

Vers la modélisation du geste par la vision

Towards gesture modeling by vision

par Hocine OUHADDI et Patrick HORAIN

Institut National des Télécommunications
Unité de projets Artemis, département Signal et Image
9, rue Charles Fourier, 91011 Evry Cedex
<http://www-sim.intevry.fr/~hand>
prénom.nom@intevry.fr

résumé et mots clés

Nous présentons nos travaux sur la conception, l'ajustement et le recalage d'un modèle 3D articulé de la main pour son suivi dans des séquences d'images monoscopiques, sans recourir à des marqueurs. Après un état de l'art, nous décrivons notre modèle 3D générique de la main et son ajustement à la morphologie de l'opérateur sur une image de sa main ouverte. Nous comparons ensuite différentes fonctionnelles et méthodes d'optimisation pour le recalage. Une originalité de notre approche réside dans la prise en compte des contraintes biomécaniques dans la procédure de recalage, ce qui permet de réduire considérablement l'espace de recherche en évitant l'obtention de configurations irréalistes. Enfin, des résultats sur des séquences d'images sont présentés. A terme, un tel système d'acquisition du geste par la vision artificielle pourrait être utilisé pour des interfaces homme-machine par le geste, pour l'animation d'humains virtuels, le codage à très bas débit du geste pour la téléprésence, ou la reconnaissance de la langue des signes.

Modèle articulé, anthropométrie, biomécanique, ajustement morphologique, recalage, suivi.

abstract and key words

In this paper, we present our work on the design, adjustment and registration of a 3D articulated model of the hand on monoscopic images sequences, without markers. After a review of the state of the art, we describe our 3D generic model of the hand and how it is adjusted to the operator's morphology on an image of his open hand. Then, we compare various cost functions and optimisation methods to register the 3D model of the hand. As an innovative contribution, we integrate biomechanical constraints into the optimisation procedure. This considerably reduces the search space by excluding unrealistic postures. Finally, we show results on image sequences. Gesture capture by artificial vision could be a valuable help for applications such as human-machine interaction by gesture, for virtual humanoid animation, low bit rate coding of gesture for telepresence, or sign language recognition.

Articulated model, anthropometry, biomechanics, morphological adjustment, registration, tracking.

1. introduction

Les gestes sont un moyen naturel de communication. Dans la vie courante, ils viennent ponctuer ou renforcer l'expression orale entre personnes. Avec la langue des signes, ils constituent le moyen de communication privilégié pour de nombreux malentendants. Bien des machines sont conçues pour être pilotées par des gestes à travers des interfaces mécaniques ou électroniques, par exemple un volant de voiture, une commande de robot télémanipulateur, ou simplement une souris d'ordinateur.

Les possibilités de communication par le geste avec les machines se sont enrichies avec l'apparition de gants instrumentés, munis de nombreux capteurs mesurant les angles des articulations et la position de la main dans l'espace. Ceux-ci ont été utilisés pour commander des acteurs virtuels, pour désigner et manipuler des objets graphiques sur des écrans d'ordinateurs [Baudel93, Bellik96] ou même pour la reconnaissance de la langue des signes [Braffort96]. Malheureusement, de tels gants restent chers et fragiles, et leurs câbles de liaison constituent une entrave. C'est pourquoi de plus en plus de travaux de recherche s'intéressent à la vision artificielle pour l'acquisition des gestes.

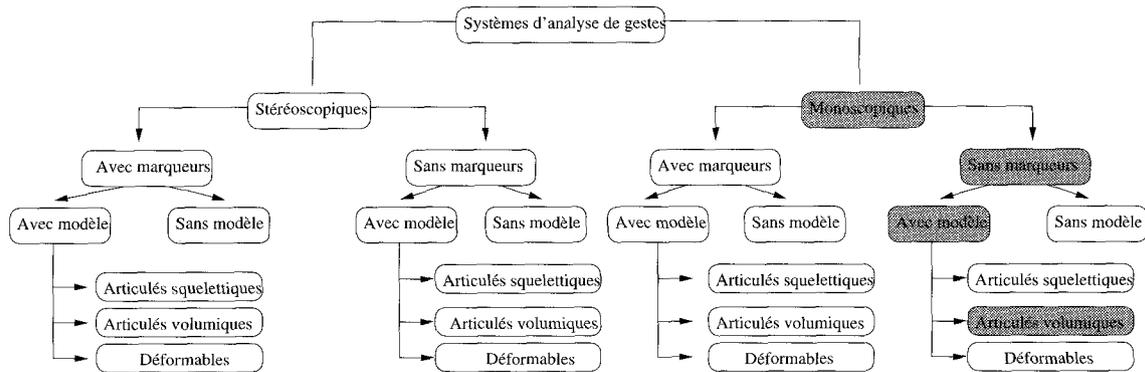


Figure 1. – Systèmes d'analyse de gestes. En gris : le type de système proposé.

L'acquisition des gestes de la main pour des applications d'interaction homme-machine ou la reconnaissance de la langue des signes avec des caméras nécessite de localiser la main dans l'image, puis d'estimer des paramètres 2D ou 3D tels que les positions des bouts des doigts ou les angles des articulations. L'analyse des gestes peut faire appel à des gants colorés ou des marqueurs sur la main [Iwai96, Hienz96], à une ou plusieurs caméras et à une analyse 2D ou 3D des gestes avec ou sans modèle pour représenter la main (figure 1).

Les méthodes d'analyse 2D de gestes se limitent à la reconnaissance d'un nombre limité de configurations prédéfinies, le plus souvent à l'issue d'une étape d'apprentissage. Par exemple, Berard *et al.* [Berard95] ont développé une application de dessin qui fournit en temps réel la position d'un doigt, par corrélation sur le bout du doigt. Cootes *et al.* [Cootes95] ont proposé une approche par modèles déformables utilisant un modèle de distribution de points, qui représente la silhouette d'une forme quelconque par un ensemble de points caractéristiques et des modes de variation qui décrivent les déplacements de ces points. Ce modèle est constitué à partir d'un ensemble d'apprentissage par une analyse en composantes principales des différences entre les formes de l'ensemble d'apprentissage et la forme moyenne. La reconnaissance est réalisée à partir du modèle de distribution des points. Cette démarche a également été proposée par Martin et Crowley [Martin97]. Toutefois, les méthodes 2D ne permettent pas un traitement simple et général des auto-occultations (figure 2).

Les méthodes par analyse 3D remplacent la connaissance *a priori* de l'ensemble des gestes attendus par une connaissance *a priori* de l'objet observé (la main). Elles permettent une analyse plus riche du geste, car de nature tridimensionnelle. Pour cette raison, nous avons choisi une approche par analyse et synthèse consistant à comparer les formes extraites des images avec celles obtenues par projection d'un modèle générique.

Nous présentons d'abord un état de l'art des méthodes d'analyse de gestes par recalage d'un modèle 3D, puis quelques rappels d'anatomie, d'anthropométrie et de biomécanique de la main. Notre modèle générique 3D articulé ainsi que son ajustement

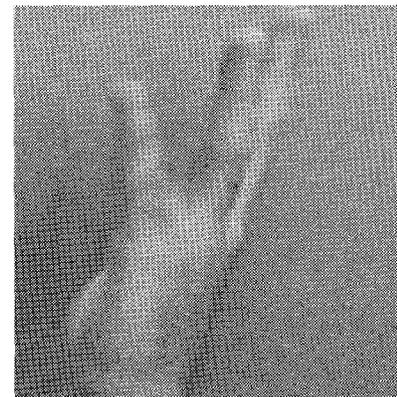


Figure 2. – Image de la main avec auto-occultations.

sur la morphologie de la main est ensuite décrit, et nous comparons différentes approches pour le recalage du modèle. Enfin, nous présentons notre méthode pour le recalage sous contraintes biomécaniques, ainsi que des résultats sur des séquences d'images.

2. la reconnaissance de gestes par recalage d'un modèle 3D

Les méthodes d'analyse de gestes par ajustement et recalage d'un modèle générique 3D à partir d'une ou plusieurs vues de la main utilisent différents types de modèles : articulés squelettiques ou volumiques, ou encore de type déformable. Les premiers représentent la main par un ensemble de primitives rigides (segments de droite ou volumes) et de repères locaux associés aux différentes articulations. Les seconds décrivent les déformations de l'enveloppe de la main.

2.1. modèles squelettiques articulés

Dorner [Dorner94] a proposé une méthode de reconnaissance de la langue des signes utilisant un modèle squelettique dont les degrés de liberté sont estimés par minimisation itérative des distances entre les projections des articulations du modèle et les points correspondants dans l'image. La détermination de ces derniers est facilitée par l'utilisation d'un gant coloré.

Une méthode similaire a été proposée par Lee et Kunii [Lee95] pour l'identification de postures de la main. Les points caractéristiques sont les bouts des doigts. Les positions 3D de points caractéristiques marqués sont obtenues par stéréovision. Les valeurs des degrés de liberté sont ensuite estimées en utilisant la cinématique inverse. Les contraintes biomécaniques permettent de choisir parmi des solutions multiples.

Millar et Crawford [Millar96] déterminent directement (sans itération) les valeurs des degrés de liberté du modèle à partir des positions 3D des bouts des doigts et du poignet désignées interactivement dans des paires d'images stéréo.

2.2. modèles surfaciques déformables

Dans le modèle de Kuch et Huang [Kuch], la surface de la main est représentée par des β -splines. La reconnaissance de gestes est effectuée par ajustement du modèle puis suivi de la main. L'ajustement fournit de manière supervisée à partir de 3 vues de la main un modèle dont les mesures morphologiques sont similaires à celles de la main de l'opérateur. Les valeurs des degrés de liberté des articulations sont estimées par une procédure itérative de minimisation de l'écart entre la silhouette de la main dans l'image et celle du modèle après projection.

Heap et Hogg [Heap96] utilisent un modèle surfacique à base de maillage simplexe. La méthode de suivi proposée procède par déformation du maillage et se limite à l'estimation des paramètres globaux de mouvement du modèle (rotations, translations et facteur d'échelle) par appariement des contours occultants du modèle et ceux extraits de l'image en utilisant une variante 3D du modèle de distribution de points dont l'ensemble des points des contours externes de la silhouette est remplacé par les points du maillage.

2.3. modèles articulés volumiques

A notre connaissance, la première utilisation d'un modèle 3D articulé volumique pour le suivi de la main dans une séquence d'images a été proposée par Regh et Kanade [Regh94]. Le modèle est composé de cylindres. L'ensemble des vues possibles de chaque doigt est discrétisé. Les degrés de liberté des articulations sont estimés par minimisation d'une fonctionnelle de coût combinant les distances entre les projections des bouts des doigts

du modèle et les points correspondants dans l'image, l'inclinaison des axes des phalanges estimés par rapport à ceux mesurés à partir de l'image, ainsi que les différences entre les doigts dans l'image et les vues correspondantes. Deux caméras sont utilisées pour traiter les auto-occultations.

La méthode proposée par Davis et Shah [Davis94] construit un modèle de type cylindre généralisé pour chaque phalange après extraction des doigts (de la main ouverte). La procédure de suivi se limite à l'estimation des paramètres globaux de mouvement (translations et rotations) en se basant sur le calcul du flot optique.

Mochimaru et Yamazaki [Mochimaru94] utilisent une procédure similaire à celle proposée par Kuch et Huang, à la différence que le modèle est composé de cônes plutôt que de β -splines.

Shimada *et al.* utilisent un modèle volumique à base de cylindres et estiment tous les degrés de libertés du modèle à partir d'une seule caméra. Leur méthode consiste à sélectionner pour chaque image les postures possibles du modèle parmi un ensemble fini, à calculer leurs projections et à prendre la posture qui recouvre au mieux la silhouette extraite de l'image. Le nombre de postures candidates est réduit par l'examen des postures dans les images précédentes dans la séquence [Shimada95]. Cette méthode ne fournit qu'une estimation grossière des degrés de liberté. L'utilisation d'un filtre de Kalman étendu et l'ajout de contraintes biomécaniques permettent d'améliorer la précision de l'estimation des paramètres et de traiter les auto-occultations [Shimada98].

2.4. conclusion

Un modèle générique 3D renferme une connaissance *a priori* sur la main. Son utilisation permet d'analyser une plus grande diversité de gestes que les méthodes 2D. La gestion des déformations du maillage des modèles déformables est plus complexe que celle des articulations d'un modèle articulé. Les méthodes par recalage d'un modèle articulé squelettique recourent à des marqueurs sur les points caractéristiques de la main ou une intervention de l'opérateur. C'est pourquoi notre choix s'est porté sur l'utilisation d'un modèle volumique articulé (figure 1) recalé sur une vue unique et sans utilisation de marqueurs en tirant profit de contraintes morphologiques et biomécaniques de la main.

3. la main

3.1. anatomie de la main

Le squelette de la main comporte les os du poignet, les métacarpes et les os des doigts. L'articulation de chaque doigt sur son métacarpe est appelée métacarpo-phalangienne (MCP). Les articulations entre phalanges sont appelées interphalangiennes proximales (IPP) et interphalangiennes distales (IPD), ou, pour le

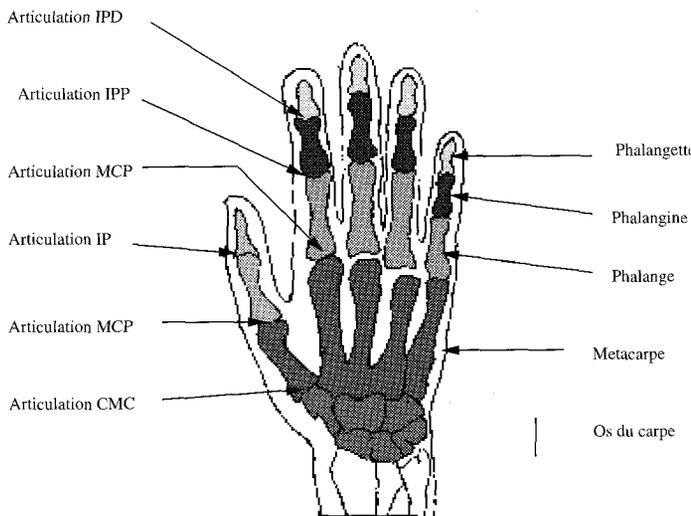


Figure 3. – Squelette de la main.

pouce, simplement interphalangiennes (IP). L'articulation à la base du pouce est dite carpo-métacarpienne (CMC) (figure 3).

3.2. anthropométrie de la main

L'anthropométrie est l'étude comparative des mesures du corps humain. Une des études les plus élaborées, à notre connaissance, est celle de Wagner [Wagner88] réalisée sur un échantillon de 127 hommes âgés de 17 à 63 ans et 111 femmes âgées de 16 à 64 ans, et qui repose sur 48 mesures morphologiques. A partir de ces mesures, un certain nombre de contraintes anthropométriques sont formulées sous forme de relations entre mesures et d'intervalles de valeurs possibles.

Pour tirer profit des mesures morphologiques, nous avons adopté une démarche similaire à celle de Buchholz [Buchholz92, Teo94]. Elle consiste à formuler, à partir de ces mesures, un certain nombre de contraintes anthropométriques sous forme de relations entre mesures. Formellement, une contrainte anthropométrique définit l'intervalle de variation d'une mesure morphologique L_i sous la forme :

$$(Rap_i - \Delta_i) \times L_M \leq L_i \leq (Rap_i + \Delta_i) \times L_M \quad (1)$$

où L_M est la longueur de la main, $Rap_i \times L_M$ est la moyenne de cette mesure et $\Delta_i \times L_M$ sa variation.

Quelques mesures morphologiques et les contraintes anthropométriques correspondantes sont par exemple :

- Largeur de la paume/longueur de la main : 0.44 ± 0.007
- Longueur de la paume/longueur de la main : 0.496 ± 0.003
- Longueur de l'index/longueur de la main : 0.449 ± 0.002
- Longueur de majeur/longueur de la main : 0.512 ± 0.004
- Longueur de l'annulaire/longueur de la main : 0.494 ± 0.007

- Longueur 1^{ère} phalange de l'index/longueur de la main : 0.245 ± 0.001
- Longueur 2^{ème} phalange du majeur/longueur de la main : 0.170 ± 0.003
- Longueur 3^{ème} phalange de l'annulaire/longueur de la main : 0.107 ± 0.003

3.3. biomécanique de la main

La main possède 6 degrés de liberté au niveau du poignet : 3 rotations et 3 translations. Deux degrés de liberté, l'abduction/adduction (figure 4) et la flexion/extension (figure 5), sont associées aux articulations MCP de chaque doigt, 1 autre (flexion/extension) au niveau de chaque articulation interphalangienne. D'après [Buchholz92], l'articulation CMC du pouce possède 3 degrés de liberté : abduction/adduction, flexion/extension et une pseudo-rotation due à l'incongruité entre les os du carpe et la base du métacarpe du pouce, et au relâchement des ligaments les reliant. Au total, la main possède donc 28 degrés de liberté.

Les mouvements des doigts de la main sont régis par des contraintes biomécaniques qui rendent certaines postures irréalisables [Kuch95, Lee95, Millar96, Teo94, Buchholz92, Rijpkema91]. Ces contraintes sont statiques ou dynamiques. Les contraintes statiques traduisent les limites des angles d'abduction/adduction ou de flexion/extension possibles des différentes articulations. Par exemple, l'angle de flexion/extension des premières phalanges des quatre doigts est compris entre -110° et 15° .

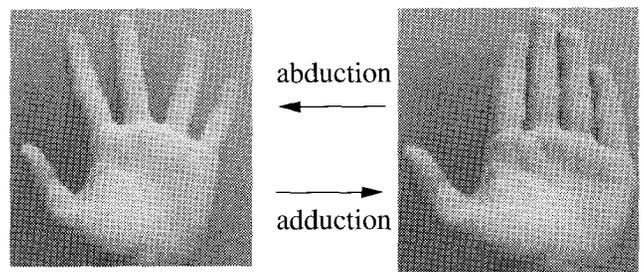


Figure 4. – Mouvements d'adduction et d'abduction.

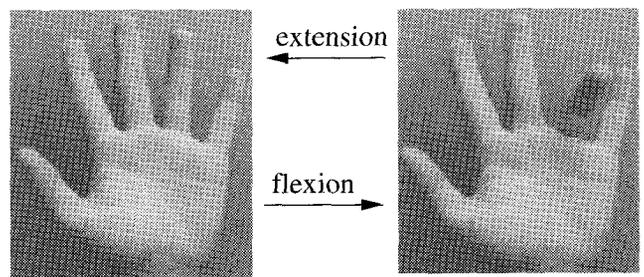


Figure 5. – Mouvements de flexion et d'extension.

Les contraintes dynamiques représentent les relations entre les degrés de liberté des articulations des doigts. Dans ce qui suit :

- $\theta_{f/e}^{\text{articulation doigt}}$ (respectivement $\theta_{a/a}^{\text{articulation doigt}}$) représente l'angle de flexion/extension (respectivement abduction/adduction) de l'articulation en question. Ainsi $\theta_{f/e}^{\text{MCP Index}}$ représente le degré de flexion de la première phalange de l'index. MCP, IPP et IPD désignent respectivement les première, seconde et troisième articulations des 4 grands doigts. Pour le pouce, les articulations sont notées CMC, MCP et IP.
- $\theta_{\text{degré de liberté min}}^{\text{articulation doigt}}$ et $\theta_{\text{degré de liberté max}}^{\text{articulation doigt}}$ représentent les valeurs minimales ou maximales des degrés de liberté en question.
- les valeurs angulaires sont toutes données en degrés.

Les contraintes dynamiques pour les doigts autres que le pouce sont les suivantes :

- La relation entre les flexions de la phalangette et de la phalangine :

$$\theta_{f/e}^{\text{IPP}} = \frac{3}{2} \theta_{f/e}^{\text{IPD}} \quad (2)$$

Cette contrainte traduit le fait que, en l'absence de force externe appliquée sur le doigt, il est généralement impossible de bouger indépendamment la phalangette et la phalangine à cause du tendon qui les relie et traverse tout le doigt.

- L'interdépendance entre la flexion/extension et l'abduction/adduction de chaque phalange. En effet, plus l'angle de flexion/extension est grand, plus l'abduction ou l'adduction des doigts est limitée (figure 6). Par exemple, la valeur maximale de l'angle d'abduction/adduction de la phalange dans le cas où l'angle de flexion/extension n'est pas nul est donnée par :

$$\theta_{a/a}^{\text{MCP doigt}} \leq \theta_{a/a \text{ max}}^{\text{MCP doigt}} - \frac{\theta_{a/a \text{ max}}^{\text{MCP doigt}}}{\theta_{f/e \text{ max}}^{\text{MCP doigt}}} \times \theta_{f/e}^{\text{MCP doigt}} \quad (3)$$

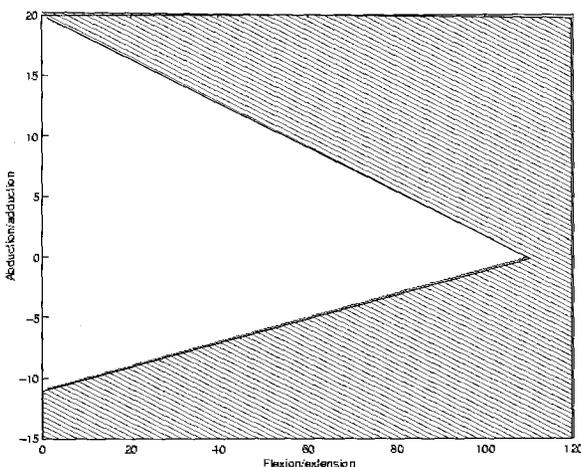


Figure 6. – Contrainte d'interdépendance entre l'abduction/adduction et la flexion/extension d'un doigt.

où $\theta_{f/e \text{ max}}^{\text{MCP}}$ est la valeur maximale de l'angle de flexion/extension de la phalange.

- La relation entre les flexions/extensions des phalanges de doigts voisins; par exemple, la flexion de la phalange de l'index ou de l'annulaire entraîne la flexion de la phalange du majeur :

$$\theta_{f/e}^{\text{MCP Majeur}} \geq \max \left(\begin{array}{l} \theta_{f/e}^{\text{MCP Index}} - 25^\circ \\ 3\theta_{f/e}^{\text{MCP Annulaire}} - 45^\circ \\ 3\theta_{f/e \text{ min}}^{\text{MCP Majeur}} \end{array} \right) \quad (4)$$

où $\theta_{f/e \text{ min}}^{\text{MCP Majeur}}$ est la valeur minimale de l'angle de flexion/extension de la phalange de majeur.

A cause de la morphologie particulière du pouce, ses mouvements sont soumis à un ensemble de contraintes différentes. Ainsi, la flexion du pouce au niveau de l'articulation CMC est comprise entre -80° et 25° . De plus, deux contraintes dynamiques relient les flexions/extensions et les abductions/adductions des deux premières articulations du pouce :

$$\theta_{f/e}^{\text{MCP}} = 2(\theta_{f/e}^{\text{CMC}} - 30^\circ) \quad (5)$$

$$\theta_{a/a}^{\text{MCP}} = \frac{7}{5} \theta_{a/a}^{\text{CMC}} \quad (6)$$

Nous avons présenté un bref rappel sur l'anatomie de la main. Nous avons vu que ses mesures morphologiques obéissent à des *contraintes anthropométriques*. Bien que la main soit fortement articulée, ses mouvements sont régis par des *contraintes biomécaniques* qui imposent des limites sur les différents degrés de liberté ou des interdépendances. Ces différentes contraintes sont intégrées dans notre modèle informatique de la main décrit ci-après.

4. modèle de la main

4.1. structure du modèle

Pour représenter les gestes, nous avons choisi de décrire la main par un modèle de type volumique articulé. Ce modèle comporte 2 parties. La première décrit la structure articulée du modèle par un graphe représentant les translations et les rotations associées aux différents degrés de liberté de la main (degrés de liberté globaux, abductions/adductions, flexions/extensions et torsions). La seconde partie représente la forme des éléments articulés composant le modèle. Cette dissociation de la structure et de la morphologie du modèle permet de changer facilement la forme des primitives du modèle, par exemple de passer facilement d'un modèle constitué de primitives géométriques simples (cube, cylindres et sphères) à un modèle plus précis, composé de maillages (figure 7). Ce changement est facilité par la séparation entre les modules de

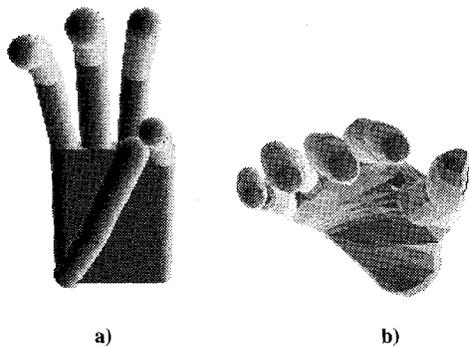


Figure 7. – Modèles de la main. a) Modèle à base de primitives simples. b) Modèle à base de primitives maillées.

morphologie et de biomécanique : seule la partie morphologie est modifiée, par opposition à d'autres modèles [Regh94].

La cohérence des mesures morphologiques est assurée par l'intégration de 21 contraintes anthropométriques : 5 sur les longueurs des doigts, 14 pour les longueurs des phalanges et 2 sur les dimensions de la paume.

4.2. degrés de liberté du modèle

Notre modèle est conforme au paramétrage MPEG-4 [Preteux98]. Ceci, permettra de se placer dans un cadre normalisé pour le codage et l'échange de fichiers d'animation. Dans le cas de l'analyse de gestes, les forces extérieures sont nulles. La torsion au niveau des articulations MCP, qui est intégrée dans le modèle MPEG-4, est donc ignorée.

Nous avons vu précédemment que les biomécaniciens dénombrent 28 degrés de liberté pour la main. Toutefois, la pseudo-rotation de l'articulation CMC et l'abduction de l'articulation MCP du pouce sont habituellement négligeables [Dorner94]. Ainsi, notre modèle possède 26 degrés de liberté à ajuster en respectant les contraintes biomécaniques. Les angles de flexion/extension des articulations interphalangiennes des quatre grands doigts d'une part, et des articulations du métacarpe du pouce d'autre part étant liés par les contraintes dynamiques vues précédemment, le nombre de degrés de liberté est ramené de 26 à 21.

4.3. animation du modèle

Pour animer le modèle, nous avons choisi une technique explicite consistant à spécifier entièrement les paramètres de chaque plan de la séquence, tout comme l'application de suivi devra les fournir. Par exemple, des séquences d'animation représentant quelques lettres de l'alphabet de la langue des signes (figure 8) ont été créées de cette façon.

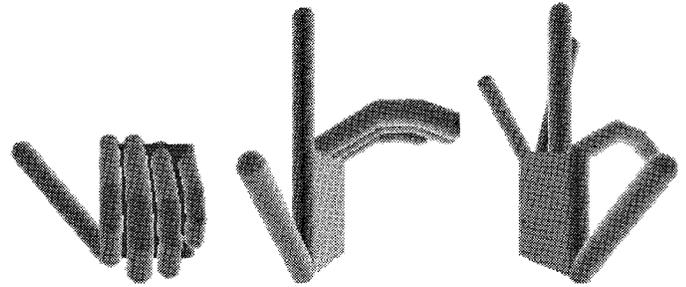


Figure 8. – Synthèse de signes de l'alphabet de la langue des signes (A, D, F respectivement).

Après ajustement, le modèle articulé est destiné à être animé pour reproduire les mouvements de la main dans des séquences vidéo. Les méthodes d'animation procédurale et déclarative [Teo94] peuvent utiliser la dynamique inverse [Rijkema91], l'analyse par éléments finis ou l'interpolation de plans [Gourret89].

Le modèle de la main a été développé en VRML et animé au moyen du logiciel Open Inventor sur station Silicon Graphics [Wernecke94].

5. ajustement morphologique

Le modèle générique 3D ne correspond pas en général à la morphologie particulière de la main d'un opérateur. L'ajustement préalable des paramètres de morphologie du modèle est nécessaire pour la robustesse du suivi de la main dans des séquences. Notre approche, non supervisée, utilise une seule image de la main ouverte (doigts écartés).

5.1. segmentation de l'image

Pour extraire de l'image la silhouette de la main, nous avons utilisé l'information de couleur. La distribution des couleurs de la peau est apprise sur un échantillon à partir de l'histogramme des teintes et des saturations (peu sensible aux conditions d'éclairage de la scène).

La silhouette de la main est considérée comme étant la plus grande composante connexe correspondant aux couleurs sélectionnées. Les trous de la silhouette dus au bruit sont remplis (figure 9). D'autres algorithmes de segmentation à base de teinte chair ou par le mouvement éventuellement en présence d'arrière-plans complexes peuvent être utilisés. On pourra consulter sur ce sujet les actes de la série de conférences : International Conference on Automatic Face and Gesture Recognition, 1995 (Zurich, Suisse), 1996 (Killington, USA), ou 1998 (Nara, Japon).

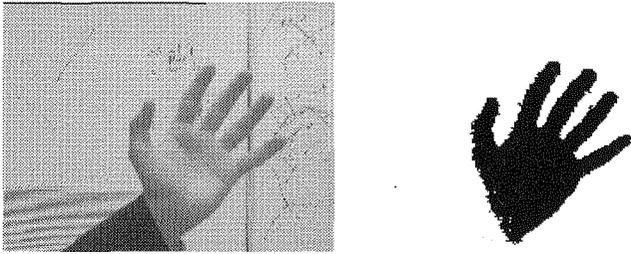


Figure 9. – Segmentation de la main par seuillage sur la couleur.

L'estimation des paramètres morphologiques du modèle requiert l'identification des doigts et de la paume dans l'image. La silhouette de la main est filtrée par des opérateurs de morphologie mathématique et les trous sont comblés par ouverture et fermeture. Les doigts sont extraits par un chapeau haut-de-forme [Serra] dont l'élément structurant est un disque de diamètre légèrement supérieur à leur largeur. Les doigts étant écartés, ils correspondent aux plus grandes composantes connexes (figure 10). Ils sont ordonnés par un balayage perpendiculaire à l'axe d'inertie de la silhouette. Le poignet est localisé approximativement comme étant le centre de gravité des points d'intersection de l'axe d'inertie de la main et ceux des 4 grands doigts qui sont écartés. La paume est obtenue par soustraction des régions correspondant aux doigts et à l'avant-bras.

5.2. ajustement morphologique

L'étape de segmentation fournit la silhouette des différentes parties de la main. Les paramètres de morphologie des doigts (rayon et longueur) et de la paume (longueur et largeur) sont calculés à partir des dimensions des régions correspondantes à un facteur d'échelle près. L'épaisseur de la paume est supposée égale au diamètre du pouce.

Les contraintes anthropométriques vues au paragraphe 3.2 permettent de valider la cohérence des mesures morphologiques effectuées sur l'image : longueur des doigts et dimensions de la paume. Elle sont également utilisées pour déterminer les longueurs des phalanges du modèle.

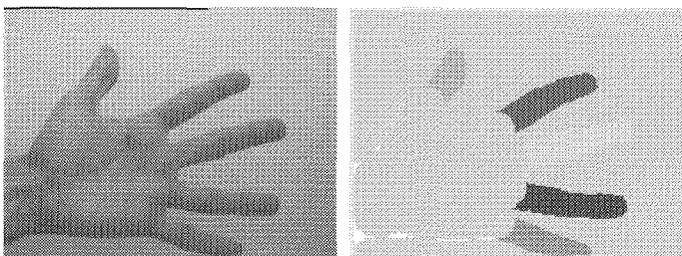


Figure 10. – Détection des doigts.

6. suivi de la main : fonctionnelles et optimisations

Après avoir ajusté les dimensions du modèle à la morphologie de l'opérateur, nous nous intéressons au suivi de geste, c'est-à-dire au recalage du modèle sur des images de la main en position quelconque. Il s'agit de trouver les paramètres de position globale et des articulations qui minimisent la différence entre le modèle projeté et les primitives correspondantes dans l'image.

Ci-après, nous présentons un état de l'art des méthodes de recalage de modèle pour l'analyse du geste. Puis, nous définissons et comparons deux fonctionnelles qui mesurent l'écart entre la projection du modèle et les primitives de l'image. Enfin, nous comparons différents algorithmes d'optimisation, et présentons des résultats de recalage.

6.1. analyse de gestes par recalage : état de l'art

Mochimaru et Yamazaki [Mochimaru94] puis Kuch et Huang [Kuch95] ont minimisé l'écart entre la silhouette du modèle projeté et celle extraite de l'image par une procédure de « perturbation locale de degrés de liberté ». Celle-ci consiste à modifier la valeur de chaque degré de liberté d'une petite quantité. Si la fonctionnelle diminue, la modification est répétée dans la même direction, sinon la direction de perturbation est inversée. Les méthodes à base de perturbation locale de degrés de liberté semblent assez efficaces en pratique. Toutefois, les différents paramètres doivent être choisis de manière empirique.

Ohya et Kishino [Ohya] utilisent un algorithme génétique pour recalculer un modèle représentant la partie supérieure du corps humain sur des paires d'images stéréo (images de synthèse), par minimisation de l'écart entre les silhouettes projetée et extraite de l'image. Les résultats présentés sont obtenus à partir d'une population d'environ 1000 solutions initiales et nécessite plusieurs centaines d'itérations, ce qui est très coûteux en temps de calcul.

Gavrila et Davis [Gavrila96] recalculent un modèle du corps humain sur des paires d'images stéréo par minimisation d'une fonctionnelle de distance entre les contours du modèle projeté et ceux extraits des images. La minimisation est réalisée après discrétisation de l'espace de paramètres et recherche hiérarchique dans un arbre représentant l'ensemble des configurations possibles. Shimada *et al.* [Shimada98] recalculent un modèle de la main sur une séquence d'images monoscopiques par génération et sélection d'un ensemble discret de postures candidates en fonction d'une part du taux de recouvrement entre la silhouette de modèle et celle extraite de l'image, et d'autre part des distances entre les contours

Tableau 1.

Méthode de minimisation/Fonctionnelle	Distances entre contours	Surface de non recouvrement
Perturbation de degrés de libertés	–	[Mochimaru94] [Kuch95]
Algorithmes génétiques	–	[Ohya94]
Approche probabiliste	[Shimada98]	[Shimada98]
Recherche arborescente	[Gavrila96]	–

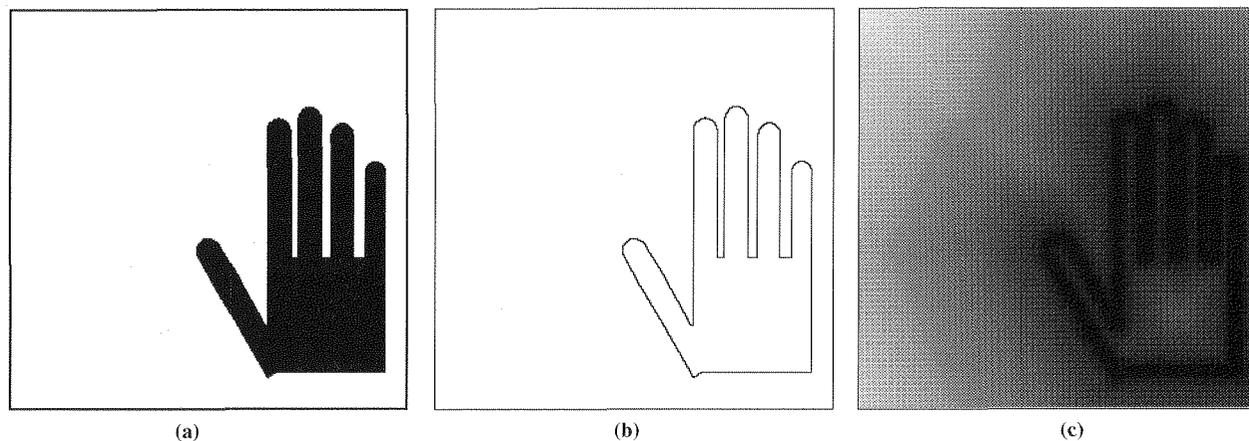


Figure 11. – Silhouette, ses contours et la carte de distance correspondante. a) image de la silhouette, b) contours de la silhouette, c) carte de distance.

des doigts dans l’image et les contours du modèle projeté. Le nombre de postures générées est réduit en fonction des postures dans les images précédentes en utilisant une loi de dépendance statistique entre configurations successives. La discrétisation de l’espace des paramètres nécessite un compromis entre la précision et le coût de recalage. (voir tableau 1)

Nous nous intéressons au recalage d’un modèle 3D de la main en utilisant une seule vue. Nous étudions ci-après la mise en œuvre de fonctionnelles utilisant la surface de non recouvrement ou les distances entre contours, et leurs minimisation dans l’espace continu des paramètres.

6.2. fonctionnelles pour le recalage

6.2.1. taux de non-recouvrement

Nous le définissons par :

$$F = \frac{SurfNonRecouv(silModèle, silImage)}{Surface(silModèle) + Surface(silImage)} + \frac{SurfNonRecouv(silImage, silModèle)}{Surface(silModèle) + Surface(silImage)} \quad (7)$$

où la surface non recouverte $SurfNonRecouv(sil1, sil2)$ est le nombre de pixels appartenant à la silhouette $sil1$ et non à $sil2$.

$Surface(sil)$ est la surface de la silhouette sil . Cette fonctionnelle étant localement plate lorsque les objets ne se recouvrent pas, le processus itératif de minimisation doit être initialisé à partir d’une position présentant un recouvrement partiel. Ceci peut être généralement réalisé en ramenant le centre de gravité de la projection du modèle au centre de gravité de la silhouette dans l’image.

6.2.2. distances entre contours

La fonctionnelle à minimiser est la somme des distances des pixels de la silhouette du modèle projeté à la silhouette extraite de l’image normalisée par les longueurs des contours. La carte de distances est calculée par un algorithme de chanfrein [Borgefors86] (figure 11).

Soit I l’ensemble des points des contours extraits de l’image et M l’ensemble des points de projections des contours occultants du modèle. La distance $DC(M, I)$ est obtenue en sommant les distances entre chaque point dans l’ensemble M et le point de I le plus proche. Formellement :

$$DC(M, I) = \frac{1}{|M|} \sum_{m \in M} \text{distance}(m, I) \quad (8)$$

où $|M|$ est le cardinal de l’ensemble M .

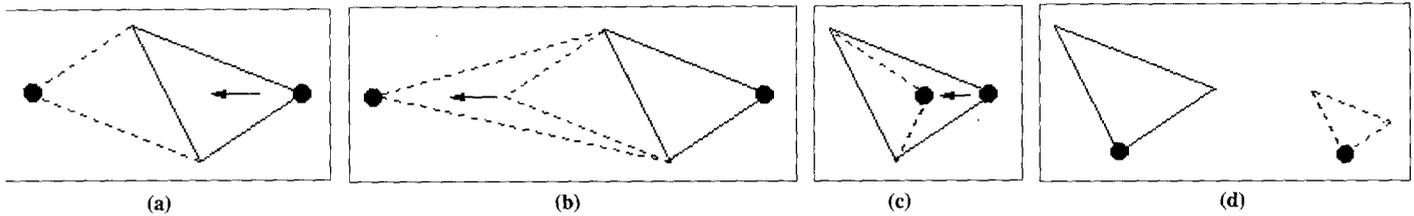


Figure 12. – Transformations géométriques du simplexe. a) Symétrie, b) Expansion, c) Contraction 1D, d) Contraction ND.

6.3. méthodes d'optimisation

Lorsque les dérivées partielles de la fonctionnelle à optimiser sont disponibles, il est possible de les utiliser pour une descente de gradient [Ouhaddi98b]. Toutefois, leur estimation pour les fonctionnelles considérées est malaisée. Nous avons donc privilégié des méthodes d'optimisation ne nécessitant pas l'estimation des dérivées.

6.3.1. descente du simplexe

La descente du simplexe (downhill simplex) est une méthode d'optimisation *itérative* due à Nelder et Mead [Nelder65].

Soit n le nombre de paramètres à optimiser. La première étape consiste à générer $n + 1$ points dans l'espace R^n à partir d'une solution initiale. Ces $n + 1$ points forment un simplexe, par exemple un triangle dans R^2 ou un tétraèdre dans R^3 .

Le simplexe initial est construit en appliquant la règle suivante :

$$p_i = \lambda_i * e_i + p_0 \quad (9)$$

où p_i est le $i^{\text{ème}}$ sommet du simplexe, e_i le vecteur unitaire dont la $i^{\text{ème}}$ composante est non nulle et λ_i est la longueur de l'arête p_0p_i . Le simplexe initial subit une suite de transformations géométriques pour atteindre un minimum de la fonctionnelle, en adaptant sa forme à la topographie de la fonctionnelle dans l'espace de recherche. Ces différentes transformations sont (figure 12) :

- une symétrie,
- une expansion,
- une contraction,
- une contraction multi-dimensionnelle.

Ces opérations peuvent être mises en œuvre suivant diverses heuristiques [McKinnon96]. Nous avons utilisé celle de Press *et al.* [Press92] qui consiste à trier les sommets du simplexe suivant la valeur de la fonctionnelle pour sélectionner trois sommets ayant respectivement le meilleur score, le plus mauvais score et l'avant-dernier. Le sommet du plus mauvais score est remplacé par symétrie du simplexe par rapport à l'hyperplan défini par les autres sommets, ce qui revient à faire une recherche dans la direction

opposée à celle qui a donné le mauvais score. Si la fonctionnelle est améliorée alors une expansion est effectuée, sinon une contraction mono-dimensionnelle est réalisée. Si malgré cela la fonctionnelle ne s'améliore pas, l'algorithme effectue une contraction multi-dimensionnelle dans plusieurs directions. L'enchaînement de ces différentes opérations est illustré dans la figure 13. Ce processus est répété jusqu'à ce que la variation relative de la fonctionnelle sur les sommets du simplexe soit proche de la précision du calculateur.

6.3.2. méthode de Powell

Il existe de nombreuses méthodes pour minimiser une fonction monodimensionnelle sans utiliser les dérivées, telles que la recherche par dichotomie, la méthode de Fibonacci ou la méthode de la section dorée [Press92].

Pour minimiser une fonction multidimensionnelle, il est possible d'effectuer une succession de minimisations unidimensionnelles pour chacun des paramètres à optimiser jusqu'à convergence. Cependant, si la fonctionnelle possède une vallée qui n'est pas orientée suivant un axe de l'espace paramétrique, cette technique multiplierait des déplacements petits qui se traduiraient par une très lente progression dans la vallée. La méthode de Powell [Press92] améliore ce processus en modifiant la direction des recherches unidimensionnelles pour effectuer des explorations mieux orientées par rapport à la direction de la vallée.

Soit X le vecteur de n paramètres à optimiser. Désignons par d_i ($i = 1 \dots n$) les directions linéairement indépendantes utilisées pour rechercher l'optimum.

- Initialisation : Lors de la première itération, les d_i sont choisis comme étant les axes de l'espace paramétrique.

- Itérations :

- Etape 1 : en partant du vecteur de paramètres X^0 trouver X^n en effectuant n minimisations unidimensionnelles successives dans les directions d_i .

- Etape 2 : déterminer X^{n+1} en effectuant une $(n + 1)^{\text{ème}}$ minimisation dans la direction :

$$d_{n+1} = X^n - X^0 \quad (10)$$

- Etape 3 : remplacer la direction ayant conduit à la plus grande amélioration du critère à l'étape 1 par d_{n+1} et itérer.

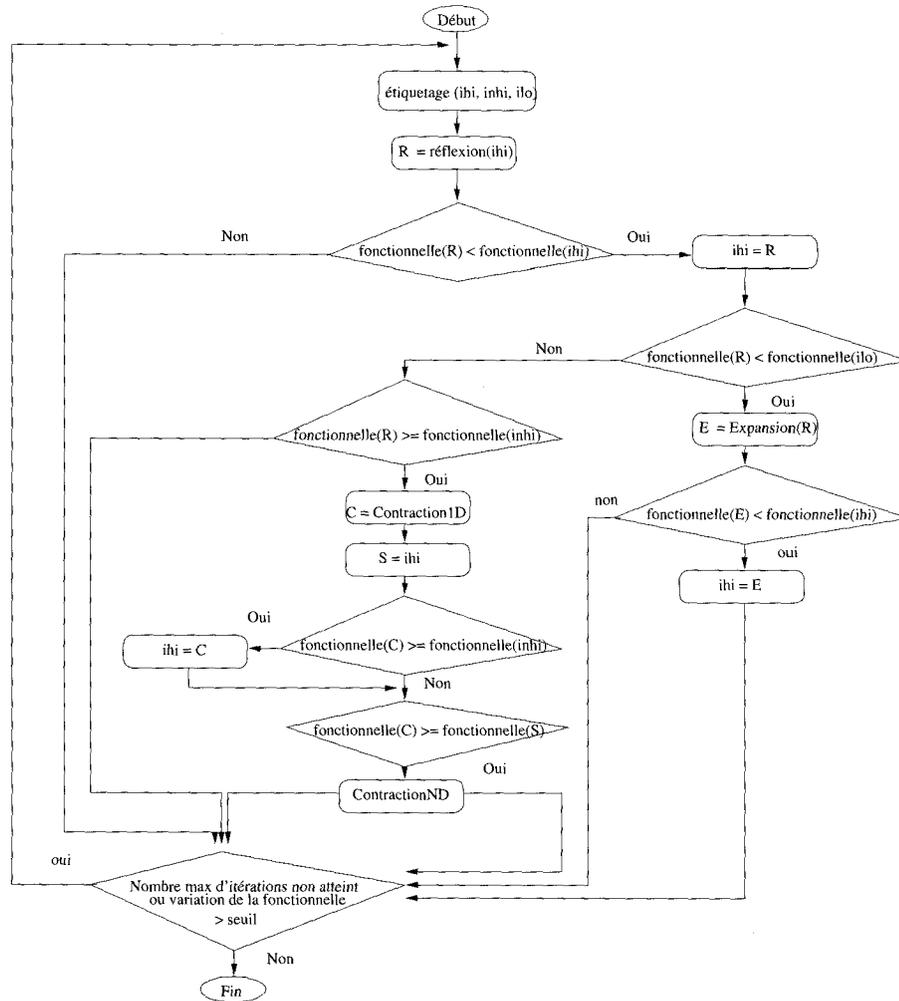


Figure 13. – Algorithme par descente du simplexe. ilo, inhi et ihi représentent les sommets ayant respectivement le meilleur score, le plus mauvais score et l'avant-dernier.

La minimisation à l'étape 2 s'effectue dans la direction moyenne des déplacements effectués à l'étape 1. Elle permet donc des déplacements mieux orientés par rapport à la vallée. Le fait d'éliminer à l'étape 3 la direction qui conduit aux meilleurs résultats a pour but d'éviter que les d_i ne deviennent linéairement dépendants au fil des itérations. Signalons que la méthode de Powell nécessite plus d'évaluations de la fonctionnelle par itération que la méthode de descente du simplexe [Press92].

La fonctionnelle de distance nécessite de déterminer et de projeter dans l'image les contours occultants du modèle. Pour notre modèle constitué de primitives géométriques simples, ces contours sont des segments de droite repérés par leurs extrémités et tracés en projection dans l'image par l'algorithme de Bresenham [Foley95]. Dans le cas d'un modèle à primitives maillées, ces contours peuvent être obtenus par adaptation d'un algorithme de Z-Buffer [Bourges97].

6.4. comparaison des méthodes d'optimisation et des fonctionnelles

Nous avons comparé expérimentalement ces fonctionnelles et ces méthodes d'optimisation pour estimer les 21 paramètres de notre modèle de main par recalage sur des images.

6.4.1. projection du modèle dans l'image

L'estimation des fonctionnelles retenues nécessitent le calcul de la projection de la silhouette du modèle ou bien des contours occultants. Ce calcul est effectué en projetant (projection orthographique) un ensemble de points caractéristiques du modèle. La position dans l'espace d'un point du modèle articulé est calculée par multiplication des matrices en coordonnées homogènes

associées aux degrés de liberté regroupés dans le vecteur p des paramètres du modèle. La matrice de transformation globale est :

$$M_{\text{globale}} = T.R.S \quad (11)$$

où T représente une translation dans l'espace, R une rotation et S une homothétie. La transformation des composantes articulées du modèle est également calculée par multiplication de coordonnées homogènes. Par exemple, la transformation pour la phalange d'un doigt peut être calculée à partir de :

- M_{doigt} la matrice de translation de la base du doigt (articulation MCP),
 - $M_{a/a}^{\text{phal1}}$ et $M_{f/e}^{\text{phal1}}$ les matrices de rotation de l'abduction/adduction et flexion/extension de la phalange, et
 - $M_{f/e}^{\text{phal2}}$ la matrice de rotation de la flexion/extension de la phalange,
- par la relation :

$$M(p) = M_{\text{globale}} \cdot M_{\text{doigt}} \cdot M_{a/a}^{\text{phal1}} \cdot M_{f/e}^{\text{phal1}} \cdot M_{f/e}^{\text{phal2}} \quad (12)$$

Désignant par X les coordonnées d'un point de la phalange dans le repère lié à celle-ci, les coordonnées absolues $X(p)$ de ce point du modèle articulé sont :

$$X(p) = M(p).X \quad (13)$$

6.4.2. initialisation du recalage

Les paramètres estimés pour une image donnée sont utilisés pour initialiser le recalage sur l'image suivante. Pour les séquences réelles, le suivi est initialisé à partir d'une image de la main ouverte après ajustement morphologique du modèle. Les paramètres du modèle sont estimés en supposant que toutes les flexions sont nulles et que l'axe optique de la caméra est perpendiculaire au plan de la main. L'orientation globale du modèle est alors fixée comme étant celle de l'axe d'inertie de la silhouette de la main. Les angles d'abduction/adduction des articulations sont estimés à partir des axes d'inertie des doigts déterminés lors de l'étape de l'ajustement. Les translations dans le plan image sont calculées en ramenant le poignet du modèle à la position du poignet estimée dans l'image. Le facteur d'échelle est déterminé ensuite à partir des positions des bouts des doigts. Cette initialisation est affinée en appliquant dès la première image le processus de recalage pour le suivi (figure 14).

6.4.3. choix d'une méthode d'optimisation et d'une fonctionnelle

Nous avons comparé la minimisation par les algorithmes de descente du simplexe et de Powell, et sur les deux fonctionnelles de coût en utilisant soit des surfaces de non recouvrement, soit des distances entre contours. Les tests ont été effectués sur des séquences d'images de synthèse et d'images naturelles.

Nous avons obtenu les résultats suivants :

- La minimisation par descente du simplexe de la fonctionnelle utilisant des surfaces de non recouvrement permet de traiter des mouvements rigides, des mouvements d'abduction/adduction (figure 15) et de flexion (figure 16). Si la procédure d'optimisation converge avec une surface de non recouvrement résiduelle alors elle a convergé vers un optimum local qui n'est pas satisfaisant. Ce problème est généralement réglé en réinitialisant la procédure d'optimisation avec un simplexe de volume plus grand.
- Les résultats obtenus en utilisant la méthode de Powell sont en général équivalents à ceux de descente du simplexe. Toutefois,

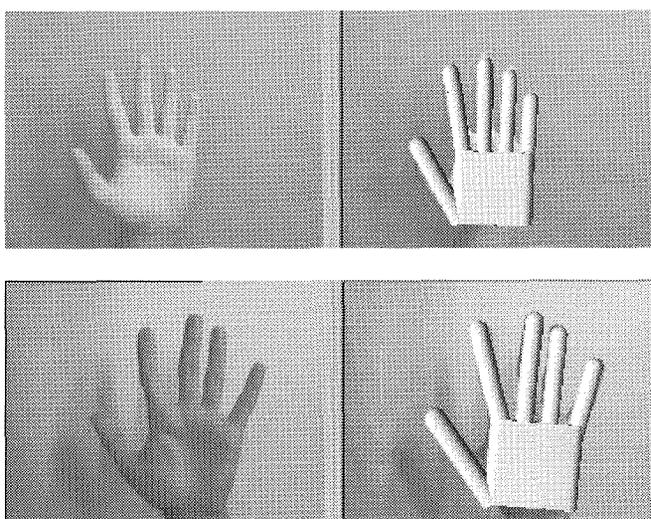


Figure 14. – Exemples d'initialisation de suivi sur la main ouverte.

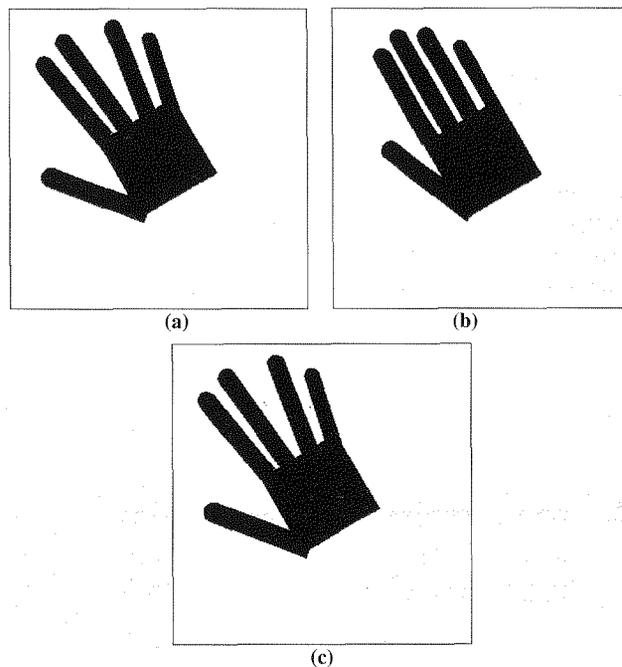


Figure 15. – Recalage par descente du simplexe sur la fonctionnelle utilisant la surface de non recouvrement. a) Image à recalage, b) Initialisation du recalage, c) Fin de recalage.

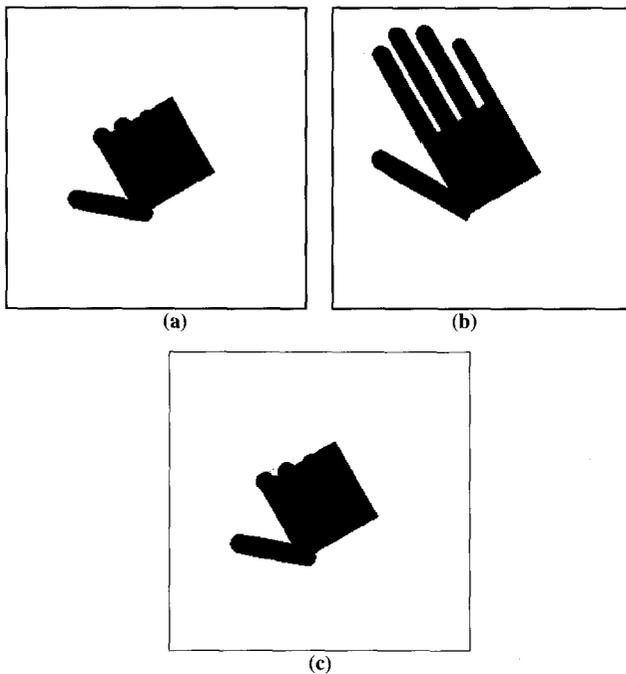


Figure 16. – Traitement flexion/extension. a) Image à recaler, b) Initialisation du recalage, c) Fin de recalage.

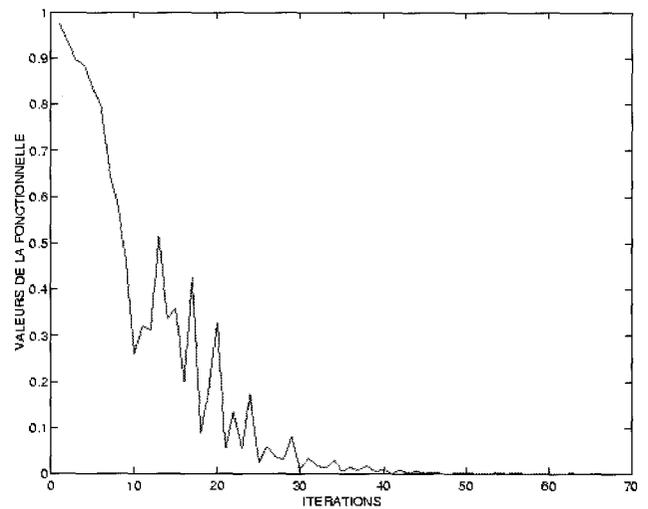
le nombre d'évaluations de la fonctionnelle est beaucoup plus grand (figure 17). En effet, la méthode de Powell fait appel à une procédure de minimisation mono-dimensionnelle qui effectue plusieurs évaluations de la fonctionnelle afin de déterminer un petit intervalle contenant un minimum de la fonctionnelle dans une direction donnée.

- Le recalage avec une fonctionnelle utilisant les distances entre contours semble très sensible à l'initialisation et au bruit dans l'image. Ceci est dû au fait que la transformation distance effectuée en chaque point un appariement implicite avec le contour le plus proche. Si la configuration initiale est très éloignée de l'optimum recherché, ou si les arêtes extraites de l'image sont bruitées, de faux appariements peuvent conduire à un mauvais recalage du modèle.

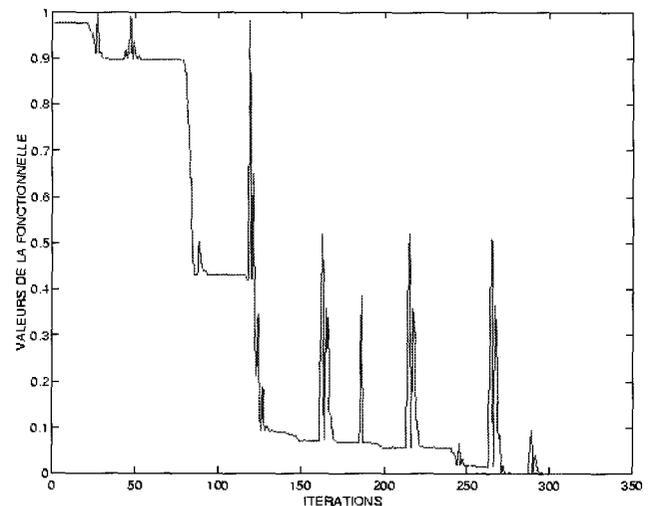
Ces résultats nous ont conduit à retenir la fonctionnelle utilisant les surfaces de non recouvrement et la méthode d'optimisation avec descente du simplexe.

6.4.4. intégration des contraintes biomécaniques

Les contraintes biomécaniques constituent une connaissance *a priori* sur les configurations possibles de la main que nous introduisons dans l'algorithme d'optimisation. Ces contraintes linéaires définissent des demi-espaces dans l'espace des paramètres, dont l'intersection est un domaine convexe. Nous imposons alors à chaque transformation du simplexe de générer un simplexe contraint dans ce domaine. Cette stratégie originale permet de respecter les contraintes biomécaniques tout en réduisant l'espace de recherche.



a) Descente du simplexe



b) Powell

Figure 17. – Exemple de décroissance de la fonctionnelle pour un recalage : la méthode de Powell nécessite 308 évaluations de la fonctionnelle pour 18 itérations, alors que la descente du simplexe est effectuée en 32 itérations pour seulement 63 évaluations.

Vu que le domaine des contraintes est convexe, chaque contraction du simplexe à l'intérieur de cette région crée un simplexe qui respecte les contraintes. Par conséquent, il suffit de tester les contraintes seulement après les symétries et les expansions. Si un sommet du simplexe est déplacé à l'extérieur de cette région, alors celui-ci est remplacé par un nouveau sommet situé à l'intersection de l'hyperplan de contrainte et de la droite qui relie ce nouveau sommet à l'ancien (figure 18).

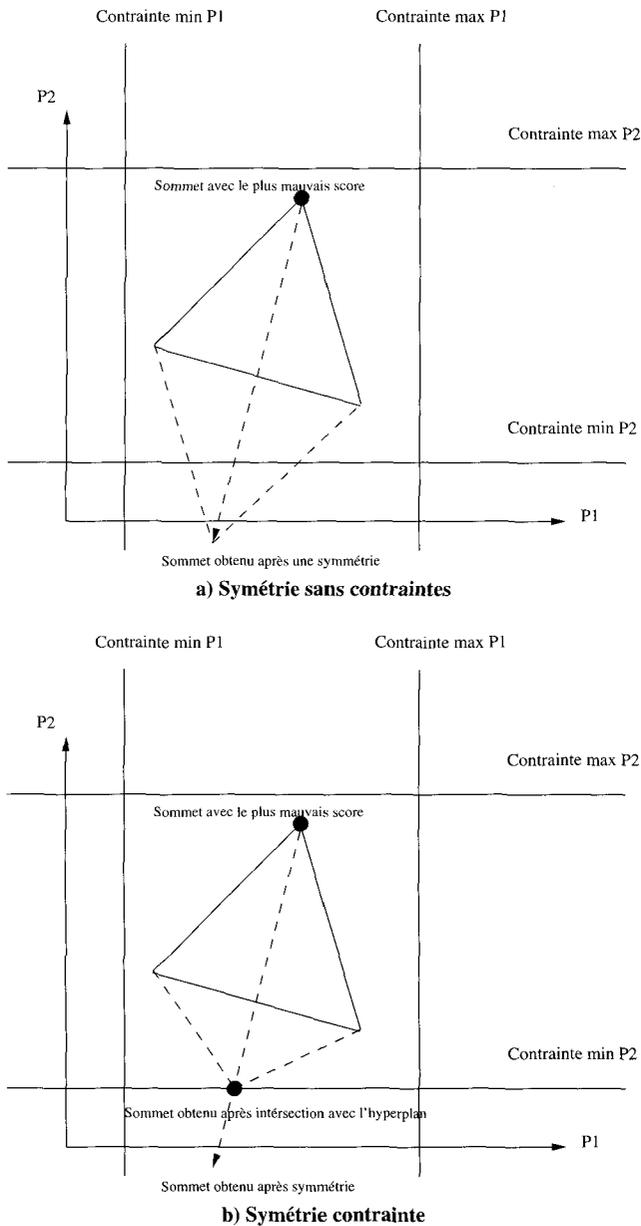


Figure 18. – Exemple de transformation contrainte du simplexe.

6.5. résultats

Nous avons testé notre méthode sur des séquences naturelles présentant différents mouvements. Notre modèle actuel ne comporte pas d'avant-bras pour ne pas introduire de biais dans l'évaluation du recalage. En particulier lorsque la fonctionnelle utilisée est le taux de non-recouvrement, nous avons interactivement enlevé l'avant-bras dans les images. Notons que la localisation du poignet pourrait être automatisée [Deimel99].

Parmi les 21 degrés de liberté du modèle, les 3 rotations globales, le facteur d'échelle et les flexions interphalangiennes n'ont

pas été pris en compte dans les résultats présentés dans les figures 19, 20, 21, 22, 23. Les résultats obtenus sont satisfaisants.

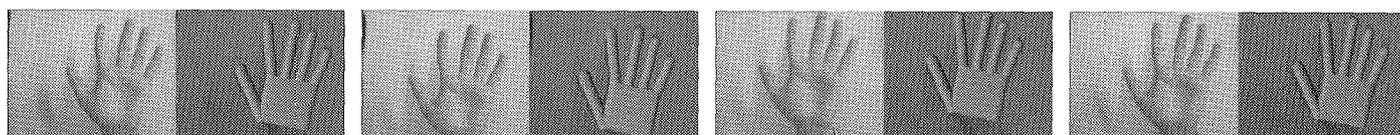
Dans le cas de la flexion, il peut arriver que la troisième et la seconde phalange soient devant la paume, et apparaissent à l'intérieur de la silhouette. Aucune des deux fonctionnelles considérées ne fournit une information sur cette partie de la main, et le recalage de ces phalanges ne peut alors être correctement réalisé. Nous envisageons d'utiliser les contours internes de l'image pour traiter les configurations où les doigts sont devant la paume (figure 22b).

Nous avons également mis en œuvre une technique hiérarchique consistant à recalculer d'abord la position globale seulement, puis à rajouter successivement le facteur d'échelle, les degrés de liberté de rotation, d'abduction et enfin de flexion MCP. Avec ces 16 degrés de liberté, le recalage est moins précis. Nous envisageons d'intégrer l'information de mouvement dominant apparent (flot optique) dans la procédure de recalage pour choisir l'ordre du recalage hiérarchique.

7. conclusion et perspectives

Nous avons présenté une méthode pour le suivi de la main dans une séquence d'images à partir d'une seule vue et sans utilisation de marqueurs. La main est représentée par un modèle 3D articulé. Les mesures morphologiques du modèle sont ajustées automatiquement sur la main ouverte de l'opérateur dans la première image. Nous avons présenté ensuite différentes fonctionnelles de coût et méthodes d'optimisation pour le recalage d'un modèle 3D de la main ainsi que la méthode proposée. Nous avons validé notre approche sur des séquences d'images de synthèse et des séquences naturelles. Les surfaces de non recouvrement donnent de meilleurs résultats que les distances aux contours. Les méthodes d'optimisation par descente du simplexe et de Powell donnent des résultats similaires, mais la descente du simplexe nécessite moins d'opérations d'évaluation de la fonctionnelle. Ceci nous a amené à retenir la fonctionnelle utilisant des surfaces de non recouvrement et la méthode d'optimisation de descente du simplexe. Une originalité de notre approche réside dans la prise en compte des contraintes biomécaniques. Ceci permet de réduire considérablement l'espace de recherche et évite l'obtention de configurations irréalistes. Des résultats sur des séquences naturelles montrent la validité de notre approche. Nous étudions actuellement la prise en compte des configurations d'auto-occultations et de l'information mouvement pour le recalage de modèle articulé.

A long terme les applications possibles concernent la transmission du geste à bas débit, la télévirtualité, l'interprétation de gestes pour des interfaces homme-machine multimodales, la reconnaissance de la langue des signes.



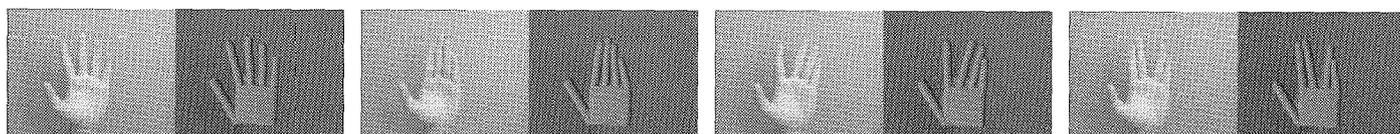
(a) image 0

(b) image 7

(c) image 38

(d) image 47

Figure 19. – Translation de la main.



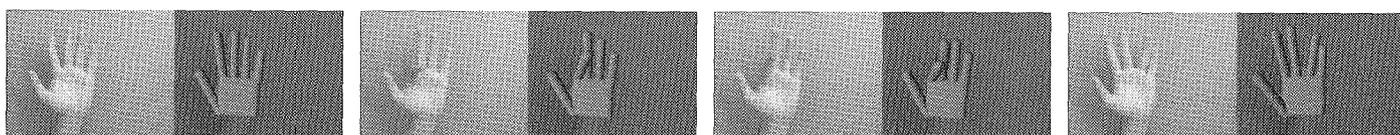
(a) image 0

(b) image 18

(c) image 55

(d) image 67

Figure 20. – Abduction/adduction des doigts



(a) image 0

(b) image 13

(c) image 34

(d) image 44

Figure 21. – Flexion de l'index.



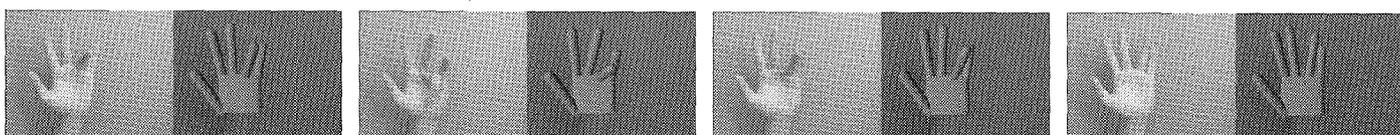
(a) image 10

(b) image 27

(c) image 35

(d) image 41

Figure 22. – Flexion du majeur.



(a) image 12

(b) image 26

(c) image 44

(d) image 50

Figure 23. – Flexion de l'annulaire.

BIBLIOGRAPHIE

- [Baudel93] T. Baudel and M. Beaudouin-Lafon, «Charade : Remote Control of objects using Free-Hand Gestures», *Communications of the ACM*, Vol. 36, no 7, 1993, p. 28-35.
- [Berard95] F. Berard, J. Coutaz et J. L. Crowley, «Finger Tracking as Input Device for Augmented Reality», *Proc. Intel Workshop on Automatic Face and Gesture-Recognition*, Zurich, Switzerland, June 1995.
- [Bellik96] Y. Bellik, «Modality Integration : Speech and Gesture», *Survey of the State of the Art in Human Language Technology*, Section 9.4, R. A. Cole ed in chief, 1996. <http://www.cse.orgi.edu/CSLU/HLTsurvey>
- [Buchholz92] B. Buchholz and T. J. Armstrong, «A Kinematic model of the human hand to evaluate its prehensile capabilities», *J. Biomechanics*, Vol. 25, no 2, 1992, p. 149-162.
- [Borgefors86] G. Borgefors, «Distance Transformations in Digital Images», *Computer Vision, Graphics and Image Processing*, Vol. 34, 1986, p. 344-371.

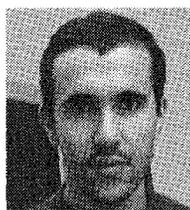
- [Bourges97] M. Bourges-Sevenier, P. Horain, F. Preteux, P. Leray, «Recalage d'un modèle générique sur une séquence d'images 2D», *Actes CORESA'97 : 3^{ème} Journées d'études et d'échanges Compression et Représentation des Signaux Audiovisuels*, Issy-les-Moulineaux, France, 26-27 March 1997, p. 163-171.
- [Braffort96] A. Braffort, *Reconnaissance et Compréhension de gestes, application à la langue des signes*, thèse de l'université de Paris XI, spécialité informatique, juin 1996.
- [Cootes95] T. F. Cootes, C. J. Taylor, D.H. Cooper and J. Graham, «Active Shape Models : their Training and Application», *Computer Vision and Image Understanding*, vol. 61, n 1, p. 38-59, 1995.
- [Davis94] J. Davis, J. and M. Shah, «Determining 3-D Hand Motion», *Proc. 28th of Asimolar Conf. Signals, Systems and Computers*, 1994.
- [Deimel99] B. Deimel and S. Schroter, «Improving Hand-Gesture Recognition via Video Based Methods for the Separation of the forearm from the Human Hand», *Gesture Workshop'99, France, March 1999*.
- [Dorner94] B. Dorner, *Chasing the color glove : visual hand tracking*, Master Thesis, Simon Fraser University, 1994.

- [Foley95] J. D. Foley, A. Van Dam, S. K. Feiner and J. F. Hughes. *Introduction à l'infographie*, Addison-Wesley France 1996.
- [Gavrila96] D. M. Gravila and L. S. Davis, «3-D model-based tracking of humans in action : a multi-view approach», *Proc. IEEE Computer Vision and Pattern Recognition*, San Francisco, 1996.
- [Gourret89] J. Gourret, N. Magnenat-Thalmann and D. Thalmann, «Simulation of Object and Human Skin in a Grasping Task», *Comp. Graphics*, Vol. 23, no. 3, Siggraph'89, 1989, p. 21-30.
- [Hienz96] H. Hienz, K. Grobel and G. offner, «Real-Time Hand Arm Motion Analysing using a Single Video Camera», *Proc. of Gesture Workshop'96*, Univ. of York, p. 323-327.
- [Heap96] T. Heap and D. Hogg, «3D Deformable Hand Models, Progress in Gestural Interaction», *Proc. of Gesture Workshop'96*, March 19th, 1996, University of York, UK, Philip A. Harling and Alstair D. N. Edwards (Eds), p. 131-139.
- [Iwai96] Y. Iwai, Y. Yagi and M. Yachida, «Gesture Recognition using Colored Gloves», *Proc. of ICPR'96*, p. 662-666.
- [Kuch95] J. J. Kuch, and T. S. Huang, «Vision Based Hand Modeling and Tracking for Virtual Teleconferencing and Telecollaboration». In *Proc. of fifth International Conference on Computer Vision*, June 20-23, 1995, MIT, Cambridge, Massachusetts, p. 666-671.
- [Kuno94] Kuno Y. Kuno, M. Sakamoto, K. Sakata and Y. Shirai, «Vision Based Human Interface with User-Centered Frame», *Proc. of IROS'94*, p. 2923-2929.
- [Lee95] J. Lee and T. Kunii, «Model-Based Analysis of Hand Posture». *IEEE Computer Graphics and Applications*, September 1995, p. 77-86.
- [Martin97] J. Martin and J. L. Crowley, «An Appearance-Based Approach to Gesture-Recognition», *Proc. of 9th Conf. on Image Analysis and Processing*, Italy, 1997.
- [McKinnon96] K. I. M. Mc Kinnon, «Convergence of the Nelder-Mead Simplex method to a non-stationary point», Technical Report MS96-000, Department of Mathematics and Statistics, University of Edinburgh, May 1996.
- [Millar96] R. J. Millar, C. F. Crawford, «A Mathematical Model for Hand-Shape Analysis», *Proc. Of Gesture Workshop'96*, March 19th 1996, University of York, UK, Philip A. Harling and Alstair D. N. Edwards (Eds), p. 235-245.
- [Mochimaru94] M. Mochimaru and N. Yamazaki, «The three-dimensional measurement of unconstrained motion using a model-matching method», *Ergonomics*, Vol. 37, no 3, 1994, p. 493-510.
- [Nelder65] J. A. Nelder and R. Mead, «A Simplex Method for Function Minimization», *Computer Journal*, Vol. 7, 1965, pp. 308-313.
- [Ohya94] J. Ohya and F. Kishino, «Human Posture Estimation from Multiple Images Using Genetic Algorithm», *Proc. of the 12th IAPR Int. Conf. On Pattern Recognition*, Vol 1, pp. 750-753.
- [Ouhaddi98a] H. Ouhaddi et P. Horain, «Conception et ajustement d'un modèle 3D articulé de la main», *Actes des 6^{èmes} Journées de Travail du GT Réalité Virtuelle*, Issy-les-Moulineaux, 12-13 Mars, 1998, p. 83-90. <http://www-sim.int-evry/Publications>.
- [Ouhaddi98b] H. Ouhaddi, P. Horain and K. Mikolajczyk, «Modélisation et suivi de la main», *Actes des 4^{èmes} Journées d'études et d'échanges «Compression et Représentation des Signaux Audiovisuels» (CORESA'98)*, Lannion, France, 9-10 1998, pp. 109-114.
- [Preteux98] F. Preteux, P. Horain, H. Ouhaddi and M. Preda, *Report on Core Experiment 3 on Hand Baps interpretation*, ISO/IEC JTC1/WG1 MPEG97/M3332, March 1998.
- [Press92] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, 1992.
- [Regh94] J.M. Regh and T. Kanade, «Visual Tracking of Height DOF Articulated Structures : an Application to Human Hand Tracking», *Proc. of 3rd ECCV*, vol 2, 1994, p. 37-46.
- [Rijpkema91] H. Rijpkema and M. Girard, «Computer Animation of Knowledge-Based Human Grasping», *Computer Graphics*, Vol. 25, no. 4, 1991, p. 339-348.
- [Serra94] J. Serra and P. Soille, *Mathematical morphology and its applications to image processing*, Kluwer Academic Publishers, 1994.
- [Shimada95] N. Shimada, Y. Shirai and Y. Kuno, «Hand Gesture Recognition Using Computer Vision Based on Model-Matching Method», *Proc. of 5th International Conference on HCI*, p. 11-16. Elsevier, 1995.
- [Shimada98] N. Shimada, Y. Shirai, Y. Kuno and J. Mivra, «Hand Gesture Estimation and Model Refinement using Monocular Camera - Ambiguity Limitations by Inequality Constraints », *Proceedings of FG'98*, p. 268-273, april 1998 Nara, Japan.
- [Teo94] C.G. Teo, *A Hybrid Procedural Knowledge-Based Approach to the Animation of Human Hand Grasping*, Master Thesis, Simon University, 1994.
- [Wagner88] C. Wagner, «The pianist's hand anthropometry and biomechanics», *Ergonomics*, Vol. 31, no 1, 1988, p. 97-131.
- [Wernecke94] J. Wernecke, *The Inventor Mentor*, Addison-Wesley Publishing Company, 1994.

Manuscrit reçu le 2 avril 1998.

LES AUTEURS

Hocine OUHADDI



Hocine Ouhaddi est ingénieur d'état en informatique de l'Institut d'Informatique de l'Université de Sciences et Technologie d'Alger depuis 1994. Après une année théorique en magistère de linguistique à l'ENS d'Alger, et une année à l'Université de Paris 6, il a obtenu un Diplôme d'Études Approfondies en Intelligence Artificielle, Reconnaissance de Formes et Applications. Il prépare depuis octobre 1996 une thèse sur l'analyse de gestes par vision monoscopique à l'Institut National des Télécommunications (INT) au sein de l'Unité de Projets ARTEMIS.

Patrick HORAIN



Patrick Horain est docteur-ingénieur en informatique de l'Institut Polytechnique National de Grenoble depuis 1984. En 1986, il a rejoint le Département Image de l'École Nationale Supérieure Télécommunications (ENST) comme Enseignant-Chercheur, puis en 1995 l'Institut National des Télécommunications (INT) au sein de l'Unité de Projets ARTEMIS. Ses intérêts de recherche sont l'analyse d'image et la modélisation 3D, ainsi que l'utilisation des nouvelles technologies pour l'enseignement du traitement d'image.