré de liberté, c'est-à-dire la direction dans laquelle vont évoluer les effecteurs si on applique une légère variation aux valeurs des degrés de liberté. Utiliser la matrice Jacobienne inverse permet une réciproque : déterminer dans quelle direction faire varier les angles de degrés de liberté pour que l'effecteur se déplace dans la direction voulue (i.e. vers la cible).

Malheureusement le calcul de l'inverse d'une matrice quelconque est un problème n'ayant pas de solution dans beaucoup de cas concernant la cinématique inverse. On utilise par conséquent d'autres méthodes dans une optique heuristique donnant des solutions approchées mais acceptables. Ces solutions sont nombreuses, on peut noter le rapport de Buss dans (Buss 2005) qui fait l'état de deux méthodes à base de Jacobienne :

- La *Jacobienne transposée* (Wolovich et al. 1984) : au lieu d'inverser la matrice, on se contente de la transposer. Le résultat donne un résultat particulièrement rapidement (en comparaison aux autres méthodes) mais de piètre qualité. La Jacobienne transposée est une méthode de résolution généralement utilisée dans le cas de problèmes ne disposant que d'un seul effecteur.
- La *pseudo-inverse* (Buss 2003) : on utilise une forme particulière de la matrice Jacobienne dite pseudo-inverse. La pseudo-inverse est une approche des moindres carrés de l'inverse de la Jacobienne. Elle donne donc une solution approchée de l'inverse et a l'avantage d'exister quelque soit la Jacobienne. Malheureusement, la pseudo-inverse est très sensible aux singularités et ne marche de manière efficace que dans peu de cas.

Par ailleurs, le rapport de Buss fait état de deux autres méthodes appelées respectivement *damped least squares*<sup>31</sup> et *selective damped least squares*<sup>32</sup>. Ces deux méthodes, basées sur la pseudo-inverse, incluent l'utilisation d'un facteur d'amortissement permettant une plus grande stabilité à l'approche des singularités, écartant ainsi le plus gros

---

<sup>31</sup> Méthode des « moindres carrés amortis »

<sup>32</sup> Méthode des « moindres carrés amortis sélectifs »

défaut de la méthode par pseudo-inverse. Ces deux méthodes sont plus adaptées à la résolution de contraintes multiples sur une même chaîne.

### 1.3. Descente de Coordonnées Cycliques (CCD<sup>33</sup>)

La *descente de coordonnées cycliques* (Canutescu et al. 2003) est une approche gloutonne<sup>34</sup> au problème de cinématique inverse. On parcourt la chaîne en optimisant les éléments de la chaîne un par un, séquentiellement, cherchant la position permettant de rapprocher au maximum chaque effecteur de sa cible. Le processus est itéré plusieurs fois jusqu'à ce que le système se soit stabilisé. Cette méthode est particulièrement indiquée pour la résolution de problèmes ne disposant que d'un seul effecteur mais marche mal dans le cas d'une résolution multiple. Le fait que la méthode soit simple la rend attrayante, néanmoins comme pour toute méthode d'optimisation itérative, la vitesse de chaque pas d'optimisation doit être réduite pour éviter toute instabilité près des solutions, ce qui résulte en une méthode relativement lente de manière générale.

### 1.4. Exploration par Monte-Carlo

Un autre type de méthode intéressant est l'utilisation d'échantillonnage Monte-Carlo pour trouver une solution. Présentée par Courty dans (Courty et al. 2008) cette méthode génère un ensemble de postures aléatoires (valeurs d'angles aléatoires), qu'elle va faire bouger selon un mouvement brownien<sup>35</sup> dirigé de façon à trouver un optimum. Cette méthode a l'avantage de faire évoluer plusieurs solutions en parallèle contrairement aux précédentes qui se concentrent sur une seule et même solution dépendant des conditions d'initialisation du système. En revanche, la taille de l'échantillon à générer est fortement dépendante de la taille de l'espace à explorer. Plus la chaîne cinématique a de degrés de liberté et plus il faut accroître le nombre de solutions en parallèle, ce qui peut alourdir le processus.

### 1.5. Descente de gradient

La descente de gradient est une méthode permettant de « tomber » dans le minimum local le plus proche, de manière rapide et simple. Etant donné que l'on utilise la notion mathématique de gradient, cette méthode reste basée sur la Jacobienne. La descente telle que proposée dans (Héloir 2008) définit le gradient du système comme une fonction de la Jacobienne transposée et d'un vecteur d'erreur (distance entre la position courante de

---

<sup>33</sup> CCD : *Cyclic Coordinates Descent*

<sup>34</sup> Approche itérative prenant la meilleure solution locale à chaque itération.

<sup>35</sup> Description mathématique d'un mouvement aléatoire.

l'effecteur et sa cible). On corrige ainsi les angles de manière itérative en pondérant ce gradient par une fonction dépendant du vecteur d'erreur, de manière à effectuer de plus petits pas lorsqu'on se rapproche d'un optimum local. Nous donnons une illustration de descente de gradient sur la Figure 44 en section 3.3.

### 1.6. Autres méthodes et méthodes hybrides

Comme indiqué dans l'introduction, cette section n'a pas pour but de faire un état de l'art des techniques de cinématique inverse (et à plus forte raison d'optimisation). Il en existe de nombreuses variantes et seules les méthodes les plus courantes ont été décrites ici. Il est possible de mentionner d'autres techniques utilisées en génération de postures telles que la méthode du simplexe (Dave et al. 1998) ou encore le recuit simulé (Kirkpatrick et al. 1983).

Toutes ces techniques ne sont cependant pas nécessairement exclusives et il est important de noter que l'on peut combiner plusieurs méthodes. Il est notamment fréquent de combiner deux méthodes d'optimisation : une première méthode va explorer l'espace pour de potentiels optimums, et une seconde méthode va approfondir cette exploration en certains points de l'espace afin de déterminer un minimum local. Enfin, la meilleure solution est choisie comme solution du problème. Ces hybrides permettent une meilleure recherche au détriment de la vitesse de calcul.

Par ailleurs, il existe aussi des algorithmes de cinématique inverse gérant plus spécifiquement des priorités anatomiques ou des priorités de contraintes (Boulic 2003). Ces priorités permettent de déterminer, lors de multiples contraintes, quel objectif est à atteindre prioritairement ou quelle partie de la chaîne bouger en priorité.

## 2. Problématiques liées à la génération automatique de langue des signes

Du fait de la spécificité du domaine dans lequel nous nous situons ainsi que l'indique le *No Free Lunch Theorem*<sup>36</sup>, il est nécessaire de spécifier les problématiques liées à la cinématique inverse de façon à pouvoir utiliser une méthode appropriée à la résolution de contraintes. On dénote trois contraintes sur le système liées aux modèles linguistique et anatomique.

### 2.1. Placements relatifs et absolus

Le modèle linguistique utilisé, comme indiqué dans la section 2.5 du premier chapitre, définit des contraintes sur le corps permettant la construction de postures. Ces contraintes sont de deux types : placement de sites corporels dans l'espace et orientation d'os, ou de vecteurs associés à des os. Il est capital de noter que le placement de sites corporels dans l'espace de signation se passe en deux temps. Le placement d'éléments ne se fait pas de manière absolue dans l'espace mais nécessite une première définition relative à l'espace de signation. Si nous reprenons l'exemple du signe [MALADE] présenté en section 1.1 du chapitre II et rappelé en Figure 37, on va effectuer deux placements, un pour chaque main. Tout d'abord, on va placer la main gauche devant le signeur au niveau de son torse. Puis, on va placer la main droite au dessus de la main gauche. Ce repérage relatif aux éléments du corps est ensuite remplacé par une réelle évaluation des localisations, afin d'établir des coordonnées absolues de l'espace en vue d'une cinématique inverse.

---

<sup>36</sup> Le « *No free lunch theorem in search and optimization* » est une métaphore proposée par Wolpert et Macready présentant les méthodes d'optimisations comme autant de restaurants proposant le même menu et dont chaque plat dispose d'un prix différent. Choisir son restaurant pour payer le moins cher sous entend de connaître ce que l'on veut commander et les prix des différents restaurants. En optimisation, cette métaphore illustre le fait qu'aucune méthode n'est meilleure globalement et que l'emploi d'une méthode particulière nécessite de connaître le problème à résoudre (ce que l'on veut commander au restaurant) ainsi que les spécificités et points forts de la méthode utilisée (les prix du restaurant).



Figure 37 [MALADE] en LSF

## 2.2. Dépendance de contraintes

L'exemple du signe [MALADE] illustre par ailleurs un autre phénomène important lié au modèle linguistique : la notion de dépendance de contraintes. La résolution de la posture doit prendre en compte le fait que la main droite est placée en fonction de la main gauche, qui elle-même est placée dans l'espace. Il n'est donc pas question de placer les éléments contraints dans le désordre sous peine de trouver une posture invalide. Il est ainsi nécessaire, dans le cas de dépendances, de bien tenir compte des positions finales de chaque effecteur.

## 2.3. Résolution multiple

Le modèle anatomique apporte lui aussi une contrainte forte sur la méthode de résolution de cinématique inverse : l'approche du confort nécessite d'avoir plusieurs solutions pour effectuer un choix de meilleure posture. Il est donc impératif que le système de cinématique inverse permette soit le calcul de solutions en parallèle, soit le calcul rapide d'un ensemble de postures en un temps limité.

Ces trois points spécifient le cœur-même du système de résolution lié aux modèles anatomique et linguistique. La section suivante décrit la solution de cinématique utilisée pour convenir à ces critères.

### 3. Elaboration d'une méthode de cinématique inverse adaptée

La méthode de cinématique inverse que nous décrivons dans cette section se décompose en quatre étapes, détaillées dans les sous-sections suivantes :

1. Le calcul des objectifs du système, lequel définit la position précise de chaque effecteur dans l'espace ainsi que les cibles du problème.
2. Une phase exploratoire lors de laquelle on détermine un grand nombre de solutions aléatoirement.
3. Pour un ensemble de solutions de la phase exploratoire, une passe d'optimisation de manière à améliorer localement chaque solution.
4. Chaque solution se voit attribuer un score de confort en accord avec le modèle présenté en section 6.4 du chapitre II. En sortie de cette dernière phase, une seule et unique solution est choisie comme résultat de la cinématique inverse.

Le processus dépend des paramètres détaillés en Annexe 3. Il est résumé sur la Figure 38.

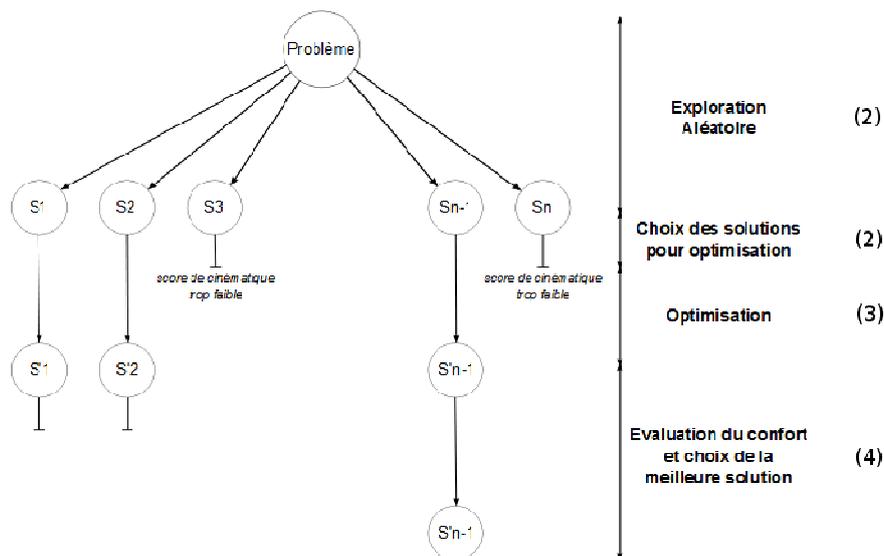


Figure 38 Processus de résolution. La première phase (définition des objectifs) n'est pas indiquée sur le schéma. Elle se déroule en amont de la résolution du problème par le système.

#### 3.1. Définition des cibles

La première tâche est de déterminer, à partir des contraintes du modèle linguistique, le positionnement souhaité des articulateurs dans l'espace. Pour cela, on établit le graphe orienté des dépendances des contraintes (cf. Figure 39). Rappelons que les contraintes sont données simultanément au système par la description issue du modèle linguistique. Pour autant, certaines contraintes peuvent être interdépendantes, nécessitant donc de prendre

en compte un certain ordre dans la résolution. Une contrainte est dépendante d'une autre si et seulement le bloc terminal auquel elle est dépendant d'un autre bloc terminal contraint. Par exemple, si une posture comporte deux contraintes (C1 et C2), telles que : C1 place la main gauche sur la main droite et C2 oriente l'avant bras droit à plat, vers l'avant, alors nous notons que C1 dépend du positionnement de la main droite, qui lui-même est dépendant de l'orientation du bras droit contraint par C2. C1 dépend donc de C2.

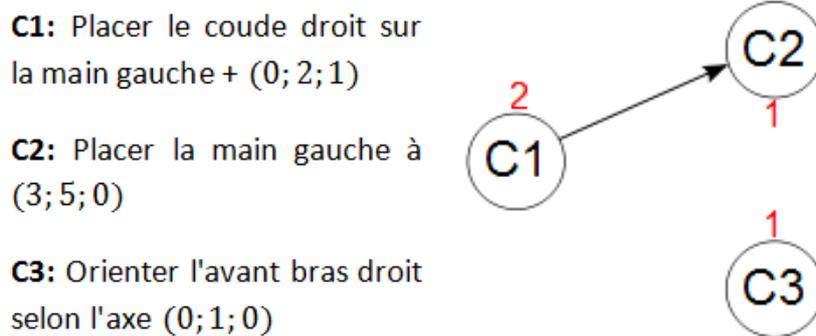


Figure 39 Exemple de résolution de contrainte. Gauche : les contraintes ; Droite : le graphe de dépendance et à côté de chaque contrainte, l'ordre de résolution

Chaque contrainte représente un nœud du graphe et on définit une arête entre des contraintes  $C_i$  et  $C_j$  si  $C_i$  dépend de  $C_j$ . Afin de définir les cibles effectives du système, on parcourt le graphe en commençant par les nœuds ayant un degré sortant nul. Pour chaque nœud parcouru, on calcule l'objectif de la contrainte en fonction des nœuds parcourus précédemment. Cette méthode de calcul des objectifs a été présentée dans (Filhol et al. 2006).

Dans l'exemple donné en Figure 39, le signe que l'on souhaite générer comporte trois contraintes. La construction du graphe se fait comme expliqué : on crée trois nœuds correspondant aux contraintes. La contrainte  $C_1$  nécessite la position de la main gauche pour être évaluée. Or, la main gauche est placée par la contrainte  $C_2$ .  $C_1$  dépend donc de  $C_2$ . On détermine en premier les nœuds de degré sortant nul :  $C_2$  et  $C_3$ .  $C_3$  est une contrainte d'orientation dont la cible ne nécessite aucun calcul.  $C_2$  ne dépend d'aucune contrainte donc ne nécessite pas non plus de calcul. Grâce à la contrainte  $C_2$ , nous savons désormais que la main gauche sera placée en  $(3;5;0)$ . On calcule ensuite les nœuds qui dépendent directement des contraintes précédentes :  $C_1$ .  $C_1$  place la main droite en un point de coordonnées  $P_{main\ gauche} + (0;2;1)$ . On prend la position de la main gauche à partir de la contrainte  $C_2$ , ce qui nous donne un placement de la main droite au point de coordonnées  $(3;7;1)$ .

Une fois que la cible de chaque contrainte est calculée, on procède à la première phase de cinématique inverse : l'exploration de l'espace des solutions.

### 3.2. Phase exploratoire

La plupart des méthodes de cinématique inverse se basant uniquement sur une optimisation sans phase exploratoire reposent sur la condition initiale du système, à partir de laquelle l'optimisation va commencer. La condition initiale peut influencer très fortement le résultat. Ainsi une mauvaise solution n'est pas nécessairement la conséquence d'une méthode de piètre qualité, mais peut être le résultat d'une mauvaise initialisation du problème. La phase exploratoire permet de contourner cet inconvénient en répondant au problème de résolution multiple énoncé en section 2.3 : cette phase détermine plusieurs conditions initiales valables à partir desquelles l'optimisation est effectuée.

### 3.2.1. Calcul des conditions initiales

Puisque l'objectif de la méthode de cinématique inverse est une optimisation locale, on souhaite obtenir les solutions les plus écartées possibles, afin de couvrir au maximum l'espace de recherche. On génère donc chaque condition initiale, comme une posture aléatoire. Chaque degré de liberté du squelette est déterminé aléatoirement sur l'intervalle de ses dépendances articulaires, puis calibré de manière à prendre en compte les distributions et dépendances dynamiques.

### 3.2.2. Distance entre deux solutions et éligibilité des solutions

Pour déterminer l'ensemble  $S$  de  $N$  solutions, on cherche à maximiser la distance respective entre les solutions de manière à couvrir un maximum l'espace de recherche. On définit la *distance* entre deux blocs fonctionnels comme étant l'angle séparant deux transformations locales. Chaque transformation locale est définie par un quaternion unitaire combinant les degrés de liberté d'un bloc. La distance entre deux quaternions  $q_1$  et  $q_2$  est donnée par l'angle séparant les deux transformations d'un même vecteur directeur :

$$\begin{aligned}v &= (1, 0, 0) \\t_1 &= q_1^* v q_1 \\t_2 &= q_2^* v q_2 \\dist(q_1, q_2) &= \cos^{-1}(t_1 \cdot t_2)\end{aligned}$$

La distance entre deux blocs est ramenée à la distance entre leurs transformations locales. Afin de déterminer la distance entre deux postures, il convient d'établir une norme permettant le calcul des distances angulaires des deux chaînes cinématiques. Pour deux postures  $p_1$  et  $p_2$  comprenant chacune  $n$  blocs (et donc  $n$  transformations locales) notées :

$$\begin{aligned}p_1 &= (q_{11}, q_{12}, \dots, q_{1n}) \\p_2 &= (q_{21}, q_{22}, \dots, q_{2n})\end{aligned}$$

On détermine la distance entre les deux postures par une norme euclidienne des distances de chaque bloc :

$$d(p_1, p_2) = \sqrt{dist(q_{11}, q_{21})^2 + dist(q_{12}, q_{22})^2 + \dots + dist(q_{1n}, q_{2n})^2}$$

L'ensemble de  $S$  est rempli au fur et à mesure que sont découvertes des solutions éligibles de meilleure qualité. Une posture est dite *éligible* si sa distance à toutes les postures de l'ensemble est inférieure à un seuil  $\varepsilon$  fixé empiriquement en fonction du nombre de blocs à traiter.

### 3.2.3. Score d'une solution

Afin de comparer les différentes solutions il est nécessaire d'établir un score de qualité. Le score que nous nous proposons d'utiliser est d'abord calculé sur chaque contrainte, puis combiné en un score général :

- Dans le cas d'un placement d'effecteur, le score est fonction de la distance entre le bloc à placer et la cible où l'on souhaite le placer. Nous choisissons une fonction exponentielle inverse de façon à obtenir un score sur l'intervalle [0 ; 1] où 1 représente un placement parfait :

$$S_{placement} = e^{-\|pos-cible\|}$$

- Dans le cas d'une orientation, le score est le produit scalaire de l'angle formé entre le bloc et l'orientation souhaitée. Il est ensuite ramené sur l'intervalle [0 ; 1] :

$$S_{orientation} = (1 + (dir \cdot cible))/2$$

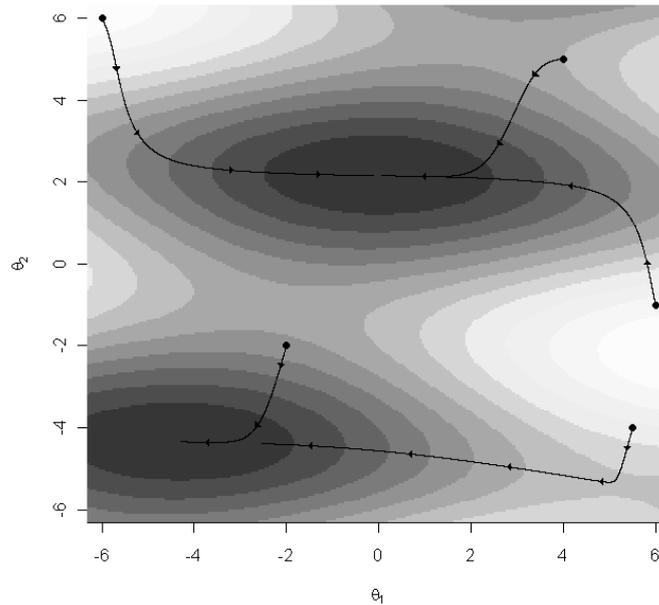
Les scores ne sont associés qu'à une contrainte. Un problème de cinématique regroupant plusieurs contraintes, le score total d'une posture donnée par cinématique inverse est le résultat d'un produit de tous les scores obtenus :

$$Score\ final = \prod_{s \in scores} s$$

Nous choisissons un produit pour obtenir un score tendant très rapidement vers zéro dans le cas où un élément de la posture ne respecte pas les contraintes. Ce score est utilisé à la fois pour la méthode exploratoire et la phase d'optimisation que nous allons maintenant détailler.

### 3.3. Optimisation locale des solutions

Des  $N$  solutions trouvées lors de la phase exploratoire, on ne garde que les  $M$  meilleures, en comparant leurs scores de cinématique inverse. Sur chacune de ces solutions, on effectue une passe d'optimisation locale afin de l'améliorer autant que possible. Cette optimisation peut être faite à partir de n'importe quelle méthode d'optimisation locale. Nous utilisons une simple descente de gradient de manière à ce que les  $M$  solutions tombent dans les minima locaux les plus proches (cf. Figure 40).



**Figure 40** Multiples résolutions par descente de gradient. Les points représentent les positions de départ déterminées par la phase d'exploration.

La descente de gradient utilisée est la même que celle présentée dans (Héloir 2008). Cette méthode optimise itérativement les solutions en corrigeant les degrés de liberté progressivement, de manière à suivre le gradient des solutions jusqu'au minimum local le plus proche. Chaque itération modifie les valeurs d'angles des degrés de liberté du squelette et est donc soumise à une correction par le modèle anatomique, notamment par le système de dépendances articulaires présenté en section 4 du précédent chapitre. L'algorithme arrête les itérations si le score dépasse un seuil limite proche de 1. Dans ce cas, on estime que la cinématique inverse a réussi et a correctement placé/orienté les effecteurs.

### 3.4. Sélection du résultat par le modèle de confort

La dernière étape de la génération de posture consiste à choisir parmi les postures optimisées la plus naturelle. Ce choix passe par la comparaison des scores de confort issus du modèle anatomique détaillé en section 6.4 du chapitre II. L'attribution de ces scores peut être considérée selon deux approches :

- Soit nous partons de l'hypothèse que *toutes les solutions optimisées sont qualitativement égales*, auquel cas on se contente d'utiliser un simple score de confort.

- Soit nous souhaitons garder une trace du score de cinématique inverse, auquel cas le score final devient une combinaison du score de confort et de celui de cinématique inverse.

Nous choisissons la deuxième solution, de manière à être sûr que le résultat obtenu corresponde à la fois aux critères linguistiques (cinématique inverse) et anatomiques (modèle de confort). Le score est tiré des courbes de confort présentées précédemment. Etant donnée une série de valeurs d'angles sur le squelette, il est calculé grâce à un produit des probabilités sur la chaîne cinématique :

$$Confort = \prod_{i \in \text{degrés de liberté}} Confort_i$$

Une fois ce score général obtenu, on doit le combiner avec le score obtenu par le système de cinématique inverse. On calcule le score final en effectuant une combinaison linéaire des deux scores.

$$Score = \omega_{confort} \cdot Confort + \omega_{cinématique} \cdot Score_{cinématique}$$

Où  $\omega_{confort}$  et  $\omega_{cinématique}$  sont des pondérations scalaires utilisées pour la combinaison.

### 3.5. Limites de l'approche

Deux difficultés résident dans l'établissement du score final de résolution :

- la normalisation du score de confort, qui a une forte variabilité sur [0 ; 1] alors que le score de cinématique inverse est généralement proche de 1 à ce stade ;
- la pondération de la combinaison linéaire.

En ce qui concerne la normalisation du score de confort, il est important de comprendre que le score final sera très variable d'un problème à un autre, du fait des articulations mises en jeu. En effet, si la chaîne articulaire varie, alors le score de confort change d'ordre de grandeur. En revanche, pour un même problème, les différents scores attribués aux solutions sont comparables entre eux, car traitant la même chaîne articulaire.

La pondération est un problème qui semble particulièrement épineux du fait de l'intentionnalité du mouvement. Le placement géométrique doit primer sur le confort afin que les postures générées répondent aux contraintes linguistiques. Cependant, donner trop d'importance au placement atténue l'effet du confort, et pour deux solutions envisageables

une moins confortables peut être sélectionnée à cause de son score de cinématique inverse plus élevé. Il est donc nécessaire de calibrer cette pondération, au moyen d'expérimentations évaluant la qualité des résultats en fonction des indices de pondération.

Ces deux difficultés seront abordées dans de futurs travaux, nécessitant une étude à part entière, non incluse dans cette thèse.

## 4. Ingénierie des modèles

Afin de pouvoir évaluer les modèles présentés au chapitre II, ainsi que le système de génération présenté dans les sections précédentes, nous avons réalisé deux bibliothèques et deux applications.

### 4.1. Bibliothèque de squelette

La gestion du modèle anatomique présenté en chapitre II est intégrée à une bibliothèque développée en C/C++. Elle est utilisable comme une API par d'autres applications et permet la manipulation du squelette à deux niveaux :

- Manipulation bas niveau : contrôle direct des blocs, positionnement, orientation.
- Manipulation haut niveau : contrôle du squelette, cinématique directe, chargement et sauvegarde de données à partir d'un fichier XML, stockage d'animations.

Cette bibliothèque est la brique de base utilisée par tous les systèmes décrits pas la suite.

### 4.2. Bibliothèque de séquences

Cette bibliothèque permet l'interface entre le modèle linguistique Zebedee et les autres modules du système. De même que la bibliothèque du squelette, elle est utilisable en tant qu'API et représente les données des signes sous la forme de structures. Elle fournit à l'utilisateur quelques outils de base pour la manipulation des signes, mais n'est pas supposée être utilisée pour la modification des données, simplement leur lecture et leur intégration à une structure adaptée au C/C++.

### 4.3. Système de génération

Le système de génération, basé sur la méthode de cinématique décrite dans ce chapitre est une brique logicielle de l'environnement de génération. Développé en C/C++ il s'appuie sur les deux bibliothèques décrites précédemment pour déterminer des ensembles de contraintes à résoudre sur le squelette et pour les résoudre. Le module de génération est un programme en ligne de commande entièrement paramétrable, de façon à faciliter par la suite les futurs travaux d'optimisation du système. Ce système est basé sur une architecture modulaire permettant de définir des classes C++ de cinématique inverse correspondant à la phase exploratoire et à la phase d'optimisation. Dans un souci d'ouverture à de futurs

travaux, ces deux classes sont modulables à volonté et peuvent être paramétrées à la volée, de façon à permettre une adaptation aux différents problèmes si nécessaire<sup>37</sup>.

#### 4.4. Gestionnaire et connexion à l'environnement du LIMSI

Un gestionnaire en C++ a été développé en même temps que le modèle de squelette. Ce gestionnaire comporte une interface visuelle de la bibliothèque de squelette. L'interface développée grâce à la librairie Qt permet une manipulation du squelette en trois dimensions, la définition des blocs et sites corporels, ainsi que la visualisation des dépendances et des courbes du modèle de confort (cf. Figure 41). Par ailleurs, il est possible de charger des animations au format BVH pour les calquer sur un squelette déjà chargé.

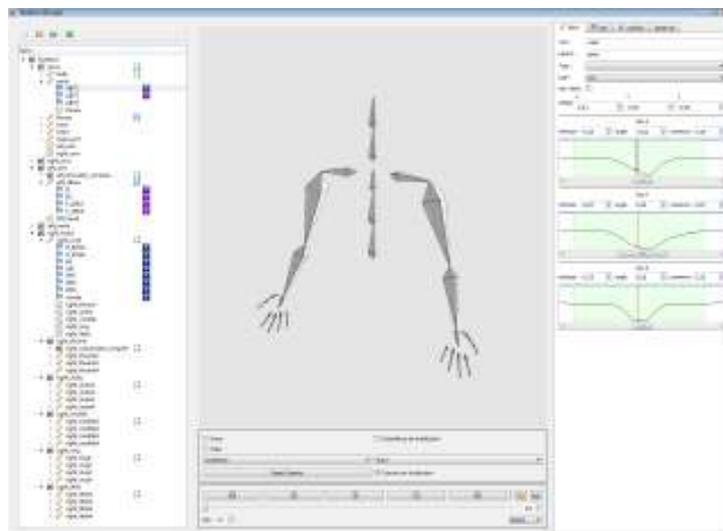


Figure 41 Gestionnaire permettant la manipulation du modèle anatomique, ainsi que la résolution des signes.

Ce gestionnaire permet en outre une connexion au système de génération et à une base de données de signes développée au LIMSI. Il est possible au travers de ce gestionnaire d'avoir accès à une liste de signes définis grâce au modèle Zebedee, de visualiser leur contenu et d'appeler directement le système de génération, de résoudre les différentes postures, de les stocker dans un fichier et de charger directement l'animation résultante sur le squelette. Ainsi, le gestionnaire permet l'intégration de tous les travaux présentés dans cette thèse.

<sup>37</sup> Nous revenons sur cette notion dans la partie « perspectives » : chapitre IV, section 2.5.3

## 5. Conclusion du chapitre

Nous avons vu dans ce chapitre un processus de génération de postures pour la langue des signes. Ce processus a été élaboré à partir de méthodes déjà existantes (Monte-Carlo et Descente de gradient), puis adapté au modèle de squelette présenté en chapitre II afin de répondre précisément aux besoins de l'étude. Chaque posture est déterminée grâce à une exploration aléatoire, une optimisation locale et enfin une sélection de la meilleure posture en fonction du modèle de confort détaillé dans le chapitre précédent. Nous présentons maintenant dans un dernier chapitre les résultats de ces générations ainsi que l'évaluation des différents modèles.

## Chapitre IV : Résultats et discussions

Ce chapitre décrit les différents résultats obtenus grâce au modèle anatomique présenté en chapitre II ainsi que son implémentation présentée en chapitre III. Chacune des sections suivantes présente les résultats, les discute et donne des perspectives à ces travaux. La première section présente les résultats liés au modèle de pouce opposable. La seconde ceux du système de génération et du modèle de confort.

## 1. Modèle de pouce opposable

Cette section détaille les résultats du modèle de pouce présenté en section 5 du chapitre II. La section 1.1 présente les résultats quantitatifs du modèle. La section 1.2 présente les apports du modèle pour la synthèse de configurations manuelles, et la dernière section (1.3) donne les perspectives du modèle.

### 1.1. Résultats

L'intégration de ce modèle de pouce au système permet la génération de configurations complexes de la main. La confrontation de ce modèle à l'expérimentation, c'est-à-dire à la génération de différentes configurations manuelles, s'est déroulée de manière objective sur l'intégralité des configurations statiques de la dactylogogie<sup>38</sup> ainsi que sur des configurations courantes utilisées dans les signes en LSF.

Le corpus d'évaluation comporte 22 configurations de mains dont 14 placent le pouce en opposition. Du fait de la nature heuristique de la cinématique inverse, chaque signe est résolu 10 fois. Pour chaque contrainte de placement imposée par le modèle, le score de cinématique est obtenu par les calculs présentés au chapitre précédent. Plusieurs scores peuvent être obtenus pour une même configuration en fonction du nombre de contraintes posées sur le pouce.

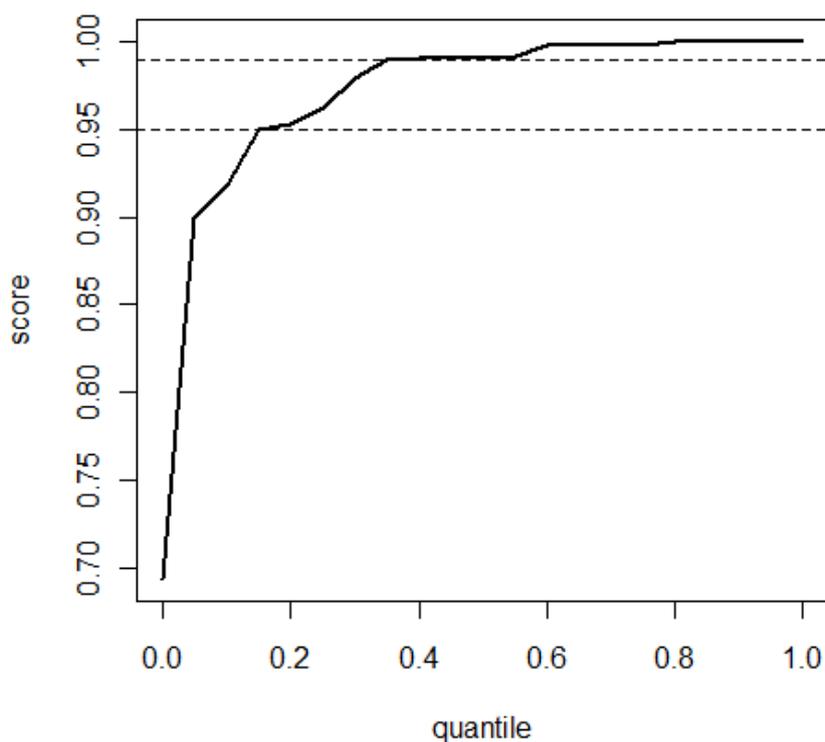
Au total, le corpus présente 390 contraintes sur le pouce et autant de scores en résultat d'évaluation. Les contraintes peuvent être de deux natures : placement ou orientation. Les scores donnés dépendent du type de contrainte comme indiqué en annexe 1.

---

<sup>38</sup> L'alphabet dactylogique est le code gestuel représentant les lettres latines. Chaque lettre est représentée par une configuration manuelle et peut être accompagnée d'un mouvement.

Les résultats de cette évaluation donnent des scores répartis entre 0.694 et 1. L'observation des quantiles (cf. Figure 42) montre que 90% des scores obtenus dépassent 0.92 et que 70% des scores dépassent 0.98<sup>39</sup>. La distribution des quantiles donne une évaluation positive de la qualité du modèle de pouce.

Figure 42 Distribution des quantiles des scores de résolution



La Figure 43 détaille la répartition des scores en fonction des tâches assignées. Des 22 signes générés, 17 disposent d'une médiane dépassant le seuil fixé de 0.99. Les tâches impliquant l'opposition du pouce (D, F, G/Q, H, I, K/P, N, O, R, S, T, U, V, W et X) sont toutes résolues avec un score minimum de 0.98. Les scores les pires sont attribués à l'extension du pouce en position latérale. La présence de ces scores faibles est due à une dépendance articulaire du scaphoïde trop contraignante, ce défaut sera pallié dans de futurs travaux.

<sup>39</sup> Ce score représente une différence de moins de 15 degrés entre l'orientation courante et l'objectif.

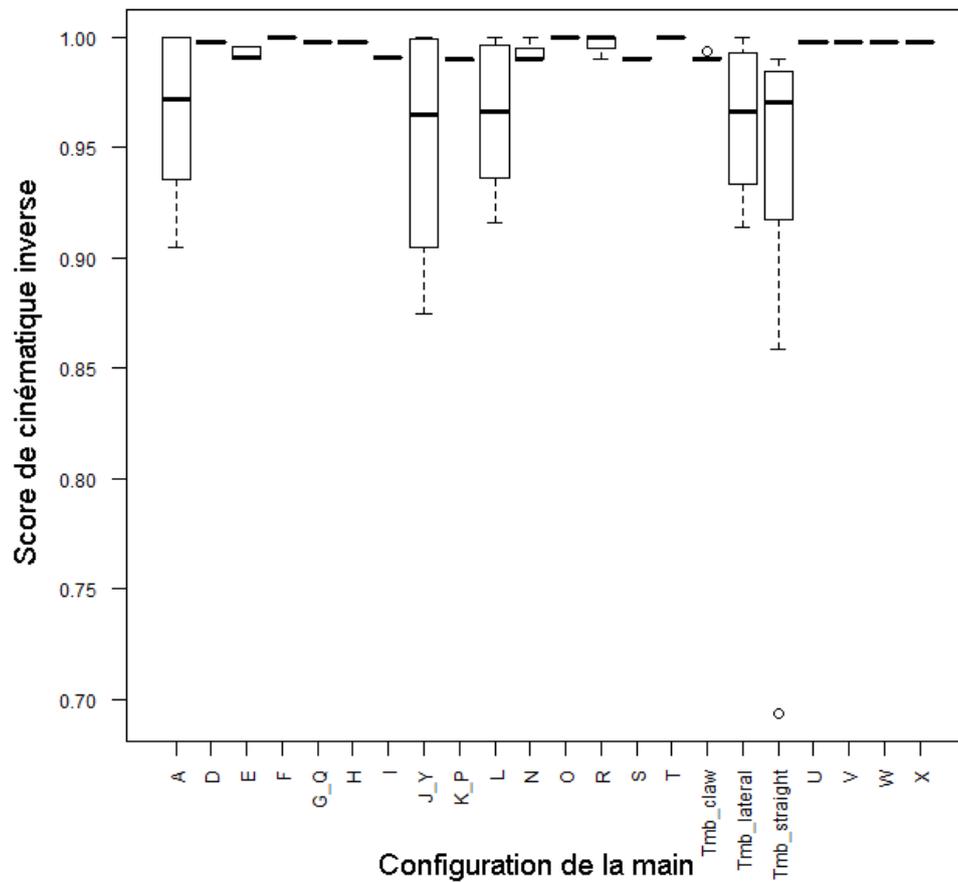


Figure 43 Distribution en boîte à moustache des scores de résolution sur les contraintes du pouce

De ces résultats, nous déduisons que le nouveau modèle permet une résolution efficace des problèmes de placement du pouce en opposition.

La Figure 44 recense quatre exemples de résolutions du pouce. Les 3 premiers (E, I et U) montrent l'opposition correcte du pouce. La dernière (A) montre la difficulté mentionnée précédemment à placer le pouce en position latérale.

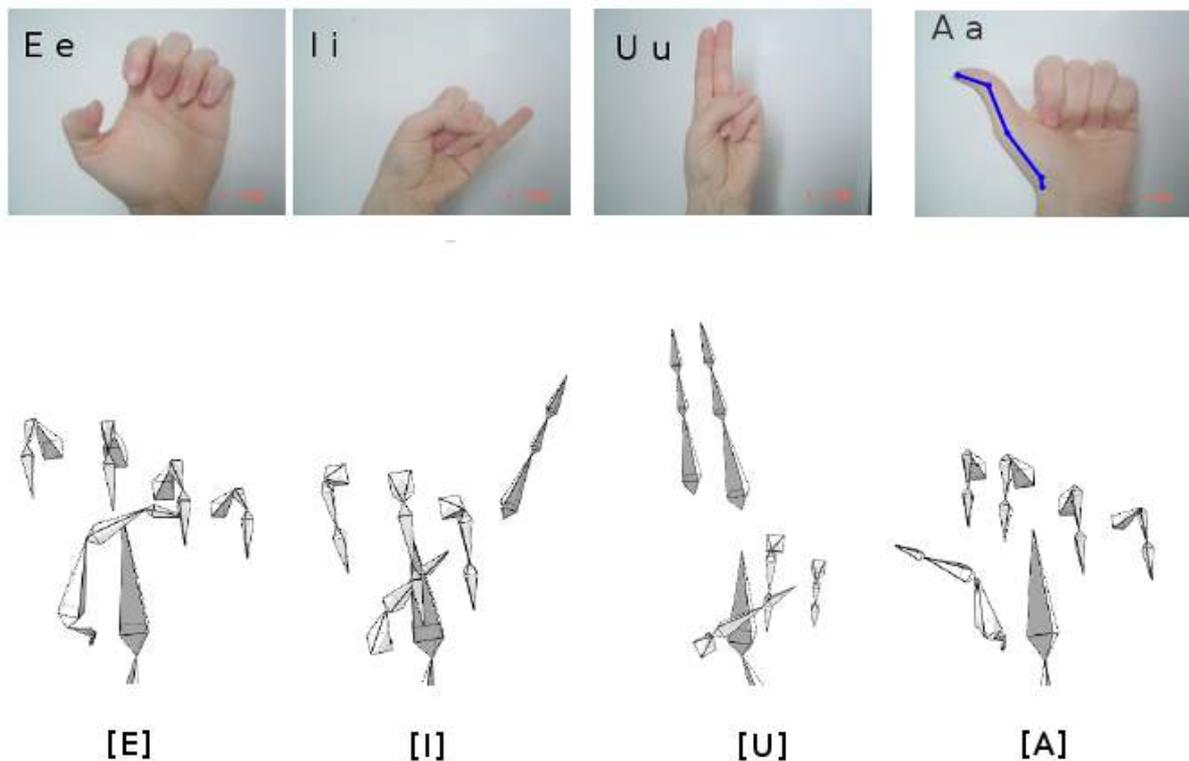


Figure 44 Exemples de résolution de configurations du pouce. La configuration de [E] pour le placement du pouce est relativement libre, laissant le choix pour la flexion du métacarpe.

## 1.2. Apports

Ce modèle de pouce ne comporte que quatre degrés de liberté : deux pour le CMC (rotation + flexion), et un pour chaque articulation inter-phalangienne. Un système disposant d'un faible nombre de degrés de liberté permet au processus de cinématique inverse de procéder rapidement et efficacement. Pour autant, le faible nombre de degrés de liberté ne limite pas de manière superflue le système, au contraire, les contraintes additionnelles imposent la rotation du pouce pour atteindre certaines configurations de la main (cf. Figure 48), ajoutant ainsi au réalisme des configurations manuelles générées.

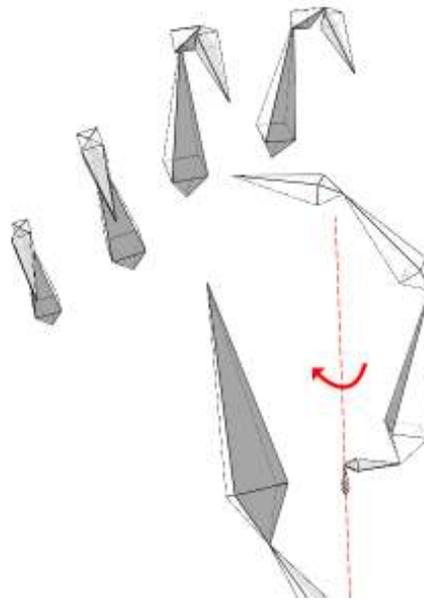


Figure 45 Axe de rotation virtuel du pouce

### 1.3. Perspectives

Plusieurs améliorations au modèle peuvent être proposées pour des études futures. Premièrement, l'axe virtuel de rotation du CMC est déterminé de manière totalement empirique et mériterait une calibration plus précise pour un meilleur réalisme. Par ailleurs, les contraintes trop fortes sur le scaphoïde, qui empêchent la latéralisation, ont besoin d'être mieux étudiées afin de permettre une amplitude de mouvement correcte autorisant à la fois l'opposition du pouce et sa latéralisation. Ces deux calibrages pourraient être étudiés plus en détail au moyen de données de captures de mouvement centrées sur la main. Une méthode alternative pourrait se baser sur d'autres études biologiques dirigées sur l'élaboration d'un modèle virtuel comme nous l'avons fait. Cependant, du fait de l'hybridation entre le modèle virtuel que nous proposons et le réel fonctionnement biologique de la main, cette solution semble être limitée à une simple approche empirique comparable à celle que nous avons présentée.

Comme nous l'avons vu dans le chapitre précédent, la définition d'un axe virtuel pour l'opposition du pouce n'est en fait qu'un palliatif au fait que nous ne puissions pas modéliser le comportement de surfaces articulaires avec un modèle centré sur des degrés de liberté et des axes de rotation. Il est cependant possible d'envisager une refonte totale du modèle de squelette, ne considérant plus le squelette comme un arbre/graphes d'os n'étant que des repères géométriques, mais comme de véritables os. Ces approches, déjà utilisées dans le domaine médical (Chao 2003) sont cependant coûteuses. Il s'agirait donc d'établir un

modèle simplifié de surface articulaire permettant une simulation plus fine et plus précise mais utilisable dans un cadre de cinématique inverse. Par ailleurs, cette approche particulièrement complexe serait *a priori* limitée à l'animation d'humanoïdes.

## 2. Système de génération et modèle de confort

Il n'est pas possible en l'état de juger séparément le système de génération et le modèle de confort. Ces deux éléments sont fortement liés et le calcul du score final dépend des deux. Quelques méthodes objectives d'évaluation de ce genre de système sont présentées dans le chapitre des perspectives.

Cette section présente et discute les résultats obtenus sur la génération et le modèle de confort. La première section (2.1) présente une étude préliminaire, sur la génération de signes. La seconde section (2.2) présente la méthode d'expérimentation, la troisième (2.1) discute la pertinence du modèle de confort. Les deux sections suivantes et présentent des résultats de génération (2.2) et de calibration grâce au modèle de confort (2.3). Enfin, les deux dernières sections et présentent les apports (2.4) et les perspectives du système (2.5).

### 2.1. Etude préliminaire

Nous disposons d'une base de 1753 signes de la LSF décrits à l'aide du modèle linguistique Zebedee. Chacun de ces 1753 signes a été résolu avec succès par le système de génération. La réussite de cette génération nous permet à la fois de confirmer que les descriptions faites grâce au modèle linguistique Zebedee donnent des systèmes de contraintes que le système arrive à générer, mais aussi que le système de génération est capable de résoudre tous les signes définis en Zebedee.

### 2.2. Protocole

Pour faire cette étude, nous choisissons sept signes répondant à deux critères. Tout d'abord, nous n'avons choisi que deux signes bimanuels ([ARBRE] et [TABLE]). Considérant que le système, capable de générer sur un bras sera capable sur l'autre, il n'est nécessaire d'évaluer les signes bimanuels que dans un seul cas : celui où un côté va dépendre de l'autre. C'est le cas de ces deux signes :

- pour [ARBRE], il est nécessaire de placer la main gauche sous le coude droit ;
- pour [TABLE] les deux mains sont placées en contact lors de la première posture, puis écartées.

Un second critère pour le choix des signes est celui de l'évaluation du confort. Nous avons montré dans la section précédente l'utilité du modèle de pouce et de son apport. Nous avons donc choisi en priorité des signes basés sur des postures faisant plus intervenir les épaules et les coudes. Les sept signes choisis sont illustrés en Figure 46.

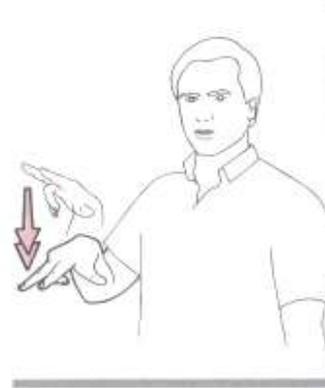
La résolution étant basée sur une phase d'exploration aléatoire, nous avons procédé à plusieurs groupes de résolution. Chaque signe à été résolu 100 fois. Un score de cinématique est déterminé pour chaque contrainte, alors qu'un score de confort n'est établi que sur les articulations concernées (tout sauf les mains). Au total, le corpus d'évaluation dispose de 19 813 scores.



[TABLE]



[MIROIR]



[APPARTEMENT]



[FROID]



[ARBRE]



[A MOI]



[20]

Figure 46 Les sept signes utilisés pour l'étude

## 2.1. Pertinence du modèle de confort

Avant d'évaluer la qualité des postures générées grâce au modèle de confort, il est intéressant de se poser la question de l'apport du modèle et de son lien avec la cinématique inverse. Il est par exemple possible que le modèle de confort donne des résultats liés à la cinématique inverse. Sous la forme d'une hypothèse formulée comme suit : « *Existe-t-il un lien fonctionnel entre la cinématique inverse et le confort ?* ». S'il est possible d'affirmer qu'un score élevé de cinématique inverse implique un score élevé de confort, alors l'utilité même du modèle de confort sera remise en question. Il est possible d'étudier cette hypothèse en confrontant les scores de cinématique inverse à ceux de confort (cf. Figure 47), ainsi qu'en étudiant la possibilité d'une régression linéaire. Cette étude est nécessaire pour affirmer l'utilité du modèle de confort. Pour autant, nous ne pouvons la présenter avant, puisque cette étude se base sur les résultats de confort pour établir un lien potentiel entre ce dernier et le score de cinématique inverse. Afin de juger de la pertinence d'un modèle de régression linéaire, nous calculons le coefficient de corrélation linéaire de Pearson (où  $x_i$  figure la série de scores de cinématique,  $y_i$  celle de scores de confort et  $N$  la taille des échantillons) :

$$r_p = \frac{\sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}$$

La conjecture indiquée par la Figure 47 est confirmée par l'analyse de ces données qui nous donne un coefficient de corrélation de 0.0113. Ce coefficient est particulièrement faible (proche de 0) et implique l'absence de relation linéaire entre le confort et la cinématique inverse. Il est envisageable qu'une relation d'un autre type existe entre le confort et la cinématique inverse, mais la figure indique tout de même une répartition sans réelle structure.

En conséquence, nous considérons qu'il n'y a aucun lien entre les deux variables et qu'il est pertinent d'utiliser un modèle de confort pour améliorer la qualité des solutions générées. Il nous reste maintenant à démontrer l'apport du modèle de confort sur la génération de postures et à estimer la qualité générale du système de génération.

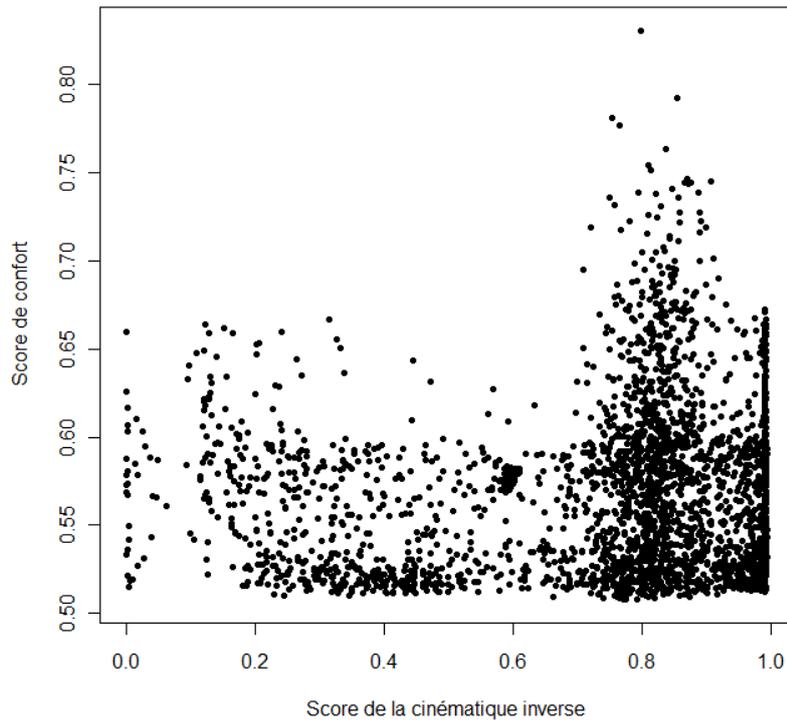


Figure 47 Répartition des scores de cinématique inverse en fonction des scores de confort

## 2.2. Résultats de génération

Comme pour l'évaluation du modèle de pouce, une possible estimation de la qualité du système de cinématique inverse passe par une étape statistique.

L'analyse des quantiles permet d'estimer la répartition des scores de cinématique inverse et de déterminer si un résultat à une solution peut être trouvé aisément ou non. Rappelons que le processus de cinématique inverse utilisé et présenté dans le chapitre précédent fonctionne sur des chaînes longues, comportant jusqu'à 20 degrés de liberté. La répartition des quantiles est donnée en Figure 48).

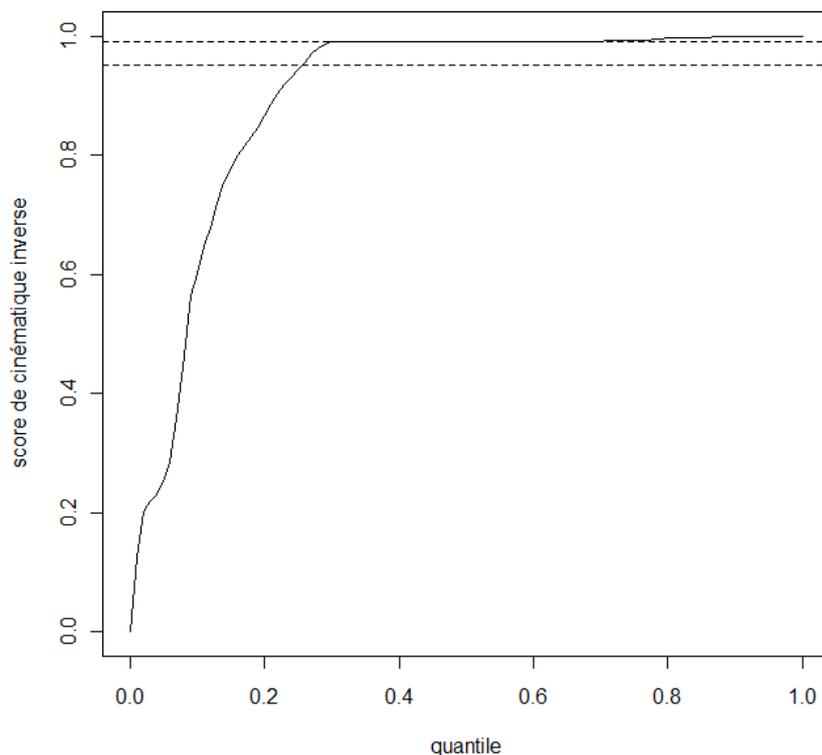


Figure 48 Répartition des quantiles des scores de cinématique inverse. Deux lignes figurent les scores 0.95 et 0.99

Le seuil d'acceptabilité des postures, comme indiqué dans le chapitre précédent, est fixé à 0.99. Ce score est atteint par 70% des résultats. En outre, 75% passent le cap des 0.95 ce que l'on peut considérer comme un score acceptable. Seuls 10% des scores ont un résultat inférieur à 0.50, ce que l'on attribue en partie à une définition linguistique trop exigeante (contraintes d'orientation trop fortes) ou à des défauts du système de cinématique inverse (mauvais tirage aléatoire par exemple).

Etant donné que la convergence de la descente de gradient ne peut dépasser un score de 0.99, on peut évaluer la proportion des postures trouvées directement par la phase d'exploration. Celle-ci est à peu près de 42% des solutions. Au total, 42% des solutions sont déduites immédiatement, et 33% tombent dans un optimum local satisfaisant ( $>0.99$ ) après descente de gradient.

### 2.3. Résultats du confort

Il n'existe pas, à l'heure actuelle, de méthode objective pour évaluer la qualité d'une posture. Bien que des perspectives à long terme puissent être envisagées, l'évaluation actuelle des signeurs virtuels est essentiellement subjective.

L'évaluation du modèle de confort se fait par l'observation des postures générées et la confrontation visuelle entre les solutions de moins bon confort et celles de meilleurs scores. Quelques exemples de confrontation suivent et permettent d'indiquer l'amélioration des postures par le modèle de confort. Dans la suite de cette section, nous rappelons que les scores de confort sont obtenus par le produit d'une ou plusieurs probabilités. Etant donné que le nombre de produits dépend de la taille de la chaîne, la valeur absolue des scores de confort n'a en soi pas de réel intérêt puisqu'elle est spécifique de la chaîne. En revanche, la relation d'ordre entre les différents scores est capitale. Ainsi un score proche de 0 n'est pas indicatif d'une mauvaise solution tant qu'il est supérieur à tous les autres scores issus de la même génération.

### 2.3.1. [ARBRE]

La posture prise par le squelette pour le signe [ARBRE] pose deux contraintes de placement (main gauche et coude droit) et trois d'orientations (les deux mains et l'avant bras droit). Voici deux exemples de générations avec deux scores de confort différents :

Notons par ailleurs que le système de génération ne gère pas les collisions. C'est pourquoi sur l'image de gauche, la main du squelette traverse l'avant-bras opposé. Le confort permet, pour un même jeu de contraintes de placer le poignet de manière correcte. L'effet du confort se voit particulièrement sur la configuration des deux clavicules et des épaules. Dans le cas où le confort est bas, les clavicules sont levées et les bras repliés, alors que sur la position avec confort, ces derniers sont dans une posture paraissant plus détendue.



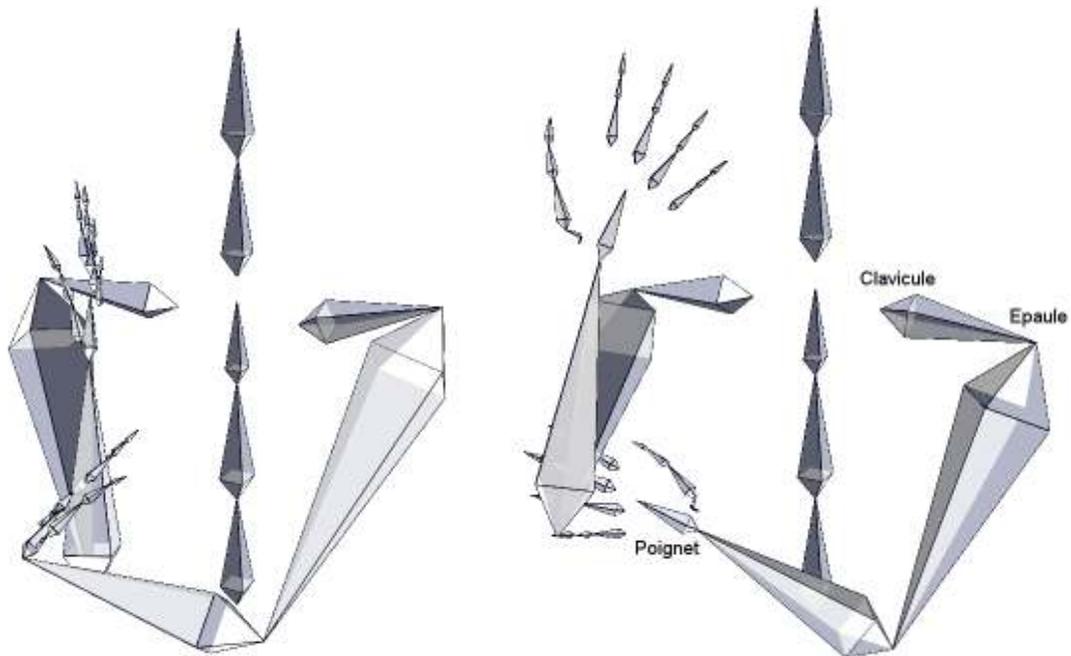
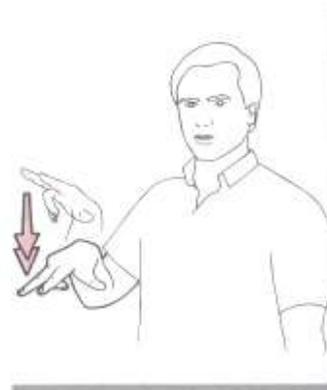


Figure 49 Génération du signe [ARBRE]. Scores de confort : 0.545 à gauche, 0.603 à droite.

### 2.3.2. [APPARTEMENT]

Le signe [APPARTEMENT] présente une configuration particulière de la main et place le poignet dans l'espace. Il s'agit d'un signe mono-manuel ne faisant intervenir que le bras droit. Les deux exemples qui suivent montrent une nouvelle fois la différence entre un score de confort faible et un score de confort élevé sur le squelette :



[APPARTEMENT]

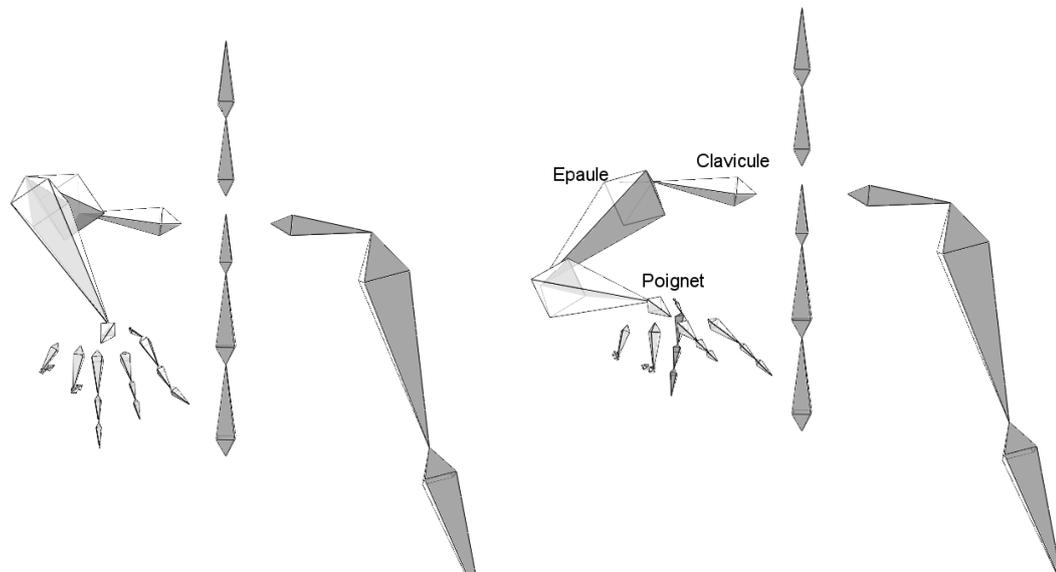


Figure 50 Génération du signe [APPARTEMENT]. Scores de confort : 0.515 à gauche, 0.632 à droite.

Comme pour le signe [ARBRE], la position la plus confortable résulte du placement des épaules et, par une configuration adéquate des épaules et du coude, permet un meilleur placement du poignet.

### 2.3.3. [A MOI]

Le signe [A MOI] est lui aussi mono-manuel. Il contraint la main à une configuration précise (celle de [V]). Le signeur place sa paume devant lui et l'oriente vers l'arrière. La Figure 51 présente deux exemples de générations de la posture de départ avant la rotation du poignet. L'orientation des doigts (vers le haut, ou à l'horizontale) n'est pas importante à ce stade du signe. Elle l'est cependant à la fin du signe, lors de la dernière posture où les doigts sont tournés vers le torse, le touchant.



[A MOI]

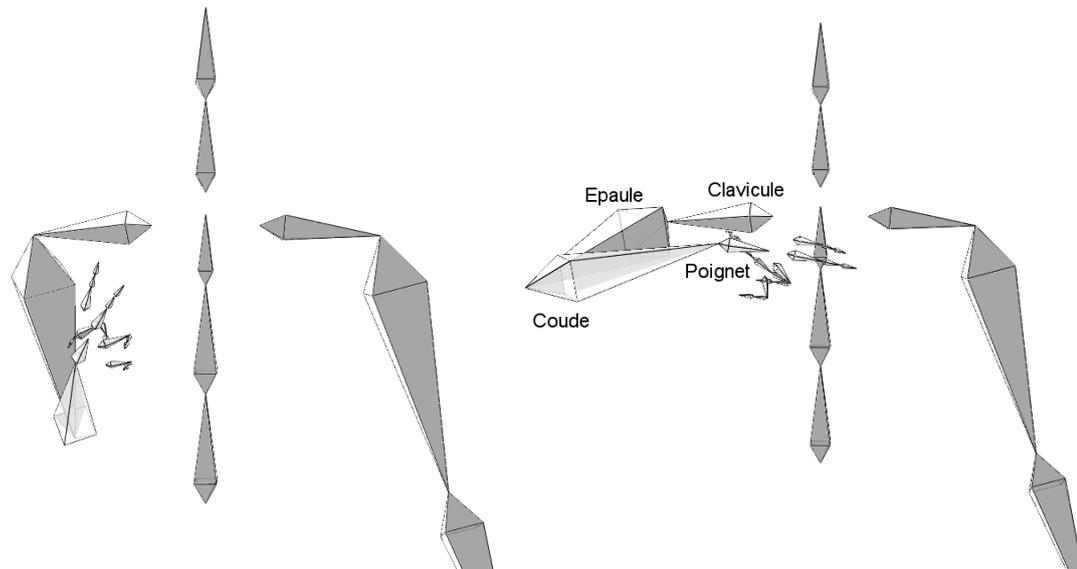


Figure 51 Signe [A MOI]. Postures de départ. Scores de confort : 0.569 à gauche, 0.588 à droite.

Encore une fois, on peut voir l'influence du confort dans la configuration de l'épaule. Sur l'image de gauche, le signeur colle son coude contre son torse et tire son bras en arrière pour placer le poignet en position alors que sur l'image de droite, il écarte son coude de façon à placer son avant bras à plat.

La Figure 52 présente par ailleurs la configuration manuelle générée en vue rapprochée. Elle montre la combinaison effective du système de génération et du modèle de pouce.

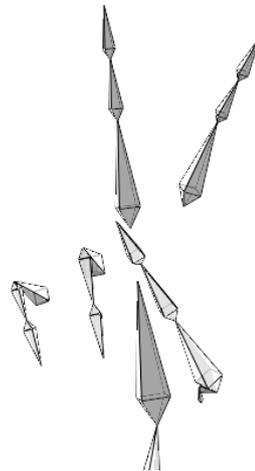


Figure 52 Configuration manuelle du signe [A MOI]

## 2.4. Apports

La différence de contexte entre un système de génération de langue des signes et d'autres systèmes rendent une comparaison des apports difficilement quantifiable. Notre système innove sur certains points par rapport à l'état de l'art.

D'abord, la résolution multiple de postures permet le choix d'une meilleure posture parmi un ensemble. Cette méthode permettra à long terme de toujours avoir le choix entre différentes postures de bonne qualité et d'avoir la possibilité de garder les meilleures. La combinaison des deux méthodes de cinématique inverse permet une approche couvrant un grand nombre de cas et un affinage des meilleures solutions.

Le modèle de confort quant à lui permet une sélection plus efficace des postures et un rendu plus naturel. L'approche est centrée sur les données, mais n'est pas dépendante de ces dernières. Ainsi le modèle est applicable à toutes sortes de mouvements, même ceux non présents à l'origine dans le corpus de capture de mouvement. Cette approche a cependant ses limites et le modèle pourra défavoriser la prise d'une posture correcte en langue des signes si elle est moins confortable qu'une autre posture un peu moins correcte. Bien que nous considérons le nombre de cas problématiques comme négligeable pour l'instant, il est important de garder à l'esprit ces cas particuliers pour les traiter à l'avenir.

## 2.5. Perspectives

Le modèle présenté dans ce chapitre n'est qu'une première étape nécessitant de nombreux approfondissements. Nous présentons dans cette section un certain nombre d'améliorations envisageables, aussi bien sur le modèle lui-même que sur son implémentation et son évaluation.

### 2.5.1. Niveaux d'abstraction dans le confort

L'étude du confort présentée dans le chapitre II se contente d'extraire des courbes de confort pour chaque degré de liberté. Comme nous l'avons vu pour le modèle de pouce, certaines articulations produisent des mouvements complexes ne se limitant pas qu'à un seul degré de liberté. Il est donc important d'étudier cette dépendance de certains degrés sur d'autres. Une amélioration possible du modèle serait d'étudier le confort, non plus au niveau de chaque degré de liberté, mais au niveau de chaque articulation.

Le raisonnement peut être poussé plus loin. Il peut être intéressant de considérer le confort d'un groupe fonctionnel complet, comme par exemple le bras entier. Empêchant ainsi les situations où deux articulations sont en configuration confortable et une troisième en configuration d'inconfort. Enfin, notons que ces différents niveaux de confort peuvent être compatibles entre eux, et évalués aux différents niveaux d'abstraction des blocs fonctionnels, donnant plus d'influence sur les degrés de confort de haut niveau d'abstraction

(« le bras est dans un état confortable ») qu'aux faibles niveaux d'abstraction (« l'os pisciforme est en configuration d'inconfort »).

### **2.5.2. Parallélisation**

L'intégralité de la phase d'optimisation est parallélisable. L'optimisation de différentes solutions à partir des résultats trouvés par l'exploration peut être réalisée par des processus parallèles. De plus, l'utilisation intensive de calcul matriciel lors de la phase d'optimisation rend cette étape particulièrement adaptée à une parallélisation sur carte graphique grâce à des méthodes de GPGPU<sup>40</sup> telles que Cuda ou OpenCL. Cela permettrait un gain de temps sur la résolution totale, tout en augmentant le nombre de solutions optimisées. Une telle parallélisation serait une plus-value pour l'amélioration de la vitesse d'exécution.

### **2.5.3. Paramétrage du système**

Le système de génération dépend d'un certain nombre de paramètres (détaillés en Annexe 3). Ces paramètres sont fortement imbriqués et nécessitent une étude à part entière afin de déterminer le meilleur rapport possible de la *qualité de résolution sur le temps d'exécution*. Nous proposons l'utilisation de plans d'expérience Taguchi<sup>41</sup> afin d'optimiser ces paramètres et d'obtenir des résultats optimaux.

### **2.5.4. Méthodologie d'évaluation du modèle de confort**

Une hypothèse que nous avons posée lors de la création du modèle de confort (section 6.3 du chapitre II) est que le corpus de capture de mouvement que nous avons étudié couvre la palette de mouvements employés en langue des signes. Cette hypothèse forte nécessite une vérification que nous n'avons pas pu faire dans le cadre de cette thèse.

Outre la correction du modèle, en appliquant la même méthodologie sur un corpus de capture de mouvement dédié à la langue des signes, il est intéressant de vérifier si l'assertion est vraie ou fausse. Pour démontrer ce point, il est possible de procéder de diverses manières. Nous proposons une solution particulière. Elle consiste à créer un modèle de confort basé sur un corpus de capture de mouvement uniquement en LSF. Les courbes

---

<sup>40</sup> GPGPU : General Purpose Graphical Processing Unit. Cet acronyme signifie l'utilisation de processeurs graphiques pour effectuer des calculs. Les processeurs de cartes graphiques, à la différence des processeurs standards d'unité centrale sont plus adaptés à certains types de calculs tels que les calculs matriciels. De plus, ils sont spécialement conçus pour permettre une parallélisation efficace.

<sup>41</sup> La méthode Taguchi, élaborée par Genichi Taguchi permet d'évaluer l'influence de plusieurs paramètres sur un même système et d'étudier leur corrélation en un nombre limité d'étapes. Cette méthode est très prisée dans le domaine des statistiques pour l'optimisation de systèmes à plusieurs paramètres.

obtenues sont ensuite comparées aux courbes du modèle présenté dans cette thèse. La méthode de comparaison des deux courbes est laissée au choix de l'expérimentateur. Nous proposons de réduire l'échantillon comparé aux mêmes bornes (si cela semble raisonnable) puis de comparer chaque série deux à deux grâce à un test de Kruskal-Wallis<sup>42</sup>.

### ***2.5.5. Méthodologie d'évaluation des résultats***

Comme indiqué dans la section 2.3, il est très difficile de quantifier la qualité d'une solution confortable par rapport à une autre. Il existe des solutions pour quantifier *a posteriori* le confort d'une posture et inférer la qualité du modèle.

Une solution serait d'introduire de réels éléments de simulation biomécanique dans le processus d'évaluation. Nous avons vu que les simulations biomécaniques sont très précises mais que leur intégration au système de génération est très coûteuse. Cependant, les utiliser dans le cadre d'un processus de validation est tout à fait envisageable et permettrait de quantifier la qualité du modèle de confort.

Enfin, une dernière piste serait l'utilisation à grande échelle d'évaluations subjectives. La présentation d'images de postures générées et l'interrogation de sujets quant au confort. Ce type d'expérimentation doit cependant être effectué avec précaution et rigueur. Il est notamment important de définir précisément les termes de l'évaluation. Il semble par exemple difficile de demander à une personne de quantifier le confort d'une posture sur une échelle.

---

<sup>42</sup> Le test de Kruskal-Wallis est un test non-paramétrique (ne nécessitant pas que les données suivent une distribution normale), permettant de détecter si deux séries statistiques sont issues d'une même population. Dans notre cas il s'agit de détecter si les données du corpus de LSF et celles de capture de mouvement standard sont équivalentes.

### **3. Conclusion du chapitre**

Nous avons présenté dans ce chapitre les résultats du modèle de pouce opposable, du système de génération et du modèle de confort. Ces systèmes apportent tous une avancée par rapport à l'état de l'art, mais disposent de limites qu'il nous faudra traiter par la suite.

Le dernier chapitre permet de présenter quelques perspectives plus générales, et de conclure sur cette thèse.



## Chapitre V : Perspectives générales et conclusion

Ce dernier chapitre commence par présenter les travaux futurs et les perspectives envisageables à la suite du modèle anatomique présenté dans les chapitres précédents. Nous reviendrons enfin sur un bref résumé des éléments du modèle ainsi que de ses apports.

## 1. Résumé et conclusion

Dans cette thèse, nous avons présenté un nouveau modèle anatomique permettant la représentation informatique du squelette, des contraintes qui s'appliquent sur ses os, de notion de confort d'articulation et d'un nouveau modèle de pouce opposable. Nous avons par ailleurs présenté une brève étude anthropométrique sur la relation entre la longueur de l'avant bras et celle des différents éléments constituant de la main.

Nous avons tout d'abord présenté le contexte dans lequel se place cette étude : le domaine de la langue des signes et de sa synthèse. Nous avons ainsi détaillé les raisons qui poussent les équipes de recherche à poursuivre des travaux dans le domaine de la synthèse de langue des signes basée sur des signeurs virtuels. Nous avons ensuite justifié le choix d'une génération automatique appuyée par un modèle linguistique de représentation des signes et établi une série de problématiques. Nous avons choisi de revenir sur la représentation du squelette en amont de la synthèse d'animation afin de fournir une base solide à la génération dédiée aux humanoïdes.

Dans un second chapitre, nous avons présenté le nouveau modèle linguistique, à commencer par la représentation du squelette en blocs fonctionnels. Chaque os de la représentation traditionnelle est modélisée par un bloc terminal. Les blocs sont organisés en niveaux hiérarchiques dits « d'abstraction » où le squelette représente le plus haut niveau d'abstraction et où l'os représente le plus bas niveau. Nous avons par la même occasion introduit les notions de bloc inconscient et de bloc conscient, définissant les unités fonctionnelles activées volontairement (et donc susceptible d'être utilisées par la représentation linguistique) ou non. Nous avons ensuite présenté une étude anthropométrique permettant d'établir une moyenne sur le rapport entre la longueur des os de la main et celle de l'avant bras de manière à créer un personnage neutre. La partie suivante s'est intéressée à la définition de dépendances articulaires de différents types (simples, dynamiques et distributives) sur les blocs fonctionnels de manière à simuler précisément l'interdépendance des os et les limitations articulaires. Dans un quatrième temps, nous avons présenté un nouveau modèle de pouce intégrant un axe virtuel représentatif de l'action des os du carpe de la main, permettant le placement du pouce en opposition. Ce modèle mieux contraint que l'approche traditionnelle à trois degrés de liberté, permet la génération de configurations complexes de la main. Enfin, la dernière partie du second chapitre présente une modélisation du confort déduite de l'étude d'un

corpus de capture de mouvement. Ce modèle de confort permet la quantification du naturel d'une posture, permettant le choix entre différentes générations d'un même signe pour choisir le plus naturel.

Le troisième chapitre détaille le processus de génération et l'intégration du modèle anatomique dans la synthèse. On revient d'abord sur les approches les plus courantes de la cinématique inverse utilisées en animation. Dans un second temps, nous avons vu les problématiques spécifiques au contexte de recherche (la langue des signes) et au modèle anatomique, puis établi une liste de critères à prendre en compte pour l'établissement d'une méthode de cinématique inverse. La troisième partie de ce chapitre nous a permis de détailler le processus de cinématique inverse élaboré pour répondre à ces critères. Cette partie décrit en détail les quatre phases de la résolution : définition des objectifs à atteindre, exploration aléatoire, optimisation locale et sélection de la solution la plus naturelle. Nous y avons aussi présenté les différentes limites de notre approche. Le troisième chapitre a aussi été l'occasion de présenter les travaux d'ingénierie réalisés pendant cette thèse, et notamment l'organisation logicielle des différents éléments détaillés dans le chapitre II.

Dans le quatrième chapitre, nous avons présenté les résultats du système de génération, et plus précisément du modèle de pouce opposable et de celui de confort. Nous avons montré par des analyses statistiques et empiriques l'apport de ces deux modèles, la nécessité d'utiliser un modèle de confort. Chaque section a été l'occasion de détailler le protocole expérimental ainsi que les résultats mais aussi de les discuter et de présenter quelques perspectives pour les travaux futurs.

Enfin, nous avons présenté dans le cinquième chapitre les perspectives plus générales du modèle anatomique et du système de génération de langue des signes. Ce chapitre a permis de montrer l'intérêt de l'ouverture du modèle anatomique non seulement à d'autres domaines de la langue des signes (comme l'analyse et la reconnaissance), mais aussi à d'autres domaines de l'animation et des humains virtuels.

Nous avons cherché dans ces travaux à mettre en avant plusieurs axes de réflexions. Tout d'abord, l'importance du paradigme dans lequel nous développons des solutions d'animation et la nécessité de le questionner. Il nous a paru fondamental de remettre en question la représentation standard du squelette et de réfléchir à des alternatives pour conduire à une modélisation à la fois plus précise et rapide que les standards de l'animation.

Par ailleurs, ces travaux ont cherché à mettre en relief l'importance de considérer l'ensemble des paramètres environnementaux à un domaine. Nous incitons par exemple à la prise en compte d'informations biologiques et statistiques dans la création de mouvement et d'animation. Il nous semble nécessaire de s'interroger sur la nature du mouvement réel, sur la notion de posture, ou encore de confort, pour s'attacher à créer des animations de plus en plus réalistes.

Ce n'est qu'en s'interrogeant sur ces bases fondamentales que nous pourrons un jour générer des animations réalistes et peut-être proposer des systèmes de génération automatique de langue des signes à la communauté sourde.

## 2. Perspectives

Outre les perspectives présentées dans le chapitre précédent, directement liées aux éléments du modèle, il existe d'autres travaux envisageables à une échelle plus globale.

### 2.1. Extension du modèle aux expressions faciales

Cette thèse est entièrement dédiée à la synthèse de postures. Elle ne s'intéresse pas aux expressions du visage qui sont pourtant des éléments clés de la compréhension du discours, servant à la fois à désambiguïser le sens des énoncés mais aussi à y ajouter de l'information. Il est impératif que le modèle de squelette de signeur virtuel intègre des points de contrôle du visage. Certains travaux présentés en chapitre I existent sur le contrôle du visage pour l'expressivité. Il s'agit désormais d'étudier non seulement l'intégration de ces méthodes au modèle, mais aussi de faire le lien entre les travaux sur l'expressivité générale de personnage et la langue des signes.

Plus qu'une réutilisation de standards déjà en place il nous faut étudier de nouvelles méthodes de contrôle du visage de manière à pouvoir faire un lien aisé entre le modèle linguistique et les expressions faciales. Ce point ne peut être assuré par la simple combinaison d'expressions faciales actuellement utilisée comme standard dans le domaine de la langue des signes. On pourra par exemple s'inspirer de travaux menés dans le domaine des agents conversationnels animés (Pelachaud 2009). Tout comme le modèle de confort présenté dans cette thèse, il peut être intéressant de se pencher sur des données de capture de mouvement et sur le domaine de la reconnaissance d'expressions faciales pour établir des modèles structurels adéquats en amont de l'animation elle-même. Enfin notons l'existence de modèles linguistiques sur les expressions faciales en langue des signes (Chételat-Pelé 2010) qui doivent aussi être incorporées aux systèmes de génération.

### 2.2. Animation du signeur

Pour l'instant cette étude ne s'intéresse qu'à la génération de postures statiques. Pourtant le contexte de la thèse présenté en chapitre I est présenté comme du domaine de l'animation. Il ne faut pas perdre de vue que le réalisme d'une animation ne se mesure pas uniquement à un enchaînement de postures naturelles. Un critère décisif d'une bonne animation réside dans sa dynamique. Un mouvement uniforme donnera lieu à un rendu trop robotique. (Lefebvre-Albaret 2010) donne une première piste intéressante de courbes de dynamique du mouvement extraites à partir de systèmes de reconnaissance de mouvement. Les informations générales quant à la dynamique du mouvement, qu'elles soient purement

biomécaniques ou pilotées par les données ont leur place dans le modèle anatomique et pourraient donc être intégrées dans les blocs fonctionnels à la manière des courbes de confort.

Par ailleurs, l'inclusion de la dynamique doit respecter une certaine cohérence sémantique. Il semble mal-approprié d'avoir une dynamique « énergétique » lorsque le signeur explique qu'il est fatigué. De même, un mouvement lent lors d'un discours « colérique ». Ces informations contextuelles n'ont cependant pas leur place dans le modèle anatomique, mais plus dans la partie linguistique. Une réflexion *a posteriori* de leur définition doit être effectuée pour déterminer comment combiner la dynamique contextuelle à la dynamique générale du mouvement. Une fois de plus, outre les données dont nous avons déjà connaissance, il peut être intéressant de calquer l'approche statistique que nous avons eu ici pour le confort sur des corpus de capture de mouvement afin d'obtenir des informations sur la dynamique du mouvement. Enfin, il existe déjà de nombreux travaux sur le mouvement de manière générale ou sur les gestes co-verbaux sur lesquels nous pouvons nous appuyer pour débiter une réflexion dans le domaine de la langue des signes.

### 2.3. Gestion des collisions

Un élément décisif du réalisme de posture et de mouvement est la gestion des collisions. L'utilisation de dépendances articulaires permet au squelette de limiter fortement la construction de postures impossibles. Pourtant, la section 0 du chapitre IV montre des exemples de génération du signe [ARBRE] où la main traverse l'avant bras (dans le cas d'un confort faible). Ces postures sont encore possibles du fait de l'absence de gestion des collisions.

Pour la prise en compte de ces collisions, il est nécessaire de considérer plusieurs aspects. Tout d'abord, les systèmes de collisions les plus simples peuvent se contenter d'empêcher que les os s'entrecroisent. Néanmoins une telle représentation ne suffit pas à prendre en compte tous les détails de la peau, notamment au niveau du visage et du torse où le volume à prendre en compte est complexe et bien différent de la simple structure des os. Le système utilisé dans (Kavan et al. 2005) détecte les collisions en utilisant un modèle de volumes approximatifs sensés représenter le corps du signeur. Cependant, encore une fois, ces volumes simplifiés ne permettent pas de tenir véritablement compte de la complexité de la peau du signeur (et par peau nous englobons aussi les vêtements). Tout comme nous

l'avons fait auparavant dans cette étude, nous pouvons aussi nous inspirer de systèmes utilisés dans les domaines médicaux, tel que celui présenté dans (Arbabi et al. 2009). Ces méthodes, trop complexes pour être utilisées en animation de langue des signes peuvent fournir un point de départ pour une réflexion ultérieure.

En marge de cette réflexion, il faut relever qu'il est parfois nécessaire que des articulations entrent en collision. Le système de génération doit donc savoir différencier les contacts volontaires et les collisions.

Enfin, il est important de noter qu'en aval de la détection de collision, il faut aussi disposer d'un système capable de résoudre le problème et de proposer une solution alternative. Le système de génération que nous avons proposé permet un contournement partiel du problème. En effet, si une solution, tant est qu'elle soit valide, est détectée en état de « collision » il est possible de l'écartier pour en prendre une meilleure. Cependant cette méthode n'est pas exempte de défauts puisque lorsque la collision est volontaire toutes les solutions valides risquent d'être écartées.

#### **2.4. Evolutions du modèle en rapport avec la peau**

Toutes les évaluations du modèle ont utilisé la représentation en octaèdres du squelette. Il est important d'effectuer les évaluations en incluant une peau au signeur afin de détecter d'autres problèmes et/ou limitations du modèle. L'évaluation du modèle doit se faire par le biais de différentes peaux afin de détecter les problèmes de mesures du squelette.

L'utilisation de peaux différentes nécessite évidemment la construction d'un squelette différent à chaque fois, premièrement pour que ce dernier ait une taille adéquate et d'autre part pour le placement précis des sites corporels qui nécessitent un calibrage. On pourra s'inspirer de travaux faits dans le domaine des agents conversationnels animés (Courgeon et al. 2008).

Par ailleurs il peut être pertinent d'établir des méthodes de placement automatique des signes corporels en fonction d'un squelette et d'une peau, de manière à les localiser plus précisément et aisément sur la peau.

#### **2.5. Mouvement ambiants**

A l'heure actuelle, le système de génération ne prend d'entrées que du modèle linguistique et du modèle anatomique. Afin d'accroître le réalisme de la génération, il est nécessaire d'inclure aussi des mouvements ambiants au signeur. Rappelons que les

mouvements ambiants sont des mouvements non caractéristiques du discours tels que certains clignements d'yeux, le balancement du torse (non sémantique), la respiration, etc. Ces éléments insignifiants pour le discours sont cependant perceptibles par le spectateur rendant le personnage plus humain, plus vivant.

Pour autant, à la différence de certains travaux existants (Glauert et al. 2011) il est impératif que ces mouvements ambiants ne soient pas inclus dans la description linguistique, mais bien inclus *a posteriori*. On peut ainsi envisager une synthèse en plusieurs étapes où une phase de post-traitement des animations générées rajouterait l'information ambiante.

## 2.6. Application à la reconnaissance de langue des signes

Le modèle anatomique et le système de génération peuvent être employés dans les domaines de reconnaissance de langue des signes. Certains systèmes de reconnaissance (Lefebvre-Albaret 2010) commencent à s'appuyer sur des modèles linguistiques et plus particulièrement sur Zebedee. Par ailleurs, ces mêmes équipes se tournent de plus en plus vers la reconnaissance assistée par la génération de signes. Ces méthodes utilisent la génération de signe pour détecter et extraire des informations de descriptions linguistiques, puis les confronter aux informations tirées de vidéos. Le système de génération et le modèle anatomique trouvent leur place dans ce cadre où il faut être capable de générer des séquences issues du modèle linguistique et que les postures/animations résultantes soient naturelles et comparables à celles des vidéos. Il existe par ailleurs une autre piste possible d'apports de la génération au domaine de la reconnaissance automatique. Les systèmes d'analyse et de reconnaissance sont souvent entraînés à partir d'échantillons. Ces échantillons pourraient être enrichis par le biais d'un système de génération, permettant ainsi un apprentissage sur une plus grande quantité de données, et donc un apprentissage plus fiable. Cela nécessiterait cependant un système de génération particulièrement robuste et efficace.

## 2.7. Applications à d'autres domaines

Le modèle anatomique n'est pas foncièrement dépendant de la langue des signes. Il est développé dans ce contexte et est clairement orienté pour une utilisation dans ce cadre. Cependant, puisque le système de génération est basé sur la définition de contraintes sur le squelette, il n'est pas du tout impossible de l'utiliser dans un cadre plus général.

Outre pour la génération d'animations comme nous l'employons ici, il nous paraît clair que l'utilisation d'un tel modèle pour assister la conception graphique d'animations peut être significative. Il existe de nombreuses utilisations de la cinématique inverse pour l'assistance au positionnement de membres dans divers logiciels de graphisme. L'utilisation d'un modèle anatomique intégré permettrait à certains logiciels un paramétrage plus simple et par la même une définition de postures et d'animations plus rapide.

Le domaine de la robotique pourrait lui aussi bénéficier des apports d'un modèle anatomique. Notamment lorsqu'il s'agit de champs de recherche sur la construction d'androïdes où le rendu humain est prioritaire, l'utilisation d'un modèle anatomique tel que celui défini ici permet encore une fois une assistance à la prise de posture.





## Bibliographie

- Alegria, J.** 2007. "Phonologie audiovisuelle: lecture, lecture labiale et lecture labiale assistée.", *Surdit  et Langage: proth ses, LPC et implants cochl aires* p. 99-149. Editions Presses Universitaires de Vincennes.
- Arbabi, A. et al.** 2009, "Fast collision detection methods for joint surfaces.", *Journal of Biomechanics*, vol.42(2), p.91-99.
- Austin, N.M.** 2005. "Chapter 9: The Wrist and Hand Complex.", *Joint Structure and Function: A Comprehensive Analysis (4e  dition)*. F.A. Davis.
- Aznar, G. et al.** 2005. "Projet «Usages de l'Internet» WebSourd.", Rapport de recherche.
- Badler, N.I. et al.** 1999. "Simulating Humans: Computer Graphics, Animation and Control". Oxford University Press.
- Bailly, G. et al.** 2009. "Parole et expression des  motions sur le visage d'humano ides virtuels", *Trait  de r alit  virtuelle (5e  dition)*. Presses de l'Ecole des Mines.
- Boulic, R.** 2003. "Cin matique inverse pour l'optimisation de postures humaines soumises   des contraintes multiples.", *Humanoid Day, Soci t  Fran aise de Biom canique*.
- Braffort, A. et al.** 2011. "Virtual signer coarticulation in Octopus, a Sign Language generation platform.", *GW 2011: The 9th International Gesture Workshop*.
- Buss, S.** 2003. *3D Computer Graphics: A Mathematical Introduction with OpenGL*. Cambridge University Press.
- Buss, S.** 2005. *Introduction to Inverse Kinematics with Jacobian Transpose, Pseudoinverse and Damped Least Squares methods*.
- Canutescu, A.A et al.** 2003. "Cyclic coordinate descent: A robotics algorithm for protein loop closure." *Protein Science*, vol.12(5), p.963-972.
- Chang, L.Y. et al.** 2008. "Method for determining kinematic parameters of the in vivo thumb carpometacarpal joint." *IEEE transactions on bio-medical engineering*, vol.55(7), p.1897-906.
- Chao, E.** 2003. "Graphic-based musculoskeletal model for biomechanical analyses and animation." *Medical Engineering & Physics*, vol.25(3), p.201-212.
- Ch telat-Pel , E.** 2010. "Les Gestes Non Manuels en Langue des Signes Fran aise; Annotation, analyse et formalisation: application aux mouvements des sourcils et aux clignements des yeux.", Th se de doctorat pr sent e devant l'Universit  de Provence.
- Courgeon, M.** 2008. "Simulation et contr le de visages virtuels expressifs en situation d'interaction.", Rapport de Master Recherche.
- Courgeon, M., et al.** 2008. "MARC: un personnage virtuel r actif expressif.", *Troisi me Workshop Francophone sur les Agents Conversationnels Anim s. WACA'08*, Citeseer.
- Courty, N. et al.** 2008. "Inverse Kinematics using Sequential Monte Carlo Methods." In *International Conference on Articulated Motion and Deformable Objects (AMDO 2008)*.

- Cox, S. et al.** 2002. "Tessa, a system to aid communication with deaf people." *Proceedings of the fifth international ACM conference on Assistive technologies - Assets '02*.
- Cuxac, C.** 2000. *Langue des signes française - les voies de l'iconicité*, Ophrys.
- Dave, U. et al.** 1998. *Linear Programming-1: Introduction*. Springer.
- Duarte, K. et al.** 2010. "Heterogeneous Data Sources for Signed Language Analysis and Synthesis: The SignCom Project.", *Seventh international conference on Language Resources and Evaluation (LREC 2010)*.
- Duarte, K. et al.** 2011. "Presentation of the SignCom Project." In *International Workshop on Sign Language Translation and Avatar Technology (SLTAT)*.
- Elliott, R. et al.** 2008. "Linguistic modelling and language-processing technologies for Avatar-based sign language presentation." *Universal Access in the Information Society*, vol.6(4), p.375-391.
- Filhol, M.** 2008. "Modèle descriptif des signes pour un traitement automatique des langues des signes.", Thèse de doctorat présentée devant l'Université Paris-Sud.
- Filhol, M. et al.** 2006. "Sign description: how geometry and graphing serve linguistic issues." In *Theoretical issues on Sign Language Research (TISLR 9)*.
- Fotineas, S-E et al.** 2008. "A knowledge-based sign synthesis architecture." *Universal Access in the Information Society*, vol.6(4), p.405-418.
- Furniss, M.** 1999. "Motion Capture.", rapport disponible sur <http://web.mit.edu/comm-forum/papers/furniss.html>
- Gillot, D.** 1998. *Le droit des sourds: 115 propositions, rapport au premier ministre*.
- Glauert, J.R.W. et al.** 2011. "Overview of Signing Avatar Work at UEA.", *International Workshop on Sign Language Translation and Avatar Technology (SLTAT)*.
- Gleicher, M.** 1998. "Retargetting motion to new characters." *Proceedings of the 25th annual conference on Computer graphics and interactive techniques - SIGGRAPH'98*.
- Hollister, A. et al.** 1992. "The axes of rotation of the thumb carpometacarpal joint." *Journal of orthopaedic research: official publication of the Orthopaedic Research Society*, vol.10(3), p.454-60.
- Héloir, A.** 2008. "Système de communication par agent virtuel, aide à la communication des personnes sourdes.", Thèse de doctorat présentée devant l'Université de Bretagne-Sud.
- Kavan, L. et al.** 2005. "Fast Collision Detection for Skeletally Deformable Models." *Computer Graphics Forum*, vol.24(3), p.363-372.
- Kirkpatrick, S. et al.** 1983. "Optimization by Simulated Annealing." *Science*, vol.220(4598), p. 671-680.
- Klopkar, N. et al.** 2007. "A kinematic model of the shoulder complex to evaluate the arm-reachable workspace." *Journal of biomechanics*, vol.40(1), p.86-91.
- Lee, J. et al.** 1995. "Model-based Analysis of Hand Posture." *IEEE Computer Graphics and Applications*, vol.15(5), p.77-86.

- Lefebvre-Albaret, F.** 2010. "Traitement automatique de vidéos en LSF, Modélisation et exploitation des contraintes phonologiques du mouvement.", Thèse de Doctorat présentée devant l'Université Paul Sabatier.
- Liddell, S.K. et al.** 1989. "American Sign Language: the phonological base" *Sign Language Studies*, vol.64, p.195-278.
- Lombardo, V. et al.** 2011. "An Avatar-based Interface for the Italian Sign Language." *International Conference on Complex, Intelligent and Software Intensive Systems (CISIS)*.
- Lombardo, V. et al.** 2010. "A Virtual Interpreter for the Italian Sign Language." In *Intelligent Virtual Agents: 10th international conference (IVA 2010)*.
- Losson, O.** 2000. "Modélisation du geste communicatif et réalisation d'un signeur virtuel de phrases en langue des signes française.". Thèse de doctorat présentée devant l'Université des Sciences et Technologies de Lille.
- Lu, P. et al.** 2009. "Accessible motion-capture glove calibration protocol for recording sign language data from deaf subjects." *Assets'09 Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*.
- L'Huillier, M-T.** 2004. "Surdité et langue des signes: d'où vient l'illettrisme chez les personnes sourdes.". Retranscription de conférence disponible sur [http://www.liturerre.org/D\\_ou\\_vient\\_l\\_illettrisme\\_chez\\_les\\_personnes\\_sourdes-Intervention\\_de\\_Marie\\_Therese\\_L\\_Huillier.pdf](http://www.liturerre.org/D_ou_vient_l_illettrisme_chez_les_personnes_sourdes-Intervention_de_Marie_Therese_L_Huillier.pdf)
- Makhsous, M. et al.** 1999. "Total shoulder and relative muscle strength in the scapular plane." *Journal of biomechanics*, vol.32(11), p.1213-20.
- Marschark, M.** 2001. "Language Development in Children Who Are Deaf: A Research Synthesis.", Rapport de Recherche.
- McDonald, J. et al.** 2001. "An improved articulated model of the human hand." *The Visual Computer*. vol.17(3), p.158-166.
- McDonald, J. et al.** 2002. "A Direct Method for Positioning the Arms of a Human Model." *Graphics Interface 2002*.
- Neff, M. et al.** 2006. "Modeling relaxed hand shape for character animation." *Lecture notes in computer science*, vol.4069, p.262-270.
- Niemi, J. et al.** 1996. "A static shoulder model based on a time-dependent criterion for load sharing between synergistic muscles." *Journal of Biomechanics*, vol.29(4), p.451-460.
- Van Nierop, O.A. et al.** 2007. "A natural human hand model." *The Visual Computer*, vol.24(1), p.31-44.
- Parzen, E.** 1962. "On Estimation of a Probability Density function and Mode." *Annals of Mathematical Statistics*, vol.33(3), p.1065-1076.
- Pelachaud, C.** 2009. "Modelling multimodal expression of emotion in a virtual agent." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol.364(1535), p.3539-48.
- Prillwitz, S.** 1989. "HamNoSys version 2.0 - Hamburg Notation System for Sign Languages, an introductory guide." *International studies on Sign Language and communication of the Deaf 5*.

- Schnepp, J. et al.** 2010. "Synthetic Corpora: A Synergy of Linguistics and Computer Animation." 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies.
- Stokoe, W.C.** 1960. "Sign language structure: an outline of the visual communication systems of the American deaf." *Journal of deaf studies and deaf education*, vol.10(1), p.3-37.
- Tsang, W. et al.** 2005. "Helping Hand: An Anatomically Accurate Inverse Dynamics Solution For Unconstrained Hand Motion." Eurographic/ACM SIGGRAPH Symposium on Computer Animation.
- Wolfe, R. et al.** 2011. "Linguistics as structure in computer animation: Toward a more effective synthesis of brow motion in American Sign Language." *Nonmanuals in Sign Language Special issue of Sign Language & Linguistics*, vol.14(1), p.179-199.
- Wolovich, W.A. et al.** 1984. "A computational technique for Inverse Kinematics." In *The 23rd IEEE Conference on Decision and Control*.
- Wuelker, N. et al.** 1995. "A dynamic shoulder model: reliability testing and muscle force study." *Journal of Biomechanics*, vol.28(5), p.489-499.
- Zhao, J. et al.** 1994. "Inverse kinematics positioning using nonlinear programming for highly articulated figures." *ACM Transactions on Graphics*, vol.13(4), p.313-336.

# Annexe 1 : Terminologie

Cette annexe pose un certain nombre de jalons terminologiques visant à clarifier le discours des chapitres. Dans un premier temps sont listés les termes techniques, liés au domaine de la biologie, particulièrement utilisés dans le second chapitre. La seconde section de ce chapitre introduit les termes techniques liés à l'animation, aux signeurs virtuels et à la cinématique inverse.

## 1. Terminologie biologique

Dans un premier temps nous présentons des considérations générales : le système de repères en anatomie. Ensuite (section 1.2), sont rappelés les os et articulations des membres supérieurs ainsi que les simplifications/réductions utilisées dans cette thèse. Enfin, nous terminons en section 1.3 sur les noms des différents mouvements possibles du corps humain.

### 1.1. Repérage en anatomie

Lorsque l'on considère le corps humain en termes géométriques on s'aperçoit de la complexité à désigner ses différentes parties. Quelques éléments simples nous permettent d'effectuer des localisations tels que « la gauche » ou « la droite », « le haut » ou « le bas », voire à partir d'éléments du corps lui-même : « sur la jambe ». Ces termes n'ont en soi aucune consistance mathématique et ne peuvent être utilisés dans le discours sans l'appui d'une terminologie bien précise. Les éléments d'anatomie utilisés dans les sections et chapitres qui suivent sont généralement placés ou organisés selon trois plans principaux : sagittal, coronal et transverse (cf. Figure 53). Ces plans sont relatifs aux os du corps mais peuvent être utilisés sans précision (« le plan sagittal ») auquel cas ils désignent les plans du corps.

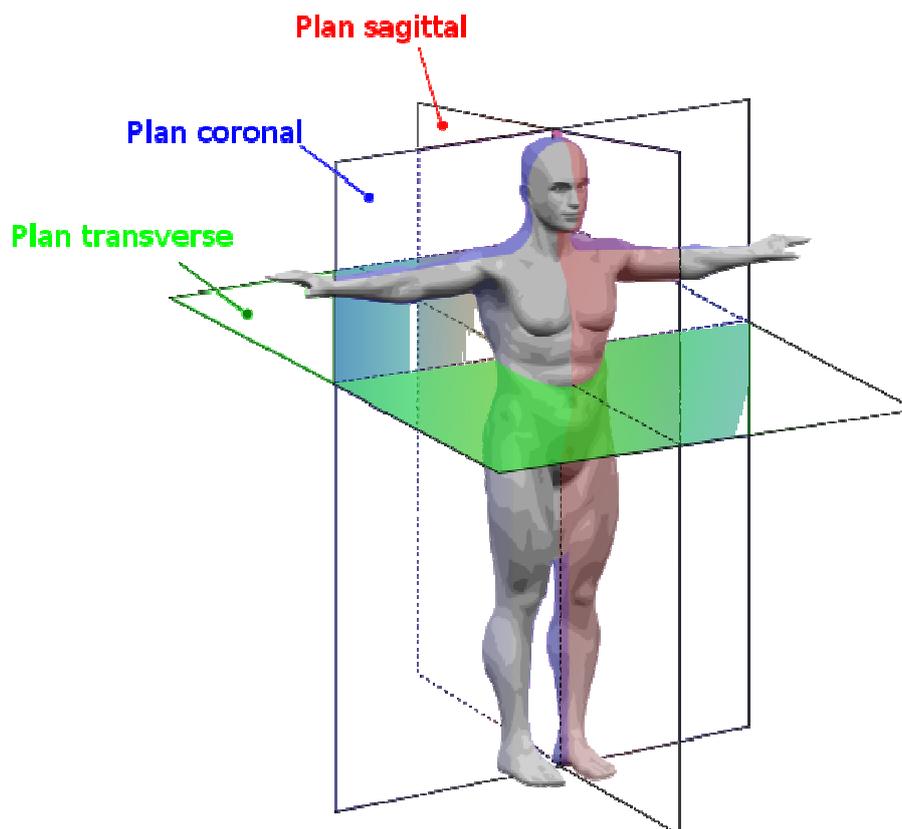


Figure 53 Les trois plans du repère anatomique. (Image sous licence creative commons CC-BY-SA, auteur Yassine Mrabet).

Ces plans sont toujours multiples : il n’y a pas un seul plan sagittal mais **des** plans sagittaux. Le plan *sagittal* correspondant au plan de symétrie droite/gauche du corps est appelé *plan médian*. Les autres plans sagittaux sont tous les plans parallèles au plan médian. Les plans *transverses* séparent le corps en deux parties : *craniale* du côté de la tête et *caudale* du côté des pieds. Il n’existe pas de plan singulier correspondant au plan médian pour les plans transverses. Enfin, les plans coronaux distinguent « l’avant » de « l’arrière ».

A chaque plan est associé un axe de référence qui lui est perpendiculaire. L’axe correspondant aux plans sagittaux est dit axe *horizontal* ou *gauche/droite*. Celui correspondant aux plans transverse est appelé axe *rostro-caudal*. Enfin, l’axe *dorso-ventral* correspond aux plans coronaux.

Enfin, lorsque l’on considère les membres du corps humain, un certain nombre d’adjectifs permettent de qualifier le placement relatif des éléments :

- *Latéral/Médial* : désigne, sur un plan coronal ce qui s’écarte du corps (latéral) ou ce qui s’en rapproche (médial). Par exemple, l’ulna est l’os médial de l’avant bras alors que le radius est l’os latéral.

- *Proximal/Distal* : sur un plan coronal, ce qui se rapproche de l'extrémité (distal) ou du corps (proximal). La pulpe du doigt est placée sur la phalange dite distale alors que la première phalange est dite proximale.

C'est grâce à ces termes que nous définirons les mouvements. Néanmoins, avant de parler des mouvements des os et des articulations, rappelons brièvement ces différentes parties du corps qui sont au cœur de l'étude du troisième chapitre.

## 1.2. Nomenclature du squelette

Puisque nous ne considérons pour la synthèse de langue des signes que la partie supérieure du corps, seuls les éléments des membres supérieurs sont détaillés dans la suite. Ces éléments sont listés en remontant de la colonne vertébrale; l'épaule et le bras; le coude et l'avant bras; et termine par la main. Cette section n'aborde cependant pas les réalités complexes des mécanismes du mouvement et n'incluse pas deux de leurs acteurs principaux : les muscles et les tendons.

### 1.2.1. Colonne vertébrale

La colonne vertébrale est un ensemble de petits os articulés entre eux jouant à la fois le rôle de squelette et de gaine permettant la transmission d'informations au corps tout entier. Les *vertèbres*, unités élémentaires de la colonne, sont au nombre de 24 et se répartissent en trois groupes : 5 vertèbres *lombaires* ; 12 *thoraciques* et 7 *cervicales*. Ces vertèbres créent une double courbure donnant sa forme en sigmoïde à la colonne.

### 1.2.2. L'épaule et le bras

On appelle ici complexe de l'épaule l'ensemble d'articulations permettant la jonction entre le torse et le bras. Il s'agit de quatre os en interaction : la *clavicule*, l'*omoplate*, la *tête humérale* et le *manubrium sternal*. La tête humérale assure la jonction avec l'humérus qui est l'os du bras alors que le manubrium sternal représente la connexion entre le complexe de l'épaule et le sternum, donc la cage thoracique. Par la suite, nous considérons simplement l'ensemble de l'épaule comme deux groupes d'articulations : la clavicule et « l'épaule ». La première regroupe effectivement la clavicule ainsi que le manubrium sternal, la seconde regroupe l'omoplate et la tête humérale.

### 1.2.3. Le coude et l'avant bras

Le coude est l'articulation joignant l'humérus avec les deux os de l'avant bras. Ces deux os, nommés *radius* et *ulna*, permettent par leur opposition la rotation du poignet. Le coude est une articulation permettant uniquement la flexion, donc le repli de l'avant bras sur le bras.

### 1.2.4. La main

La main est probablement l'élément le plus complexe en termes d'anatomie fonctionnelle du corps humain. La jonction des deux os de l'avant bras (*radius* et *ulna*) avec les os de la main se fait grâce au *carpe*. Le carpe est un ensemble de huit os répartis en deux rangées donnant au poignet sa mobilité et au pouce sa capacité à s'opposer. Le carpe est articulé avec les *métacarpes*, longs os de chaque doigt contenus dans la paume de la main. L'ensemble de ces os est résumé sur la Figure 54.

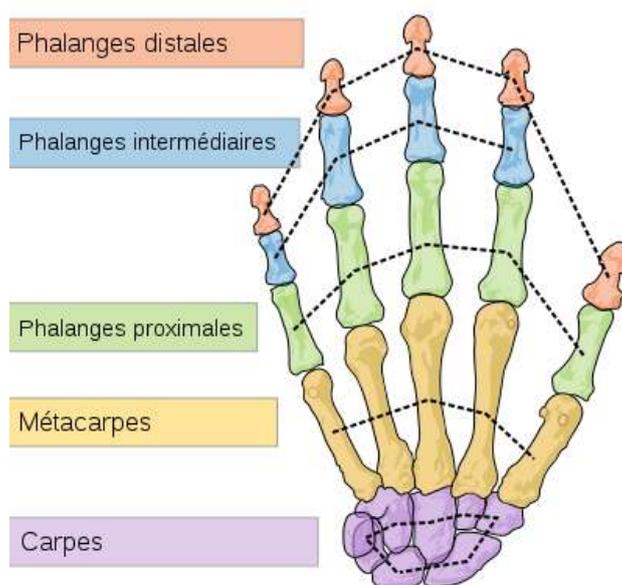


Figure 54 Les différents os constituant le squelette de la main.

Les métacarpes sont articulés avec les phalanges. Chaque doigt dispose de trois phalanges (proximale, intermédiaire et distale) à l'exception du pouce qui ne dispose pas de phalange intermédiaire. Les articulations inter-phalangiennes sont toutes nommées en fonction des os qu'elles connectent :

- *CMC* pour articulation *Carpo-MetaCarpienne* : l'articulation de chaque doigt connectant le carpe au métacarpe ;
- *MCP* pour articulation *Metacarpo-Phalangienne* : l'articulation connectant le métacarpe à la phalange proximale ;

- *PIP* pour articulation *Proximale-InterPhalangiennes* : l'articulation entre les phalanges proximale et intermédiaire ;
- *DIP* pour articulation *Distale-InterPhalangiennes* : l'articulation entre les phalanges intermédiaire et distale.

Chaque articulation permet un ou plusieurs mouvements que nous allons maintenant décrire.

### 1.3. Mouvements du corps humain

Les mouvements du corps humain sont spécifiques des articulations et utilisent très souvent le système de repérage décrit en section 1.1.

De manière générale, on peut caractériser trois types de mouvements selon les trois types de plans du repérage anatomique : coronaux, sagittaux et transverses. Le mouvement concerne toujours une articulation et l'os dont elle est l'origine. On décrit donc le déplacement des os sur chaque plan comme suit :

- *Abduction/Adduction* : un mouvement le long d'un plan coronal. Par exemple l'abduction de la colonne vertébrale. Le mouvement d'abduction écarte l'extrémité du membre du plan sagittal de référence, le mouvement d'adduction l'en rapproche.
- *Flexion/Extension* : un mouvement sur le plan sagittal. Par exemple la flexion des doigts ou du poignet. La flexion caractérise une diminution de l'angle formé entre l'os concerné et son prédécesseur, l'extension caractérise l'augmentation de cet angle.
- *Rotation* : un mouvement qui se déroule sur un plan transverse en faisant tourner l'os selon son axe directeur.

Ces descripteurs généraux ne s'appliquent pas pour toutes les articulations, auquel cas il faut redéfinir d'autres mouvements. C'est le cas notamment de la clavicule, de l'épaule et du coude :

- Le mouvement le long des plans sagittaux de la clavicule sont dits mouvements *d'élévation* et *d'abaissement*.
- Les mouvements qualifiés d'abduction/adduction de la clavicule sont dits mouvements *d'antéimpulsion* et de *réimpulsion*.

- De même, le mouvement le long des plans coronaires de l'épaule sont dits mouvements d'*élévation* et d'*abaissement*.
- En lieu et place de flexion et d'extension de l'épaule on utilisera les termes d'*antépropulsion* et de *répropulsion*.
- Enfin, le mouvement de rotation opposant l'ulna et le radius dans l'avant bras sera qualifié de *pronosupination*. Tourner son poignet de manière à ce que les deux os soient « parallèles » est qualifié de mouvement de *supination* alors que l'inverse est dit mouvement de *pronation*.

## 2. Terminologie d'animation

Cette section s'attache à définir plus précisément certains éléments techniques et propres à l'animation. La première sous-section s'attache au squelette du personnage virtuel, la seconde détaille l'outil mathématique utilisé pour symboliser les rotations : le quaternion. Enfin, la troisième section détaille la notion de cinématique inverse.

### 2.1. Notions sur le squelette

Le squelette du signeur est toujours représenté d'une manière symbolique en marge de son équivalent anatomique. Lorsque l'on considère le squelette en animation, on le décrit généralement en fonction de deux types d'entités : les articulations et les os connectant ces articulations. On détermine une articulation comme le point de départ d'un os si le positionnement de cet os dépend de la configuration de l'articulation. De même, on nommera une articulation comme une *terminaison* d'un os si ce dernier n'est pas sujet à modification par cette articulation.

On aura tendance, en animation, à considérer uniquement les articulations qui sont les parties mobiles du corps. Repérés dans un espace en trois dimensions, les os ne représentent en fait que des vecteurs de translation, alors que les articulations sont des éléments de rotation. Chaque articulation définit son propre repère rotationnel *local*, et contribue à l'orientation *globale* des os et des articulations plus distales. Ce repère de rotation peut faire tourner l'articulation selon un, deux ou trois axes. Chaque axe sur lequel on peut appliquer une rotation est appelé *degré de liberté*. Une articulation dispose donc d'un à trois degrés de liberté, définissant autant d'axes de rotations orthogonaux et par conséquent autant de mouvements comme définis en section 1.1.3 de ce chapitre. Ces degrés de liberté représentent le cœur des méthodes de cinématique inverse. Cependant,

chaque bloc fonctionnel terminal (cf. section 2.1 du chapitre II) regroupe plusieurs degrés de liberté, définissant ainsi un repère propre pour chaque « os ». L'orientation de l'os est déterminée grâce à un outil mathématique nommé quaternion, que nous allons maintenant présenter.

## 2.2. Quaternions

L'ensemble des quaternions est un ensemble de nombres étendant la définition de nombre complexe à un espace en trois dimensions. Alors que les complexes sont des couples de réels (partie réelle ; partie imaginaire), les quaternions sont des quadruplets de réels. La manipulation de seulement 4 réels pour représenter une rotation en lieu et place de 9 dans le cas d'une matrice rend leur utilisation particulièrement prisée du domaine de l'animation puisqu'elle permet un calcul bien plus rapide. Cette section n'a pas pour vocation une définition étendue des quaternions, elle se contente de présenter succinctement les éléments et opérateurs importants dans le cadre de la représentation d'orientations. Voici une liste de propriétés qui nous intéressent :

- Un quaternion  $Q$  est représenté en informatique par un quadruplet de réels  $(a, b, c, d)$  ou par l'union d'un scalaire et d'un vecteur  $(a, v)$  où  $v = (b, c, d)$ .
- On définit la norme euclidienne d'un quaternion comme étant  $\|Q\| = \sqrt{a^2 + b^2 + c^2 + d^2}$ .
- Un quaternion est dit *unitaire* si sa norme vaut 1.
- Tout comme les nombres complexes, chaque quaternion dispose d'un conjugué, écrit  $Q^* = (a, -v)$ .
- Les quaternions ne sont pas commutatifs. Ainsi  $Q_1 Q_2$  est différent de  $Q_2 Q_1$ .
- La multiplication de deux quaternions  $Q_1$  et  $Q_2$  se note :

$$Q_r = (a_1 a_2 - v_1 \cdot v_2, a_1 v_2 + a_2 v_1 + v_1 \wedge v_2)$$

La valeur prise par ces réels peut être calculée de manière à représenter n'importe quelle rotation de l'espace. Une rotation quelconque de l'espace peut être définie par un axe de rotation (un vecteur  $v' = (x, y, z)$ ) et un angle de rotation ( $\theta$ ). Le quaternion correspondant à cette rotation est défini par :

$$\begin{cases} a = \cos\left(\frac{\theta}{2}\right) \\ v = v' \sin\left(\frac{\theta}{2}\right) \end{cases}$$

Partant donc d'une orientation de base (colinéaire au vecteur  $(1, 0, 0)$ ), il est possible de définir l'orientation d'un os grâce à un quaternion. Un vecteur peut être converti en quaternion en laissant nulle sa partie scalaire ( $a$ ) et en affectant chaque coordonnée du vecteur  $(x, y, z)$  au triplet de réels restant  $(b, c, d)$ . Considérant un vecteur  $V$  sous forme de quaternion, et un quaternion de rotation  $Q$ , le résultat de la rotation de  $V$  par  $Q$  est obtenu en calculant  $V' = QVQ^*$ . De plus, deux rotations sous la forme de quaternion peuvent être composées en les multipliant.

Toutes les rotations utilisées dans cette thèse sont représentées sous la forme de quaternion. Il est important de garder à l'esprit que le positionnement et l'orientation des os du squelette lors des processus de cinématique directe ou inverse font intervenir des quaternions.

### 2.3. Cinématique inverse

Pour comprendre ce qu'est la cinématique inverse il est d'abord nécessaire de comprendre ce qu'est la *cinématique directe*. Comme indiqué dans la section précédente, les principaux acteurs de la prise de posture du corps sont les articulations. Ces dernières, par un jeu de rotations successives permettent l'orientation des différents os du corps humain de façon à lui faire prendre une posture particulière. Pour toute articulation  $a$  du corps, on note  $a_i$  le  $i$ -ème degré de liberté de l'articulation  $a$  avec  $1 \leq i \leq 3$ . Enfin, on note  $t(a)$  la rotation locale créée par les degrés de liberté de l'articulation  $a$ .

Le seul moyen d'agir sur le corps humain est de définir un jeu d'angles s'appliquant aux degrés de liberté du corps. L'application des transformations successives sur une *chaîne cinématique* – c'est-à-dire une suite d'articulations connectées par des os – nous donne les coordonnées cartésiennes  $(x, y, z)$  de n'importe quelle articulation à un instant donné. On note  $f(a_1, a_2, \dots, a_n)$  la fonction permettant d'obtenir la position de l'articulation  $a_n$  grâce à la série de transformations des articulations  $a_1, a_2, \dots, a_n$ . Cette fonction est dite fonction de *cinématique directe*. De manière générale, la cinématique directe permet à partir d'un ensemble de transformations, d'obtenir la position dans l'espace d'un *effecteur* (i.e. d'une partie mobile du corps).

Cependant, l'animation d'un personnage considère généralement le problème sous l'angle opposé : Le système a une connaissance des positions désirées des effecteurs, mais ne connaît pas précisément les angles à affecter aux degrés de liberté pour atteindre une telle position. On appelle fonction de cinématique inverse la fonction  $f(x, y, z) = (t(a_1), t(a_2), \dots, t(a_n))$  qui, à la position d'un effecteur dans l'espace, associe un ensemble de transformations (et plus précisément d'angles de rotation) à affecter à la chaîne cinématique. Les méthodes de cinématique inverse permettent donc de déterminer les angles à attribuer au squelette pour le placer dans une configuration particulière. Pour la suite de cette thèse, nous étendons cette définition à la résolution d'un ensemble d'angles sur le corps permettant de placer différentes articulations du corps dans l'espace mais aussi d'orienter certains os selon des vecteurs précis. De plus, cette définition basique montre le cas où l'on souhaite placer un seul effecteur. Il est courant que le système doive en fait résoudre un ensemble de contraintes (placements et/ou orientations).

Les méthodes de résolution de cinématique inverse peuvent être de deux types : *analytiques* ou *numériques*. Les méthodes analytiques reposent sur une résolution du problème de manière purement trigonométrique. Les solutions ne sont cependant trouvables que sur des chaînes de cinématique comportant peu de degrés de libertés. Les solutions analytiques les plus courantes s'appliquent sur des chaînes cinématiques allant jusqu'à six degrés de liberté (Zhao et al. 1994). Passé ce seuil, les équations deviennent trop complexes et les solutions analytiques n'existent plus.

Les méthodes numériques quant à elles reposent sur des heuristiques considérant un problème de cinématique comme un problème d'optimisation. On cherchera alors à minimiser de manière itérative :

- La distance entre la position courante de l'effecteur et la position où l'on souhaite le placer (la *cible*) dans le cas d'un placement.
- L'angle formé entre l'os en position courante et le vecteur selon lequel on souhaite l'aligner dans le cas d'une orientation.

Les méthodes de cinématique inverse sont nombreuses et variées. Elles sont présentées plus en détail dans la section 1 du chapitre III. Du fait de leur nature heuristique ces méthodes déterminent toujours une solution approchée. Cette solution peut donc être de

mauvaise qualité dans le cas d'un problème mal conditionné (par exemple dans le cas où le squelette doit atteindre une position en dehors de sa portée).

## Annexe 2 : Définition du squelette et dépendances

### 1. Normes des blocs terminaux du squelette

Groupe fonctionnel	Nom du bloc terminal	Translation par rapport au parent
<b>Colonne vertébrale</b>	Bassin	-
	Taille	2.040
	Thorax	2.040
	Cou	1.740
	Tête	1.740
<b>Bras</b>	Clavicule	1.500
	Epaule/bras	2.240
	Coude/avant bras	5.000
<b>Main</b>	Poignet	3.500
	Scaphoïde	0.320
	Trapèze	0.100
	Métacarpe pouce	0.200
	Proximale pouce	0.630
	Distale pouce	0.450
	Métacarpe index	1.390
	Proximale index	0.570
	Intermédiaire index	0.297
	Distale index	0.294
	Métacarpe majeur	1.257
	Proximale majeur	0.655
	Intermédiaire majeur	0.361
	Distale majeur	0.312
	Métacarpe annulaire	1.060
	Proximale annulaire	0.598
	Intermédiaire annulaire	0.332
	Distale annulaire	0.313
	Métacarpe auriculaire	1.050
	Proximale auriculaire	0.440
Intermédiaire auriculaire	0.240	
Distale auriculaire	0.280	

Ces données ne tiennent pas compte de la latéralisation gauche/droite.

## 2. Dépendances articulaires connues

- Flexion de la phalange distale par rapport à la phalange intermédiaire (Losson 2000):

$$\theta_{dist} = \frac{2}{3}\theta_{inter}$$

- Distribution des rotations entre l'épaule et la clavicule (Badler et al. 1999). En fonction d'une antépulsion/rétropulsion totale ( $\theta$ ) et d'une élévation/abaissement total ( $\varphi$ ) :

$$\text{Élévation de la clavicule : } \varphi_c = \cos(\theta)\beta_1 + (1 - \cos(\theta))\beta_2 - 90$$

$$\text{Antépulsion de la clavicule : } \theta_c = 0.2\theta$$

$$\text{Élévation de l'épaule : } \varphi_s = \varphi - \varphi_c$$

$$\text{Antépulsion de l'épaule : } \theta_s = \theta - \theta_c$$

Avec  $\beta_1$  et  $\beta_2$  deux réels définis par :

$$\beta_1 = \begin{cases} 0.2514\varphi + 91.076 & \text{si } 0 \leq \varphi \leq 131.4 \\ -0.035\varphi + 128.7 & \text{si } \varphi > 131.4 \end{cases}$$

$$\beta_2 = \begin{cases} 0.21066\varphi + 92.348 & \text{si } 0 \leq \varphi \leq 130.0 \\ 120.0 & \text{si } \varphi \geq 130.0 \end{cases}$$



## Annexe 3 : Paramètres du système de génération

Cette section regroupe les paramètres utilisés par le système de génération, leur valeur utilisée dans le chapitre IV et leur rôle.

Paramètre	Rôle
<b>Configurations d'exploration</b>	Le nombre de valeurs aléatoires tirées par le système Monte-Carlo lors de la phase d'exploration
<b>Distance angulaire minimale</b>	Lors de la phase d'exploration, les valeurs tirées sont choisies de manière à être les plus écartées les unes des autres. Cette valeur permet de choisir une distance seuil minimale à respecter
<b>Configurations d'optimisation</b>	Le nombre de configurations sélectionnées après la phase d'exploration pour être optimisées
<b>Itérations</b>	Le nombre d'itérations maximum lors de la phase d'optimisation
<b>Pas de correction</b>	Un scalaire permettant une mise à l'échelle de chaque pas d'optimisation, de manière à avancer plus ou moins vite
<b>Seuil d'acceptabilité</b>	Seuil au-delà duquel les solutions sont considérées comme acceptables.



## Annexe 4 : Comparatif des systèmes de génération

Cette annexe présente un comparatif des différents systèmes de génération utilisés par les équipes de recherche dont nous avons connaissance.

Méthode de génération	Laboratoire/Projet	Projet en activité	Génération automatique des mains	Utilisation d'un modèle anatomique
<b>Automatique</b>	<i>LIMSI/CNRS</i>	<i>Oui</i>	<i>Oui</i>	<i>Oui</i>
	<i>UEA (Angleterre)</i>	<i>Oui</i>	<i>Non</i>	<i>Non</i>
	<i>ILSP (Grèce)</i>	<i>Non</i>	<i>Non</i>	<i>Non</i>
	<i>DFKI (Allemagne)</i>	<i>Oui</i>	<i>Non</i>	<i>Non</i>
	<i>ZCU (République Tchèque)</i>	<i>Oui</i>	<i>Non</i>	<i>Non</i>
	<i>HCTLab (Espagne)</i>	<i>Oui</i>	<i>Non</i>	<i>Non</i>
<b>Capture de mouvements</b>	<i>SignCom</i>	<i>Oui</i>	-	<i>Non</i>
	<i>NYU (EUA)</i>	<i>Oui</i>	-	<i>Non</i>
	<i>ATLAS (Italie)</i>	<i>Oui</i>	-	<i>Non</i>
<b>Rotoscopie</b>	<i>LIMSI/CNRS</i>	<i>Oui</i>	-	<i>Non</i>
<b>?</b>	<i>De Paul (EUA)<sup>43</sup></i>	<i>Oui</i>	<i>Inconnu</i>	<i>Inconnu</i>

<sup>43</sup> Malgré un grand nombre de publications, l'université de De Paul n'a jamais véritablement détaillé son principe de génération. Nous savons cependant qu'ils utilisent un modèle paramétrique pour la génération des configurations des mains (présentés dans (McDonald et al. 2001)) mais tout porte à croire qu'ils utilisent des configurations pré-calculées des mains.



## Annexe 5 : Exemple de descriptions linguistiques

### 1. HamNoSys

Séquence [IMMEUBLE] :

•• ⊃ △ 0 ☐ ↑ ] [

**Explication :**

*HamNoSys* est un modèle paramétrique, c'est-à-dire que la définition d'un signe est le contrôle de quatre paramètres : position, configuration, orientation et mouvement de la main. Par ailleurs, un cinquième paramètre d'expression du visage peut être spécifié en dehors de la notation du signe. Les valeurs de ces paramètres sont choisies dans un ensemble fini d'éléments, simplifiant ainsi la notation mais limitant la portée du modèle aux ensembles prédéfinis.

Le symbole •• indique un signe bi-manuel où les mains vont se déplacer symétriquement par rapport au plan sagittal. Le symbole ⊃ indique la configuration manuelle en U formée par le pouce et l'index. Les symboles △ et 0 indiquent l'orientation de la main. Le premier spécifie l'axe directeur de la main (la paume), et le second l'axe normal. Le symbole suivant, ☐, indique où se place le signe devant le signeur. Ici, le symbole indique que le signe se place au milieu du torse. Enfin, les deux derniers symboles caractérisent le mouvement : ↑ « vers le haut » et ] [ « s'arrête au niveau du cou ». En guise de comparaison, le signe [IMMEUBLE] est représenté sur la Figure 55.

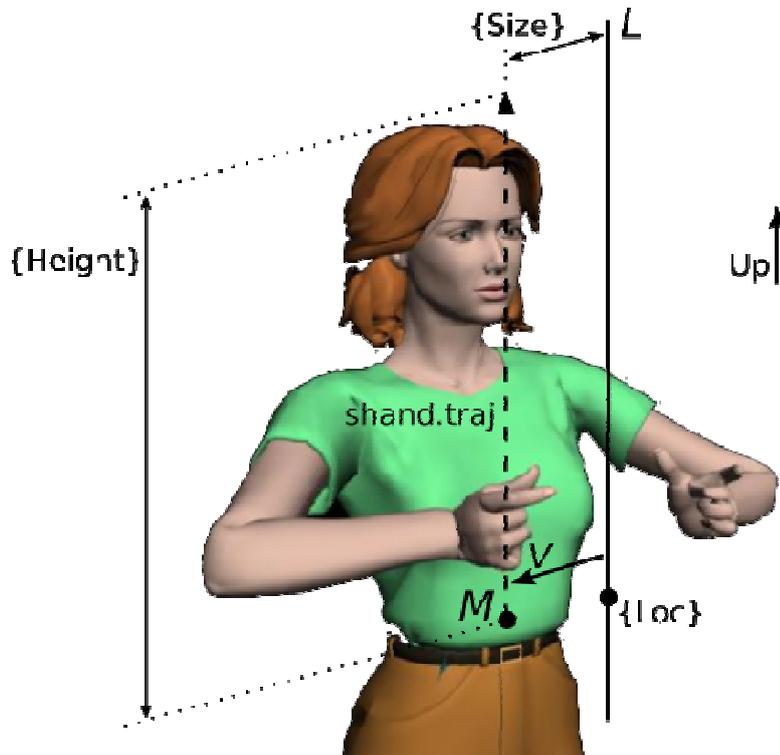


Figure 55 Le signe [IMMEUBLE] en LSF.

## 2. Zebedee

### Séquence [IMMEUBLE] :

**SEQUENCE** "immeuble"

```
DEP loc = @ABST + <FWD | medium>
DEP height = large
DEP foundation = medium
```

**KEY\_POSTURE**(0) {

**KEEP :**

```
For $h=s,w
  #L_closed($h)
  #R_closed($h)
  #M_closed($h)
  #wrench($h, small)
  Orient !palm($h) _|_ UP
End
```

```
Orient NRM!palm(w) along -NRM!palm(s)
```

**HERE:**

```
For $h=s,w
  Place @T_TIP($h) at [loc] - <DIR!index($h,3) | [foundation]>
End
```

}

**TRANSITION**(10) {

**Accel** 1

**For** \$h =s,w

```
Arc @PA($h) : <0,0,0>
```

**End**

}

```

KEY_POSTURE (5) {
  HERE:
    For $h=s,w
      Place @T_TIP($h) at [loc] - <DIR!index($h,3) | [foundation]> +
<UP| [height]>
      End
    }
End "immeuble"

```

### Explication :

Ce signe est composé de deux postures-clé (`Key_posture`) et d'une transition. Avant le début de la description, on définit trois dépendances contextuelles (`loc`, `height` et `foundation`). Ces trois dépendances ont des valeurs par défaut (par exemple `large` pour `height`). Ces valeurs peuvent changer au moment de la résolution du signe si par exemple on souhaite modifier l'emplacement où le signe se déroule (`loc`) ou la taille virtuelle de l'immeuble (`height`).

Chaque élément de la séquence dure un certain temps (respectivement 0, 10 et 5 unités de temps). Ces unités de temps sont abstraites et ne servent qu'à indiquer la durée relative des éléments entre eux.

Analysons la première posture clé :

```

KEY_POSTURE (0) {
  KEEP :
  For $h=s,w
    #L_closed($h)
    #R_closed($h)
    #M_closed($h)
    #wrench($h, small)
    Orient !palm($h) _|_ UP
  End

  Orient NRM!palm(w) along -NRM!palm(s)

  HERE:
  For $h=s,w
    Place @T_TIP($h) at [loc] - <DIR!index($h,3) | [foundation]>
  End
}

```

Tout d'abord, rappelons que la résolution d'une posture équivaut à la résolution d'un système de contraintes géométriques appliquées sur le corps. Ces contraintes peuvent être appliquées ponctuellement (`HERE`) ou sur toute la durée de la posture et de la transition qui suit (`KEEP`). Les contraintes sont divisées en deux types : placements d'un site articulaire dans l'espace (`Place`) et orientation de l'axe d'un os selon un vecteur (`Orient`). Par ailleurs,

il est possible d'appeler des « macros » (e.g. `#L_closed($h)`), qui sont des morceaux de posture enregistrés et utilisées comme raccourcis pour la définition d'autres postures. L'exemple `#L_closed($h)` correspond au placement de l'annulaire en position « fermé » (i.e. contre la paume).