

Estimation mono-vue de la posture du bras, méthode et évaluation

Pose Reconstruction of a Human Arm, method and evaluation

Boris LENSEIGNE¹

Frederick GIANNI¹

Patrice DALLE¹

¹ IRIT, Université Paul Sabatier

118, route de Narbonne
31062 TOULOUSE cedex 4

lenseign@irit.fr, gianni@irit.fr, dalle@irit.fr

Résumé

L'étude des gestes d'une personne humaine à partir de séquences d'images mono-vue, est un domaine du traitement d'images qui fait l'objet de recherches actives. D'autre part l'étude des gestes et des mouvements corporels a conduit à l'émergence de la biomécanique qui a pour but de modéliser la mécanique de ces gestes et des composants permettant de les effectuer (squelette, muscles, etc ...). Nous associons ces deux disciplines pour proposer un nouvel algorithme de reconstruction de la posture du bras à partir d'une image du locuteur prise par une caméra calibrée. La mise en œuvre de cette technique se fait en deux étapes : d'une part la validation du principe et des modèles à partir d'une scène et d'un dispositif d'acquisition virtuels, d'autre part, l'étude sur des images réelles. La première étape nous permet de dégager les contraintes et tolérances sur les opérateurs de traitement d'images à mettre en œuvre, la seconde nécessite l'emploi d'opérateurs de traitement d'images dont les performances concordent avec ces contraintes. Dans cet article, nous proposons une résolution purement géométrique du problème qui nous permet de produire un ensemble de solutions valides, puis un modèle biomécanique du bras permettant de sélectionner la bonne. Nous montrons ensuite comment appliquer cette méthode sur des images réelles tout en levant un certain nombre des hypothèses faites précédemment, la reconstruction se faisant alors à un facteur d'échelle près. Enfin nous proposons une étude expérimentale de la robustesse de l'algorithme vis-à-vis du bruit sur les données et nous donnons quelques pistes permettant de faire le lien entre la précision désirée pour la reconstruction et celle sur les données initiales extraites par traitement d'images.

Mots Clef

Vision par ordinateur, analyse du geste, espace atteignable, paramètres de Denavit Hartenberg

Abstract

Human gesture analysis using a single camera is an active field of interest in Computer Vision research. On the other hand, human gesture and body motion have been widely studied in biomechanics. The aim of biomechanics is to build models of the mechanical structure (bones, muscles, etc...) involved in gesture production. We use methods combining these two disciplines, to propose a new simple technique to build the posture of the arm from a calibrated single-view video sequence. We implement this method in two steps : in a first time, the method is geometrically validated on a virtual scene watched with a virtual camera. This stage allows us to define the constraints and tolerances on the image processing operators we need. The second stage concerns the implementation of the method on real images, which needs the use of image processing tools whose performances agree with the constraints defined previously. We propose, in this paper, a geometric method which allows us to produce a set of valid solutions and a biomechanical model of the arm which is used to choose, among them, the right solution. After that we show how to use this algorithm on real images in order to avoid some of the hypothesis made previously. Finally, we present an experimental study of the robustness of the algorithm and we show how to connect the quality of the reconstruction and of the input data's provided by image processing.

Keywords

Computer Vision, gesture analysis, reachable space, Denavit Hartenberg parameters

1 Problématique

Notre travail se situe dans le cadre de l'étude des composantes langagières du geste. Plus exactement, nous nous intéressons aux cas où le geste est produit avec une intention de communication, tels que les langues des Signes ou encore certains gestes coverbaux.

1.1 Étude des gestes de communication

Les langues des Signes sont les langues naturelles des sourds à travers le monde. Dans ces langues, le message transmis passe par le canal gestuel et non plus oral. Le locuteur fait donc un usage important de l'espace autour de lui non seulement pour réaliser les Signes, mais également pour y placer les différents éléments du discours et y faire référence [6]. Il est donc indispensable d'intégrer cette notion dans les systèmes informatiques qui analysent cette langue. Pour l'heure, la majorité des systèmes de capture des Signes ont été étudiés dans le cadre de la traduction automatique et font appel à des dispositifs spécifiques [2][5]. D'autres utilisent des caméras mais travaillent directement dans l'espace image [15], [18]. Mais les outils permettant d'accéder à l'information 3D sur le geste à partir d'une caméra unique restent rare [14]. Pour notre part, nous souhaitons accéder à ces informations dans le cadre de la conception d'un système d'analyse de cette langue [4]. Les gestes coverbaux, quant à eux, accompagnent un acte de communication orale, souvent sans même que le locuteur en ait conscience. Ils complètent - ou remplacent - l'information orale. À la différence des langues des Signes, il n'existe pas de vocabulaire standard ; les gestes coverbaux sont le plus souvent dépendants du contexte du discours et de l'environnement du locuteur. De plus, leur compréhension demande de la part des deux interlocuteurs une certaine culture commune, permettant par exemple d'attribuer le même sens à une catégorie de gestes. Du point de vue informatique, peu de travaux spécifiques ont été menés à ce jour, mais des projets récents¹ font le lien entre ces gestes et ceux produits dans la langue des Signes. Nous retrouvons donc tout naturellement les mêmes besoins en informations spatiales dans les deux domaines.

1.2 Analyse et représentation des gestes

L'étude du geste en tant qu'acte de communication passe par l'étude du mouvement des membres supérieurs. La description de ces mouvements a fait l'objet d'un grand nombre de travaux dont un état de l'art peut être trouvé dans [8]. D'autre part, lorsque le périphérique d'entrée est un système de Vision, plusieurs configurations sont possibles. Ainsi, dans [16], les auteurs utilisent plusieurs caméras non calibrées pour estimer la posture du corps complet. D'autres applications n'utilisent qu'une seule caméra, calibrée pour reconstruire le mouvement d'un bras [7]. On peut également noter les travaux de [19] qui estime la conformation d'une chaîne articulée à partir d'une caméra non calibrée, en estimant simultanément une matrice de projection perspective. Dans [3] les auteurs modélisent les mouvements d'un corps articulé (une chaîne cinématique) afin de calculer dans une image les valeurs angulaires des articulations à partir de celles connues dans l'image précédente. Enfin [17] fournit un modèle du bras humain à quatre degrés de liberté et l'utilise pour estimer la pose d'un bras à partir d'images fixes. Nous utilisons pour notre

part une unique caméra calibrée associée à un modèle mécanique du bras. Ce modèle nous permet de lier les coordonnées articulaires du bras et sa position dans l'espace et de prendre en compte les valeurs limites sur les articulations. Du point de vue de la modélisation du bras humain, un modèle proche du notre, associé à des butées articulaires, est également utilisé pour la reconstruction et la prédiction de postures de bras manipulateurs dans [12], qui reprend les travaux d'Abdel-Malek sur le calcul d'espace de travail de bras articulés [11]. Enfin, un modèle du bras humain basé sur les paramètres de Denavit-Hartenberg (utilisé en robotique) est exploité dans [13] pour piloter un bras robot anthropomorphe. L'auteur donne également les valeurs des butées pour chaque articulation. Nous reprenons ces résultats pour mener à bien le filtrage des solutions fournies par l'algorithme de vision (Cf.3).

1.3 Organisation de l'article

Nous proposons pour notre part un algorithme simple de vision par ordinateur qui nous permet de définir un ensemble des configurations possibles (au maximum quatre) pour le bras étudié. Cet algorithme est basé sur l'utilisation d'une seule caméra calibrée, associée à des informations a priori sur l'espace atteignable par les différentes articulations (Cf.2.2). Un modèle biomécanique, inspiré de ceux utilisés en robotique pour les bras manipulateurs, est utilisé ensuite pour réduire encore cet espace. Ce modèle nous donne également accès à une représentation de la posture et, dans le cas de séquences d'images, de la trajectoire de chaque articulation, à la fois dans l'espace objet cartésien et dans l'espace des configurations angulaires de chaque articulation. Les aspects géométriques de cette démarche (Cf.2.3) ont été, dans un premier temps évalués sur une scène de synthèse, en utilisant les images générées par une caméra virtuelle. Nous utilisons ensuite le modèle biomécanique pour vérifier la validité des solutions trouvées de façon géométrique (Cf.3). Lors de la mise en œuvre de la méthode sur des images réelles, nous revenons sur les hypothèses faites au préalable (Cf.2.1) et nous montrons comment lever certaines d'entre elles (Cf.4). Cette démarche nous permet en outre de mener un étude théorique de la sensibilité de l'algorithme au bruit (Cf.2.5), de valider ces résultats sur des données réelles (Cf.4.3) et de faire le lien entre la précision des résultats fournis en entrée par des opérateurs d'analyse d'images et la précision obtenue dans la reconstruction 3D.

Notations : Dans les sections traitant de la reconstruction géométrique, nous utiliserons les notations suivantes :

- Les points 3D dans le repère caméra sont notés en majuscules : $P_i = (X_i, Y_i, Z_i)^T$
- Les points images en coordonnées image sont notés : $\tilde{p}_i = (u_i, v_i, 1)^T$
- Les points images exprimés dans le repère caméra sont notés : $\tilde{\mathbf{p}}_i$

¹www.irit.fr/AS-IG/

- les points entachés de bruit sont surmontés d'une astérisque : $P_i^*, \tilde{P}_1^* \dots$

2 Vers une solution géométrique

Nous décrivons dans cette section comment reconstruire, à partir d'une image d'un locuteur faisant des gestes devant une caméra calibrée, la posture 3D de l'un de ses bras. Nous appelons posture la configuration du bras à un instant donné, cette configuration étant définie, soit par les valeurs angulaires des articulations (épaule, coude) du bras, soit par les positions des extrémités de chaque segment (épaule, coude, poignet) dans l'espace 3D. Nous montrons comment nous pouvons ramener l'espace des postures possibles à quatre configurations seulement à partir d'un modèle de la scène et du dispositif d'acquisition, ceci moyennant certaines hypothèses. Nous démontrons la validité de cette solution sur des images de synthèse.

2.1 Pré-requis

Notre technique repose sur un certain nombre d'hypothèses et de connaissances préalables, parfois contraignantes dans le cadre d'une exploitation en environnement non contrôlé. Elles peuvent toutefois être en partie levées si nous nous limitons à une estimation à un facteur d'échelle près. Il faut de plus noter qu'une telle reconstruction reste tout à fait utilisable dans la majorité des applications d'analyse des gestes, notamment en ce qui concerne la résolution de problèmes tels que l'estimation de la position relative.

Dispositif d'acquisition : Le dispositif d'acquisition est constitué d'une seule caméra, calibrée de façon à pouvoir calculer l'équation d'un rayon de projection passant par un pixel donné. Ceci suppose la connaissance de la matrice de calibrage C .

Suivi des articulations : Nous faisons également l'hypothèse que nous sommes capables d'identifier dans une image les positions des trois articulations du bras. En l'absence de marqueurs, le problème peut s'avérer d'autant plus difficile à résoudre que les points recherchés correspondent en réalité aux axes de rotation de chaque articulation. Leur position devra alors probablement être déduite du résultat de la recherche de primitives plus facilement détectables dans l'images, par exemple les différents segments du bras. Enfin le problème des occultations (bras tendu vers l'avant par exemple) devra également être pris en compte. De façon générale, le choix d'un algorithme de détection de ces points devra donc être guidé par les contraintes d'exécution de l'application et par la précision des résultats souhaités.

Posture du bras : Nous ne traitons ici que de la reconstruction de la posture d'un bras, sans tenir compte de la main, soit, si ces coordonnées sont exprimées dans le repère caméra :

- pour l'épaule : $P_1 = (X_1, Y_1, Z_1)^T$
- pour le coude : $P_2 = (X_2, Y_2, Z_2)^T$

- pour le poignet : $P_3 = (X_3, Y_3, Z_3)^T$

Cette représentation permet de ramener le problème de l'estimation de la posture à celui de l'estimation de la position de trois points dans l'espace.

2.2 Modèle géométrique du bras

Toujours dans le cadre de la résolution géométrique, nous modélisons le bras par le système articulaire reliant l'épaule au poignet. Ce système est constitué d'articulations (liaison pivot pour l'épaule et rotule pour le coude) reliant des segments rigides (bras et avant bras) notés l_i , le segment l_i reliant les articulations P_i et P_{i+1} . La position de l'organe terminal correspond à celle du poignet, c'est-à-dire à P_3 , extrémité du segment l_2 . Ces articulations ne permettant que des rotations pures du segment l_i autour de l'articulation P_i , nous pouvons définir l'ensemble des positions atteignables par l'articulation P_j ($j = 2, 3$) comme étant une sphère centrée sur l'articulation précédente P_{j-1} et dont le rayon est $\|l_i\|^2$ (fig.1).

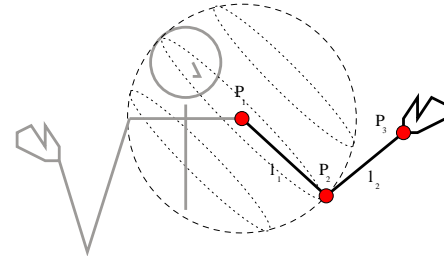


Figure 1: Modèle du système articulaire du bras. La sphère représente l'ensemble des positions possibles pour le coude.

Grâce à cette modélisation, l'espace de recherche pour chaque articulation se trouve ramené à une sphère dont les paramètres sont connus dès lors que l'on a déterminé la position de l'articulation précédente et la longueur de chaque segment du bras. Ce qui nécessite donc de connaître dès le début la position 3D de l'épaule. Il faut toutefois noter que, dans la mesure où nous utilisons un dispositif d'acquisition calibré, le fait de connaître la position de l'épaule dans l'image et une de ses coordonnées 3D, permet de calculer de façon simple les coordonnées manquantes. De la même façon, la position de l'épaule pourra être initialisée à un facteur d'échelle près si l'on fixe à l'avance une de ses coordonnées (par exemple la distance à la caméra) (Cf.4).

2.3 Algorithme

La méthode que nous proposons exploite un modèle géométrique simple de la scène et spécialement de la structure - le bras - dont nous souhaitons déterminer les paramètres (Cf.2.2). Nous supposons connues les coordonnées, dans l'image, des points correspondants aux articulations, notées en coordonnées homogènes :

- pour l'épaule : $\tilde{p}_1 = (u_1, v_1, 1)^T$

² $\|l_i\|$ est la norme du segment l_i

- pour le coude : $\tilde{p}_2 = (u_2, v_2, 1)^T$
- pour le poignet : $\tilde{p}_3 = (u_3, v_3, 1)^T$

Grâce au calibrage de la caméra, nous pouvons alors associer à chaque point image un rayon de projection, c'est-à-dire la droite (passant par le centre optique et le point image considéré) sur laquelle se trouve l'homologue 3D de ce point.

Espace des configurations possibles pour le coude : Étant donné P_1 la position dans l'espace de l'épaule (éventuellement arbitraire), nous pouvons définir l'espace des positions possibles pour le coude comme étant la sphère S_1 centrée sur l'épaule et dont le rayon est la longueur $\|l_1\|$ du bras. L'équation cartésienne de cette sphère est la suivante :

$$(X_1 - x)^2 + (Y_1 - y)^2 + (Z_1 - z)^2 - \|l_1\|^2 = 0 \quad (1)$$

Équation du rayon de projection : Soit \tilde{p}_1 la position de l'épaule dans l'image, exprimée en coordonnées homogènes. Le calibrage de la caméra nous donne la matrice de projection perspective pour les directions de l'espace C (matrice de calibrage). L'inverse de cette matrice nous donne pour un point de l'image \tilde{p}_i l'équation du rayon de projection associé. Le point 3D recherché se trouve nécessairement sur cette droite :

$$R_i(\lambda) = \lambda \tilde{p}_i \quad (2)$$

Il s'agit d'une équation paramétrique où λ est un simple coefficient multiplicateur et où \tilde{p}_i représente les coordonnées du point image considéré exprimées dans le repère caméra par :

$$\tilde{p}_i = C^{-1} \tilde{p}_i \text{ avec } C^{-1} = \begin{bmatrix} \frac{1}{fk_u} & 0 & \frac{u_0}{fk_u} \\ 0 & \frac{1}{fk_v} & \frac{v_0}{fk_v} \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

$$\text{Soit : } \tilde{p}_i = \begin{bmatrix} \frac{(u_i - u_0)}{fk_u} \\ \frac{(v_i - v_0)}{fk_v} \\ 1 \end{bmatrix} \quad (4)$$

La position 3D recherchée se trouve alors à l'intersection de la surface de la sphère S_1 définissant l'ensemble des configurations possibles pour le coude, et du rayon de projection $R_i(\lambda)$. Le calcul de ces intersections dans le repère caméra revient donc à déterminer des valeurs de λ telles que :

$$(X_1 - \lambda \frac{(u_i - u_0)}{fk_u})^2 + (Y_1 - \lambda \frac{(v_i - v_0)}{fk_v})^2 + (Z_1 - \lambda)^2 - \|l_1\|^2 = 0 \quad (5)$$

Cela revient à calculer les solutions du polynôme du second degré $a\lambda^2 + b\lambda + c = 0$ avec pour coefficients :

$$\begin{aligned} a &= (\frac{(u_i - u_0)}{fk_u})^2 + (\frac{(v_i - v_0)}{fk_v})^2 + 1 \\ b &= 2[(\frac{(u_i - u_0)}{fk_u})(-X_1) + (\frac{(v_i - v_0)}{fk_v})(-Y_1) - Z_1] \\ c &= X_1^2 + Y_1^2 + Z_1^2 - l_1^2 \end{aligned} \quad (6)$$

La résolution de ce polynôme nous donne donc deux valeurs possibles pour λ (éventuellement une seule), dont nous pouvons aisément déduire des positions $\hat{P}_{2,j}$ ($j = 1, 2$) possibles pour le coude puisque $R_i(\lambda) = \lambda \tilde{p}_i$.

Nous pouvons ensuite calculer les positions $\hat{P}_{3,j}$ ($j = 1, \dots, 4$) du poignet, dans le repère caméra. Nous considérons pour cela les deux sphères dont le centre est donné par chaque position estimée du coude et le rayon par la longueur de l'avant-bras. Suivant le même principe, nous pouvons alors calculer, pour chaque valeur de la position du coude, deux positions possibles pour le poignet, soit en tout quatre configurations possibles pour la posture du bras. Cet algorithme permet donc de ramener l'espace des configurations possibles pour un bras à quatre possibilités. L'étude de séquences d'images exploitant des hypothèses sur la continuité du mouvement (Cf.2.4) et l'apport du modèle biomécanique (Cf.3.1) nous permettent de déterminer, parmi ces quatre postures, la plus probable.

2.4 Analyse de séquences d'images

Dans le cas d'une séquence d'images, nous calculons à partir de chaque image l'ensemble des points 3D candidats. Nous constituons ensuite les trajectoires. Il y a à chaque étape (sauf configuration particulière) deux points à répartir sur deux trajectoires. Toutefois, dans la mesure où il n'est pas possible de savoir a priori quel point rajouter à quelle trajectoire, nous devons faire appel à un critère de linéarité afin de les reconstruire. Comme un simple critère de distance ne nous permettrait pas de prendre en compte les cas de trajectoires candidates se croisant, les points concernés étant alors très proches, nous devons utiliser un critère permettant de prendre en compte la distance entre les points mais aussi la variation de la courbure de chaque trajectoire afin d'empêcher les changements trop brusques. Nous calculons pour cela la norme du produit vectoriel $\|\vec{v}_{i,j,k} \wedge \vec{v}_{i,j,k+1}\| = \|\vec{v}_{i,j,k}\| \|\vec{v}_{i,j,k+1}\| \sin(\alpha)$, où $\vec{v}_{i,j,k}$ est défini par les points $\hat{P}_{i,j,k-1}$ et $\hat{P}_{i,j,k}$, et $\vec{v}_{i,j,k+1}$ par $\hat{P}_{i,j,k}$ et $\hat{P}_{i,j,k+1}$. $\hat{P}_{i,j,k}$ est la j^{eme} ($j = 1, 2$) estimée des coordonnées spatiales de l'articulation i à l'étape k (fig.2). Le point candidat pour lequel cette norme est la plus faible est alors affecté à la trajectoire pour laquelle cette valeur minimale a été calculée. Le second point de la branche est alors affecté à la trajectoire restante.

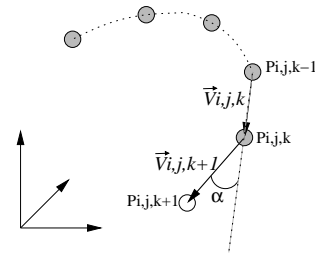


Figure 2: Calcul du produit vectoriel entre le dernier vecteur directeur de la trajectoire considérée et un nouveau vecteur directeur construit à partir du point candidat à affecter (en blanc). Le critère de linéarité des trajectoires utilisé consiste à affecter un premier point à la trajectoire correspondant au plus petit produit vectoriel et le point restant à la seconde trajectoire.

Configurations particulières : La construction des trajectoires décrite ci-dessus s'effectue correctement dans le

cas général où l’algorithme donne deux intersections du rayon de projection avec chaque sphère. Il existe cependant des configurations où cela ne sera pas le cas. Elles doivent être prises en compte dans l’algorithme ; elles peuvent également servir de détecteur pour des mouvements particuliers.

Il existe deux catégories de configurations particulières :

1. Le polynôme (6) n’a qu’une seule solution. C’est le cas où le segment considéré (bras ou avant-bras) se situe dans un plan parallèle au plan image. Dans ce cas, le rayon de projection arrive de façon tangentielle à la sphère. Il n’y aura donc qu’une seule “intersection”. Ce point est alors concaténé aux deux trajectoires considérées, dans la mesure où cela correspond effectivement à un cas où ces deux trajectoires se croisent.
2. Le polynôme (6) n’a pas de solution. En l’absence de bruit, ce cas ne peut se produire que pour le poignet. En effet, après calcul des deux positions possibles pour le coude, nous définissons les deux sphères qui forment l’ensemble des positions possibles du poignet. Il peut alors se produire des cas où, lorsque l’on se base sur la “mauvaise” position du coude, la sphère ainsi définie n’a pas d’intersection avec le rayon de projection. De telles configurations permettent directement d’éliminer une branche complète de l’arbre des solutions.

2.5 Prise en compte de données bruitées

Lors de la mise en œuvre de cet algorithme sur des images réelles, la principale source de bruit proviendra de la détection des articulations dans l’image. Nous cherchons donc à étudier de quelle façon un bruit ϵ_i sur la position de chaque point \tilde{p}_i se répercute sur l’ensemble des positions 3D calculées. De façon intuitive, une erreur sur la position du point \tilde{p}_i entraînera une déviation du rayon de projection, et donc un décalage de l’intersection avec la sphère d’autant plus important que l’angle incident est éloigné de la normale à la surface. Dans les cas extrêmes, c’est-à-dire lorsque le rayon de projection devient tangent à la sphère, le rayon peut même “manquer” éventuellement la sphère. Nous étudions dans cette section la façon dont les erreurs sur la détection de l’articulation \tilde{p}_i influenceront le calcul de la position de l’articulation P_{i+1} . Une validation expérimentale de ces résultats sera également proposée (Cf.4.3).

Propagation du bruit : Afin d’étudier l’effet de l’accumulation du bruit lors du calcul des positions successives de chaque articulation, nous supposons la position 3D de l’épaule connue avec précision. Nous faisons ensuite l’hypothèse que la position \tilde{p}_i de chaque point est entachée d’un bruit ϵ_i uniforme dans toutes les directions et de faible amplitude par rapport à la taille de l’image, soit $\tilde{p}_i^* = \tilde{p}_i + \epsilon_i$ la position bruitée du point \tilde{p}_i . Si nous reprenons les équations données en section 2.3, nous pou-

vons poser l’équation du rayon de projection :

$$R(\lambda) = \lambda \tilde{p}_i^* \text{ avec } \tilde{p}_i^* = C^{-1}((u_i + \epsilon_i), (v_i + \epsilon_i), 1)^T \quad (7)$$

L’équation (5) définissant a priori l’ensemble des configurations possibles du coude reste inchangée. Nous pouvons alors réécrire les coefficients du polynôme (6) donnant les valeurs possibles de λ et en déduire l’expression du discriminant :

$$\begin{aligned} \Delta = & 8\left[\left(\frac{(u_i + \epsilon_i) - u_0}{fk_u}\right)\left(\frac{(v_i + \epsilon_i) - v_0}{fk_v}\right)X_1 Y_1 \right. \\ & \left. + \left(\frac{(u_i + \epsilon_i) - u_0}{fk_u}\right)X_1 Z_1 + \left(\frac{(v_i + \epsilon_i) - v_0}{fk_v}\right)Y_1 Z_1\right] \\ & - 4\left[\left(\frac{(u_i + \epsilon_i) - u_0}{fk_u}\right)^2(Y_1^2 + Z_1^2 - l_1^2) \right. \\ & \left. + \left(\frac{(v_i + \epsilon_i) - v_0}{fk_v}\right)^2(X_1^2 + Z_1^2 - l_1^2) \right. \\ & \left. + (X_1^2 + Y_1^2 - l_1^2)\right] \end{aligned} \quad (8)$$

D’après cette expression, les solutions générales $\lambda = \frac{-b \pm \sqrt{\Delta}}{2a}$ varient suffisamment peu en fonction du bruit pour que les erreurs commises lors de la détection du coude ne rendent pas impossible l’estimation de la position du poignet. Cette observation se vérifie expérimentalement (Cf.4.3). Dans le cas où Δ devient nul, la solution double s’écrit :

$$\lambda_{1,2} = \frac{\left(\frac{(u_i + \epsilon_i) - u_0}{fk_u}\right)X_1 + \left(\frac{(v_i + \epsilon_i) - v_0}{fk_v}\right)Y_1 + Z_1}{\left(\frac{(u_i + \epsilon_i) - u_0}{fk_u}\right)^2 + \left(\frac{(v_i + \epsilon_i) - v_0}{fk_v}\right)^2 + 1} \quad (9)$$

À ce moment là, la présence dans cette expression d’un terme constant additif au dénominateur nous garantit que les valeurs trouvées pour λ resteront bornées. D’autre part, les cas où Δ devient nul correspondant à des configurations particulières (Cf.2.4), ces dernières peuvent être détectées à l’avance et les postures correspondantes déduites de celles calculées précédemment. Enfin, des cas où le polynôme n’a pas de solution peuvent se produire en présence de bruit, il s’agit de cas où le rayon de projection est dévié au point de “manquer” la sphère. Comme par construction, il ne s’agit que de rayons initialement proches de la tangente à la sphère, une approximation valable consiste à se ramener au cas où $\Delta = 0$.

3 Filtrage des solutions

La résolution géométrique proposée donne 4 configurations possibles pour le bras. Ces configurations sont données dans l’espace objet 3D. Toutefois, lors du calcul de ces configurations, nous ne tenons pas compte des possibilités réelles de mouvement d’un bras humain. En effet, l’espace de travail défini précédemment par des sphères est en réalité réduit à des portions seulement de ces sphères par les butées présentes sur les articulations. Nous aurons donc, parmi les configurations candidates définies géométriquement, un certain nombre de configurations aberrantes, au sens des valeurs angulaires sur les articulations. D’autre part, en plus des positions aberrantes, nous aurons également un certain nombre de configurations plausibles mais peu probables du fait de leur inconfort. Cette notion est appelée “zone de travail adroite” par

i	θ_i	d_i	α_i	a_i
1	θ_1	0	$-\pi/2$	0
2	θ_2	0	$\pi/2$	0
3	θ_3	l_1	$-\pi/2$	0
4	θ_4	0	$\pi/2$	0
5	0	l_2	0	0

Table 1: Paramètres de DH définissant le modèle du bras utilisé

A. Malek dans [11]. Afin d'exploiter ces informations, nous complétons notre modélisation par un second modèle du bras, similaire à ceux utilisés en robotique.

3.1 Modèle biomécanique

Le modèle paramétrique du bras humain proposé est basé sur les paramètres de Denavit-Hartenberg (DH) modifiés [10]. Ce modèle décrit le système par une matrice de transformation homogène permettant de passer du repère objet au repère associé à l'organe terminal - dans notre cas le poignet. Cette matrice est paramétrée par les valeurs angulaires de chaque articulation et par les dimensions des segments (bras et avant-bras) les reliant. Il s'agit là du modèle géométrique direct. Réciproquement, le modèle géométrique inverse permet d'associer aux coordonnées cartésiennes des articulations les valeurs angulaires correspondantes.

Paramètres de Denavit-Hartenberg modifiés : La méthode décrite dans [10] nécessite une utilisation judicieuse des repères associés à chaque articulation du bras considéré. Les repères R_i associés à chaque articulation i (fig.6) sont donc définis selon les règles suivantes :

1. Les axes x_{i-1}, x_i sont les perpendiculaire communes aux axes z_{i-1} et z_i .
2. Les axes y_i sont tels que: $y_i = z_i \wedge x_i$
3. Les paramètres définissant la configuration relative du repère R_i par rapport à R_{i-1} sont :
 - α_i : angle entre les axes z_{i-1} et z_i en tournant autour de x_{i-1} .
 - d_i : la distance entre z_{i-1} et z_i , mesurée selon x_{i-1} .
 - θ_i : l'angle entre les axes x_{i-1} et x_i en tournant autour de z_i .
 - a_i : la distance entre x_{i-1} et x_i , mesurée selon z_i .

Les repères 0 et 1 étant confondus, le modèle du bras se décrit alors par la table (1) des paramètres de DH.

L'utilisation des paramètres de DH permet d'écrire la matrice de transformation homogène ${}^{i-1}T_i$, qui décrit le passage du repère R_{i-1} au repère R_i :

$${}^{i-1}T_i = \begin{bmatrix} \cos \theta_i & -\sin \theta_i & 0 & a_i \\ \cos \alpha_i \sin \theta_i & \cos \alpha_i \cos \theta_i & -\sin \alpha_i & -r_i \sin \alpha_i \\ \sin \alpha_i \sin \theta_i & \sin \alpha_i \cos \theta_i & \cos \alpha_i & r_i \cos \alpha_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10)$$

Où les $\theta_i, \alpha_i, d_i, a_i$ sont les paramètres de DH.

Modèle géométrique direct : Le modèle géométrique direct nous permet de passer des valeurs angulaires des articulations du bras à leurs coordonnées cartésiennes. La transformation homogène 0T_5 permettant de passer du repère 0 de l'épaule au repère 5 du poignet est obtenue en multipliant successivement les matrices ${}^{i-1}T_i, i = 1, \dots, 5$. La matrice 0T_5 (13) résultante est donnée en annexe. Ce modèle est paramétré par les valeurs θ_i des articulations ; il définit l'espace de travail du bras. Pour effectuer le filtrage en fonction des butées sur les valeurs angulaires de chaque articulation, nous avons besoin d'un modèle permettant de passer d'une position 3D donnée de l'organe terminal à l'ensemble de valeurs angulaires correspondant.

Modèle géométrique inverse : Le modèle géométrique inverse permet, à partir des coordonnées cartésiennes de position de l'organe terminal (dans notre cas le poignet), de calculer les valeurs angulaires θ_i correspondantes des articulations (fig.3). La première solution envisageable est de calculer l'inverse de 0T_5 (13), mais l'extraction des valeurs de θ_i n'est alors pas réalisable directement. Nous décomposons alors la chaîne cinématique afin de calculer dans un premier temps $\theta_{1,2}$ à partir de la position du coude exprimée dans le repère R_0 puis $\theta_{3,4}$ à partir de la position du poignet exprimée dans le repère R_3 . Nous définissons pour cela les transformations inverses permettant de passer, de la position cartésienne du poignet aux valeurs angulaires du coude 5T_3 , et de la position du coude aux valeurs angulaires de l'épaule 3T_0 .

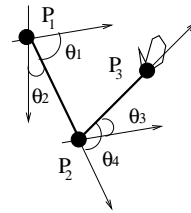


Figure 3: Le modèle inverse du bras nous donne les valeurs angulaires des articulations $\theta_1, \theta_2, \theta_3, \theta_4$ en fonction des coordonnées spatiales de P_1, P_2, P_3 (l'épaule, le coude, le poignet).

En ne considérant que l'articulation de l'épaule, les paramètres de Denavit-Hartenberg sont donnés par les trois premières lignes de la table (1). On peut alors déterminer les transformations homogènes permettant de passer de R_{i-1} à R_i pour $i = 1, 2, 3$, et de calculer la transformation correspondante suivante ($\cos \theta_{12}$ signifie $\cos \theta_1 * \cos \theta_2$): Le modèle géométrique direct est donné par la quatrième colonne de 0T_3 soit (les termes définissant la rotation ne sont pas pris en compte pour le calcul du MGI et ne sont

pas détaillés) :

$${}^0T_3 = \begin{bmatrix} r_{11} & r_{12} & r_{13} & -l_1 \cos \theta_1 \sin \theta_2 \\ r_{21} & r_{22} & r_{23} & -l_1 \cos \theta_1 \sin \theta_2 \\ r_{31} & r_{32} & r_{33} & l_1 \cos \theta_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Nous pouvons alors en déduire les valeurs de $\theta_{1,2}$ qui constituent la première partie du modèle géométrique inverse du système :

$$\begin{cases} \theta_1 = \arctan(y_2/x_2) \\ \theta_2 = \arccos(z_2/l_1) \end{cases} \quad (11)$$

De même, d'après la position du poignet dans le repère R_3 et l'expressions de 3T_5 , nous pouvons calculer $\theta_{3,4}$:

$$\begin{cases} \theta_3 = \arccos(z_3/l_2) \\ \theta_4 = \arctan(y_3/x_3) \end{cases} \quad (12)$$

Grâce à ce modèle, nous pouvons, pour une position calculée à l'aide de l'algorithme décrit précédemment (Cf.2), calculer l'ensemble des valeurs angulaires correspondantes. Ces valeurs angulaires vont maintenant nous permettre de trier les trajectoires ou points candidats calculés géométriquement.

3.2 Exploitation du modèle

En utilisant le modèle géométrique inverse décrit dans la section précédente, il devient possible, connaissant les valeurs en butée de chaque articulation, de déterminer de façon instantanée les positions aberrantes parmi celles disponibles. En effet, en utilisant les butées articulaires d'un bras humain, il est possible de tester les valeurs des θ_i et de les trier selon leurs limites. Ces valeurs sont données dans [13] :

$$40 < \theta_1 < 270, 40 < \theta_2 < 270, 0 < \theta_3 < 180, 0 < \theta_4 < 160$$

La figure (4) montre les domaines atteignables de chaque articulation après prise en compte des valeurs en butée des articulations.

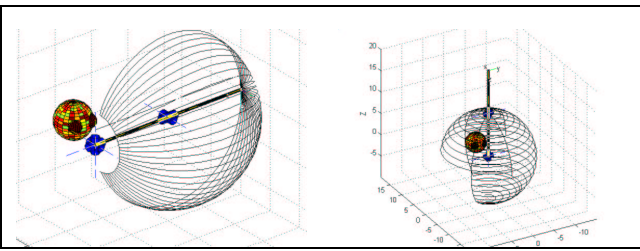


Figure 4: Domaine atteignable par les articulations du coude (à droite) et du poignet (à gauche). Dans le second cas, l'épaule reste fixe. Ces domaines prennent en compte les valeurs de butées sur les articulations.

D'autre part ce modèle est aisément différentiable. Le modèle cinématique inverse ainsi obtenu permet de relier les vitesses et les accélérations mesurées dans l'espace aux valeurs angulaires correspondantes. Ces informations peuvent être recoupées avec des résultats issus des neurosciences décrivant la cinématique d'un bras humain [1] afin de définir, parmi les trajectoires candidates, celle(s) dont le mouvement(s) est (sont) les plus "naturels".

4 Application sur des images réelles

La mise en œuvre de cet algorithme sur des séquences réelles nous impose de pouvoir vérifier les hypothèses faites précédemment (Cf.2.1) sur le calibrage de la caméra, les mensurations du locuteur et la configuration de la scène. Dans la mesure où cela n'est pas possible dans les corpus de langue des Signes existants réalisés à l'usage des linguistes, nous avons dû réaliser nos propres séquences. Celles-ci ont été réalisées à l'aide d'une caméra tri-CCD Canon XM2 dont la focale a été maintenue fixe. Le calibrage a été réalisé à l'aide d'une mire plane de dimensions connues selon la méthode décrite dans [9]. En revanche aucune mesure n'a été faite sur le locuteur ou sur la scène. En effet, dans le cas de mesures qualitatives, nous pouvons nous contenter d'une estimation de ces dernières à un facteur d'échelle près à partir du moment où celui-ci est constant. Nous avons donc mis au point une méthode permettant d'estimer ces mesures à un même facteur d'échelle près à partir d'une image où les segments du bras du locuteur sont parallèles au plan image.

4.1 Initialisation

Lorsque les différents segments du bras, sont parallèles au plan image de la caméra, leur longueur apparente est : $\|\hat{l}_i\| = k\|l_i\|$ et la profondeur Z_i pour chaque articulation est constante. Une façon simple d'obtenir les coordonnées 3D des articulations dans l'espace à un même facteur d'échelle près consiste alors à fixer à l'avance cette profondeur soit : $Z_i = Z_0$. Les autres coordonnées estimées \hat{P}_i pour chaque articulation sont alors données en recherchant sur la droite $R(\lambda) = \lambda \hat{p}_i$ le point \hat{P}_i tel que $\hat{P}_i = (\hat{X}_i, \hat{Y}_i, Z_0)^T$. Ces points 3D servent également au calcul de la matrice de transformation homogène permettant le passage du repère de la caméra où se fait la reconstruction au repère lié à l'épaule où se fait le calcul des valeurs articulaires (Cf. 3.1). Dans la pratique, une telle recherche peut être réalisée sans difficultés par dichotomie. Elle nous fournit les coordonnées 3D estimées des articulations à un même facteur d'échelle près ce qui permet ensuite de calculer avec ce même facteur d'échelle les longueurs des différents segments du bras. En revanche, en l'absence d'informations sur les dimensions réelles de la scène, le facteur d'échelle k reste, lui, inconnu.

4.2 Résultats

Étant donné la matrice de calibrage de la caméra et après calcul des mensurations du locuteur, nous pouvons, après avoir désigné à la main les coordonnées image des articulations, appliquer l'algorithme décrit en 2, et reconstruire l'ensemble des trajectoires candidates (fig.5). En revanche, à ce stade de nos travaux, le filtrage des trajectoires n'a pas encore été intégré au système ; cette manipulation nous a donc permis essentiellement de valider les aspects géométriques de la méthode, mais également de disposer d'une référence pour étudier la sensibilité de l'algorithme par rapport au bruit sur les entrées et d'en dé-

duire des contraintes sur les performances des opérateurs d'analyse d'images qui seront utilisés par la suite pour la détection des articulations.

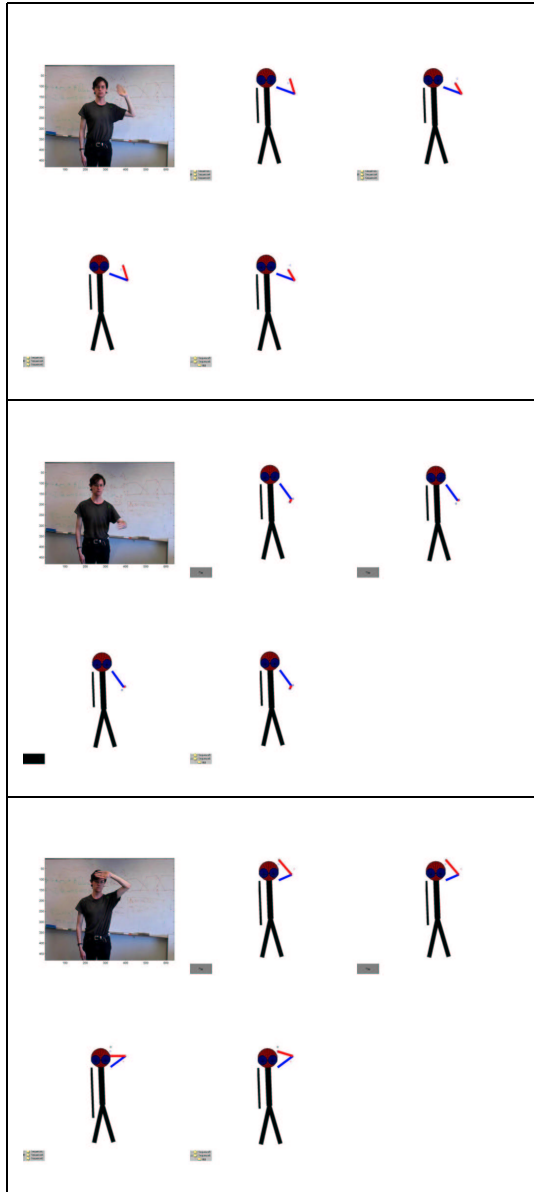


Figure 5: Quelques ensembles de solutions reconstruits géométriquement. (en haut) Configuration initiale où les segments du bras sont parallèles au plan image. Dans ces images, outre l'image originale, la première ligne correspond aux configurations reconstruites à partir de la première position du coude et celles du bas aux positions reconstruites à partir de la seconde.

4.3 Étude sur la robustesse

Comme cela a été décrit précédemment (Cf.2.5), si l'on excepte les erreurs, sur le calibrage de la caméra et l'estimation des mensurations du locuteur, la principale source de bruit dans cet algorithme résulte de la détection des articulations dans les images, qui est un problème non trivial de traitement d'images. Nous avons donc cherché à évaluer comment vont évoluer les coordonnées 3D des

points reconstruits en fonction de l'erreur sur la position des articulations dans les images.

Calcul d'erreur : Afin de mener à bien notre étude, nous utilisons comme référence les points 3D calculés après désignation manuelle des articulations. Nous entachons ensuite les points 2D d'une erreur uniforme dans toutes les directions et dont l'amplitude varie de 1 à 64 pixels. Cette amplitude maximale doit être comparée avec l'épaisseur du bras dans nos images qui est d'environ 30 pixels (la longueur maximale des segments du bras est proche de 100 pixels). Ce qui signifie que dans le cas d'une détection automatique des articulations, l'erreur sera généralement inférieure à cette valeur (au moins dans une direction) dès lors que le point détecté se trouve dans le bras. Nous calculons ensuite les différences de position sur chaque axe entre les points P_i^* bruités et les points P_i de référence (fig.7a).

Interprétation : Les courbes obtenues (fig.7a) nous donnent une idée de la façon dont l'erreur évolue sur les 3 axes du repère objet en fonction de l'amplitude de l'erreur de désignation. À partir de ces dernières, il est possible de faire plusieurs remarques concernant le comportement de l'algorithme :

- L'amplitude des erreurs de reconstruction varie de façon similaire sur les trois axes pour le coude et le poignet. Les erreurs de reconstruction pour le coude n'ont donc que peu d'influence sur l'estimation de la position du poignet. Ce qui confirme les résultats théoriques (Cf.2.5)
- L'amplitude des erreurs de reconstruction sur les axes X et Y varie de façon linéaire avec l'amplitude de l'erreur de désignation. En revanche l'erreur de reconstruction sur l'axe des Z (profondeur) ne dépend pas que de l'amplitude de l'erreur de désignation, mais également de l'angle entre le rayon de projection et la normale à la surface au points d'intersection. En effet, l'erreur sera d'autant plus importante que le rayon de projection sera proche d'une normale à la sphère.

D'après ces observations, nous pouvons conclure à la robustesse de cet algorithme dans le cadre de mesures qualitatives et relationnelles. En effet, nous voyons que d'importantes erreurs sur les coordonnées image des articulations entraînent des déviations relativement faibles des positions 3D des articulations par rapport à la longueur des segments du bras. D'autre part, il devient possible, de définir l'amplitude maximale en pixels de l'erreur de désignation acceptable en fonction de la précision souhaitée pour la reconstruction. En revanche, il faudra tenir compte à ce moment là du fait que l'erreur de reconstruction en profondeur est maximale lorsque le bras est proche d'une configuration singulière (Cf.2.4).

5 Conclusion, perspectives

Nous proposons une méthode simple de reconstruction de la posture d'un bras basée sur des techniques de Vision par ordinateur. Cette méthode nous permet, lors de l'analyse de séquences d'images, de reconstruire un ensemble de trajectoires possibles pour le coude et le poignet. Dans le pire des cas, cet algorithme nous fournit deux trajectoires candidates pour le coude, et quatre pour le poignet, moins si le bras passe par certaines configurations particulières. Nous disposons d'autre part d'un modèle biomécanique du bras qui permet d'associer à une configuration des articulations, donnée en coordonnées cartésiennes, les valeurs angulaires correspondantes pour chaque articulation. L'exploitation de ce modèle nous permet d'éliminer, parmi les trajectoires, celles contenant des positions aberrantes, c'est-à-dire des positions impossibles au sens des valeurs prises par les articulations. Cette méthode a été validée sur des images de synthèse dans lesquelles la position des articulations était connue. Lors de l'application sur des images réelles, nous passons par une phase d'initialisation qui permet d'estimer à un facteur d'échelle près les mensurations du locuteur, la reconstruction se faisant alors également à ce facteur d'échelle près. Dans ces conditions, l'algorithme nous permet d'obtenir de façon simple et rapide des mesures qualitatives et relationnelles sur les positions des différentes articulations du bras. Nous avons également étudié l'impact des erreurs sur la position des articulations dans les images sur la reconstruction. Nous avons notamment pu montrer que cette méthode était robuste à ce type d'erreurs (une importante erreur dans la position 2D de l'articulation entraînant une variation limitée de sa position dans l'espace). Enfin, nous définissons les conditions permettant d'estimer la précision nécessaire dans la détection des articulations pour garantir une précision donnée dans la reconstruction du bras. La suite de ces travaux portera sur l'exploitation du modèle biomécanique. En effet, celui-ci nous permet d'avoir une vision "proprioceptive" et articuloire du mouvement ; nous pouvons alors envisager de prendre en compte, en plus des limites sur les valeurs angulaires des articulations, des notions de confort ou de cible qui permettent, en exploitant ce modèle en commande, d'obtenir l'ensemble des configurations angulaires que prendra le bras pour passer d'une configuration donnée à une autre. Un autre aspect de cette étude porte sur l'intégration de connaissances propres aux geste étudiés pour améliorer le filtrage des solutions mais aussi pour construire des représentations de haut niveau des gestes observés. Enfin, la nécessité de calibrer la caméra reste une contrainte importante dans le cadre de l'exploitation de corpus existants. L'étude d'une méthode d'autocalibrage permettant d'éviter cette étape constitue donc un dernier aspect de nos travaux futurs.

References

[1] A. Berthoz. Le sens du mouvement. Odile Jacob, ISBN : 2738104576, Paris., 1997.

- [2] A. Braffort. *Reconnaissance et compréhension de geste, application à la langues des signes*. PhD thesis, Université Paris XI, UFR Sciences, LIMSI, juin 1996.
- [3] J. Malik C. Bregler. Video motion capture. In *SIGGRAPH 1998*, 1998.
- [4] P. Dalle C. Hudelot, B. Lenseigne. Apports d'un système d'analyse d'images pour l'étude de la langue des signes. In *Actes du colloque "Recherches sur la langue des Signes"*, IRIT, UPS Toulouse III, novembre 2001.
- [5] D. Metaxas C. Vogler. Parallel hidden models for american sign language recognition. *Proceedings of the International Conference on Computer Vision, kerkyra, Greece*, septembre 1999.
- [6] C. Cuxac. La langue des signes française. les voies de l'iconocité. *Faits de Langue, Ophrys*, 2000.
- [7] L. Goncalves et al. Monocular tracking of the human arm in 3d. *Fifth International Conference on Computer Vision*, page 764, 2000.
- [8] D.M. Gravila. The visual analysis of human movement: A survey. *Computer Vision and Image Understanding*, 73(1), janvier 1999.
- [9] J. Heikkilä. Geometric camera calibration using circular control points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1066–1077, octobre 2000.
- [10] R.S. Hartenberg J. Denavit. A kinematic notation for lower-pair mechanisms based on matrices. *Journal of Applied Mechanics*, pages 215–221, June 1955.
- [11] H. Yeh K. Abdel-Malek. Local dexterity analysis for open kinematic chains. *Mechanism and Machine Theory*, 35:131–154, 2000.
- [12] M. Jaber K. Abdel-Malek, W. Yu. Realistic posture prediction. (submitted) SAE Digital Human Modeling and Simulation.
- [13] T. Huntsberger M. Soltys. Inverse kinematic analysis of articulated figure using the fussy self-organizing feature map, 1993.
- [14] M. Bomb P. Horain. 3d model based gesture acquisition using a single camera. In *Proceedings of the IEEE Workshop on Applications of Computer Vision (WACV 2002)*, pages 158–162, Orlando, décembre 2002.
- [15] M. Turk R. Cutler. View-based interpretation of real-time optical flow for gesture recognition. *IEEE International Conference on Automatic Face and Gesture Recognition*, avril 1998.

- [16] Romer Rosales, Matheen Siddiqui, Jonathan Alon, and Stan Sclaroff. Estimating 3D body pose using uncalibrated cameras. Technical Report 2001-008, 5 2001.
- [17] E. Granum T. Moeslund. Pose estimation of human arm using kinematics constraints. In *12th Scandinavian conference on image analysis (SCIA'01)*, Bergen, Norway, June 2001.
- [18] A. Pentland T. Starner. Real-time american sign language recognition from video using hidden markov models. Technical Report TR-375, M.I.T Media Laboratory Perceptual Computing Section, 1995.
- [19] C.J. Taylor. Reconstruction of articulated objects from points correspondences in a single uncalibrated image. *Computer Vision and Image Understanding*, 80(3):349–363, Decembre 2000.

Annexe

La matrice de passage 0T_5 du repère 0 au repère 5. ($\cos \theta_{ij}$ signifie $\cos \theta_i * \cos \theta_j$):

$${}^0T_5 = \begin{pmatrix} R_1 & R_2 & R_3 & T \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R_1 = \begin{pmatrix} \cos \theta_4 (\cos \theta_{123} - \sin \theta_{13}) - \sin \theta_{42} \cos \theta_1 \\ \cos \theta_4 (\sin \theta_1 \cos \theta_{23} + \cos \theta_1 \sin \theta_3) - \sin \theta_{124} \\ -\cos \theta_{34} \sin \theta_2 - \cos \theta_2 \sin \theta_4 \\ -\sin \theta_3 \cos \theta_{12} - \sin \theta_1 \cos \theta_3 \end{pmatrix}$$

$$R_2 = \begin{pmatrix} -\sin \theta_{13} \cos \theta_2 + \cos \theta_{13} \\ \sin \theta_{23} \\ \sin \theta_4 (\cos \theta_{123} - \sin \theta_{13}) + \cos \theta_{14} \sin \theta_2 \\ \sin \theta_4 (\cos \theta_{23} \sin \theta_1 + \sin \theta_3 \cos \theta_1) + \sin \theta_{12} \cos \theta_4 \\ -\sin \theta_{24} \cos \theta_3 + \cos \theta_{24} \end{pmatrix}$$

$$R_3 = \begin{pmatrix} \sin \theta_4 (\cos \theta_{123} - \sin \theta_{13}) + \cos \theta_{14} \sin \theta_2 \\ \sin \theta_4 (\cos \theta_{23} \sin \theta_1 + \sin \theta_3 \cos \theta_1) + \sin \theta_{12} \cos \theta_4 \\ -\sin \theta_{24} \cos \theta_3 + \cos \theta_{24} \end{pmatrix}$$

$$T = \begin{pmatrix} l_2 (-\sin \theta_4 (\cos \theta_{123} - \sin \theta_{13}) - \cos \theta_{14} \sin \theta_2) + l_1 \cos \theta_1 \sin \theta_2 \\ l_2 (-\sin \theta_4 (\cos \theta_{23} \sin \theta_1 + \sin \theta_3 \cos \theta_1) - \cos \theta_4 \sin \theta_{12}) + l_1 \sin \theta_{12} \\ -l_2 (-\sin \theta_{24} \cos \theta_3 - \cos \theta_{24}) + l_1 \cos \theta_2 \end{pmatrix} \quad (13)$$

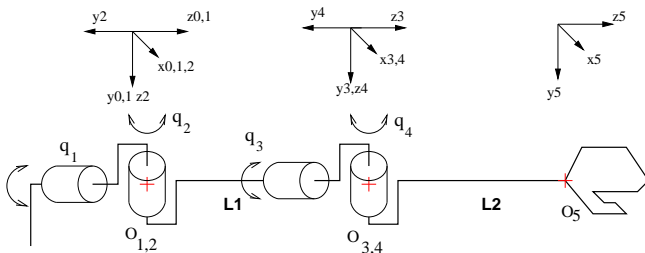


Figure 6: Modèle du bras avec les repères utilisés pour le calcul du modèle géométrique direct.

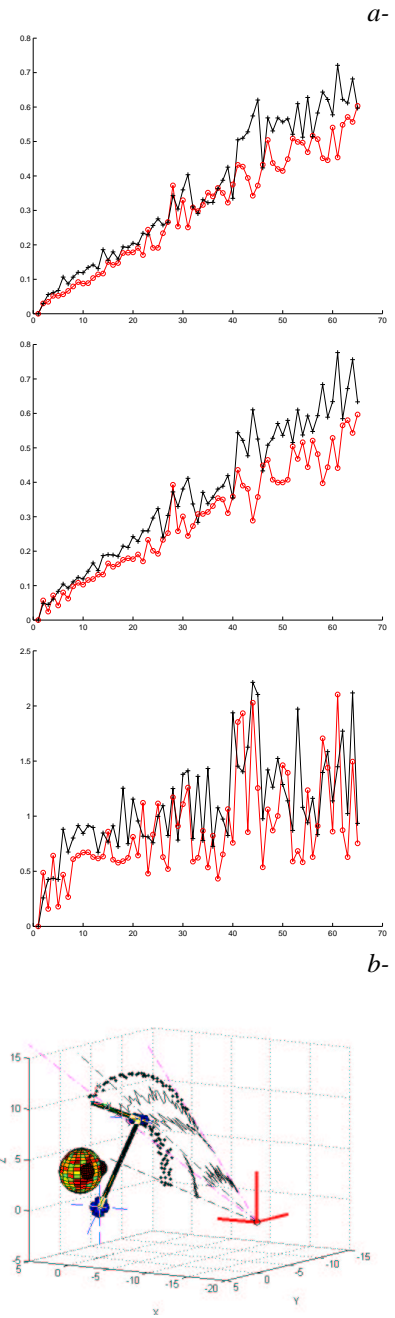


Figure 7: *a-* Évolution de l'erreur d'estimation 3D de la position du coude (o) et du poignet (+) en fonction de l'amplitude du bruit sur les points image (De haut en bas erreur sur X, Y et Z, Z étant dans notre repère la profondeur). L'erreur sur ces points est donnée en pixels. L'amplitude moyenne de l'erreur (donnée en valeur absolue) est sans dimension car nous ne disposons pas de référence pour calculer le facteur d'échelle. En revanche elle peut être comparée aux longueurs estimées des segments du bras soit 2.3 pour le bras et 2.15 pour l'avant-bras. *b-(en bas)* Reconstruction des trajectoires dans une scène à partir d'images de synthèse. Du bruit a été rajouté sur les points image à hauteur de 5 % de la taille de l'image. Le filtrage des trajectoires a été effectué manuellement. Il est à noter que la position relative des points calculés reste inchangée. De tels résultats restent donc exploitables dans le cadre de mesures qualitatives sur les positions relatives des différents segments du bras.