

# Suivi du corps humain par vision monoculaire

## Human Body Tracking by Monocular Vision

par Frédéric LERASLE, Gérard RIVES, Michel DHOME, Ali YASSINE

Laboratoire des Sciences et Matériaux pour l'Electronique, et d'Automatique (LASMEA)  
URA 1793 du CNRS, Université Blaise Pascal de Clermont-Ferrand 63177 Aubière cedex.  
lerasle@lasmea.univ-bpclermont.fr, tél : 04.73.40.72.30, fax : 04.73.40.72.62

### *résumé et mots clés*

Cet article décrit une méthode de suivi d'objets 3D articulés, de forme complexe telle que les membres du corps humain, à partir d'une séquence monoculaire d'images perspectives. Ces objets devront être préalablement modélisés ainsi que les articulations associées. Le principe de la méthode est basé sur l'interprétation de points caractéristiques dans l'image comme étant les projections perspectives de points 3D liés au modèle de l'objet et d'un processus itératif fondé sur la méthode de Levenberg-Marquardt pour calculer l'attitude de ce modèle conforme à l'image analysée.

Cette attitude et la covariance associée sont fournies à un filtre de Kalman de haut niveau pour prédire l'attitude du modèle relative à l'image suivante de la séquence. L'extraction de primitives s'effectue localement par prise en compte de la prédiction obtenue.

Des expérimentations de suivi, illustrées dans cet article par une séquence de pédalage, ont été effectuées, prouvant ainsi la validité de l'approche.

Vision monoculaire, modèle polyédrique articulé, appariement, localisation, suivi.

### *abstract and key words*

This article describes a tracking method of 3D articulated complex objects (for example, the human body), from a monocular sequence of perspective images. These objects and their associated articulations must be modelled. The principle of the method is based on the interpretation of image features as the 3D perspective projections points of the object model and an iterative Levenberg-Marquardt process to compute the model pose in accordance with the analysed image.

This attitude is filtered (Kalman filter) to predict the model pose relative to the following image of the sequence. The image features are extracted locally according to the computed prediction.

Tracking experiments, illustrated in this article by a cycling sequence, have been conducted to prove the validity of the approach.

Monocular vision, Articulated polyhedric model, Matching, Localization, Tracking.

## 1. introduction

Les physiologistes, au lieu de mesurer directement le travail mécanique effectué, ont l'habitude de passer par la thermochimie (mesures des échanges gazeux) pour en déduire indirectement le travail produit par le sujet. Or, le travail dissipé à un quelconque ergomètre (bicyclette ou tapis roulant) ne représente que le travail externe. Il s'avère que la non mesure de l'énergie « cinétique » interne sous-estime grandement le travail mécanique total. Il s'en suit alors des approximations grossières dans le calcul du rendement des mouvements humains.

Les travaux portant sur l'analyse cinématographique, puis plus récemment ceux mettant en oeuvre des techniques associant « vidéo » et marqueurs ont prouvé qu'il était possible d'obtenir à partir de telles analyses des informations quantifiables permettant d'accéder à la mesure de cette énergie interne.

Dans cette même problématique, les travaux présentés ici portent sur l'analyse automatique du mouvement d'un sujet observé par un système de vision. Comme très souvent pour ce genre de méthodes, nous supposons connu *a priori* : l'objet observé sous forme de son modèle articulé volumique, le type de projection utilisée pour former les images (perspective dans notre cas), un ensemble d'appariements entre primitives 2D extraites de l'image et éléments 3D du modèle. Pour illustrer ces travaux,

nous nous sommes intéressés à une séquence de pédalage. Pour s'affranchir des problèmes d'occultation, une des deux jambes sera volontairement masquée durant toute la séance de prise de vue.

## 2. état de l'art

Durant ces dernières années, certains chercheurs ont travaillé sur la localisation des objets articulés par vision monoculaire. Les méthodes basées sur le modèle de l'objet se classent suivant trois types d'approches.

Les méthodes à division ([Sha91],[Grim87]) considèrent chacune des parties de l'objet articulé séparément et estiment leurs positions indépendamment les unes des autres. Dans [Grim87], Grimson valide le paradigme sur une paire de ciseaux : il localise chacune des branches séparément puis vérifie la contrainte rotatoire de mouvement définie entre ces deux parties. Cette approche est attractive par sa simplicité. Cependant, nous n'exploitons pas directement les contraintes entre parties du même objet. Par conséquent, chacune des parties est localisée à partir de ses propres mesures.

Les méthodes à fusion de contraintes tiennent compte des mesures et des contraintes dans le processus d'estimation. L'idée est de traiter les mesures de la même manière que les contraintes. Dans [Hel94], les mesures sont considérées comme des contraintes molles tandis que les contraintes liées au modèle sont considérées comme fortes. La fusion entre mesures et contraintes est réalisée pendant le processus d'estimation à partir d'un filtre de Kalman. Voici, à titre d'exemples, quelques contraintes utilisées : distances constantes, coplanarité entre points du modèle, invariance de centres de rotation...

Une dernière approche possible ([Mul89],[Low91]) pour cette problématique est de ne pas prendre en compte directement les contraintes, ceci en réduisant le nombre de paramètres décrivant l'attitude de l'objet. On introduit alors la notion de degré de liberté. Dans [Mul89], Mulligan reprend cette approche pour localiser le bras d'une pelleuse défini par trois degrés de liberté. La méthode qu'il propose détermine les degrés de liberté les uns après les autres.

Comme Mulligan et Lowe, nous avons choisi cette dernière d'approche, même si notre méthode se différencie très nettement de celle de Mulligan car nous estimons conjointement les paramètres extrinsèques et les degrés de liberté internes. Elle est basée sur l'interprétation de caractéristiques de l'image comme étant les projections perspectives de primitives (arêtes ou points surfaciques) associées au modèle. Nous recherchons l'attitude globale de l'objet articulé superposant au mieux les projections perspectives des primitives concernées du modèle et les caractéristiques associées dans l'image. Une approche similaire est décrite dans [Low91] où Lowe propose une méthode pour superposer un modèle 3D paramétré à une image perspective. Cependant, notre méthode diffère de celle de Lowe dans le critère à minimiser : Lowe utilise un critère 2D calculé dans le plan image alors que notre critère est 3D ce qui permet de simplifier les calculs dans notre processus.

Concernant le problème de la localisation du corps humain, Winter [Wi79] a décrit une méthode d'analyse possible à base de marqueurs, le corps étant modélisé par 14 segments indéformables et articulés entre eux. Ces marqueurs sont placés aux extrémités de chaque segment (cuisse, bras, tronc ...), c'est-à-dire aux articulations. Cet ajout de primitive dans la scène permet une mise en correspondance non ambiguë entre projections 2D de ces marqueurs et le modèle de référence.

Actuellement, la plupart des techniques concernant l'analyse du mouvement humain, utilisent des marqueurs ([Yea89],[Els93]). Nous pouvons citer le système complet **ELITE** de Bioengineering Technology ([Bor88]) : les caméras sont munies de flashes infra-rouge qui permettent leur utilisation quelles que soient les conditions d'éclairage. Les marqueurs de type catadioptrique, placés au niveau des articulations du sujet, réfléchissent la lumière infra-rouge des flashes stroboscopiques. Les caméras détectent les marqueurs et le système calcule leurs positions 3D.

Le système **VICON** ([Mol85]) est un analyseur de mouvement utilisant 30 marqueurs passifs réfléchissants et 7 caméras synchrones cadencées à 200 images/seconde. Par des conditions d'éclairage particulières et par un pré-traitement approprié sur les signaux vidéo, le système permet l'acquisition d'images très contrastées où l'emplacement des marqueurs est facilement repérable.

Le principe du système **CODA-3** [Ath84] est de scanner l'espace de l'objet. Un dispositif de miroirs mobiles est utilisé pour produire des rayons de lumière blanche balayant la scène. Quand un rayon croise un marqueur réfléchissant, une impulsion lumineuse est réfléchiée via le chemin inverse. Cette impulsion est alors détectée par des photodiodes dans l'unité de balayage. la position du marqueur est déterminée en mesurant l'intervalle de temps entre le départ du balayage et la détection de l'impulsion.

Le système **SELSPOT** [Ath84] place sur le sujet observé des diodes émettant dans les infra-rouges. Ceci permet de s'affranchir d'une quelconque condition d'éclairage. Des détecteurs de lumière infra-rouge sont positionnés à des endroits pré-déterminés dans la scène. Ceux-ci délivrent alors des signaux électriques proportionnels à la position (horizontale et verticale) du point d'impact d'un rayon infra-rouge sur la surface du détecteur.

L'emploi de ces marqueurs pose quelques problèmes. En premier lieu, la non rigidité de l'enveloppe corporelle, durant le mouvement, entraîne un déplacement relatif corps/marqueurs et engendre une incertitude sur les résultats. De plus, l'installation est simple pour les chevilles et les poignets, mais pose des difficultés pour les articulations complexes tels que genoux, épaules et hanches. Pour ce type de méthodes, le mouvement idéal vu par la caméra doit être dans un plan perpendiculaire à son axe optique, puisque les mouvements estimés sont supposés planaires. Il est alors évident que même si le déplacement vérifie cette condition, certains membres ne sont pas dans ce plan. Enfin, la présence de marqueurs passifs ou actifs provoque inévitablement chez le sujet observé un effet psychologique voire une rigidité dans le mouvement. Il est donc souhaitable de réduire le plus possible les contraintes sur le sujet lui-même.

Geurtz [Ge91] se démarque de ces méthodes en n'utilisant aucun marqueur. Mais sa méthode d'analyse se limite à la représentation bidimensionnelle du corps. Les segments corporels du modèle 2D

sont restreints à des ellipses décrivant les contours des segments. Les attributs de l'image sont constitués seulement de contours. Blake [Bla93] a également développé une méthode de suivi à partir des contours. Le contour de l'objet (supposé planaire) est modélisé par une  $\beta$ -spline passant par un ensemble limité de points de contrôle. L'estimation du contour s'effectue au voisinage de la position des points de contrôle prédit par un filtre de Kalman. Bien sûr, cette approche ne peut donner lieu qu'à une estimée du déplacement 2D.

Pour lever l'ambiguïté mouvement-profondeur, certains chercheurs se sont orientés vers les modèles volumiques déformables ou statiques, issus de connaissance *a priori* de l'objet observé. Les méthodes basées sur des modèles déformables sont généralement associées à des techniques de décomposition modale « mécanique » ([Pen91],[Ter91],[Na93]). Les modes de fréquences les plus basses représentent les modes de rotation et de translation d'un solide rigide tandis que les solides déformables à forte dissymétrie géométrique peuvent avoir substantiellement des modes de fréquences très élevées. Dans [Pen91], Pentland utilise pour sa modélisation un maillage de l'objet représenté par ses noeuds (éléments finis). L'équation cinématique du système s'interprète comme l'assignation d'une certaine masse à chacun des noeuds et d'une certaine élasticité entre ces noeuds. Terzopoulos [Ter91] reprend cette approche mais il modélise l'objet par un ensemble de quadriques. Le problème crucial de ce type de méthode réside dans la modélisation du comportement élastique des matériaux.

Pour limiter la complexité du problème, certains chercheurs ([Reh94],[Wa92]) ont considéré le corps humain comme un solide rigide dans ses dimensions. Wang [Wa92] modélise les différents membres corporels par des primitives géométriques simples (cylindres, surfaces planes) articulées entre elles. Les différentes images de la séquence sont segmentées en région au sens du mouvement. A partir de ces régions et d'un modèle affine du mouvement, Wang en déduit le mouvement 2D dans chacune des images de la séquence pour remonter aux paramètres 3D de positionnement de ce modèle. L'erreur de modélisation, due à la représentation par de simples cylindres peut avoir pour effet de biaiser le résultat de l'interprétation, voire de conduire à une interprétation fautive. Une modélisation plus élaborée, améliorerait certainement les résultats de l'analyse.

Pour notre part, nous avons également choisi un modèle volumique statique. Cependant, ce modèle est plus réaliste car il est construit à partir de coupes I.R.M. La section suivante décrira de manière détaillée notre modèle volumique articulé.

### 3. description du modèle

La modélisation du corps humain s'inspire des travaux réalisés au laboratoire sur les objets articulés. Nous les rappelons brièvement ici.

Le modèle articulé est composé de plusieurs modèles C.A.O. (un pour chaque partie de l'objet articulé) connectés entre eux par des articulations qui décrivent les mouvements relatifs possibles entre ses différentes parties.

Chaque modèle C.A.O. correspond à une approximation polyédrique de la partie réelle associée. Pour la partie concernée, il est composé des coordonnées des sommets 3D, exprimées dans un repère dit repère modèle, des arêtes construites à partir de ces sommets et des surfaces délimitées par ces arêtes.

Concernant nos expérimentations sur la séquence de pédalage, il nous a fallu construire les modèles C.A.O. de la cuisse et du mollet du sujet visionné. Ils ont été réalisés à partir des images acquises lors d'un examen I.R.M. Cet examen consistait en 34 coupes de la jambe distantes de 21 mm. Un traitement bas niveau approprié, lissage et détection de contours, nous a permis d'extraire les contours périphériques de chaque coupe. Les coordonnées  $x$  et  $y$  des points de contour de chaque coupe, associées à sa hauteur  $z$ , nous ont permis de déterminer les sommets 3D du modèle. La détermination et le classement de ces sommets 3D sont réalisés par un programme spécifique.

Chaque articulation décrivant le mouvement relatif entre deux parties est caractérisée par un ou plusieurs degrés de liberté. Chaque degré de liberté est défini par son type, rotation ou translation, et l'axe 3D qui le régit.

Dans nos expérimentations, il nous a fallu modéliser l'articulation du genou. Après avoir consulté des orthopédistes au C.H.R.U de Clermont-Ferrand, nous avons opté pour un modèle d'articulation à 3 rotations, correspondant à la rotation dite flexion-extension (axe  $Oy$ ), à la rotation dite interne (axe  $Oz$ ) et enfin à la rotation dite valrus-valgus (axe  $Ox$ ).

La figure 1 représente le modèle complet de la jambe et les degrés de liberté associés. Nous reconnaissons horizontalement les différentes coupes I.R.M. qui nous ont permis de déterminer les sommets du modèle.

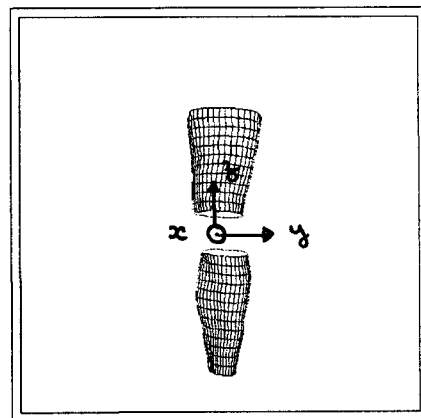


Figure 1. – Modèle de la jambe.

Pour animer le modèle ainsi constitué, il nous a fallu définir des opérateurs permettant de positionner dans le repère observateur un sommet 3D du modèle, initialement défini dans le repère modèle.

Par exemple, le  $i^{\text{ème}}$  sommet appartenant à la  $j^{\text{ème}}$  partie subira lors de l'animation de l'objet articulé, la transformation permettant de positionner dans l'espace la partie de référence de l'objet,

puis successivement toutes les transformations liées aux degrés de liberté situés entre la partie de référence et la  $j^{ème}$  partie. Dans les équations suivantes, l'exposant  $c$  fera référence au repère caméra ou observateur, et l'exposant  $m$  au repère modèle.

$$P_i^c = [M_{a_q}]^* \dots [M_{a_x}]^* \dots [M_{a_1}]^* \cdot [M] \cdot P_i^m \quad (1)$$

$[M]$  représente la transformation due aux paramètres extrinsèques de positionnement de la partie de référence. Cette transformation correspond à 3 rotations et 3 translations notées respectivement  $\alpha, \beta, \gamma, u, v, w$ .

$[M_{a_x}]^*$  représente la transformation du  $\chi^{ème}$  degré de liberté interne de l'objet. L'astérisque \* signifie que cette transformation est égale à la matrice identité si le degré de liberté  $a_x$  n'intervient pas entre la partie de référence et la  $j^{ème}$  partie de l'objet.

En réalité,  $[M_{a_x}]^*$  est une matrice de rotation ou de translation.

Pour une matrice de rotation,  $[M_{a_x}]^*$  vaut :  $[M_{a_x}]^* = [T_{OC_x}] \cdot [R_{a_x}]^* \cdot [T_{OC_x}]^{-1}$

où  $[T_{OC_x}]$  est la matrice de translation de l'origine du repère modèle  $O$  au centre de la liaison  $C_x$  par lequel passe l'axe de la rotation  $[R_{a_x}]^*$ .  $[R_{a_x}]^*$  est exprimé sous forme d'une matrice de quaternions prenant en compte  $a_x$  l'angle de rotation et  $\vec{\Lambda}_{a_x}$  le vecteur directeur de l'axe de rotation.

Dans le cas d'une matrice de translation,  $[M_{a_x}]^*$  est donnée par :  $[M_{a_x}]^* = [T_{a_x}]^*$

où  $[T_{a_x}]^*$  est une matrice prenant en compte  $a_x$  l'amplitude de la translation et  $\vec{\Lambda}_{a_x}$  le vecteur directeur de l'axe de translation.

## 4. appariements 2D-3D

Avant tout processus de localisation, il faut extraire un nombre restreint d'indices image et les appairer avec les primitives géométriques correspondantes du modèle articulé. Dhome [Dho89] calcule la position d'un objet polyédrique simple à partir de mises en correspondance entre arêtes du modèle et segments de droite extraits d'une image de luminance. Le processus de couplage s'effectue sur la totalité des arêtes visibles du modèle. Dans notre application, où les modèles C.A.O. propres à chaque parties ont une géométrie répétitive (correspondant à une surface gauche maillée), il est impossible d'appliquer une telle démarche. Nous ne travaillons que sur un sous ensemble spécifique d'arêtes qui sont les limbes du modèle. Un limbe est une arête commune à une surface visible et à une surface invisible, après projection du modèle dans le plan image. Il est clair que la non rigidité de l'enveloppe corporelle durant le mouvement entraîne une incohérence inévitable par rapport au modèle statique défini et donc par rapport aux limbes détectés. Cependant, pour des mouvements non saccadés de type pédalage, les déformations sont relativement faibles.

Pour apporter des contraintes supplémentaires, des points tirés aléatoirement à la surface du modèle sont appariés à leurs attributs visuels détectés dans l'image. Ce suivi de points spécifiques est assimilable à une technique classique de suivi de marqueurs, mais

les avantages en sont évidents : le nombre et l'emplacement de ces points caractéristiques ne sont pas prédéfinis comme pour les marqueurs. Ces points peuvent être oubliés ou retirés au cours du processus de suivi.

Les primitives extraites de l'image sont donc du type segment de droite et du type point. Du fait des déformations corporelles pouvant intervenir durant le mouvement, il est évident que le modèle articulé employé ne correspondra qu'approximativement aux contenus des images de luminance. Pour obtenir une attitude moyenne cohérente avec le contenu de l'image traitée, il faut donc se donner un maximum de contraintes. L'emploi conjoint de points surfaciques et d'arêtes du modèle, dans le processus d'appariement, va dans ce sens.

### 4.1. appariements à partir de points

Nous proposons une méthode de mise en correspondance de points 3D du modèle par corrélation dans les images de luminance. Nous disposons de l'attitude du modèle conforme à la première image de la séquence. Cette attitude a été recalée manuellement sur la première image par un processus interactif. Nous sélectionnons dans cette première image, des points 2D  $p_i^1(u_i^1, v_i^1)$  et nous cherchons leurs équivalents  $p_i^k(u_i^k, v_i^k)$  dans les images successives (notées  $I^k$ ) de la séquence par analyse d'une fenêtre de corrélation.

*A priori*, il est nécessaire de rechercher le point  $p_i^k$  dans une zone de l'image  $I^k$  suffisamment grande pour englober le déplacement existant entre les points homologues des deux images. Pour réduire la combinatoire, nous utiliserons la position prédite estimée pour l'image  $I^k$ . Cette prédiction sera calculée par un filtre de Kalman haut niveau (filtrant les attitudes 3D successivement calculées). La recherche du point  $p_i^k$  dans l'image  $I^k$  se fera au voisinage du point 3D projeté ( $u_i^k \text{ pred}, v_i^k \text{ pred}$ ) du modèle ayant subi la transformation prédite. La prédiction permet aussi d'estimer la rotation planaire existant entre les fenêtres de corrélation liées aux images  $I^1$  et  $I^k$ .

Pour schématiser la méthode, nous associons des points 3D du modèle à des motifs texturés de l'image de luminance initiale et nous les cherchons, par corrélation sur les niveaux de gris, dans les images successives de la séquence. Nous prendrons en compte la rotation éventuelle du motif dans l'image  $I^k$  en faisant tourner la fenêtre de corrélation liée à cette image. Connaissant  $n$  la demi-taille de la fenêtre de corrélation et  $t$  la demi-taille de la zone de recherche de l'image  $I^k$ , on définit le processus de corrélation utilisé :

Initialisation :  $x_1 = u_i^k, y_1 = v_i^k, Cmax = 0$

Recherche de la corrélation maximum :

pour  $-t \leq dx \leq t$  et  $-t \leq dy \leq t$ ,

$x_2 = u_i^k \text{ pred} + \cos a \cdot dx + \sin a \cdot dy$ ,

$y_2 = v_i^k \text{ pred} - \sin a \cdot dx + \cos a \cdot dy$

somXY = 0, somX2 = 0, somY2 = 0

(a) pour  $-n \leq ax \leq n$  et  $-n \leq ay \leq n$ ,

$x_1^t = x_1 + ax, \quad y_1^t = y_1 + ay$

$x_2^t = x_2 + \cos a \cdot ax + \sin a \cdot ay$ ,

$$y_2^t = y_2 - \sin a \cdot ax + \cos a \cdot ay$$

$$\text{som}XY = \text{som}XY + I_1(x_1^t, y_1^t) \cdot I_k(x_2^t, y_2^t)$$

$$\text{som}X2 = \text{som}X2 + I_1(x_1^t, y_1^t)^2,$$

$$\text{som}Y2 = \text{som}Y2 + I_k(x_2^t, y_2^t)^2$$

$$(b) C = \frac{\text{som}XY}{\text{som}X2 \cdot \text{som}Y2}$$

si  $C > C_{max}$  alors  $C_{max} = C$

$$x2_{max} = x_2 \text{ et } y2_{max} = y_2$$

$$u_i^k = x2_{max}, v_i^k = y2_{max}, \text{ fin}$$

Le point 2D dans l'image  $k$  peut donc être associé au point 3D correspondant du modèle. La localisation de ce point 2D est d'autant meilleure que la texture contenue dans les images est riche. Pour ce faire, durant la séance de prise de vue, le sujet observé est vêtu d'un juste au corps à texture non répétitive.

## 4.2. appariements à partir de limbes

Pour l'attitude propre à l'image initiale, une méthode de Z-buffer permet d'extraire les arêtes du modèle qui correspondent aux limbes.

La connaissance *a priori* d'un collant sombre texturé sur un arrière plan blanc est alors exploitée. Les points extrémités de la projection des limbes sont recalées sur les transitions blanc/noir de l'image de luminosité de départ. Ce recalage a lieu dans la direction perpendiculaire à la projection de la limbe traitée. Les deux points retenus sont ceux qui rendent maximale la fonction de corrélation  $C$  simplifiée car monodimensionnelle. Le signal de référence  $I_1(i, j)$  est une marche idéale représentée sur la figure 2 (coin bas droite).

Soient  $P_{11}$  et  $P_{12}$  les points extrémités recalés dans l'image  $I^{k-1}$  et  $P_{21}$  et  $P_{22}$  leurs correspondants trouvés dans l'image  $I^k$  après corrélation. Les segments  $P_{11}P_{12}$  et  $P_{21}P_{22}$  sont échantillonnés chacun en  $n$  points, notés respectivement  $p_1^i$  et  $p_2^i$ ,  $i = 1..n$ . Pour chaque échantillon  $i$ , le segment de longueur  $l$  centré sur  $p_1^i$  et perpendiculaire à  $P_{11}P_{12}$  est corrélé avec un segment de même longueur, centré sur  $P_2^{ik}$  et glissant sur la droite support, passant par  $P_2^i$  et perpendiculaire à  $P_{21}P_{22}$ . Le point  $P_2^{ik}$  ayant obtenu le meilleur score de corrélation est retenu.

L'ensemble  $P$  des points  $P_2^{ik}, i = 1..n$  doivent théoriquement se trouver sur une droite. Pour calculer les paramètres de cette droite, la méthode des moindres carrés n'est pas appropriée, une technique robuste dite de RANSAC lui a été adjointe [Fis81].

## 5. localisation, erreurs associées

Dans cette section, nous poserons le problème tel qu'il est formulé, les équations mathématiques résultantes seront explicitées ainsi que l'algorithme employé pour les résoudre. Cet algorithme a été validé par Yassine [Ya95] sur des objets articulés polyédriques type bras manipulateur. Un calcul de covariance associée à l'attitude finale sera développé en fin de section.

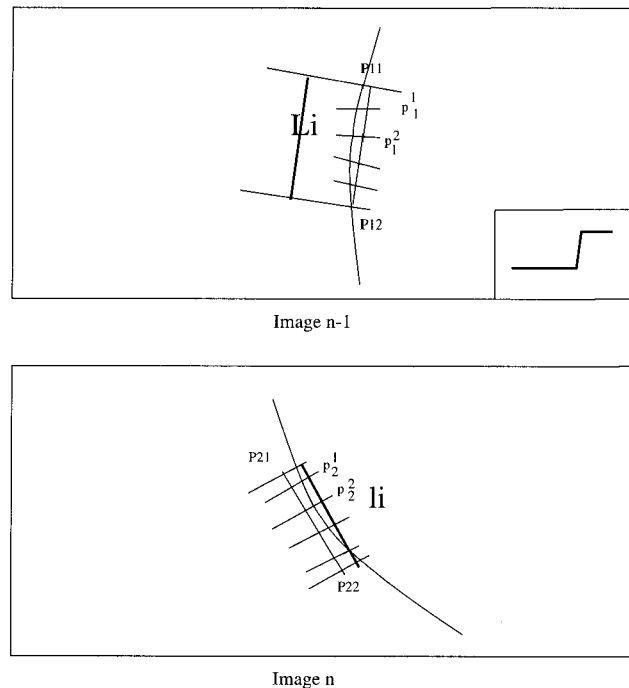


Figure 2. – Appariement d'une limbe  $L_i$  avec un segment  $l_i$  dans l'image.

### 5.1. formulation générale

La localisation d'un objet articulé à partir d'une image dépend de  $10 + k$  paramètres.

Les 4 premiers sont les paramètres intrinsèques de la caméra et représentent les coordonnées  $(u_0, v_0)$  de l'intersection de l'axe optique avec le plan image, la longueur  $f$  de la focale et le rapport  $d_x/d_y$  de la taille horizontale et verticale de la cellule élémentaire du capteur CCD de la caméra. Ces paramètres sont prédéterminés par une étape de calibrage.

Les 6 paramètres suivants sont les paramètres extrinsèques de positionnement, dans le repère caméra, de la partie de référence de l'objet articulé (3 rotations + 3 translations).

Les  $k$  paramètres suivants correspondent aux  $k$  degrés de liberté internes liés aux articulations, chaque degré de liberté ayant au maximum 6 paramètres.

Il nous faut donc estimer le vecteur de position ayant  $6 + k$  paramètres de positionnement de l'objet articulé en conformité avec l'image de luminosité.

Ces  $6 + k$  paramètres devront minimiser la distance de points 3D du modèle à des plans d'interprétation. Le plan d'interprétation d'un segment ( $l$  dans la figure 3) de l'image est le plan incluant le segment et passant par le centre optique de la caméra. Ainsi, pour les primitives du modèle de type arête, seules les extrémités des arêtes ( $E1$  et  $E2$ ) seront prises en compte. Pour chaque arête appariée, nous minimiserons la distance des deux extrémités de l'arête à un même plan d'interprétation ( $\pi$  dans la figure 3). Pour les appariements sur des points, les points appariés dans l'image ( $p$  dans la figure 3) seront remplacés par deux

segments perpendiculaires dans l'image. Nous minimiserons donc la distance du point 3D du modèle ( $P$  dans la figure 3) aux deux plans d'interprétation perpendiculaires ainsi formés ( $\pi_1$  et  $\pi_2$  dans la figure 3).

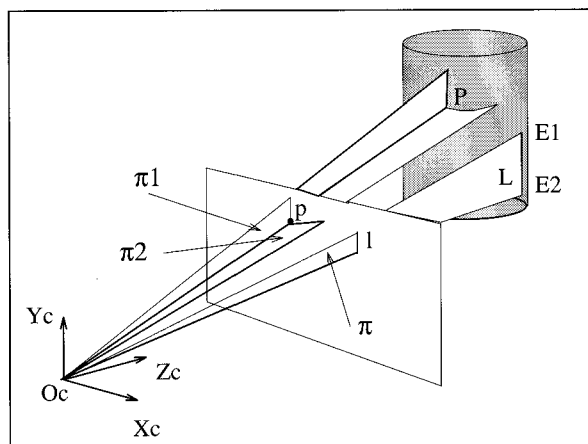


Figure 3. – Exemple de plans d'interprétation.

## 5.2. distances aux plans d'interprétation

Soient  $n$  segments image  $l_i$  appariés avec  $n$  points  $P_i^m$  du modèle. Supposons que tous les vecteurs et tous les points soient exprimés dans le système de coordonnées de la caméra ( $O, X_c, Y_c, Z_c$ ), (voir figure 5).

Chaque segment  $l_i$  de l'image est caractérisé par son vecteur directeur  $\vec{v}_i$  et un point  $p_i^c$ . Nous pouvons alors calculer aisément le vecteur normal  $\vec{N}_i$  au plan d'interprétation  $\Pi_i$  de  $l_i$ , soit :

$$\vec{N}_i = \frac{\vec{v}_i \wedge \overrightarrow{Op_i^c}}{\|\vec{v}_i \wedge \overrightarrow{Op_i^c}\|}$$

Nous recherchons le vecteur de positionnement à  $6 + k$  composantes noté  $A$ , qui minimise la somme des distances entre les points 3D appariés du modèle et les plans d'interprétation des segments correspondants dans l'image. Une distance à un plan d'interprétation  $\pi_i$  est représentée par le produit scalaire du vecteur  $\overrightarrow{Op_i^c}$  avec la normale au plan  $\pi_i$  ( $P_i^c$  étant le transformé de  $P_i^m$  par les transformations paramétrées par le vecteur  $A$  - voir équation 1). Soit la fonction de distance :

$$F(A, \vec{N}_i, P_i^m) = (\vec{N}_i \cdot \overrightarrow{Op_i^c})$$

où  $i = 1..n$  est l'indice du point apparié ( $n$  étant le nombre total de points appariés).

## 5.3. résolution du système

Une des méthodes de résolution d'un système non linéaire est de linéariser les équations qui le constituent, localement autour

d'une valeur connue. L'approximation au premier ordre par le développement de Taylor de la fonction  $F(A, P_i^m)$  autour de la valeur  $(A)_k$  s'écrit :

$$F(A, P_i^m) \approx F(A_k, \vec{N}_i, P_i^m) + \frac{\partial F(A_k, \vec{N}_i, P_i^m)}{\partial A} (A - A_k)$$

Or nous savons que, pour le vecteur  $A$  recherché, la fonction  $F(A, \vec{N}_i, P_i^m)$  est nulle (au bruit près des primitives image). Ceci conduit à l'approximation suivante :

$$-F(A, \vec{N}_i, P_i^m) = \frac{\partial F(A_k, \vec{N}_i, P_i^m)}{\partial A} (A - A_k)$$

pour  $i = 1..n$ . L'ensemble des points 3D appariés nous permet de constituer un système de  $n$  équations linéaires :

$$(E)_k = [J]_k \cdot (\Delta A)_k$$

et donc :

$$[J]_k^T \cdot (E)_k = [J]_k^T \cdot [J]_k \cdot (\Delta A)_k$$

avec :

$$(\Delta A)_k = \begin{pmatrix} \alpha - \alpha_k \\ \beta - \beta_k \\ \vdots \\ w - w_k \\ a_1 - a_{1k} \\ \vdots \\ a_q - a_{qk} \end{pmatrix}, \quad (E)_k = \begin{pmatrix} -F(A_k, P_1^m) \\ -F(A_k, P_2^m) \\ \vdots \\ -F(A_k, P_n^m) \end{pmatrix}$$

$$(J)_k = \begin{pmatrix} \frac{\partial F(A_k, P_1^m)}{\partial \alpha} & \dots & \frac{\partial F(A_k, P_1^m)}{\partial a_1} & \dots & \frac{\partial F(A_k, P_1^m)}{\partial a_q} \\ \frac{\partial F(A_k, P_2^m)}{\partial \alpha} & \dots & \frac{\partial F(A_k, P_2^m)}{\partial a_1} & \dots & \frac{\partial F(A_k, P_2^m)}{\partial a_q} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial F(A_k, P_n^m)}{\partial \alpha} & \dots & \frac{\partial F(A_k, P_n^m)}{\partial a_1} & \dots & \frac{\partial F(A_k, P_n^m)}{\partial a_q} \end{pmatrix}$$

Pour résoudre ce système, nous utilisons la méthode itérative de Levenberg-Marquardt. A travers cette dernière, nous cherchons à minimiser le critère global :

$$\text{Erreur} = \sum_{i=1}^n (\vec{N}_i \cdot \overrightarrow{Op_i^c})^2$$

Partant d'une position connue  $(A)_0$  de l'objet articulé, nous déterminons à chaque étape  $k$  du processus itératif la valeur du vecteur correctif  $(\Delta A)_k$  à appliquer au vecteur d'état  $(A)_k$  afin de faire décroître le critère global :

$$(A)_{k+1} = (A)_k + (\Delta A)_k$$

Le processus itératif s'arrête lorsque le critère à minimiser est inférieur à un seuil fixé Erreur  $< \epsilon$ .

La méthode de Levenberg-Marquardt nécessite le calcul préalable des dérivées partielles de la fonction de distance  $F(A, \vec{N}_i, P_i^m)$ . Le calcul de ces dérivées est développé dans les deux paragraphes qui suivent. Une annexe en fin d'article détaille le calcul des dérivées partielles relatives aux paramètres de rotations et de translations.

### 5.4. covariance associée à l'attitude du modèle

Cette section montre le calcul des erreurs de localisation associées à notre méthode. Ne pouvant quantifier les déformations du modèle, nous ne prenons pas en compte les incertitudes sur les points 3D du modèle. Seules les incertitudes sur les mesures (liées aux images) interviennent dans notre calcul d'erreurs.

Une étude comparable a déjà été détaillée [Dau93]. Le calcul de la matrice de covariance  $[W]$  sur les  $n$  normales  $\vec{N}_i$  correspondantes aux  $n$  segments  $l_i$  est posé.

Comme il a été vu, la position finale du modèle est donnée après convergence d'un processus itératif minimisant la somme des distances des points 3D du modèle aux plans d'interprétation passant par les points appariés de l'image. Il est impossible de calculer analytiquement la matrice de covariance des paramètres de position (6 paramètres extrinsèques et  $k$  degrés de liberté). Cette matrice  $[\sigma_{loc}]$  est donc estimée après la dernière itération du processus d'optimisation,

Soit le système général défini par :  $\Delta A = ([J]^T \cdot [J]^{-1}) \cdot [J]^T \cdot E$

Soit  $[J]^+$  le pseudo inverse de  $[J]$  tel que  $\Delta A = [J]^+ \cdot E$

Nous pouvons estimer l'espérance mathématique  $[S]$  de  $(\Delta A) \cdot (\Delta A)^T$  connaissant l'espérance mathématique  $[W]$  de  $(E) \cdot (E)^T$ , soit :

$$(\Delta A) \cdot (\Delta A)^T = (J^+ \cdot E) \cdot (J^+ \cdot E)^T = J^+ \cdot E \cdot E^T \cdot (J^+)^T$$

d'où :

$$S = J^+ \cdot W \cdot (J^+)^T$$

La matrice de covariance  $[S]$ , obtenue par le calcul précédent, concerne le positionnement relatif (calcul en  $A = 0$ ). L'expression de cette matrice dans le repère caméra est obtenue à partir d'une matrice jacobienne  $J_{pos}$ . Cette matrice est facile à calculer. Si l'attitude du modèle est donnée par  $(A) = (\alpha, \beta, \gamma, u, v, w, a_1, \dots, a_k)$  où  $(\alpha, \beta, \gamma, u, v, w)$  est le six-uplet représentant les paramètres extrinsèques et  $(a_1, a_2, \dots, a_k)$  les degrés de liberté internes liés au modèle, nous pouvons écrire la matrice  $J_{pos}$  comme suit :

$$[J_{pos}] = \begin{pmatrix} \frac{\cos \gamma}{\cos \beta} & \frac{\sin \gamma}{\cos \beta} & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ -\sin \beta & \cos \gamma & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \cos \gamma \tan \beta & \sin \gamma \tan \beta & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & w & -v & 1 & 0 & 0 & 0 & \dots & 0 \\ -w & 0 & u & 0 & 1 & 0 & 0 & \dots & 0 \\ v & -u & 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Finalement, la matrice de covariance exprimée dans le repère caméra est donnée par :  $[\sigma_{loc}] = [J_{pos}] \cdot S \cdot [J_{pos}]^T$

## 6. expérimentations, résultats sur suivi

Une des extensions directes de toute méthode de localisation est le suivi du mouvement du sujet observé. Le suivi n'est autre qu'une localisation successive effectuée pour les différentes images de la séquence. Bien sûr, cette application suppose le développement de techniques d'appariements automatiques entre primitives présentes dans une image de luminance et primitives du modèle.

A chaque étape, la méthode de localisation donne l'attitude spatiale du modèle et la matrice de covariance associée, compatible avec l'image courante. Ces résultats sont fournis à un filtre de Kalman de haut niveau qui nous prédit la position du modèle pour l'image suivante. Le formalisme du filtre de Kalman tel que nous l'utilisons a été longuement décrit par Daucher [Dau93] et Ayache [Aya87]. La position prédite permet de restreindre le domaine de recherche pour la corrélation et donc d'en réduire son coût en temps de calcul. La prédiction permet d'estimer la rotation existant entre fenêtre de référence (liée à l'image  $I^1$ ) et fenêtre de recherche (liée à l'image  $I^k$ ) pour la corrélation point à point.

Les appariements entre primitives 3D du modèle dans l'attitude propre à une image de la séquence et primitives 2D dans l'image suivante sont automatiques sauf pour la première image de la séquence.

L'initialisation du processus de suivi comporte deux étapes : manuellement, l'opérateur positionne le modèle conformément au contenu de la première image de la séquence;

pour cette attitude initiale, l'opérateur choisit manuellement des points visibles du modèle. Par projection perspective inverse, nous remontons aux points 3D du modèle dans son repère de définition. Une bonne répartition spatiale de l'ensemble de ces points augmente les contraintes apportées par ces derniers. Cet ensemble de primitives appartenant au modèle forment la liste de points 3D à appairer dans les images suivantes.

Pour gérer la disparition éventuelle de ces points 3D et donc des motifs de texture associés dans l'image initiale, cette liste de points 3D du modèle sera mise à jour tout au long du suivi. D'une image à l'autre, à partir de l'attitude calculée, certains points sont dépilés de la liste et remplacés par d'autres tirés aléatoirement dans leur voisinage pour respecter la répartition spatiale initiale. Ce remaniement est agencé en pile L.I.F.O et paramétré de la sorte que après les  $n$  localisations correspondant aux  $n$  premières images de la séquence, tous les points de la liste initiale ont été retirés. C'est, en quelque sorte, un système de marqueurs (où plutôt de motifs de texture) glissants. Dans nos expérimentations, nous avons retiré tous nos points après 10 localisations.

Nous avons délibérément travaillé en monoculaire (une seule caméra) pour limiter les contraintes matérielles. Contrairement aux méthodes à marqueurs qui nécessitent un matériel assez lourd (caméras, fils de liaison, stroboscopes...), notre objectif est de développer une méthode « portable » et peu coûteuse.

Le processus de suivi a été validé sur des séquences d'images réelles. La première séquence est constituée d'une séquence de

pédalage où nous avons fixé sur le vélo 2 tuyaux de P.V.C. articulés entre eux. La deuxième séquence est aussi une séquence de pédalage avec, cette fois, une jambe de cycliste. Pour chacune des séquences, nous présentons les résultats obtenus.

Les figures 4 (gauche) et 5 (gauche) représentent les projections des attitudes calculées pour le point de vue de la caméra. Les figures 4 (droite) et 5 (droite) représentent les projections des attitudes calculées pour un point de vue situé face au vélo. Nous constatons que l'ensemble de ces projections définit bien un plan correspondant au plan de pédalage.

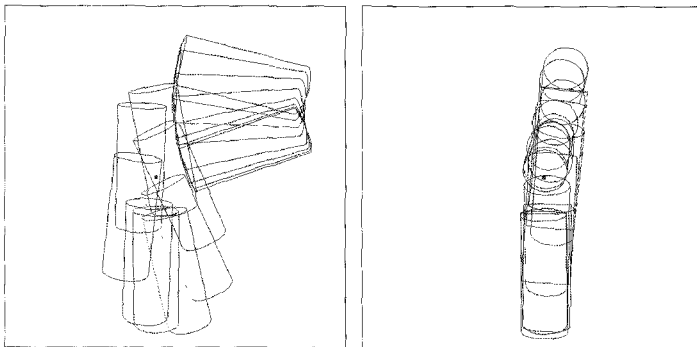


Figure 4. – Projection des attitudes calculées pour différents points de vue.

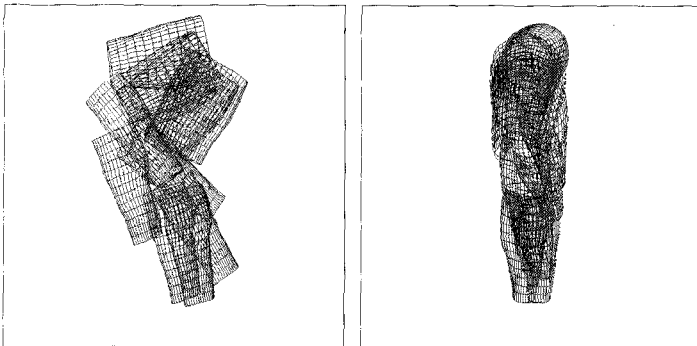


Figure 5. – Projection des attitudes calculées pour différents points de vue.

Les six images suivantes (figure 7 et figure 6) montrent la superposition de la projection du modèle articulé sur différentes images de la séquence. Pour la jambe réelle, nous constatons que les zones du modèle les moins en conformité avec le contenu de l'image de luminance sont des zones dites de déformation telles que le mollet et l'arrière de la cuisse.

## 7. conclusion et perspectives

Nous avons présenté une méthode permettant de résoudre le problème de suivi d'objets 3D articulés (à surfaces complexes) à partir d'une séquence monoculaire d'images perspectives.

Le principe est essentiellement basé sur la connaissance *a priori* de l'objet visualisé sous forme de son modèle C.A.O et sur

l'interprétation de points caractéristiques dans l'image comme étant les projections perspectives de points 3D liés au modèle de l'objet. C'est une méthode originale au sens où elle se démarque des méthodes classiques à base de marqueurs.

Les imprécisions de la méthode sont dans l'initialisation du processus et notamment l'estimation approximative de l'attitude du modèle propre à l'image de départ. Cette première position influence beaucoup la qualité des localisations obtenues durant le suivi. De plus, le modèle articulé statique défini n'est pas toujours cohérent avec le contenu des images à analyser car l'enveloppe corporelle n'est pas constante au cours du mouvement.

Nos efforts futurs devront donc porter sur l'initialisation et la modélisation ainsi que sur l'analyse de mouvement générant des occultations, par exemple l'occultation d'une jambe par l'autre. Le formalisme de Kalman déjà implémenté devrait nous aider à résoudre ce genre de problème.

Ces résultats seront exploités par la suite, par des chercheurs de l'UFR-STAPS (Sciences et Techniques des Activités Physiques et Sportives) de l'Université Blaise Pascal (Clermont-Ferrand) pour des calculs d'énergie cinétique interne et des estimations de rendement. A notre connaissance, ce serait la première fois qu'un processus sans marqueurs donnerait lieu à ce type de mesures.

### Annexe : Dérivées partielles relatives aux paramètres de localisation

La dérivée partielle de  $F(A_k, \vec{N}_i, P_i^m)$  par rapport à un des paramètres de position, fait intervenir l'équation (1). Le calcul direct de ces dérivées partielles est relativement complexe. Pour le simplifier, à chaque étape du processus itératif, nous modifions la position du modèle dans son repère de définition. Chaque sommet du modèle  $P_i^m$  est transformé en  $P_{ik}^m$  en lui appliquant le vecteur d'état  $(A)_k$  trouvé à l'étape  $k$ . Le calcul des dérivées partielles est alors grandement simplifié car la position des points  $P_{ik}^m$  est caractérisée, dans le nouveau repère de référence du modèle, par le vecteur d'état  $(A = 0)$ . Les dérivées partielles seront donc calculées pour ce vecteur d'état.

#### Paramètre de type rotation

La dérivée partielle de  $F(A_k, \vec{N}_i, P_{ik}^m)$  par rapport à  $\chi$  est :

$$\frac{\partial F(A_k, \vec{N}_i, P_{ik}^m)}{\partial \chi} \Big|_{A=0} = \vec{N}_i \cdot \frac{\partial \vec{OP}_i^c}{\partial \chi} \Big|_{A=0} = \vec{N}_i \cdot (\vec{\Lambda}_\chi \wedge \vec{OP}_{ik}^m) \quad (2)$$

Dans (2),  $\chi$  est un paramètre de  $A$  relatif à une rotation,  $\vec{\Lambda}_\chi$  est le vecteur directeur autour duquel nous effectuons cette rotation tandis que  $[R_\chi]$  est la matrice de rotation associée.  $[T_{O_c C_\chi}]$  est la matrice de translation du centre de repère modèle  $O$  au point  $C_\chi$ .

#### Paramètre de type translation

La dérivée partielle de  $F(A_k, \vec{N}_i, P_{ik}^m)$  par rapport à  $\chi$  est :

$$\frac{\partial F(A_k, \vec{N}_i, P_{ik}^m)}{\partial \chi} \Big|_{A=0} = \vec{N}_i \cdot \frac{\partial \vec{OP}_i^c}{\partial \chi} \Big|_{A=0} = (\vec{N}_i \cdot \vec{\Lambda}_\chi) \quad (3)$$

Dans (3),  $\chi$  est un paramètre de  $A$  relatif à une translation et  $[T_\chi]$  est la matrice de translation associée à  $\chi$ .  $\vec{\Lambda}_\chi$  est le vecteur directeur le long duquel nous effectuons la translation.



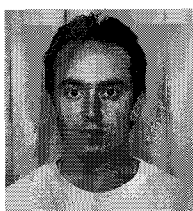
## BIBLIOGRAPHIE

- [Ath84] J. Atha. Current Techniques for Measuring Motion. *Applied Ergonomics*, 15(4) : 245-257, 1984.
- [Aya87] N. Ayache. *Vision stéréoscopique et perception multisensorielle*. Inter Editions Paris, 1987.
- [Bla93] A. Blake, R. Curwen, and A. Zisserman. A framework for spatio-temporal control in the tracking of visual contours. *Int. Journal of Comp. Vision*, 11(2) : 1265-1278, 1993.
- [Bor88] N.A. Borghese, G. Ferrigno, and A. Pedotti. Three-dimensional movement detection : a hierarchical approach. In *Int. Conf. on Systems, man and cybernetics*, pages 303-306, Tucson, 1988.
- [Dho89] M. Dhome, M. Richetin, J.T. Lapresté, and G. Rives. Determination of the attitude of 3d objects from a single perspective image. *I.E.E.E Trans. Pattern. Anal. Machine. Intell.*, 11(12) : 1265-1278, December 1989.
- [Els93] F. Elsner, K. Hassine, and J.C Angue. The use of image sequences to track human motion by spatio-temporal segmentation. In *Workshop on Comp. Vision for Space Applications*, pages 208-219, Antibes, September 1993.
- [Fis81] M.A Fischler and R.C Bolles. Random sample consensus : a paradigm for model fitting with applications to image analysis and automated cartography. *Com of A.C.M.*, 24(6) : 381-395, June 1981.
- [Ge91] A. Geurtz. Three-dimensional human motion estimation : An image processing approach. In *Int. Symposium on 3D Analysis of Human Movement*, pages 19-22, Canada, July 1991.
- [Grim87] L. Grimson. Recognition of object using parameterized models. In *Int. Conf. on Comp. Vision*, volume 3, pages 93-101, London, June 1987.
- [Hel94] Y. Hel-Or and M. Werman. Model based pose estimation of articulated and constrained objects. In *Europ. Conf. on Comp. Vision*, volume 1, pages 262-273, Stockholm, May 1994.
- [Low91] D.G. Lowe. Fitting parameterized three-dimensional models to images. *IEEE trans. Pattern Anal. Machine Intell.*, 13(5) : 441-450, May 1991.
- [Mol85] R. Mollard, A. Coblentz, and E. Fossier. Contribution of Infrared Strophotogrammetry in Movement Analysis. In *Proc. of Application on Biostereometrics*, volume 602, pages 23-30, Cannes, 1985.
- [Mul89] I.J. Mulligan, A.K. Mackworth, and P.D. Lawrence. A model based vision system for manipulator position sensing. In *Proc. of Workshop on Interpretation of 3D scenes, Austin, Texas*, pages 186-193, Austin (Texas), November 1989.
- [Na93] C. Nastar and N. Ayache. Fast segmentation tracking and analysis of deformable models. In *4th Int. Conf. Comp. Vision*, pages 275-279, Berlin, May 1993.
- [Dau93] N.Daucher, M.Dhome, J.T. Lapresté, and G.Rives. Modelled object pose estimation and tracking by monocular vision. In *British Machine Vision Conference*, volume 1, pages 249-258, Guildford, July 1993.
- [Pen91] A.P. Pentland and B. Horowitz. Recovery of non-rigid motion and structure. *IEEE trans. Pattern Anal. Machine Intell.*, 13(7) :730-742, July 1991.
- [Reh94] T.M. Rehg and T. Kanade. Visual tracking of high dof articulated structures : an application to human hand tracking. In *European Conf. Comp. Vision*, pages 35-46, Stockholm, May 1994.
- [Sha91] T. Shakunaga. Pose estimation of jointed structures. In *Conf. on Comp. Vision and Pattern Recognition*, pages 566-572, 1986.
- [Ter91] D. Terzopoulos and D. Metaxas. Dynamic 3d models with local and global deformations : deformable superquadrics. *IEEE trans. Pattern Anal. Machine Intell.*, 13(7) :703-714, July 1991.
- [Wan92] J. Wang. *Analyse et suivi de mouvements 3D articulés : application à l'étude du mouvement humain*. PhD thesis, IFSIC, Université Rennes I, Octobre, 1992.
- [Wi79] A.D. Winter. A new definition of mechanical work done in human movement. *J. Appli. Physiol.*, 46 :79-83, 1979.
- [Yas95] A. Yassine. *De la localisation et du suivi par vision monoculaire d'objets polyédriques articulés modélisés*. PhD thesis, Université Blaise Pascal de Clermont-Ferrand, Novembre, 1995.
- [Yea89] M.R. Yeadon. A method for obtaining three-dimensionnal data on ski jumping using pan and tilt camera. *Int. Journal Sports Biomech.*, 5 :238-247, 1989.

Manuscrit reçu le 10 Juillet 1996.

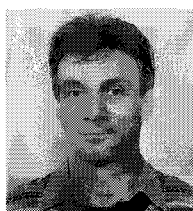
## LES AUTEURS

## Frédéric LERASLE



Ingénieur du CUST de Clermont-Ferrand (section Génie Electrique) et titulaire d'un DEA Electronique et Systèmes de l'Université Blaise Pascal option Vision pour la Robotique. Il poursuit actuellement une thèse de doctorat au LASMEA. Ses travaux concernent le suivi du geste sportif par vision.

## Michel DHOME



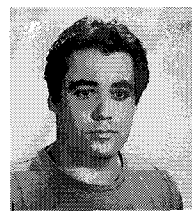
Directeur de Recherche au Centre National de la Recherche Scientifique et exerce au sein du LASMEA à Clermont-Ferrand. Il a obtenu une thèse de l'université Blaise Pascal en 1984 avant de séjourner un an, comme boursier post-doctoral, au Centre National de Recherche du Canada d'Ottawa. Ses activités de recherche concernent le domaine de la Vision Artificielle et plus particulièrement la localisation, la modélisation et la reconnaissance d'objets volumiques.

## Gérard RIVES



Maître de conférence au département Mesures Physiques de l'IUT de Clermont-Ferrand de l'université d'Auvergne. Chercheur au LASMEA de l'université Blaise Pascal URA 1793 du CNRS. Ses recherches portent sur la Vision Dynamique Monoculaire.

## Ali YASSINE



Ingénieur du CUST de Clermont-Ferrand (section Génie Electrique) et titulaire d'un DEA Electronique et Systèmes de l'Université Blaise Pascal option Vision pour la Robotique. Il a effectué une thèse en 1995 sur le thème : Localisation d'Objets Articulés par Vision Monoculaire.

## Suivi du corps humain par vision monoculaire

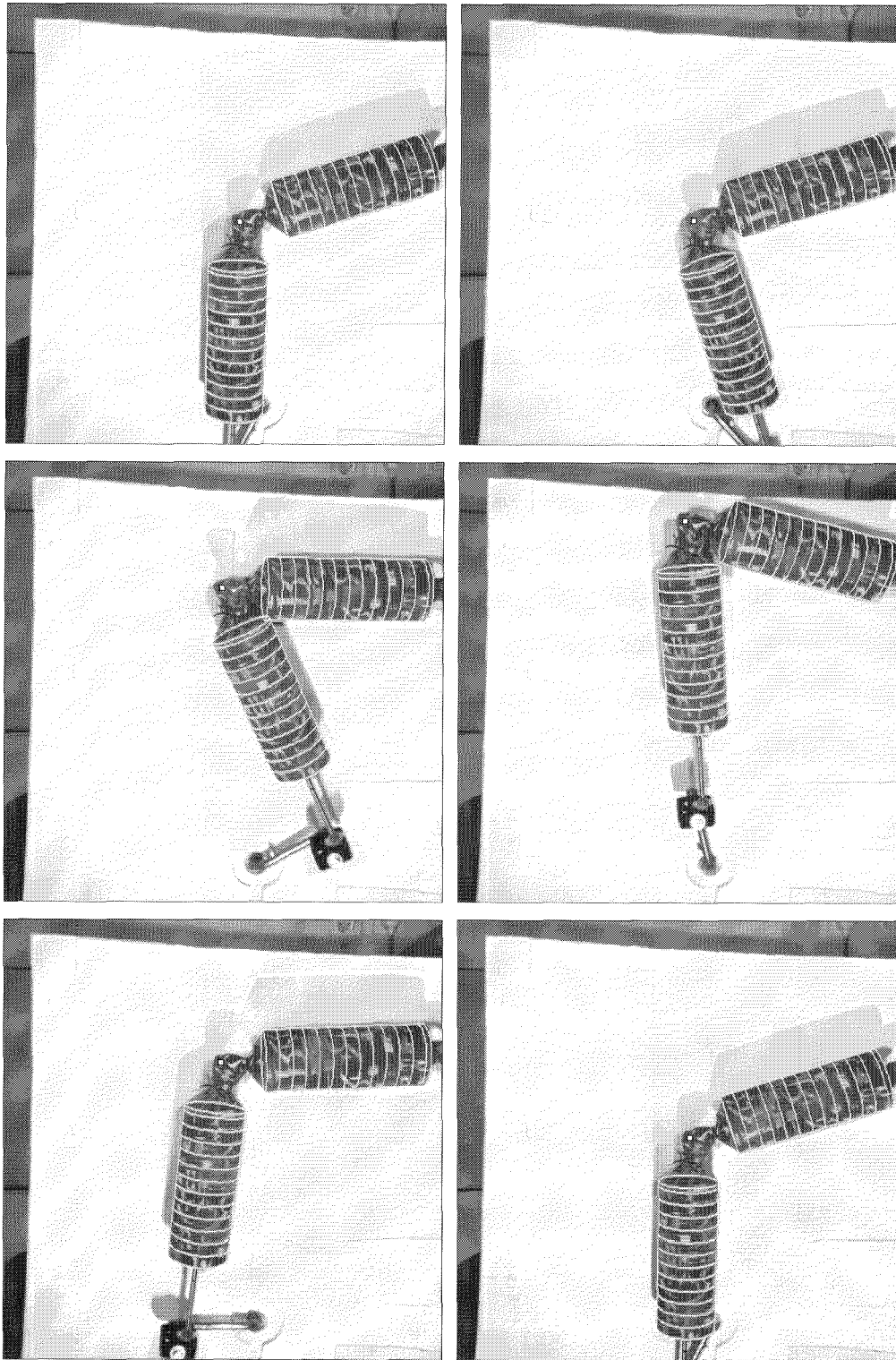


Figure 6. – Séquence de pédalage (images 1, 10, 20, 30, 40, 50) et projection du modèle superposé aux images de luminance.

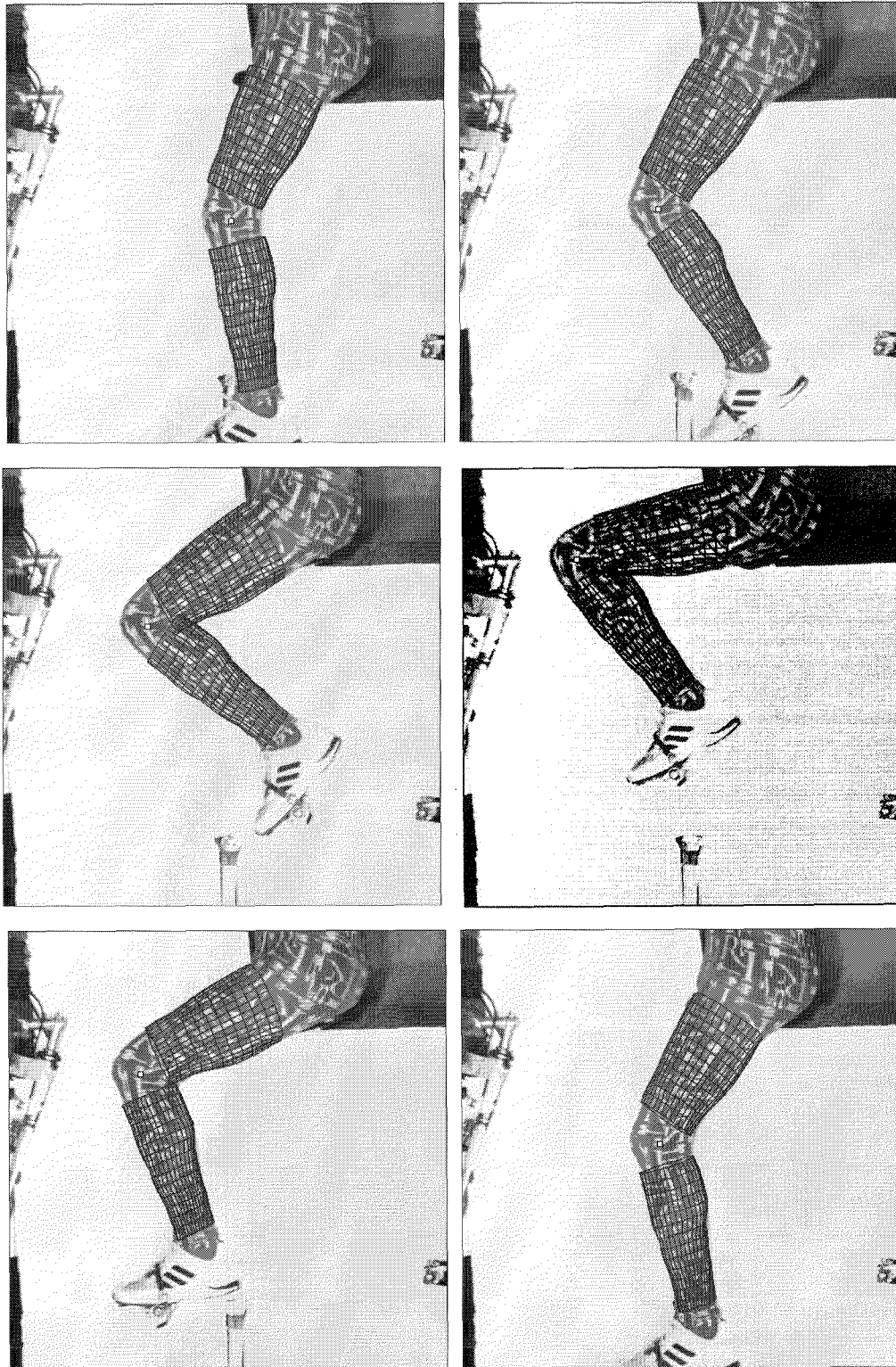


Figure 7. – Séquence de pédalage (images 1, 10, 20, 30, 40, 50) et projection du modèle superposé aux images de luminance.