

Suivi efficace d'objets 3D basé sur l'apparence

Efficient tracking of 3D objects from appearance

par Florent DUCULTY, Michel DHOME et Frédéric JURIE

LASMEA - UMR 6602 du CNRS, Université Blaise Pascal, F-63177 Aubière - FRANCE
email : {duculty,dhome,jurie}@lasmea.univ-bpclermont.fr

résumé et mots clés

Dans cet article, nous proposons un algorithme efficace de suivi d'un objet 3D dans une séquence d'images. Pour cela, l'objet 3D est représenté par une collection d'images de référence. L'originalité de cette méthode est de ne pas utiliser des primitives de haut niveau (points d'intérêt) pour suivre le déplacement de l'objet dans l'image mais plutôt la différence de vecteurs de niveaux de gris entre le motif de référence suivi et le motif courant échantillonné dans une zone d'intérêt de l'image. Le problème du suivi se ramène alors à l'estimation des paramètres qui caractérisent les mouvements possibles de l'objet dans l'image par la détermination de matrices dites d'interaction apprises lors d'une phase d'apprentissage hors ligne, et cela pour chacune des vues de référence. La première matrice lie les variations d'intensité lumineuse du motif de référence 2D de l'objet suivi à son déplacement fronto parallèle (déplacement parallèle au plan image). Sous l'hypothèse d'un tel mouvement, l'aspect apparent de l'objet suivi n'est pas modifié. Toutefois, sa position, son orientation planaire et sa taille peuvent changer. La deuxième matrice relie les variations d'apparence du motif suivi suite à un changement d'orientation par rapport au capteur (modification des angles de site et d'azimut). Nous montrons que l'utilisation en ligne de ces matrices pour la correction de la position prédite de l'objet dans l'image et de l'estimation des variations d'aspect du motif suivi correspond à un coût algorithmique très faible (multiplication d'une matrice par un vecteur) permettant une mise en œuvre temps réel. De plus, nous évoquons le problème des occultations lors du suivi par une méthode de seuillage adaptatif.

vision par ordinateur, suivi 3D, apparence, temps réel, occultation.

abstract and key words

In this article, we propose an efficient tracking algorithm to follow 3D objects in image sequences. 3D objects are represented by a collection of reference images. The originality of this method is not to use high-level primitives (points of interest) to follow the movement of the object in the image but rather the difference between the vectors of gray-levels of the tracked reference pattern and the current pattern sampled in an area of interest. The tracking problem is reduced then to the estimate of the parameters representing the possible movements of the object in the image by the determination of interaction matrices learned during an off-line training stage, and that for each reference view. The first one relates the variations of intensity of the 2D current pattern to be tracked to its fronto parallel movement (parallel movement to the image plane). The aspect of the pattern representing the tracked object is not modified by this movement. However, the position, the orientation and the size of the pattern can change. The second matrix relates the variations of appearance of the currently tracked pattern to a change of attitude between the object and the camera (modification of the angular values in rolling and pitching). We show that the on-line use of these interaction matrices for the correction of the predicted position of the object in the image and the estimate of the variations of aspect of the tracked pattern allows a real time implementation of this algorithm (a matrix multiplied by a vector). Moreover, we also show how the problem of occlusions can be managed.

computer vision, 3D tracking, appearance, real time, occlusion.

1. introduction

Nous nous intéressons, dans cette étude, au problème de suivi d'objets mobiles 3D dans des séquences d'images, en se basant sur l'apparence. Pour simplifier le problème, la localisation de l'objet dans la première image est réalisée par l'opérateur lors de l'initialisation du processus (notion d'interactivité). À l'origine de ces travaux, se trouve un algorithme, développé au laboratoire, qui permet de suivre le déplacement d'un motif visuel donné dans un flot vidéo. Ce dernier a fait l'objet de publications [10] [11] dans lesquelles la méthode de suivi est justifiée théoriquement et validée expérimentalement.

Le domaine applicatif visé est celui de la robotique manufacturière où un bras manipulateur, muni d'une caméra sur son effecteur doit naviguer autour d'un objet [15] et optimiser la trajectoire d'approche permettant de réaliser sa saisie. Ces objets peuvent être de forme et de texture plutôt complexes (figure 1). Le suivi d'objet 3D est une tâche importante pour de nombreuses applications en vision artificielle. Deux approches principales sont généralement distinguées. Les *approches basées sur la mise en correspondance de primitives visuelles* utilisent des caractéristiques locales comme des points, des segments de droite, des arêtes ou des régions. Avec ces techniques, il est possible de localiser l'objet [12] dans l'image courante et de prédire les positions des caractéristiques dans les images suivantes, selon un modèle de mouvement [9][24] et un modèle d'incertitude [14]. Nous citons également les travaux de Strom *et al.* [20] et Basu *et al.* [1]. Ils décrivent un système temps réel de suivi et une modélisation 3D. L'idée directrice est de sélectionner un ensemble dense de points caractéristiques. Ils sont ensuite mis en correspondance d'images en images pour mettre à jour la pose du modèle 3D. Pour cela, un modèle générique 3D polygonale de l'objet est nécessaire. D'une manière générale, les techniques de recherche de pose sont naturellement moins sensibles aux occultations. En effet, elles sont basées sur des correspondances locales. Si plusieurs correspondances sont manquantes, la pose peut encore être calculée.

D'autre part, les *approches globales ou basées sur un motif* utilisent le modèle dans sa totalité. Le point fort de ces méthodes

est leur capacité à traiter des motifs complexes qui ne peuvent être modélisés par des caractéristiques locales. Elles sont très robustes et ont été énormément utilisées. Elles sont aussi appelées *sum-of-square-difference* puisque elles consistent à minimiser la somme des carrés des différences entre un modèle de référence et une région de l'image. Une norme L_2 est généralement utilisée pour mesurer cette erreur. Historiquement, une recherche exhaustive était utilisée. Mais cette stratégie n'est pas applicable dans le cas de transformations plus complexes que des translations 2D, qui nécessitent des espaces de paramètres de dimensions supérieures. Des méthodes plus récentes posent le problème comme un problème de minimisation non linéaire, utilisant des algorithmes du type Newton ou Levenberg-Marquard.

Darell *et al.* [4], Brunelli *et al.* [3] proposent de maximiser un critère de corrélation entre un vecteur caractérisant le modèle de référence et le contenu de l'image. Les temps de calcul, significatifs dans ce cas, peuvent être réduits en travaillant dans des sous-espaces de la représentation initiale de l'image [23][16] [17]. La limitation principale de ces approches est leur manque de résistance au regard des occultations. Black and Jepson [2] ont surmonté cette limitation en reconstruisant les parties occultées. Ils remplacent la norme quadratique généralement utilisée pour construire l'approximation de l'image dans *l'espace propre* par une norme d'erreur robuste. Cette reconstruction revient à une minimisation d'une fonction non linéaire, optimisée en utilisant une méthode de descente de gradient simple. Ils utilisent la même stratégie pour trouver la transformation paramétrique alignant le motif sur l'image. Des travaux similaires reposant sur l'utilisation d'*espaces propres* ont été réalisés par Shree K. Nayar *et al.* [19], K. Deguchi *et al.* [22][5] pour le suivi d'objets et du positionnement d'un robot par vision.

Plus récemment, de nouvelles méthodes efficaces de suivi ont été proposées : le problème du suivi est formulé comme un problème de recherche du meilleur ensemble de paramètres (au sens des moindres carrés) décrivant le mouvement et la déformation de la cible au cours de la séquence. Dans ce cas, les variations des paramètres sont écrites comme une fonction linéaire d'une image de différence (la différence entre l'image de référence et l'image courante). Cette approche est très



Figure 1. – Exemple d'objets que l'on souhaite suivre.

efficace car le mouvement peut être facilement déduit de l'image de différence. Cootes, Edwards et Taylor [21] l'utilisent pour estimer dynamiquement les paramètres d'un modèle de visage en se basant sur l'apparence (modèle 2D). Hager et Belhumeur [8] l'utilisent dans un contexte général pour le suivi d'objet, pour des mouvements planaires affines. Seuls, quelques travaux utilisent cette approche avec des transformations projectives [7] [13], car ces dernières sont non linéaires et la taille de l'espace des paramètres est trop importante.

Dans cet article, nous proposons un algorithme efficace de suivi et de gestion des changements d'aspect d'un objet 3D dans une séquence d'images. Pour cela, l'objet 3D est représenté par une collection d'images 2D appelées vues de référence. Cette technique de suivi 3D comprend deux étapes.

Une phase d'apprentissage hors ligne est dédiée au calcul de deux matrices d'interaction. La première lie les variations d'intensité lumineuse du motif de référence 2D de l'objet suivi dans une zone d'intérêt à son déplacement fronto parallèle. Par définition, un mouvement fronto parallèle est un mouvement tel que l'objet se déplace dans des plans parallèles au plan image. Sous l'hypothèse d'un tel mouvement, l'aspect apparent de l'objet suivi n'est pas modifié; toutefois, sa position, son orientation planaire et sa taille peuvent changer. Pour cela, nous supposons également que la profondeur de l'objet par rapport à la distance objet-caméra reste faible (utilisation d'une caméra à focale longue). Ainsi, une translation 3D de l'objet dans la scène (mouvement fronto parallèle) correspondra à une translation 2D dans l'image. La deuxième matrice d'interaction, quant à elle, relie les variations d'apparence du motif suite à un changement d'orientation de l'objet par rapport au capteur (modification des angles de site et d'azimut).

Une étape en ligne consiste à prédire la position de l'objet dans l'image (en position, échelle et orientation), à multiplier la différence entre le motif observé à l'endroit prédit avec le motif de référence qui doit être suivi par la première matrice d'interaction pour corriger les erreurs sur les mouvements fronto parallèles de l'objet dans l'image (figure 2). Le problème du suivi du motif dans l'image se ramène alors à la correction des paramètres d'une transformation géométrique planaire par la détermination d'un vecteur d'offset.

Une nouvelle différence entre le motif courant corrigé et le motif de référence multiplié par la deuxième matrice d'interaction nous donne les variations d'aspect du motif suivi par rapport au motif de référence le plus proche dans la collection d'images dues aux orientations 3D relatives (site et azimut) de l'objet (figure 3). Nous pouvons ainsi changer de motif de référence pour continuer à suivre l'objet dans l'image en temps réel vidéo (moins de 20 millisecondes par itération).

Compte tenu de la rapidité des traitements (multiplication d'une matrice par un vecteur) par rapport à la vitesse de déplacement des objets dans les séquences d'images, nous n'avons pas besoin d'utiliser d'algorithme de prédiction de mouvement. En effet, l'écart de position du motif entre deux images successives reste

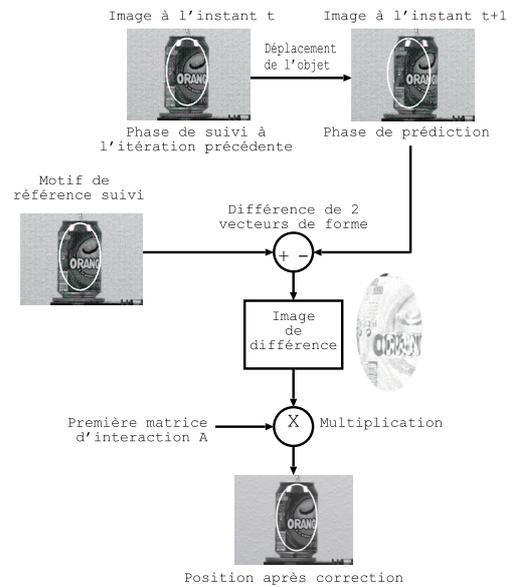


Figure 2. – Correction des mouvements fronto parallèles du motif suivi dans l'image.

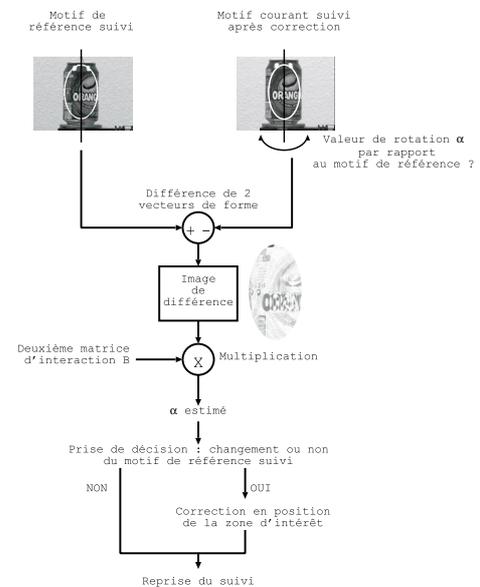


Figure 3. – Calcul de l'angle de site α en fonction du changement d'aspect du motif courant suivi.

compatible avec les variations apprises lors de la phase d'apprentissage. Quant au choix de décomposer le déplacement 3D de l'objet en plusieurs mouvements, il sera justifié par la suite, lors du développement théorique de la phase d'apprentissage hors ligne des matrices d'interaction.

Cet article se décompose en quatre parties. Dans un premier temps, nous voyons comment modéliser l'objet 3D et son apparence, puis nous développons le suivi 3D en plusieurs points. Tout d'abord, nous définissons les paramètres qui caractérisent les mouvements possibles de l'objet dans l'image et leurs

interprétations géométriques dans le suivi, puis nous introduisons la notion de matrices d'interaction calculées lors d'une phase d'apprentissage hors ligne ainsi que l'évaluation des performances de la méthode et terminons par la gestion du passage d'un motif de référence à l'autre. Dans une troisième partie, nous présentons des exemples concrets de suivi 3D d'objets volumiques sans occultation. Les occultations sont ensuite prises en compte par une méthode de seuillage adaptatif, qui fait l'objet de la dernière section.

2. modélisation de l'apparence de l'objet 3D

L'apparence d'un objet rigide dans une image dépend de sa forme, de sa réflectance [18], de sa pose et des conditions d'illumination de la scène qui varient au cours du suivi.

2.1. construction du modèle d'un objet 3D

Dans notre approche de suivi, un objet 3D est représenté par une collection d'images 2D appelées vues de référence. Chacune de ces images représente un des motifs de référence de l'objet 3D à un instant du suivi pour une position caméra/objet donnée. Ces images permettent d'effectuer le suivi 2D du motif. L'acquisition de vues intermédiaires nous permettra alors d'estimer les variations d'aspect du motif courant dans l'image par rapport au motif de référence suivi.

En pratique, nous utilisons une table à déplacement micrométrique pour photographier nos objets 3D et créer nos collections d'images 2D (figure 4). Cette table, commandée à distance, permet de contrôler précisément la pose de l'objet dans l'image.

Par exemple, pour les résultats présentés dans cet article, les vues de référence de la canette de soda sont acquises tous les 10 degrés en site sur 360 degrés et les vues intermédiaires tous les degrés (36 vues de référence pour un total de 361 images dans la base). Les vues de référence de la figurine sont acquises

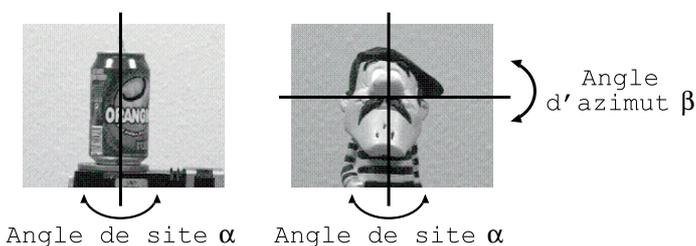


Figure 4. – Acquisition d'une collection d'images 2D pour la modélisation d'objets.

aussi tous les 10 degrés en site et azimut sur une portion de sphère ($\alpha = -40$ à $+40$ degrés et $\beta = -30$ à $+30$ degrés), les vues intermédiaires tous les 2 degrés (63 vues de référence pour un total de 1 911 images).

La disposition des différentes vues de référence dans la collection d'images dépend fortement de la forme et du volume de l'objet entraînant une variation plus ou moins importante de l'aspect du motif courant sans pour autant perdre des points caractéristiques du motif de référence suivi.

2.2. représentation de l'apparence d'un motif

Nous souhaitons représenter le motif à suivre par un vecteur de forme (vecteur de niveaux de gris de dimension N où N est le nombre de points échantillonnés) et que cette représentation soit indépendante de la position, de l'orientation et de l'échelle du motif dans l'image. Pour cela, nous proposons d'échantillonner le motif à l'intérieur d'une zone elliptique (figure 5).

Les points où sont prélevés les niveaux de gris (points blancs sur la figure 5) sont répartis sur un ensemble d'ellipses concentriques échantillonnées de la plus petite à la plus grande. L'ensemble des valeurs échantillonnées (niveaux de gris toujours numérotés dans le même ordre) est alors stocké dans le vecteur de forme. Pour notre application, le vecteur de forme comprend 373 points échantillonnés sur 15 ellipses.

Ainsi quelle que soit la position, l'orientation et la taille du motif, sa représentation vectorielle après échantillonnage sera sensiblement la même, puisque les valeurs de niveaux de gris enregistrées sont positionnées dans un repère lié à l'ellipse et par conséquent au motif. Nous réalisons donc à la fois un échantillonnage local d'une région d'image et global du motif. Le nombre de points échantillonnés dépend plus particulièrement de la méthode de sélection envisagée et de l'apparence de l'objet que l'on désire suivre.

La position et la forme de l'ellipse sont définies par un vecteur à cinq paramètres correspondant à la position du centre (X_c, Y_c), l'orientation (θ) et les longueurs du grand et du petit axe (R_1, R_2). Par la suite, nous poserons $R_2 = k * R_1$ où k est

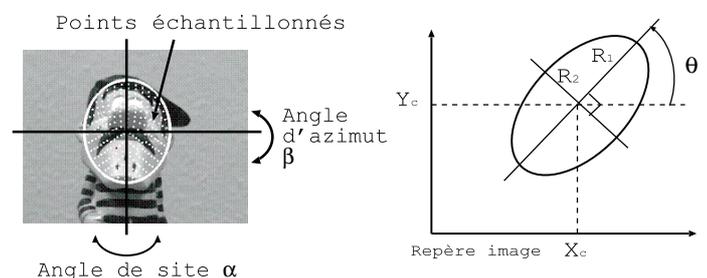


Figure 5. – Échantillonnage du motif à l'intérieur d'une ellipse.

un ratio connu et fixé lors de la phase d'apprentissage pour avoir un seul facteur d'échelle. La représentation géométrique de ce vecteur est rappelée figure 5. Le choix de prendre une ellipse comme zone d'échantillonnage du motif suivi sera justifié dans la section suivante.

De plus, pour garantir une certaine insensibilité aux changements de conditions d'éclairage de la scène, le vecteur de forme, une fois échantillonné, est alors centré et normé. Ceci permet de compenser des variations affines de la luminance entre l'image de référence et l'image courante. La figure 6 illustre la robustesse de notre algorithme de suivi pour différentes valeurs d'ouverture et de fermeture du diaphragme de la caméra.

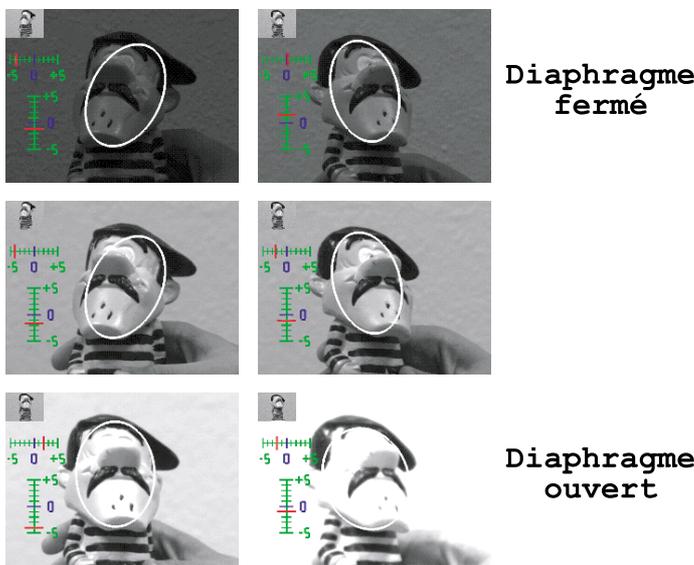


Figure 6. – Robustesse de l'algorithme de suivi aux variations de luminance.

3. suivi 3D d'un objet

Nous développons dans cette partie l'aspect théorique du suivi 3D d'objet volumique. Tout d'abord, nous devons être capable de paramétrer les mouvements de l'objet pour le suivre dans l'image.

3.1. paramétrisation des mouvements possibles d'un objet dans l'image

Les paramètres estimés lors du suivi doivent traiter tous les mouvements possibles de l'objet devant la caméra à savoir trois translations (T_x, T_y, T_z) et trois rotations (R_x, R_y, R_z) soit six degrés de liberté.

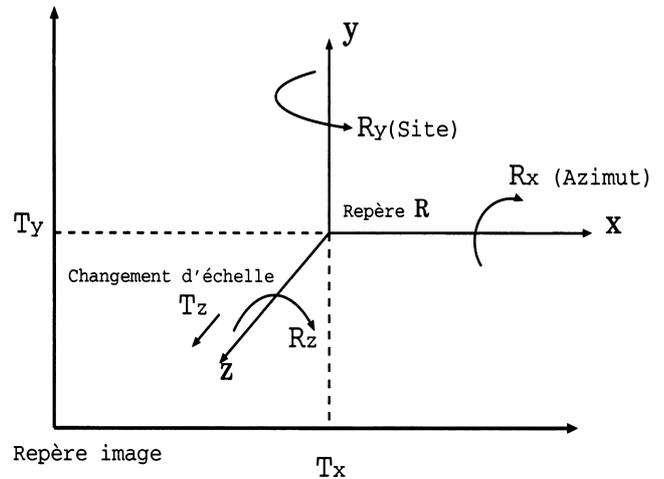


Figure 7. – Mouvements possibles d'un objet dans l'image.

Nous pouvons les classer en deux catégories :

1) ceux qui provoquent un mouvement fronto parallèle de l'objet par rapport au plan de l'image (un tel mouvement ne modifie pas l'aspect du motif suivi). Ils sont au nombre de quatre :

- T_x, T_y : translations axiales en x et y équivalentes aux coordonnées du centre de l'ellipse (X_c, Y_c),
- T_z : translation axiale en z ou changement d'échelle défini par les rayons de l'ellipse ($R_1, R_2 = k * R_1$),
- R_z : rotation autour de l'axe z équivalente à l'orientation de l'ellipse (θ).

2) ceux qui provoquent des changements de l'apparence du motif suivi. La prise en compte de la modification d'aspect du motif se fait à l'aide de deux paramètres :

- R_x : rotation autour de l'axe x équivalente à une variation de l'angle d'azimut β du motif courant dans l'image par rapport au motif de référence suivi,
- R_y : rotation autour de l'axe y équivalente à une variation de l'angle de site α du motif courant dans l'image par rapport au motif de référence suivi.

Ceci est illustré par la figure 7 où le repère \mathcal{R} correspond au repère caméra translaté vers le barycentre de l'objet (l'axe des x étant parallèle aux lignes de l'image, l'axe des y parallèle aux colonnes de l'image, l'axe des z parallèle à l'axe optique).

La combinaison de ces six paramètres ($X_c, Y_c, R_1, \theta, \alpha$ et β) nous permet de suivre un objet volumique dans une image tout en gérant ses variations d'aspect ($R_2 = k * R_1$ par définition).

3.2. interprétation géométrique du suivi 3D

Nous venons de voir que le motif que l'on désire suivre est inscrit dans une ellipse dont la forme et la position dans l'image sont données par le vecteur de paramètres μ de dimension p (ici $p = 4$) avec $\mu = (X_c, Y_c, R_1, \theta)^t$ et $R_2 = k * R_1$.

Notons μ_p le vecteur de paramètres prédit, μ_r le vecteur de paramètres de la position réelle du motif et posons $\Delta\mu$ la différence de ces deux vecteurs :

$$\Delta\mu = \mu_r - \mu_p \quad (1)$$

De plus, le motif visuel à l'intérieur de l'ellipse prédite est échantillonné pour donner le vecteur de forme courant VI_c . Le vecteur de forme du motif de référence à suivre est quant à lui noté VI_{ref} . Posons maintenant ΔVI_p la différence de ces deux vecteurs de niveaux de gris :

$$\Delta VI_p = VI_{ref} - VI_c \quad (2)$$

Il est alors intéressant de savoir si l'on peut déterminer $\Delta\mu$ en fonction de ΔVI_p . Si c'est le cas, cela signifie qu'en mesurant la différence ΔVI_p entre le motif de référence suivi et le motif courant prédit, on peut déterminer la correction $\Delta\mu$ à apporter à la prédiction pour obtenir la position réelle du motif :

$$\mu_r = \mu_p + \Delta\mu \quad (3)$$

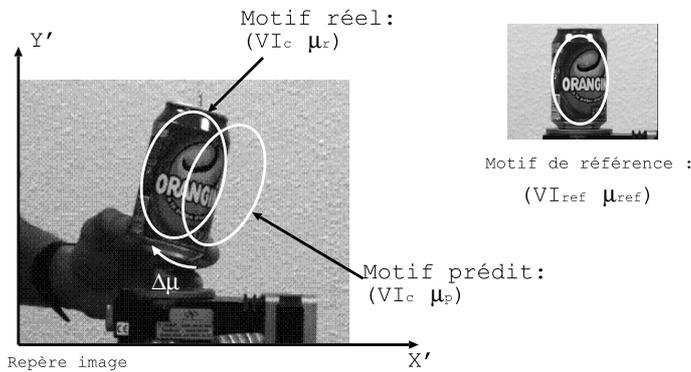


Figure 8. – Principe du suivi d'un motif de référence.

Nous repositionnons donc le problème du suivi comme la détermination d'un vecteur d'offset $\Delta\mu$, en supposant que les variations de position de l'objet dans l'image correspondent aux variations des paramètres d'une transformation géométrique f (figure 8). Dans notre cas particulier, nous utilisons une transformation affine rigide où les paramètres de l'ellipse sont les paramètres de la transformation géométrique (figure 9).

Un point de coordonnées (x, y) dans le référentiel région ou ellipse a pour coordonnées (x', y') dans le référentiel image par la transformation géométrique $f(\mu)$ tel que :

$$\begin{cases} x' = R_1 x \cos\theta - kR_1 y \sin\theta + X_c \\ y' = R_1 x \sin\theta + kR_1 y \cos\theta + Y_c \end{cases} \quad (4)$$

L'un des avantages de ce suivi est que l'on peut appliquer plusieurs types de transformation (affine non rigide, homographie).

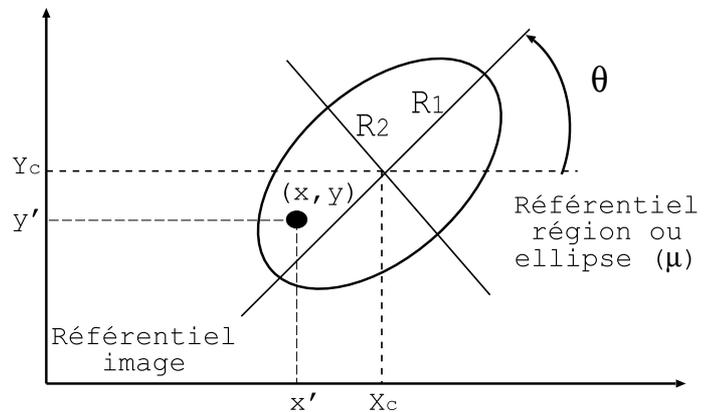


Figure 9. – Interprétation géométrique des paramètres de l'ellipse dans le suivi 3D.

Il faut simplement se définir une stratégie d'échantillonnage du motif tout en tenant compte de l'aspect et du volume de l'objet. Cette différence $\Delta\mu$ peut s'écrire sous la forme matricielle suivante :

$$\Delta\mu = A\Delta VI_p \quad (5)$$

où A est une matrice dite d'interaction ($p * N$) correspondant au calcul d'une relation linéaire entre un ensemble de différences de niveaux de gris ΔVI_p et une correction $\Delta\mu$ des paramètres du vecteur μ lors d'une phase d'apprentissage hors ligne.

Nous avons vu également que les variations de l'aspect du motif courant par rapport au motif de référence suivi pouvaient être caractérisées par deux paramètres angulaires en site et azimut (α et β). Notons le vecteur $\psi = (\alpha, \beta)^t$ de dimension q (ici $q = 2$) et ΔVI_r la différence entre le vecteur de référence du motif suivi VI_{ref} et le vecteur courant VI_c dans l'ellipse prédite après correction (motif recalé). Comme précédemment, il est alors intéressant de savoir si l'on peut déterminer ψ en connaissant ΔVI_r . Si c'est le cas, cela signifie qu'en mesurant la différence ΔVI_r , on est capable :

- de positionner en site et azimut le motif courant dans l'image par rapport au motif de référence suivi et de ses plus proches voisins dans la collection de vues de référence,
- de pouvoir décider quand changer de motif de référence.

Ceci est illustré figure 10, lors du suivi de figurine, pour une valeur α en site de 2 degrés par rapport au motif de référence 0 et une valeur en azimut β égale à 10 degrés. Dans cet exemple simple, le changement en site du motif de référence est décidé lorsque l'angle $\alpha > |\pm 6|$ pour éviter ainsi, un basculement permanent entre 2 motifs pour $\alpha = \pm 5$.

- 1) Si ($\alpha \leq 6$ et $\alpha \geq -6$) alors motif suivi = motif actuel 0,
- 2) Si ($\alpha > 6$) alors motif suivi = motif suivant 10,
- 3) Si ($\alpha < -6$) alors motif suivi = motif précédent - 10.

Ce calcul de ψ peut s'écrire sous la forme matricielle suivante :

$$\psi = B\Delta VI_r \quad (6)$$

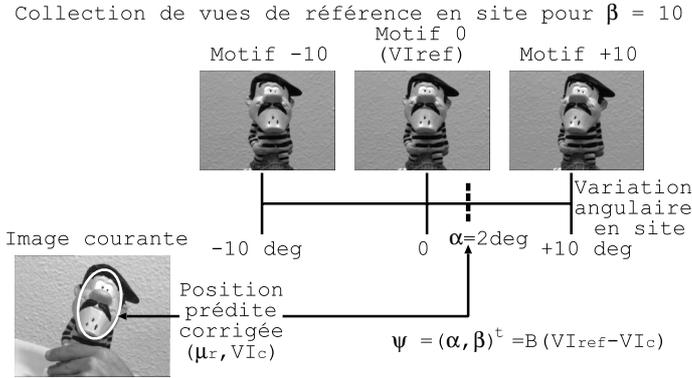


Figure 10. – Positionnement du motif courant dans la collection d'images 2D suivant les variations d'aspect.

où B est une matrice dite d'interaction ($q * N$) correspondant au calcul d'une relation linéaire entre un ensemble de différences de niveaux de gris ΔVI_r et une estimation des paramètres angulaires du vecteur ψ par rapport au motif de référence suivi lors d'une phase d'apprentissage hors ligne.

Comme les objets 3D sont représentés par une collection d'images 2D, les matrices d'interaction A et B sont calculées pour chacune des vues de référence.

3.3. estimation des matrices d'interaction pour un motif de référence donné

Le calcul des deux matrices d'interaction A et B se fait lors d'une phase d'apprentissage hors ligne. Une des originalités de la méthode de calcul proposée est que nous n'utilisons pas de matrices jacobiniennes de la vue de référence comme dans les travaux de Gregory D. Hager et Peter N. Belhumeur [8] ou Frank Dellaert et Robert Collins [6]. Nous estimons les matrices A et B par une minimisation au sens des moindres carrés en utilisant un algorithme basé sur une décomposition en valeurs singulières. Nous avons observé que dans ce cas, le domaine de convergence était beaucoup plus important [11]. Par similitude avec la commande référencée vision (CRV), nous parlons ici de « matrice d'interaction ». Son estimation est une approximation linéaire correspondant à un développement limité du premier ordre d'une fonction [10]. Durant cette phase d'apprentissage, nous cherchons à minimiser un critère dans l'image (une différence de niveaux de gris). Nous n'avons donc jamais la certitude d'avoir la bonne estimation des paramètres 3D. De plus, notre méthode par hyperplans [11] ne permet pas d'estimer l'ensemble des paramètres 3D par une seule matrice d'interaction. C'est pourquoi, nous avons décomposé le déplacement 3D d'un objet devant la caméra en plusieurs mouvements classés en deux catégories (avec ou sans modification de l'aspect du motif suivi).

3.3.1. calcul de la matrice d'interaction A

Cette matrice permet la mise à jour des paramètres de l'ellipse ou de la transformation affine rigide lors du suivi de l'objet dans une séquence d'images.

Une ellipse est placée manuellement par l'utilisateur sur le motif de référence puis échantillonnée pour donner le vecteur de forme de référence VI_{ref} de dimension N ($N = 373$ points échantillonnés). Cette initialisation nous permet également de fixer le rapport k entre les deux rayons de l'ellipse ($k = R_2/R_1$). La position de l'ellipse est perturbée M fois aléatoirement autour de sa position de référence ($M = 1000$) tout en gardant le coefficient k constant (figure 11). Les amplitudes des variations des paramètres de l'ellipse sont de 15 % de la longueur des axes pour la position du centre (rappelons que les variations de ce dernier sont réalisées le long des axes principaux de l'ellipse), de 15 % sur la longueur des axes et de 15 degrés sur la rotation.

Pour chaque perturbation j , les variations des paramètres de la transformation $\Delta \mu^j = (\Delta X_c^j, \Delta Y_c^j, \Delta R_1^j, \Delta \theta^j)^t$ ainsi que le vecteur différence $\Delta VI^j = (\Delta i_1^j, \Delta i_2^j, \dots, \Delta i_N^j)^t$ entre le motif de référence VI_{ref} et le motif courant VI_c^j sont mémorisés. Il est alors possible d'estimer A si $M \geq N$. Cela revient donc à résoudre un système surdimensionné de M équations à N inconnues pour chacun des paramètres de la transformation soit quatre systèmes. En réalité, la résolution d'un seul système linéaire, ou plus exactement, le calcul d'une seule matrice pseudo-inverse est nécessaire. En notant la matrice d'interaction A sous la forme $A = (AX_c, AY_c, AR_1, A\theta)^t$, nous obtenons la ligne $A\theta$ de la matrice d'interaction relative à l'orientation de l'ellipse à l'aide du système linéaire suivant :

$$\begin{pmatrix} \Delta i_1^1 & \Delta i_2^1 & \dots & \Delta i_N^1 \\ \Delta i_1^2 & \Delta i_2^2 & \dots & \Delta i_N^2 \\ \vdots & \vdots & \dots & \vdots \\ \Delta i_1^M & \Delta i_2^M & \dots & \Delta i_N^M \end{pmatrix} \begin{pmatrix} A\theta_1 \\ A\theta_2 \\ \vdots \\ A\theta_N \end{pmatrix} = \begin{pmatrix} \Delta \theta^1 \\ \Delta \theta^2 \\ \vdots \\ \Delta \theta^M \end{pmatrix} \quad (7)$$

Pouvant se mettre sous la forme matricielle suivante :

$$M_{\Delta VI} * A\theta = \Delta \theta \quad (8)$$

La solution est alors obtenue par :

$$A\theta = (M_{\Delta VI}^t M_{\Delta VI})^{-1} M_{\Delta VI}^t \Delta \theta = M_{\Delta VI}^+ \Delta \theta \quad (9)$$

La matrice $M_{\Delta VI}^+$ est la matrice pseudo-inverse de la matrice $M_{\Delta VI}$.

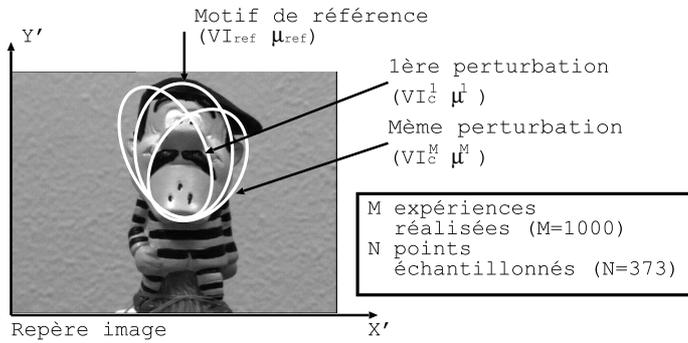


Figure 11. – Perturbations des paramètres de l'ellipse pour l'estimation de la matrice d'interaction A .

Le calcul des trois autres lignes de la matrice d'interaction A utilise le produit de la même matrice avec des vecteurs de perturbations différents $(\Delta X_c, \Delta Y_c, \Delta R_1)$:

$$\begin{cases} AX_c = M_{\Delta VI}^+ \Delta X_c \\ AY_c = M_{\Delta VI}^+ \Delta Y_c \\ AR_1 = M_{\Delta VI}^+ \Delta R_1 \end{cases} \quad (10)$$

3.3.2 résultats expérimentaux sur l'estimation de la matrice A

Nous allons ici chercher à caractériser les performances de la matrice d'interaction A sur une image statique. Dans ce cas, l'objectif est d'observer comment et dans quelles limites la matrice d'interaction A permet de revenir sur le motif sélectionné lorsque l'on écarte l'ellipse de la position de référence.

L'image utilisée pour cet exemple est présentée figure 12. L'ellipse tracée sur cette dernière englobe le motif visuel de référence choisi. Ce dernier est échantillonné sur N points ($N = 373$) à l'intérieur de l'ellipse selon la procédure décrite précédemment.

Une série de M déformations aléatoires de l'ellipse de référence ($M = 1\,000$) est réalisée pour estimer la matrice d'interaction A . Les amplitudes des variations des paramètres de l'ellipse sont



Figure 12. – Image utilisée pour l'étude de la convergence sur image statique. L'ellipse indique le motif de référence.

de 15 % de la longueur des axes pour la position du centre, de 10 % sur la longueur des axes et de 10 degrés sur la rotation. Nous présentons dans les sous-sections suivantes des courbes donnant la valeur de la correction estimée (par le calcul de $A\Delta V I_p$) en fonction de la variation $\Delta\mu$ réellement effectuée.

Afin de simplifier la représentation, nous avons différencié deux cas : un cas où les erreurs de translation sont prépondérantes sur les erreurs liées à la longueur des axes et l'orientation de l'ellipse et un cas inverse.

3.3.3. translations

La figure 13 présente différentes ellipses, déduites de la position de référence principalement par une translation du centre (X_c, Y_c) . Afin de ne pas surcharger la figure, toutes les positions intermédiaires ne sont pas représentées. Toutefois, pour chacune d'entre elles, la matrice d'interaction A est utilisée pour calculer la correction à apporter aux paramètres de l'ellipse déformée afin de retomber sur l'ellipse de référence.

Ainsi, les courbes présentées figure 14 indiquent la translation estimée en fonction de la perturbation réelle de l'ellipse.

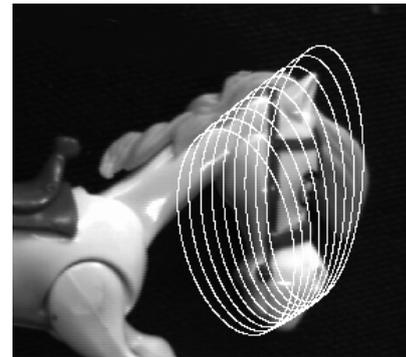


Figure 13. – Tests de translation : différentes ellipses sont calculées en traduisant essentiellement le centre de l'ellipse de référence.

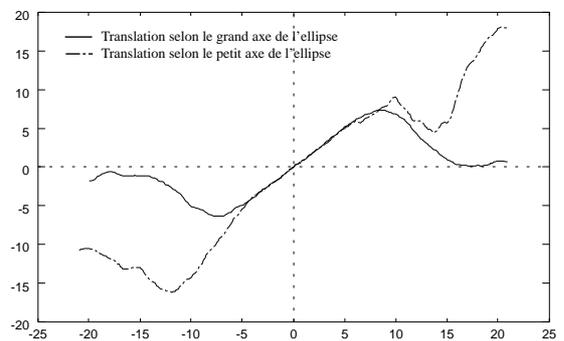


Figure 14. – Translation estimée en fonction de la translation réelle de l'ellipse (en pixels).

3.3.4. rotations et changements d'échelle

La figure 15 présente différentes ellipses, éloignées de la position de référence principalement par une combinaison de rotations et changements d'échelle (les positions intermédiaires ne sont pas toutes représentées, mais elles ont été prises en compte pour les calculs).

Les courbes présentées figure 16 et figure 17 indiquent les rotations et changements d'échelle estimés en fonction des variations réelles.

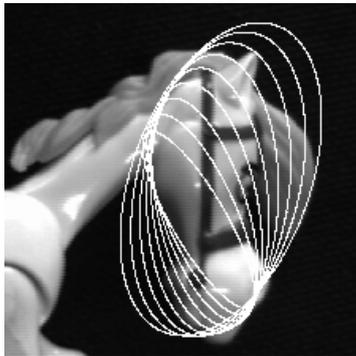


Figure 15. – Tests de rotation/changement d'échelle : différentes ellipses déformées majoritairement en tournant et grossissant l'ellipse de référence.

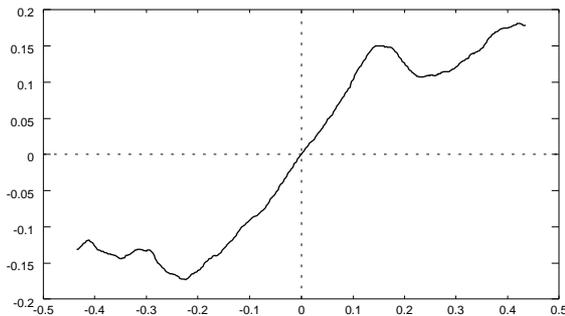


Figure 16. – Rotation estimée en fonction de la perturbation en rotation de l'ellipse (en radians).

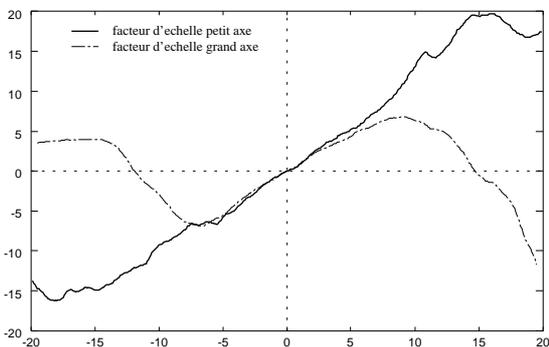


Figure 17. – Longueur des axes estimée en fonction de la perturbation de la longueur des axes (en % de la longueur).

3.3.5. analyse des courbes – domaine de convergence

Autour des petites variations, la variation prédite correspond bien à la variation réalisée. C'est la partie linéaire centrale de la courbe. Dans ce domaine, l'approximation du premier ordre est satisfaisante. Il suffira d'une itération pour passer de la position prédite du motif à sa position réelle dans l'image. Ce domaine correspond à des variations d'environ ± 10 pixels en translation, $\pm 10\%$ de la longueur des axes et de ± 10 degrés en rotation (notons que ceci correspond approximativement à l'ordre de grandeur de l'écart type utilisé pour perturber l'ellipse de référence lors de l'apprentissage de la matrice d'interaction A).

Il nous faut toutefois mentionner que le domaine de convergence est plus vaste que cette partie linéaire de la courbe. En effet, il suffit que le signe soit correct pour que la correction s'opère dans le bon sens. Il faudra dans un tel cas, itérer quelques fois le processus pour se ramener dans la zone linéaire, où une dernière itération placera l'ellipse sur le motif de référence. Le domaine de convergence observé est environ deux fois plus grand que le domaine de linéarité. Bien entendu, à chaque itération, il est nécessaire de calculer le $\Delta V I_r$ correspondant à la nouvelle position de l'ellipse.

3.3.6. calcul de la matrice d'interaction B

Cette matrice permet la mise à jour des paramètres angulaires pour traiter les changements d'aspect de l'objet lors de son suivi dans une séquence d'images.

Comme précédemment, il s'agit de déterminer puis de multiplier successivement une matrice pseudo-inverse $M_{\Delta V I}^+$ par les variations en site $\Delta\alpha$ et en azimut $\Delta\beta$ pour calculer les lignes de la matrice d'interaction $B = (B_\alpha, B_\beta)^t$:

$$B_\alpha = M_{\Delta V I}^+ \Delta\alpha \quad \text{et} \quad B_\beta = M_{\Delta V I}^+ \Delta\beta \quad (11)$$

Pour cela, il faut utiliser les n vues intermédiaires entourant le motif de référence dont le nombre varie selon les règles de décision prises pour changer de motif. Dans le cas de la figurine, illustrée par la figure 18, nous avons décidé de changer de motif de référence courant quand l'une des deux valeurs angulaires α ou β est supérieure en valeur absolue à 6 degrés. Pour avoir un calcul fiable des valeurs angulaires, il est nécessaire d'acquérir des images intermédiaires entre chaque motif de référence et nous étendons la zone de calcul à des variations de ± 8 degrés par pas de 2 degrés autour d'une vue de référence soit un total de n vues de travail ($n = 81$ puisque nous avons neuf images différentes en site par azimut et neuf variations d'azimut).

Pour chaque vue intermédiaire de variations en site $\Delta\alpha_n$ et d'azimut $\Delta\beta_n$, nous réalisons M perturbations des paramètres de l'ellipse définie sur le motif de référence ($M = 20$), puis calculons les corrections à apporter aux paramètres de la transformation géométrique ou de l'ellipse en multipliant la matrice d'interaction A de la vue de référence par la différence entre le

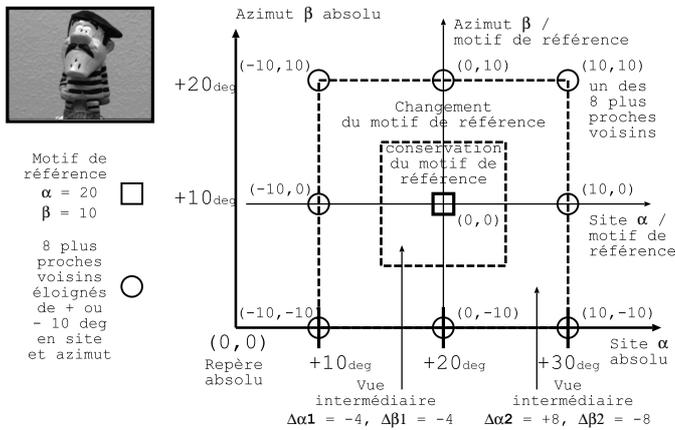


Figure 18. – Sélection des vues nécessaires au calcul de la matrice d'interaction B .

vecteur de référence VI_{ref} et le motif courant perturbé VI_c échantillonné dans l'image intermédiaire. L'ellipse étant corrigée et repositionnée sur le motif de référence à suivre dans l'image intermédiaire, la nouvelle mesure de différence de niveaux de gris entre le vecteur de référence VI_{ref} et le motif courant corrigé VI_c ainsi que les variations en site $\Delta\alpha_n$ et en azimut $\Delta\beta_n$ sont mémorisées pour calculer la matrice d'interaction B (figure 19). Cette estimation de B n'est possible que si $n * M \geq N$.

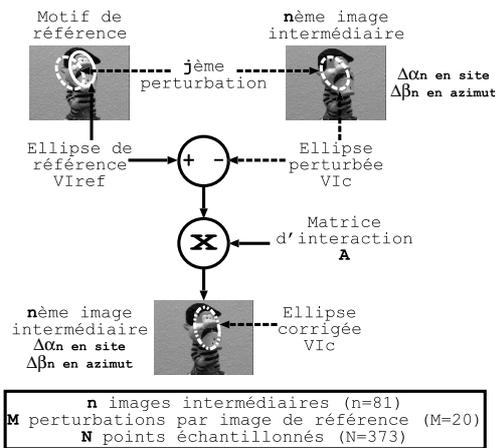


Figure 19. – Perturbations et corrections des paramètres de l'ellipse pour l'estimation de la matrice d'interaction B .

3.4. gestion du passage d'un motif de référence à l'autre

Nous avons vu précédemment que le calcul de la matrice d'interaction B dépend des règles de décision de passage d'un motif de référence à un autre, qui dépendent elles même de la méthode d'acquisition ou de disposition des vues de référence les unes par rapport aux autres. En fonction de l'objet à modéliser, chacun peut se définir des règles de décision propres à son application.

Toutefois, un problème doit être pris en compte lors du changement du motif suivi : c'est la correction des paramètres de la transformation du nouveau motif en fonction de ceux de l'ancien motif. Ceci est à effectuer lorsque les paramètres de la transformation géométrique ou de l'ellipse sont différents entre les motifs de référence modélisant l'objet 3D (figure 20).

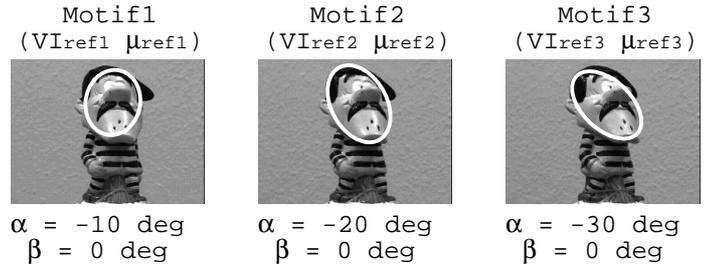


Figure 20. – Exemple de 3 motifs de référence consécutifs pour des paramètres de transformation géométrique différents.

Pendant cette étape intermédiaire du suivi, il s'agit d'exprimer les paramètres de l'ellipse du prochain motif μ'_{su} en fonction des paramètres de l'ellipse du motif courant μ'_{ac} et ceux correspondant dans la base d'apprentissage (μ_{su} et μ_{ac}). Ces calculs sont développés maintenant et illustrés par la figure 21.

Nous rappelons que par définition $R_2 = k * R_1$ où le coefficient k est fixé lors de la phase d'apprentissage.

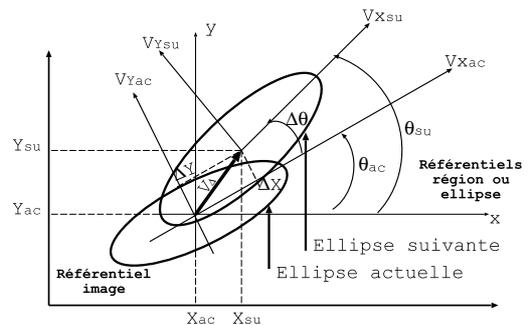
– longueurs des axes de l'ellipse pour le changement d'échelle :

$$R'_{1su} = \frac{R_{1su}}{R_{1ac}} R'_{1ac} \text{ et } R'_{2su} = k * R'_{1su} \quad (12)$$

Phase d'apprentissage

Motif actuel $\mu_{ac} = (X_{ac}, Y_{ac}, R_{1ac}, \theta_{ac})$

Motif suivant $\mu_{su} = (X_{su}, Y_{su}, R_{1su}, \theta_{su})$



Phase de suivi

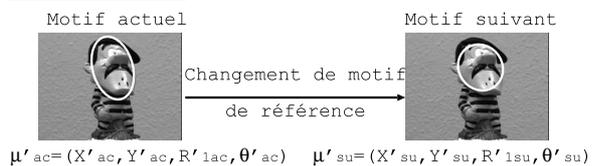


Figure 21. – Correction intermédiaire des paramètres de l'ellipse lors du changement de motif pendant le suivi (objet centré).

– orientation de l'ellipse :

$$\theta'_{su} = \theta'_{ac} + \Delta\theta = \theta'_{ac} + (\theta_{su} - \theta_{ac}) \quad (13)$$

– coordonnées du centre de l'ellipse.

Dans le référentiel (x, y) de la phase d'apprentissage, les vecteurs V_{Xac} , V_{Yac} et V_{Δ} ont respectivement pour coordonnées $(\cos\theta_{ac}, \sin\theta_{ac})$, $(-\sin\theta_{ac}, \cos\theta_{ac})$ et $((X_{su} - X_{ac}), (Y_{su} - Y_{ac}))$. Les écarts relatifs ΔX et ΔY entre les coordonnées des centres des ellipses considérées sont estimés à partir des équations suivantes :

$$\begin{cases} \Delta X = V_{\Delta} \cdot V_{Xac} = (X_{su} - X_{ac})\cos\theta_{ac} + (Y_{su} - Y_{ac})\sin\theta_{ac} \\ \Delta Y = V_{\Delta} \cdot V_{Yac} = -(X_{su} - X_{ac})\sin\theta_{ac} + (Y_{su} - Y_{ac})\cos\theta_{ac} \end{cases} \quad (14)$$

En notant $k_1 = \frac{R'_{1ac}}{R_{1ac}}$ et $k_2 = \frac{R'_{2ac}}{R_{2ac}}$ les facteurs d'échelle pour passer de la phase d'apprentissage à la phase de suivi, les coordonnées de la nouvelle ellipse dans le référentiel image sont obtenues par :

$$X'_{su} = X'_{ac} + k_1\Delta X \quad \text{et} \quad Y'_{su} = Y'_{ac} + k_2\Delta Y \quad (15)$$

En fait, ces calculs ne sont valables que si l'objet 3D reste centré dans l'image pendant la phase d'acquisition de la collection d'images 2D. En effet, lors de la phase d'apprentissage, nous sauvegardons, pour chacune des vues de référence, le vecteur de forme associé ainsi que les paramètres de l'ellipse qui correspondent à des variations angulaires en site et azimuth (α et β) nulles. Or, durant la phase de suivi, la correction des paramètres d'ellipse lors du changement de motif de référence s'opère pour des variations angulaires non nulles. Comme les écarts relatifs ΔX et ΔY entre les coordonnées des centres des ellipses considérées sont calculés à partir des paramètres d'ellipse de référence de la base d'apprentissage, cela engendre une erreur de position du centre de la nouvelle ellipse dans la phase de suivi. Cette erreur, restant compatible avec les variations apprises lors de la phase d'apprentissage de la matrice d'interaction A , est corrigée lors de l'itération suivante.

Lorsque l'objet 3D n'est pas centré dans l'image lors de l'acquisition de la collection d'images 2D, cette erreur de position est trop importante pour pouvoir être corrigée à l'itération suivante. Pour résoudre ce problème, il faut sauvegarder d'autres informations durant la phase d'apprentissage hors ligne. En particulier, pour chaque motif de référence, il faut repositionner et mémoriser les nouveaux paramètres de l'ellipse associée à l'aide de la matrice d'interaction A dans chacune des vues intermédiaires utilisées dans le calcul de la matrice d'interaction B . Ainsi, nous connaissons la forme et la position de l'ellipse de référence pour une variation angulaire non nulle. En pratique, lors de la phase de suivi, nous appliquons trois fois cette correction aux paramètres de l'ellipse du motif actuel μ'_{ac} pour obtenir

les paramètres de l'ellipse du motifs uivant μ'_{su} comme le montre la figure 22 pour une variation en site uniquement :

– à partir des paramètres d'ellipse du motif courant μ'_{ac} pour une valeur de site α_{actuel} , la première correction permet de calculer les paramètres d'ellipse du motif actuel pour un angle de site α_{ref} nul,

– la deuxième correction assure le passage d'un motif de référence à l'autre,

– à partir des paramètres d'ellipse du motif suivant pour une valeur de site α_{ref} nulle, la troisième correction permet le calcul des paramètres d'ellipse du motif suivant μ'_{su} pour un angle de site $\alpha_{suivant}$.

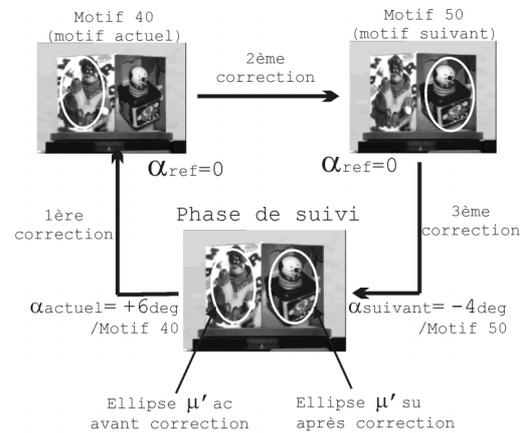


Figure 22. – Corrections intermédiaires des paramètres de l'ellipse lors du changement de motif pendant le suivi (objet non centré).

4. expérimentations

Dans ce paragraphe, nous présentons tout d'abord la phase d'initialisation du suivi ou de reconnaissance de motif, puis la phase de suivi en ligne sous forme d'un algorithme et d'une illustration. Les programmes sont implantés sur une station de travail Silicon Graphics O_2 pour un temps d'exécution inférieur à 20 millisecondes.

4.1. initialisation du suivi

Durant cette étape, nous supposons que l'objet à suivre reste immobile dans l'image. Tout d'abord, l'opérateur sélectionne à l'aide d'une ellipse le motif courant dans la première image. L'algorithme de reconnaissance d'aspect calcule alors, pour chaque image de référence de la base d'apprentissage, l'erreur quadratique de la différence de niveaux de gris entre le motif de

référence testé et le motif courant échantillonné dans l'ellipse après correction de ses paramètres (c'est-à-dire après avoir recalé le motif à l'aide de la matrice d'interaction A associée). Le motif de référence donnant alors l'erreur quadratique la plus faible sera reconnu comme le motif courant à suivre à la prochaine itération. Cette phase d'initialisation est illustrée par la figure 23.

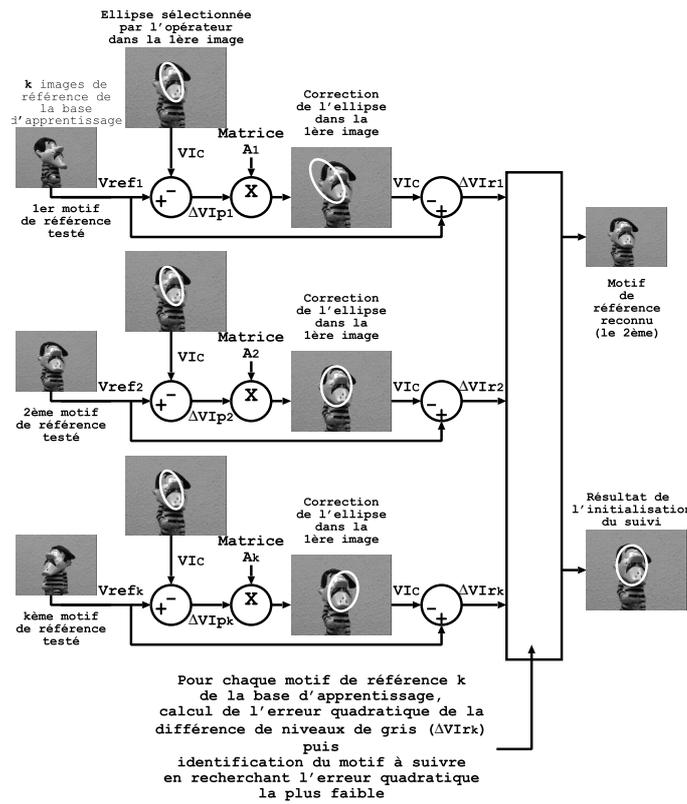


Figure 23. – Principe d'initialisation de la phase de suivi.

4.2. suivi d'objets texturés 3D

Compte tenu de son efficacité, l'implémentation de l'algorithme proposé permet de suivre en temps réel un objet 3D avec changement d'aspect (moins de 20 millisecondes par itération). Cet algorithme peut s'écrire de la manière suivante :

DEBUT PROCEDURE SUIVI ($\mu, Image$)

$$\mu_p = \mu \text{ de l'itération précédente}$$

$$V_{Ic} = \text{EchantillonnerMotif}(\mu_p, Image)$$

$$\mu_r^{actuel} = \mu_p + A^{actuel}(V_{I_{ref}}^{actuel} - V_{Ic})$$

$$V_{Ic} = \text{EchantillonnerMotif}(\mu_r^{actuel}, Image)$$

$$\psi^{actuel} = B^{actuel}(V_{I_{ref}}^{actuel} - V_{Ic})$$

$Correction = \text{ChangerMotifDeRéférence}(\psi^{actuel})$

Si $Correction = 0$

alors pas de correction des paramètres ellipse

$\mu = \mu_r^{actuel}, \psi = \psi^{actuel}$

sinon correction des paramètres ellipse

$\mu'_{su} = \text{CorrigerParamètresEllipse}(\mu'_{ac} = \mu_r^{actuel}, \mu_{ac}^{base}, \mu_{su}^{base})$

$\mu_p = \mu'_{su}$

$V_{Ic} = \text{EchantillonnerMotif}(\mu_p, Image)$

$\mu_r^{suivant} = \mu_p + A^{suivant}(V_{I_{ref}}^{suivant} - V_{Ic})$

$V_{Ic} = \text{EchantillonnerMotif}(\mu_r^{suivant}, Image)$

$\psi^{suivant} = B^{suivant}(V_{I_{ref}}^{suivant} - V_{Ic})$

$\mu = \mu_r^{suivant}, \psi = \psi^{suivant}$

Fin Si

FIN PROCEDURE SUIVI

Dans l'exemple présenté, deux objets sont utilisés : une canette de soda et une figurine. Étant donné la nature des données traitées (flot d'images vidéo), les résultats sont difficiles à représenter. Nous avons donc choisi de visualiser les résultats du suivi à des instants différents de l'expérience. Les images retenues ont

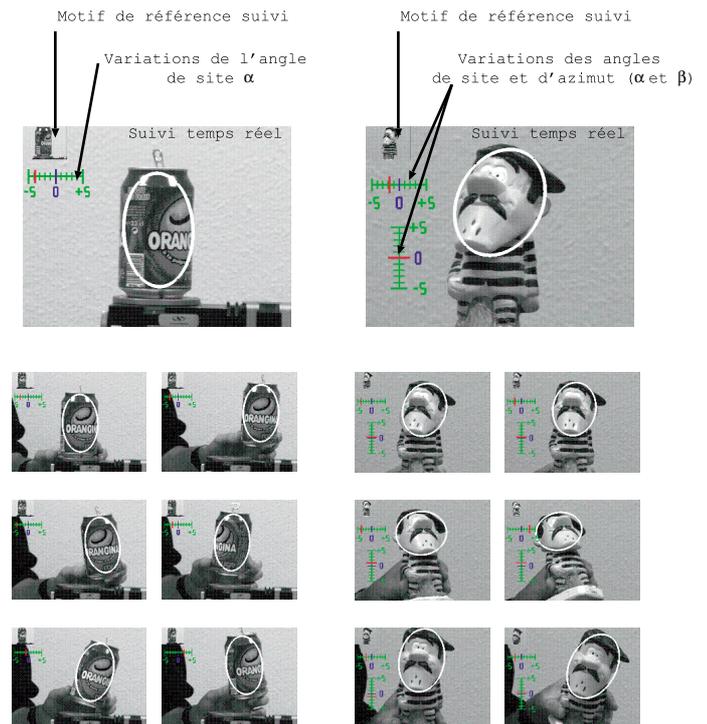


Figure 24. – Exemples de suivi d'objets 3D.

été sélectionnées de manière à montrer la robustesse du suivi (figure 24).

Cet algorithme de suivi d'objets 3D nous fournit en temps réel trois informations :

- le motif de référence actuellement suivi représenté par une image dans le coin supérieur gauche de la fenêtre de visualisation,
- les valeurs angulaires en site et azimut (α et β) données par deux curseurs sur des échelles graduées,
- le résultat du suivi du motif à l'intérieur de l'ellipse.

Les premiers essais de suivi 3D ont donné des résultats positifs. Le passage d'un motif de référence à un autre s'effectue correctement avec une grande précision. Pour pouvoir être robuste aux changements de fond texturé dans l'image, il faudrait s'assurer que pendant la phase d'apprentissage hors ligne de la matrice d'interaction A les ellipses perturbées autour de leur position de référence restent dans l'objet. Mais ceci limiterait l'amplitude des mouvements autorisés entre deux images consécutives. De plus, comme toute différence de niveaux de gris se traduit par un déplacement, nous restons sensible à certaines perturbations pouvant se produire lors du suivi comme l'apparition d'occultations. C'est pourquoi, nous proposons dans la section suivante une solution temps réel traitant le problème des occultations.

5. traitement des occultations

Un des problèmes rencontrés lors du suivi est l'apparition d'occultation sur le motif durant le déplacement de l'objet.

5.1. méthode de seuillage adaptatif

La solution mise en œuvre est basée sur les travaux de Gregory D. Hager et Peter N. Belhumeur [8]. La méthode de détection et de traitement des occultations est une méthode de seuillage appliquée au vecteur de différence de niveaux de gris ΔVI entre le vecteur de référence suivi VI_{ref} et le motif courant dans l'image VI_c . Ces seuils, calculés pour chaque élément du vecteur de résidus, sont aussi regroupés dans un vecteur dit vecteur de seuil adaptatif VI_s . Ce vecteur VI_s est appris lors de la phase d'apprentissage pour le calcul de la matrice d'interaction A . À partir du vecteur de référence choisi $VI_{ref} = (i_{ref1}, i_{ref2}, \dots, i_{refN})^t$ et des différents vecteurs courants $VI_c^j = (i_{c1}^j, i_{c2}^j, \dots, i_{cN}^j)^t$ échantillonnés sur N points ($N = 373$) après M perturbations des paramètres de la transformation ($M = 1000$), nous estimons $VI_s = (i_{s1}, i_{s2}, \dots, i_{sN})^t$ comme un vecteur moyen. Un élément k du vecteur VI_s est donné par la formule suivante :

$$i_{sk} = \frac{1}{M} \sum_{p=1}^M (i_{refk} - i_{ck}^p) \quad (16)$$

Nous estimons alors le vecteur des écarts types $V\sigma_{Is} = (\sigma_{i_{s1}}, \sigma_{i_{s2}}, \dots, \sigma_{i_{sN}})^t$. Un élément k du vecteur $V\sigma_{Is}$ est donné par la formule suivante :

$$\sigma_{i_{sk}} = \sqrt{\frac{1}{M} \sum_{p=1}^M [(i_{refk} - i_{ck}^p) - i_{sk}]^2} \quad (17)$$

Nous traitons donc les problèmes d'occultations en supposant que pour chaque point échantillonné, la variation de niveau de gris suit une distribution gaussienne (moyenne et écart type). Dans cette méthode, les occultations détectées doivent entraîner de fortes variations de niveaux de gris dans le vecteur de résidus ΔVI dont la normalisation est fonction de la mise à jour de la matrice diagonale de masquage W pondérant les erreurs dues aux occultations. Le nouvel algorithme de suivi pour un motif de

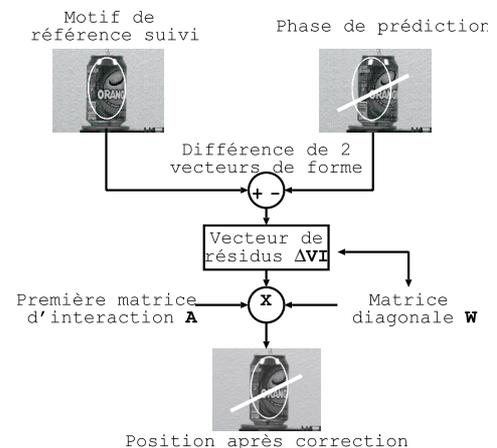


Figure 25. – Traitement des occultations pour des mouvements fronto parallèles du motif suivi dans l'image.

référence donné traitant le problème d'occultations pour des mouvements fronto parallèles de l'objet dans l'image est illustré figure 25.

5.2. essais

Dans la figure 26, nous présentons maintenant quelques images d'une séquence de suivi avec détection et suppression des occultations. Les fortes variations dans le vecteur de différence ΔVI sont bien détectées et traitées comme une apparition d'occultation (points noirs sur l'image). Le positionnement de l'ellipse sur le motif est correct.

Le mérite de cette approche est de pouvoir suivre en temps réel l'objet dans la séquence d'images puisqu'elle est peu coûteuse

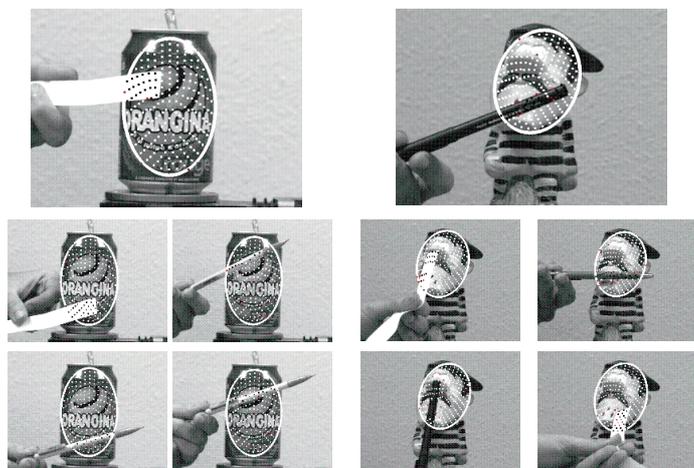


Figure 26. – Quelques exemples de détection d'occultations.

en temps de calcul (< à 30 millisecondes avec la détection des occultations) tout en minimisant l'influence des occultations. Bien entendu, comme nous travaillons avec un vecteur de résidus ΔVI pour corriger la position prédite de l'objet dans l'image, les occultations ne peuvent recouvrir que partiellement l'aspect du motif. Trop de mesures aberrantes dans le vecteur de résidus ΔVI ne nous permettraient plus d'assurer un suivi correct de l'objet dans l'image car même filtrées, la richesse des informations recueillies serait trop faible.

6. conclusion et perspectives

Dans cet article, nous avons présenté une solution de suivi d'objets 3D, temps réel, basée sur l'apparence et qui gère les changements d'aspect du motif. Pour cela, six paramètres sont nécessaires. Les quatre premiers paramètres (X_c , Y_c , R_1 et θ) caractérisent les mouvements fronto parallèles de l'objet dans l'image où l'aspect du motif suivi n'est pas modifié. Toutefois, la position, l'orientation et la taille du motif peuvent changer. Les deux derniers paramètres (α et β) caractérisent les variations en site et azimut de l'objet dans l'image modifiant alors l'aspect du motif suivi.

Notre méthode présente des originalités comme le calcul des matrices d'interaction A et B sans utiliser les matrices jacobiniennes de l'image, ou une phase d'exploration autour de la prédiction supprimée lors du suivi. C'est une méthode généraliste car les variations de position de l'objet dans l'image correspondent aux variations des paramètres d'une transformation géométrique dont le choix est à l'initiative du programmeur suivant l'aspect et le volume de l'objet à suivre. De plus, c'est un algorithme très peu coûteux en temps de calcul (multiplication d'une matrice par un vecteur).

Notre souhait, maintenant, est de développer un second algorithme de suivi 3D à partir des coordonnées 3D des points de l'objet où sont échantillonnés les niveaux de gris et de pouvoir ainsi représenter l'apparence de l'objet par un seul motif de référence

BIBLIOGRAPHIE

- [1] S. Basu and A. Pentland, « A three-dimensional model of human lip motions trained from vidéo », Technical Report 441, M.I.T. Media Laboratory Perceptual Computing Section, 1999.
- [2] M.J. Black and A.D. Jepson, « Eigenttracking : Robust matching and tracking of articulated objects using in view-based representation », *International Journal of Computer Vision*, 26(1), pp. 63-84, 1998.
- [3] R. Brunelli and Poggio T. Template matching : Matched spatial filter and beyond. A.I. », Memo 1549, M.I.T., October 1995.
- [4] T. Darell, I.A. Essa and A.P. Pentland, « Task-specific gesture analysis in real-time using interpolated views », *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(12), pp. 1236-1242, 1996.
- [5] K. Deguchi and T. Noguchi, « Visual servoing using eigenspace method and dynamic calculation of interaction matrices », In *ICPR96*, pp. A7E.3, 1996.
- [6] F. Dellaert and R. Collins, « Fast image-based tracking by selective pixel integration », In *ICCV Workshop on Frame-Rate Vision*, Greece, September 1999.
- [7] M. Gleicher, « Projective registration with difference decomposition », In *CVPR97*, pp. 331-337, 1997.
- [8] G.D. Hager and P.N. Belhumeur, « Efficient region tracking with parametric models of geometry and illumination », *IEEE Trans. on PAMI*, 20(10), pp. 1025-1039, 1998.
- [9] T. M. Caelli, J. J. Wu, R. E. Rink and V. G. Gourishankar, « Recovery of the 3-d location and motion of a rigid object through camera image », *International Journal of Computer Vision*, 3, pp. 373-394, 1989.
- [10] F. Jurie and M. Dhome, « Un algorithme efficace de suivi d'objets dans des séquences d'images » In *Congrès francophone RFIA*, Vol. 1, pp. 537-546, Paris, February 2000.
- [11] F. Jurie and M. Dhome, « Real time template matching : an efficient approach » In *the 12th International Conference on Computer Vision*, Vancouver, Canada, July 2001.
- [12] H. Kollnig and H.H. Nagel, « 3d pose estimation by directly matching polyhedral models to gray value gradients » *International Journal of Computer Vision*, 23(3), pp.283-302, 1997.
- [13] M. La Cascia, S. Sclaroff and V. Athitsos, « Fast, reliable head tracking under varying illumination: An approach based on registration of textured-mapped 3d models » *PAMI*, 22(4), pp. 322-336, April 2000.
- [14] D. Lowe, « Robust model-based motion tracking through the integration of search and estimation » *International Journal of Computer Vision*, 8(2), pp. 113-122, 1992.
- [15] E. Marchand and G.D. Hager, « Dynamic sensor planning in visual servoing », In *the IEEE International Conference on Robotics and Automation (ICRA)*, Vol. 3, pp. 1988-1993, Leuven, Belgium, May 1998.
- [16] B. Moghaddam and A. Pentland, « A subspace method for maximum likelihood target detection », Technical report, M.I.T. Media Laboratory Perceptual Computing Section, 1995.
- [17] Hiroshi Murase and Shree K. Nayar, « Visual learning and recognition of 3-d objects from appearance », *International Journal of Computer Vision*, 14, pp.5-24, 1995.
- [18] Shree K. Nayar and Ruud M. Bolle, « Reflectance based object recognition », *International Journal of Computer Vision*, 17(3), pp. 219-240,

1996.

- [19] Sameer A. Nene, Shree K. Nayar and Hiroshi Murase, « Subspace methods for robot vision », *IEEE Transactions on Robotics and Automation on Vision-Based Control of Robot Manipulators*, 12(5), pp.

750-758, October 1996.

- [20] J. Strom, T. Jebara, S. Basu and A. Pentland, « Real time tracking and modeling of faces : An ekf-based analysis by synthesis approach » , Technical Report 506, M.I.T. Media Laboratory Perceptual Computing Section, 1999.

LES AUTEURS

Florent DUCULTY



Titulaire du diplôme d'Ingénieur Génie Electrique CUST, il prépare actuellement une thèse de doctorat au Laboratoire des Sciences et Matériaux pour l'Electronique, et d'Automatique de Clermont-Ferrand (LASMEA UMR 6602 du CNRS). Son thème de recherche est le développement d'algorithmes de suivi d'objets 3D basés sur l'apparence dans une séquence d'images vidéo.

FRÉDÉRIC JURIE



Actuellement Chargé de Recherche au CNRS au sein du département STIC (Sciences et Technologies de l'Information et de la Communication), il mène ses activités de recherche au Laboratoire des Sciences et Matériaux pour l'Electronique, et d'Automatique (LASMEA), UMR 6602 du CNRS, à l'Université Blaise Pascal de Clermont-Ferrand. Ses activités de recherche concernent la reconnaissance et le suivi d'objets dans des séquences d'images vidéo.

Michel DHOME



Michel DHOME est Directeur de Recherche au Centre National de la Recherche Scientifique et exerce au sein du LASMEA de 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 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.