

Recalage de modèles tridimensionnels courbes sur des images stéréoscopiques

Pose recovery of curved 3D models based on stereoscopic images

par Didier MAMAN, Fawzi NASHASHIBI, Philippe FUCHS, Jean-Claude BORDAS

Centre de Robotique de l'École Nationale Supérieure des Mines de Paris,
60, bd Saint-Michel, F-75006 Paris, {maman,fuchs,nashashibi}@caor.ensmp.fr
Groupe Téléopération et Robotique, Département Surveillance Diagnostic Maintenance,
Direction des Études et Recherche d'EDF,
6, quai Watier, F-78401 Chatou, Jean-Claude.Bordas@edfgdf.fr

résumé et mots clés

Cet article traite de l'utilisation d'algorithmes de vision artificielle et de techniques de réalité augmentée dans le but d'améliorer les fonctionnalités des postes de téléopération déjà utilisés à EDF. Au cours d'une opération de maintenance sur site, la connaissance exacte de la géométrie de l'environnement distant sur lequel le robot intervient est cruciale puisque ce modèle constitue le seul monde connu sur lequel s'appuient les assistances proposées par le système informatique. Le système décrit ici combine des outils d'analyse d'images et de réalité augmentée au sein d'un environnement interactif contrôlé par un opérateur humain. L'utilisation d'un algorithme de recalage automatique permet à l'opérateur de spécifier interactivement un modèle géométrique et sémantique du site distant en s'appuyant sur les images du site retournées par des caméras embarquées. Cet algorithme permet en outre au système informatique de maintenir à jour le modèle global de l'environnement lors de changements de configuration ou de point de vue mal estimés. L'approche retenue, qui sera détaillée dans ce papier, opère sur des objets paramétriques à surface courbe.

Recalage 3D/2D, modèles paramétriques, modélisation interactive, téléopération.

abstract and key words

This paper deals with the use of mixed reality as a new assistance and training tool for performing teleoperation tasks in hostile environments. During a maintenance operation, a telerobotic task needs a perfect knowledge of the remote scene in which the robot operates because this local model is the world over which the robot is working. We will describe here a new interactive approach combining image analysis and mixed reality techniques for assisted 3D geometric and semantic modeling. We also tackle the problem of pose recovery using a stereoscopic system mounted on a robot arm. The proposed model-based approach is used for accurate locating of complex parametric curved objects. It therefore constitutes a unified tool for building (during an interactive modeling session) and maintaining (by correcting badly estimated displacements of the robots) the local geometric model of the remote environment. This augmented reality and vision-based system is being developed in order to improve the teleoperation already in use at EDF.

3D/2D fitting, parametric models, interactive modeling, teleoperation.

1. introduction

Cet article traite du problème du recalage d'objets tridimensionnels complexes basé sur un retour vidéo stéréoscopique, dans un contexte de téléopération en milieu hostile.

Les travaux menés actuellement s'inscrivent dans le cadre d'un partenariat entre le Centre de Robotique de l'École des Mines de Paris et la Direction des Études et Recherches d'EDF. Il s'agit d'exploiter les nouvelles techniques de réalité augmentée afin de fournir des outils d'aide à la téléopération. Typiquement, la maintenance téléopérée nécessite une connaissance parfaite (éventuellement partielle) de l'environnement dans lequel évolue le robot placé sous le contrôle d'un opérateur humain. En effet, toutes les assistances proposées par le système informatique (éviter d'obstacles, maintien d'une trajectoire particulière, réduction du nombre de degrés de liberté à contrôler, simulation et planification de mission, etc.) se basent sur ce modèle géométrique estimé.

Afin de construire et de maintenir à jour ce modèle, nous proposons un système coopératif combinant d'une part les capacités de reconnaissance et d'interprétation de l'opérateur humain, et d'autre part, l'assistance d'un système informatique pour permettre une modélisation précise, la mise à jour automatique de cette modélisation, ainsi que la réalisation de certaines tâches en mode automatique ou semi-automatique.

Au cours d'une session de modélisation interactive s'appuyant sur les images vidéos du site distant et à l'aide d'une interface adaptée, l'opérateur est en mesure de fournir une description géométrique, mécanique et sémantique de l'environnement téléprésent, en indiquant quels objets se trouvent dans la scène, quelles sont leurs position et orientation approximatives et quelles relations géométriques les lient. La figure 5 présente le dispositif utilisé et illustre le principe de fonctionnement de la modélisation interactive. L'opérateur sélectionne un modèle dans une bibliothèque et le superpose simplement à l'objet correspondant dans l'image stéréoscopique à l'aide d'un capteur de localisation à six degrés de liberté. Des relations géométriques entre les modèles peuvent être spécifiées grâce à l'interface, mais dans la majorité des cas, elles sont activées automatiquement en fonction du contexte (nature et géométrie du modèle) et de critères de proximité avec les objets déjà modélisés.

L'utilisation d'un algorithme de recalage automatique permet d'améliorer la précision de la modélisation en la rendant indépendante du placement approximatif fait par l'opérateur. D'autre part, ce même algorithme peut être utilisé pour actualiser le modèle global de l'environnement lors d'un changement de configuration mal estimé (erreurs issues des capteurs odométriques¹ et dérive due aux capteurs inertiels² pour les robots mobiles par exemple).

1. Capteurs odométriques : capteurs mesurant le nombre de tours effectués par une roue.

2. Capteurs inertiels : capteurs mesurant les mouvements (accéléromètre, gyromètre, inclinomètre).

De manière plus générale, l'algorithme de recalage est capable de préciser la configuration de tout système paramétrique ayant une répercussion visuelle prévisible dans l'image, comme par exemple les paramètres de calibration extrinsèques et intrinsèques des caméras, si un modèle paramétrique peut en être donné. Dans ce cas, il sera naturellement nécessaire de disposer d'objets de référence parfaitement connus dans la scène.

L'objectif du module de recalage proposé est donc de localiser des objets complexes en temps rapide³ et de préciser leur configuration propre. Il s'appuie sur l'information photométrique contenue dans les images du site observé et sur la donnée d'une position initiale du modèle, qui doit être relativement proche de la position effective de l'objet correspondant dans la scène. Les modèles des objets doivent être connus mais peuvent être paramétriques et déformables. En particulier, leurs dimensions peuvent être mal connues et déterminées par l'algorithme dans le cas où ce dernier dispose de plusieurs images (stéréoscopie ou plus). L'algorithme doit d'autre part respecter les contraintes géométriques entre les différents modèles et être complètement générique vis-à-vis du nombre de vues dont nous disposons du site (éventuellement une seule), ainsi que du nombre et de la nature des paramètres à déterminer (paramètres de positionnement, de forme, de taille, de liaison géométrique entre les modèles ou de calibration des caméras, qui sont spécifiés de manière interactive).

Dans cet article, nous commencerons par situer l'approche proposée relativement aux systèmes existants (section 2). Nous détaillerons ensuite le fonctionnement de notre algorithme de recalage qui constitue la principale assistance du système puisqu'elle permet d'obtenir un modèle géométrique précis de l'environnement et de le maintenir à jour tout au long de la mission.

Dans un premier temps, une extraction d'indices dans les images réelles est effectuée et des indices comparables du modèle sont déduits des structures qui le représentent. Les indices utilisés sont les contours de discontinuité de la normale à la surface ainsi que les contours d'occlusion. Cette partie est détaillée dans la section 4.

Nous cherchons ensuite à établir un ensemble d'associations entre les contours extraits des images et ceux du modèle par une recherche à travers l'espace des correspondances. La complexité de cette recherche est réduite en ordonnant le parcours de manière à considérer les hypothèses les plus vraisemblables en priorité. Pour des raisons de performance, nous utilisons pour cette partie une approximation polyédrique du modèle à recalage. La stratégie d'établissement des associations est décrite dans la section 6.

Pour finir, une minimisation des écarts entre les contours associés est réalisée. Elle peut être basée sur l'approximation polyédrique des modèles pour un recalage plus rapide (section 5) ou bien s'appuyer sur leurs contours courbes calculés grâce à leur représentation surfacique exacte. Elle permet dans ce cas d'obtenir une modélisation plus précise. Son fonctionnement est expliqué dans la section 7.

3. Temps rapide : de l'ordre de quelques secondes, compatible avec des applications interactives.

2. état de l'art

La présence d'un opérateur humain dans les missions téléopérées facilite grandement le travail du système de vision et permet de le rendre robuste et rapide. La problématique posée dans ce contexte applicatif n'est donc pas la reconnaissance et la localisation automatique des objets d'intérêt pour la mission. Pour l'opérateur, la reconnaissance est immédiate et sélective et ce dernier peut également, au moyen d'une interface spécifique, indiquer la position des objets dans l'image.

Parmi les systèmes existants, certaines approches étudient les solutions au problème de la détermination des six paramètres de positionnement pour un modèle d'un type particulier et à partir d'un nombre minimal de correspondances de points ou de segments. Les correspondances sont données par l'opérateur par pointages sur les images. Les indices pointés dépendent parfois du type de modèle considéré. Le problème de la mise en correspondance est ainsi évité mais cette technique de désignation de la position de l'objet dans l'image est fastidieuse pour l'opérateur et la précision de la modélisation qui en découle n'est pas optimale. De plus, l'algorithme peut difficilement être utilisé pour maintenir à jour la base de données lors de changements de configuration mal estimés (déplacement des caméras, du robot manipulateur ou d'un des objets d'intérêt sur le site d'intervention). En effet, tous les pointages sont à refaire dans ce cas. Toutes ces approches sont très spécifiques et si une autre classe de modèles venait à être considérée, ou si le nombre et la nature des paramètres à déterminer était différents, la solution à cet « autre » problème devrait être recherchée à nouveau. Dans [Dho89] et dans [Kim96], les objets considérés sont polyédriques. Dans [Ric91], seuls les cylindres généralisés homogènes et à axe droit sont traités. Dans ces trois cas, les modèles ne peuvent pas contenir de paramètres propres (dimensions ou forme) et les éventuelles contraintes géométriques entre composants de la scène ne sont pas représentables. Le système PYRAMIDE [Pyr95, Eve93] permet de représenter diverses primitives volumiques simples comme les parallélépipèdes rectangles, les cylindres et les cônes. Dans [Coo94], seuls les parallélépipèdes rectangles et les cylindres sont gérés. Dans les deux derniers systèmes mentionnés, les primitives peuvent en outre avoir des dimensions variables. La plupart des systèmes mentionnés ici s'appuient sur deux vues ou plus (sauf [Dho89] et [Ric91] qui n'en utilisent qu'une seule).

D'autres approches, plus génériques, ont été rencontrées. Dans [Kri90, Pon92], les auteurs proposent une méthodologie générale permettant de caractériser les contours en projection d'une très grande variété d'objets courbes utilisés en CAO. Leur approche suppose cependant une mise en correspondance manuelle et la détermination des équations permettant le recalage n'a pas été rendue automatique dans le cas général (que ce soit du point de vue de la forme de l'objet ou du point de vue du nombre et de la nature des paramètres). La méthode proposée par Lowe [Low91, Low92] est très générique vis-à-vis du nombre et de la

nature des paramètres à recalculer. Elle permet ainsi de représenter des objets composés présentant des éventuelles articulations. Chaque composant peut en outre contenir des paramètres propres faisant varier ses dimensions ou même sa forme. Ce système propose de plus une mise en correspondance automatique pour des modèles suffisamment proches de leur solution. Il ne s'applique malheureusement qu'à des objets polyédriques et l'exploration de l'espace des correspondances n'est pas complet, ce qui laisse à l'algorithme la possibilité de manquer la solution. D'autre part, la méthode ne s'applique qu'à une seule image.

Le système que nous proposons s'appuie sur un nombre quelconque de vues du site à modéliser. Si plusieurs vues sont disponibles, les indices extraits des différentes images contiennent implicitement de l'information tridimensionnelle. Le recalage simultané des contours des modèles sur ces indices permet alors de résoudre automatiquement l'indétermination taille-profondeur qui existe pour les images simples. L'algorithme que nous présentons réalise une mise en correspondance automatique par une exploration optimisée de l'espace des correspondances. L'approche retenue pour résoudre ce problème est une variante de la méthode décrite par Lowe dans [Low91, Low92]. La principale différence provient du fait que dans notre approche, l'exploration est exhaustive et assure qu'une solution sera trouvée si elle existe. Cette mise en correspondance automatique permet à l'opérateur de spécifier la position des objets de façon très naturelle et conviviale (par superposition approximative du modèle virtuel sur l'objet correspondant grâce à des techniques de réalité augmentée [Mam98a, Mam98b]). Elle est d'autre part compatible avec une mise à jour simplifiée du modèle géométrique global après des modifications incertaines. Les modèles gérés par l'algorithme sont paramétriques et des contraintes géométriques peuvent exister entre les différents composants. La généralité de l'approche permet d'autre part à l'algorithme de recalculer indifféremment la position des objets par rapport aux caméras ou bien la position et certains paramètres intrinsèques des caméras relativement à des objets de référence connus. Les modèles appartiennent à une très grande variété d'objets courbes. La méthodologie utilisée pour le recalage s'inspire de celle qui a été proposée par Kriegman et Ponce dans [Kri90]. L'approche a cependant été généralisée et les équations des contours en projection peuvent être calculées en ligne quelques soient le nombre et la nature des paramètres à déterminer. De plus, ces techniques ont été intégrées à un système réalisant la mise en correspondance de manière automatique et travaillant sur un nombre de vues quelconque.

3. les indices utilisés

Les indices utilisés pour recalculer le modèle sur les images sont les contours de discontinuité de la normale à la surface (comme les arêtes d'un cube) et les contours d'occlusion (comme les bords d'un cylindre ou le contour d'une sphère). Ils sont relativement

simples à obtenir à partir des modèles ou à extraire des images et sont stables par changement d'éclairage et de point de vue.

Pour le modèle virtuel utilisé lors de la phase de mise en correspondance et du recalage polyédrique, les contours s'obtiennent à l'aide des arêtes de son approximation polyédrique. Bien que le modèle global de la scène soit paramétrique, on connaît la valeur numérique des matrices représentant les transformations qui apparaissent dans l'arbre de construction. Par un parcours de cet arbre et à l'aide de simples multiplications de matrices, il est possible de calculer la position des sommets du modèle en projection dans les images.

En ce qui concerne le recalage basé sur les contours courbes, les contours exacts des modèles s'obtiennent grâce aux équations paramétriques des surfaces qui constituent leur frontière. Le calcul des équations cartésiennes de ces contours en projection sera détaillé en section 6.

Enfin, pour l'objet réel, les contours sont obtenus par une suite de traitements d'image que nous allons détailler ici. Pour les besoins de la mise en correspondance, ces traitements doivent fournir des chaînes de points de contours connexes. Chacune de ces chaînes doit par ailleurs représenter en partie ou en totalité un contour de l'objet réel mais ne doit en aucun cas s'étendre sur plusieurs contours. En effet, la stratégie de mise en correspondance détaillée dans la section 4 cherchera à associer chacune de ces chaînes à un contour unique du modèle. Les traitements d'images employés devront donc permettre un découpage des chaînes aux endroits où deux contours distincts se rejoignent. Ceci est obtenu dans la majorité des cas en testant les variations brutales de la courbure de la chaîne.

Les traitements effectués sont les suivants : nous commençons par transformer l'image couleur en une image de niveaux de gris car nous ne sommes intéressés que par les variations de l'intensité. Ensuite, les maxima locaux du gradient sont extraits par un filtrage récursif utilisant les opérateurs de Deriche [Der87]. Un seuillage par hystérésis adaptatif est alors appliqué aux points de contours obtenus afin de ne conserver que les contours les plus marqués et ainsi réduire l'espace des données. Les points de contours restants sont chaînés avant d'être structurés en approximation polygonale [Hor93]. Pour finir, nous scindons les chaînes obtenues afin d'obtenir des sections n'appartenant qu'à un seul contour, comme nous l'avons mentionné précédemment. Un exemple de résultat obtenu par ces traitements est donné dans la figure 6.

4. mise en correspondance des indices

La mise en correspondance, qui s'appuie sur une approximation polyédrique du modèle, se décompose en deux problèmes

distincts. Dans un premier temps, nous expliquons comment établir et parcourir les hypothèses d'associations entre les indices extraits du modèle et des images. Nous précisons ensuite les tests de validité et d'acceptation utilisés pour juger les hypothèses de correspondance et éventuellement arrêter l'exploration.

4.1. les hypothèses d'associations

Chaque contour extrait des images est une liste chaînée de segments, correspondant à son approximation polygonale. Les contours dérivés de l'approximation polyédrique du modèle sont également des listes de segments.

Les associations que nous allons tenter d'établir entre ces deux ensembles de contours se feront au moyen d'appariements, qui doivent être vus comme des correspondances entre des points particuliers des deux contours associés. Les appariements ne sont pas réalisés segments à segments puisque les découpages des contours courbes en segments ne sont pas uniques et n'ont aucune raison de se correspondre. Ils se feront plutôt entre les sommets de l'approximation polygonale des contours image et les points les plus rapprochés du contour modèle, comme l'illustre la figure 1.

Nous disposons, pour chaque contour du modèle en projection, d'un certain nombre de contours image candidats. Le problème de la mise en correspondance va être de choisir le bon candidat. Ce choix passe par l'attribution d'un coefficient de confiance, basé sur des critères de proximité et de ressemblance. Ce coefficient dépend de l'écartement moyen et de la différence d'orientation moyenne entre les deux contours. Les mesures considérées sont effectuées au niveau des appariements retenus. Le coefficient de confiance dépend également de la similarité de longueur entre les deux contours ainsi que des éventuelles concurrences d'association. Deux contours images sont dits *concurrents* vis-à-vis d'une association avec un contour du modèle si leurs projections orthogonales sur ce contour se recouvrent. On considère dans ce cas que les deux contours images ne peuvent pas être simultanément

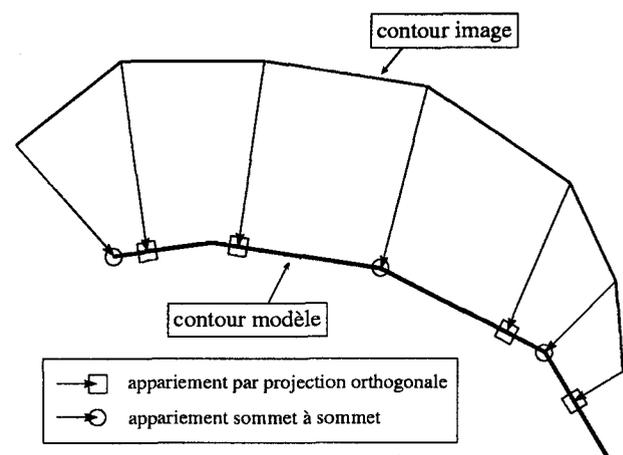


Figure 1. – Mesures d'écart établies entre deux contours associés.

associés au contour modèle. Si un contour modèle propose de nombreuses associations concurrentes, leurs coefficients de confiance respectifs sont réduits car ces associations se « partagent » la probabilité d'être l'association effective. La figure 2 représente un contour modèle avec quatre associations potentielles, dont trois qui sont concurrentes entre elles. La fonction que nous avons choisi d'utiliser afin d'atténuer les probabilités d'un groupe A d'associations concurrentes est la suivante :

$$\forall k \in A, \quad P_{conc}^k = P^k \cdot \frac{P^k}{\sum_{i \in A} P^i}$$

où P_{conc}^k représente le coefficient de confiance corrigé de l'association k et P^k son coefficient de confiance *a priori*. Cette fonction fait intervenir la confiance relative de l'association considérée parmi l'ensemble des associations concurrentes ainsi que sa confiance absolue. Ceci permet de pénaliser les associations d'autant plus fortement qu'elles avaient un faible coefficient de confiance *a priori*.

Comme nous l'avons vu, le nombre d'hypothèses de correspondance envisagé est réduit du fait de l'existence d'associations concurrentes au sein d'une même image. De la même manière, la stéréoscopie et certaines contraintes liées à la géométrie épipolaire permettent d'écarter des correspondances comportant des associations clairement incompatibles. Lorsque deux contours image provenant de deux images différentes sont associés à un même contour du modèle, la reconstruction tridimensionnelle de ce couple de contours image fournit une information essentielle. En effet, la reconstruction 3D basée sur deux contours image ne correspondant pas au contour modèle envisagé a peu de chance d'en être proche dans l'espace. Par contre, si les hypothèses d'associations sont correctes, la reconstruction 3D et le contour modèle doivent être proches l'une de l'autre.

Cependant, même en choisissant des associations compatibles entre elles et en outre les plus vraisemblables, la mise en correspondance n'est pas nécessairement correcte. Dans ce cas, il faut pouvoir détecter l'incohérence de cette correspondance et

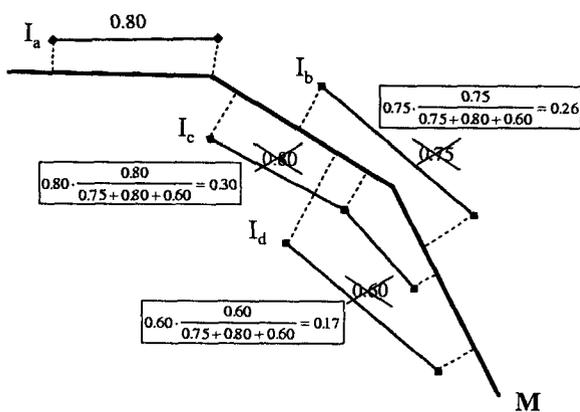


Figure 2. – Coefficients de confiance « a priori » et « corrigés » pour un groupe d'associations concurrentes.

disposer d'une stratégie permettant d'en proposer une alternative. Les associations étant constituées de manière indépendante pour chaque contour du modèle, leur cohérence globale sera vérifiée par la procédure de minimisation de l'écart entre le modèle et l'objet réel (cf. section 5). En fait, si le résidu après minimisation est de l'ordre de l'écart type de l'erreur dans les données, il est peu probable qu'une des associations soit erronée. En général, un mauvais ensemble d'associations ne pourra pas être correctement recalé sur l'objet réel. Il est en ce sens possible d'en vérifier la validité, comme l'explique la section 4.2.

La stratégie de mise en correspondance est alors la suivante : on commence par recalé le modèle sur les meilleures associations trouvées. Ce recalage partiel ayant rapproché le modèle et l'objet réel, d'autres associations vont pouvoir être établies et ajoutées à l'ensemble initial. Un nouveau recalage est alors effectué sur cet ensemble augmenté, et ainsi de suite, jusqu'à ce que le modèle soit suffisamment proche de l'objet réel et qu'aucune nouvelle association ne puisse être ajoutée. Si un des recalages partiels révèle une incohérence globale de l'ensemble des associations, cela signifie qu'une mauvaise association a été introduite dans l'ensemble au dernier ajout. On réduit alors les coefficients de confiance attribués à ces associations et on reconsidère l'ajout.

4.2. les tests de validité et d'acceptation d'une correspondance

Les recalages partiels réalisés à chaque étape de l'exploration sont obtenus par minimisation d'une fonction F qui représente la somme de N termes, chacun étant le carré d'une quantité que l'on peut considérer normalement distribuée et de variance 1. Cette hypothèse est approfondie dans [Mam98b]. Une fois que les M paramètres ont été ajustés pour atteindre le minimum de la fonction, les termes de la somme ne sont pas tous statistiquement indépendants. Il peut être montré que pour une fonction linéaire dans ses paramètres, la valeur de F à son minimum est distribuée selon une « loi du χ^2 » à $N - M$ degrés de libertés [Abr68]. Ceci reste en général vrai même si la fonction n'est pas strictement linéaire comme dans notre cas [Pre92].

La fonction densité de probabilité d'une loi du χ^2 à n degrés de libertés vaut :

$$p(x) = \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})}$$

où Γ est définie, pour tout $i > 0$, par :

$$\Gamma(i) = \int_0^{+\infty} t^{i-1} e^{-t} dt$$

La probabilité d'obtenir une valeur de F à son minimum (F_{min}), qui soit supérieure à un seuil S donné, s'obtient alors en intégrant

la fonction densité de probabilité entre S et plus l'infini :

$$Q_n(F_{min} > S) = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \int_S^{+\infty} t^{\frac{n}{2}-1} e^{-\frac{t}{2}} dt$$

Plus S est grand, plus l'obtention d'une valeur de F_{min} supérieure à S est improbable. Cette distribution statistique de F_{min} peut alors être utilisée pour juger de la validité d'une mise en correspondance. En effet, dans le cas où certains des contours image utilisés dans la minimisation ne correspondraient pas réellement aux contours modèles auxquels ils ont été associés, la fonction F que l'on cherche à minimiser ne pourrait plus être identifiée à une somme de carrés de variables aléatoires gaussiennes centrées et réduites et la distribution du minimum obtenu ne suivrait donc plus une loi du χ^2 . La correspondance utilisée à chaque niveau de la recherche étant douteuse, une valeur de F_{min} trop importante – correspondant à une occurrence très improbable dans l'hypothèse d'une correspondance exacte – devrait donc plutôt être interprétée comme une erreur dans cette correspondance. De manière plus intuitive, il est très fréquemment possible de recalcr précisément (c'est-à-dire avec $F_{min} < S$) le modèle géométrique considéré sur des contours images appartenant à l'objet réel correspondant; ne pas y parvenir laisse par conséquent présumer que l'hypothèse d'une correspondance exacte était vraisemblablement fautive. Dans l'autre sens bien sûr, l'obtention d'un F_{min} inférieur au seuil ne permet pas d'affirmer que l'hypothèse de la correspondance était exacte. En d'autres termes,

$$F_{min} > S \Rightarrow \text{correspondance vraisemblablement fautive}$$

Le choix du seuil S à utiliser passe naturellement par la notion de « vraisemblance » mentionnée précédemment. Si on choisit un seuil trop grand, on est presque assuré que seules des correspondances invalides seront écartées de notre recherche. D'un autre côté, beaucoup de correspondances incorrectes mais jugées valides vont être considérées et la recherche dans l'arbre des correspondances sera très coûteuse. Si au contraire le seuil retenu est trop petit, les correspondances incorrectes seront bien éliminées, mais certaines correspondances valides ayant par accident mené F à un minimum un peu élevé, seront également écartées de la recherche. Dans ce cas, si la proportion de bonnes correspondances injustement considérées invalides est trop importante, on risque de manquer la solution. Le compromis que nous avons choisi d'utiliser est un seuil de rejet permettant de classer correctement 84% des bonnes mises en correspondances (nous ne détaillerons pas l'origine de ce seuil).

En fait, le test réalisé ici ne permet pas réellement de valider une correspondance mais plutôt d'écartier les correspondances invalides qui, elles, ne peuvent pas (statistiquement et étant donné une certaine limite de vraisemblance) être des solutions. En effet, il existe des situations où il est possible de « parfaitement » recalcr le modèle sur des contours n'appartenant pas tous à l'objet recherché. Ces cas sont particulièrement fréquents lorsque

l'image est très bruitée ou contient beaucoup de contours proches des contours recherchés, ou encore lorsque les contours du modèle ne permettent pas de résoudre l'ensemble des degrés de liberté du système. Pour décider si une correspondance valide constitue effectivement une solution recherchée, un test supplémentaire doit donc être utilisé.

Un tel test d'acceptation peut par exemple exiger qu'un nombre minimal de contours du modèle soient impliqués dans les associations retenues ou encore que la longueur totale des contours image choisis représente au moins une certaine fraction de la longueur totale des contours visibles du modèle. La difficulté d'établir ces conditions et leurs seuils provient de la grande variété des cas de figure possibles. Les conditions à appliquer à un cube qui comporte beaucoup de contours et celles qu'il faudrait définir pour une sphère qui n'en comporte qu'un sont de natures très différentes. Le degré d'occultation des objets de la scène peut d'autre part être très variable.

Ainsi, il n'est pas possible de définir des critères sûrs permettant de savoir si une correspondance représente effectivement une solution ou non. Il est plutôt possible d'estimer ses chances d'en être une. La solution idéale serait d'explorer tout l'arbre à la recherche de la solution répondant le mieux aux critères d'acceptation choisis. Malheureusement, l'espace des correspondances est souvent trop grand pour que cette recherche exhaustive soit envisageable. Il faut rappeler que dans l'application envisagée, un échec dans la procédure de mise en correspondance ne constitue pas une situation désespérée puisque l'opérateur humain peut à tout moment intervenir et corriger l'erreur en utilisant des techniques plus classiques de désignation directe de certaines correspondances.

Le test d'acceptation que nous avons retenu considère un premier seuil sévère qui permet d'arrêter l'exploration avec la quasi certitude de détenir une solution. D'autre part, afin de répondre au risque de ne jamais atteindre ce seuil, un autre seuil plus tolérant permet de détecter des solutions potentielles. On se fixe alors un temps d'exploration limité et on retient, parmi les solutions potentielles envisagées jusqu'ici, la plus vraisemblable. Le fait que l'exploration des correspondances soit organisée par ordre de confiance décroissante augmente les chances de trouver une solution durant ce laps de temps. Cette limitation de la recherche correspond au temps maximum acceptable pour l'application en question. Il peut par exemple être de l'ordre d'une dizaine de secondes pour la modélisation interactive.

Les deux critères retenus pour le test d'acceptation imposent que la correspondance considérée implique une proportion suffisante de contours modèle en terme de nombre et de longueur. Le seuil « haut » utilisé dans notre implémentation exige qu'au moins un quart des contours visibles du modèle apparaissent dans la correspondance et que la longueur des contours images retenus représente plus d'un tiers de la longueur totale des contours modèle. Le seuil « bas » considère que si les contours images recouvrent plus de 15% des contours modèles impliqués, la correspondance a déjà des chances d'être correcte.

5. minimisation de l'écart basée sur l'approximation polyédrique

Cette minimisation, basée sur l'approximation polyédrique des modèles à recalage, a deux rôles dans le fonctionnement de l'algorithme. Premièrement, elle constitue le test de validité d'une hypothèse de mise en correspondance (cf. section 4.2); elle est donc utilisée à chaque étape du processus d'exploration. Deuxièmement, elle permet le recalage proprement dit des modèles polyédriques voire même des modèles courbes si la précision souhaitée par l'opérateur n'est pas maximale.

Comme nous l'avons déjà mentionné, l'association entre deux contours est réalisée au moyen d'appariements entre les sommets du premier et les segments du second (et non pas segment à segment). Un sommet du contour image est ainsi apparié avec le segment du contour modèle duquel il est le plus proche en projection orthogonale. Une association définit de cette manière autant de mesures d'écart qu'il y a de sommets dans le contour image.

La fonction considérée dans la minimisation sera la somme quadratique de ces multiples écarts, pris sur les différentes associations choisies pour un recalage partiel donné. Son expression est donc relativement simple et dépend directement des coordonnées en projection sur l'écran (X_e, Y_e) des extrémités des segments des contours du modèle.

Pour cette raison, la méthode d'optimisation non linéaire employée, qui s'apparente à une descente de gradient avec convergence forcée de type Levenberg-Marquardt, ne nécessite rien d'autre que la valeur de ces coordonnées et de leurs dérivées partielles en tout point de l'espace des configurations.

Considérons donc les coordonnées (X, Y, Z) d'un sommet S dans le repère propre du modèle. On appellera S_c , de coordonnées (X_c, Y_c, Z_c) , ce même sommet dans le repère de la caméra et (X_e, Y_e) les coordonnées de sa projection S_e sur le plan CCD.

Notons $M_i, 0 \leq i < n$, les matrices des transformations apparaissant dans la branche de construction de l'objet considéré (et permettant de passer du repère local de l'objet au repère caméra). Notons enfin $\partial M_i / \partial p$ (où p est un paramètre de forme ou de positionnement sur lequel opère le recalage) leur dérivée partielle terme à terme par rapport à p . Ces transformations sont connues et les matrices associées ainsi que leurs dérivées partielles peuvent être programmées hors-ligne. On a alors :

$$S_c = \prod_{i=0}^{n-1} M_i \cdot S = M_0 \cdot M_1 \cdots M_{n-1} \cdot S$$

$$\frac{\partial S_c}{\partial p} = \left[\sum_{k=0}^{n-1} \left(\prod_{i=0}^{k-1} M_i \frac{\partial M_k}{\partial p} \prod_{j=k+1}^{n-1} M_j \right) \right] \cdot S + \left[\prod_{i=0}^{n-1} M_i \right] \cdot \frac{\partial S}{\partial p}$$

D'autre part, la projection sur le capteur CCD nous donne :

$$\begin{cases} X_e = f \frac{X_c}{Z_c} \\ Y_e = f \frac{Y_c}{Z_c} \end{cases} \Rightarrow \begin{cases} \frac{\partial X_e}{\partial p} = \frac{f}{Z_c^2} \left(Z_c \frac{\partial X_c}{\partial p} - X_c \frac{\partial Z_c}{\partial p} \right) \\ \frac{\partial Y_e}{\partial p} = \frac{f}{Z_c^2} \left(Z_c \frac{\partial Y_c}{\partial p} - Y_c \frac{\partial Z_c}{\partial p} \right) \end{cases}$$

Nous remarquons donc que le calcul des coordonnées en projection des sommets du modèle et celui de ses dérivées partielles s'obtient à l'aide des seules matrices de transformations locales et de leurs dérivées partielles, supposées connues.

6. minimisation de l'écart basée sur les contours exacts

Dans cette phase du recalage, les associations établies par la procédure de mise en correspondance ne sont pas remises en cause. La formulation du problème est donc plus simple puisqu'il ne s'agit cette fois que d'une minimisation. Cependant, les indices utilisés ne sont pas aussi élémentaires et l'expression des écarts entre les primitives appariées est plus difficile à obtenir. Les techniques utilisées s'inspirent de [Kri90].

La suite de cette partie commence par définir les classes d'objets représentables puis montre comment obtenir les équations cartésiennes des contours apparents de ces objets en projection. Elle explique ensuite comment dériver de ces équations une mesure possible de l'écart entre un point image et les contours modèles correspondants. Pour finir, le calcul de ces écarts et de ses dérivées partielles (nécessaires à la méthode d'optimisation utilisée) est détaillé.

6.1. les surfaces utilisées

Ici, la frontière de l'objet n'est pas décrite, par un ensemble de facettes planes approximantes, mais par un ensemble de surfaces courbes. Les surfaces utilisées sont des surfaces rationnelles de la forme :

$$\begin{cases} X = \frac{P_x(s, t)}{D(s, t)} \\ Y = \frac{P_y(s, t)}{D(s, t)} \\ Z = \frac{P_z(s, t)}{D(s, t)} \end{cases}$$

où P_x, P_y, P_z et D sont des polynômes en s et t . En prenant des polynômes de degré 2, on peut représenter les primitives les plus courantes comme la sphère ou le cylindre. Avec des polynômes de degré 3, on obtient des surfaces bicubiques classiquement employées en CAO. La méthode peut d'autre part être étendue aux cylindres généralisés, aux superquadriques et aux modèles CSG.

6.2. obtention de l'équation cartésienne d'un contour en projection

Comme pour le recalage polyédrique, nous allons rechercher les contours d'occlusion et les contours de discontinuité des surfaces de l'objet. Ces contours seront obtenus à l'aide de trois polynômes, dérivés des équations paramétriques introduites précédemment.

6.2.1. la projection sur l'écran

Que ce soit pour les contours de discontinuité de surface ou pour les contours d'occlusion, les deux premiers polynômes sont obtenus grâce aux relations de projection de la surface de l'objet sur le capteur CCD. Si f est la focale de la caméra, (X_c, Y_c, Z_c) les coordonnées d'un point de la surface de l'objet exprimées dans le repère de la caméra et (X_e, Y_e) le point projeté correspondant dans le repère du capteur CCD, on a :

$$\begin{cases} X_e = f \frac{X_c}{Z_c} \\ Y_e = f \frac{Y_c}{Z_c} \end{cases} \iff \begin{cases} fX_c - Z_cX_e = 0 \\ fY_c - Z_cY_e = 0 \end{cases}$$

En multipliant ces deux équations par leur dénominateur commun $D(s, t)$ (qui provient des X, Y et Z et apparaît donc aux dénominateurs des X_c, Y_c et Z_c), nous obtenons les deux premiers polynômes cherchés :

$$\begin{cases} P_1(s, t, X_e, Y_e, p_1, \dots, p_n) = 0 \\ P_2(s, t, X_e, Y_e, p_1, \dots, p_n) = 0 \end{cases}$$

Les paramètres intervenant dans le recalage (p_1, \dots, p_n) apparaissent ici à travers les transformations élémentaires qui permettent de passer des coordonnées exprimées dans le repère propre de l'objet (X, Y, Z) aux coordonnées exprimées dans le repère de la caméra (X_c, Y_c, Z_c) .

6.2.2. les contours d'occlusion

Les points qui forment un contour d'occlusion correspondent aux points de la surface où sa normale \vec{N} est perpendiculaire au vecteur vision $\vec{V} = (X_e, Y_e, f)$. Naturellement, la normale \vec{N} doit être calculée dans le repère de la caméra.

En développant la condition $\vec{V} \cdot \vec{N} = 0$, nous obtenons une fraction rationnelle dont le dénominateur est $D(s, t)^4$. Il suffit alors de

multiplier l'expression obtenue par $D(s, t)^4$ afin de constituer le troisième polynôme recherché :

$$P_3(s, t, X_e, Y_e, p_1, \dots, p_n) = 0$$

6.2.3. les contours de discontinuité de surface

Les contours de discontinuité de surface se situent à l'intersection de deux des surfaces de l'objet. Dans ce cas, c'est la condition d'intersection qui va nous fournir le troisième polynôme recherché. En effet, en multipliant les trois équations paramétriques de la deuxième surface par leur dénominateur commun, on obtient un système de trois équations polynomiales en s et t . Par une double élimination, on en déduit alors une équation cartésienne de cette deuxième surface. Il suffit ensuite d'injecter dans cette équation les coordonnées paramétriques de la première surface et de supprimer les dénominateurs du résultat. Le troisième polynôme recherché est ainsi obtenu :

$$P_3(s, t, X_e, Y_e, p_1, \dots, p_n) = 0$$

6.2.4. l'équation cartésienne du contour en projection

Finalement, à partir des trois polynômes obtenus précédemment, on peut éliminer s et t et ainsi obtenir un polynôme $C(X_e, Y_e, p_1, \dots, p_n) = 0$ qui représente l'équation cartésienne du contour en projection sur l'écran, dépendant des paramètres de recalage.

6.3. mesure de l'écart entre un point de l'image et un contour en projection

Dans cette partie, nous cherchons à dériver une mesure de l'écart entre un point de l'image et un contour du modèle à partir de l'équation implicite de ce contour en projection, obtenue précédemment. Pour simplifier les expressions, nous omettrons par la suite de préciser les paramètres de recalage. Ainsi, $C(X_e, Y_e, p_1, \dots, p_n)$ s'écrira $C(X_e, Y_e)$.

Une première idée consiste à considérer l'expression $C(X_I, Y_I)$ directement comme une mesure de distance e_I entre un point de l'image $I(X_I, Y_I)$ et le contour d'équation implicite $C(X_e, Y_e) = 0$. Nous appelons « courbe de niveau de C passant par I » l'ensemble des points (X_e, Y_e) qui vérifient $C(X_e, Y_e) = e_I$. Cette courbe de niveau correspond donc à l'ensemble des points qui se verront attribuer la même mesure de distance que I . Malheureusement, l'écartement des courbes de niveau n'est en général pas constant et de plus, les valeurs retournées pour différents contours n'ont aucune raison d'être du même ordre. Cette mesure ne se comporte donc pas précisément comme le minimum de la distance euclidienne et ceci pourrait avoir des conséquences négatives sur le recalage. En effet, certaines parties de l'objet seraient alors mieux recalées que d'autres.

Une autre mesure de distance, plus fiable mais plus longue à obtenir, est la distance perpendiculaire, construite à partir de la normale à la courbe de niveau passant par le point image I considéré (et non pas à partir de la normale au contour modèle, comme l'exigerait le minimum de la distance euclidienne). Elle s'obtient en calculant le vecteur unitaire \vec{u}_I , normal à la courbe de niveau passant par I . Ce vecteur est colinéaire au gradient de C en (X_I, Y_I) . On recherche alors un λ tel que le point $(X_I, Y_I) + \lambda \vec{u}_I$ annule l'équation cartésienne. Le λ ainsi trouvé est la distance recherchée. Ce λ , que l'on sait proche de 0, peut être calculé à l'aide d'un algorithme d'optimisation non linéaire. La figure 3 illustre la construction permettant d'obtenir cette pseudo-distance.

Nous pouvons également calculer une approximation de cette pseudo-distance en considérant que la surface $Z_e = C(X_e, Y_e)$ est localement linéaire autour du point I . Dans ce cas, son expression est $\frac{C(X_I, Y_I)}{\|\nabla C(X_I, Y_I)\|}$ et son évaluation en chaque point de contour image ne nécessite plus l'utilisation coûteuse d'un algorithme d'optimisation non linéaire. La figure 4 représente

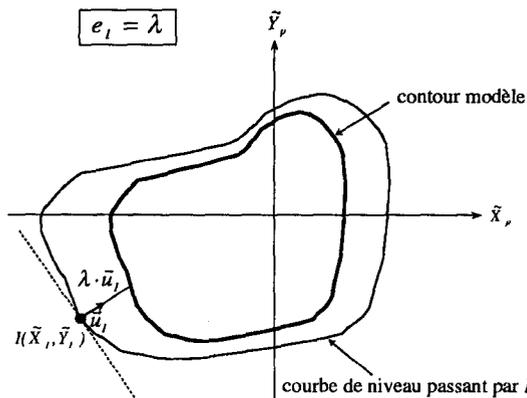


Figure 3. - Calcul approché de la pseudo distance s'appuyant sur une hypothèse de linéarité locale.

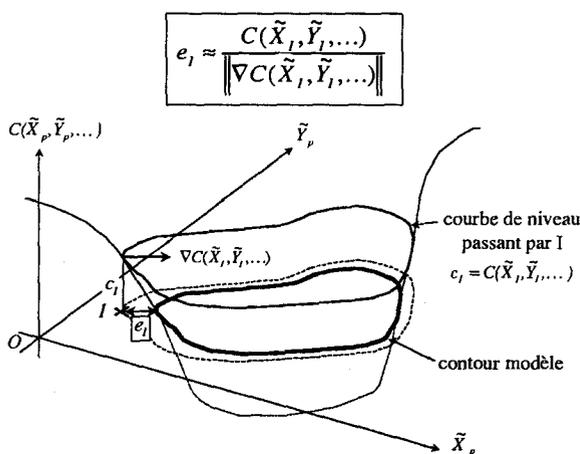


Figure 4. - Calcul approché de la pseudo distance s'appuyant sur une hypothèse de linéarité locale.

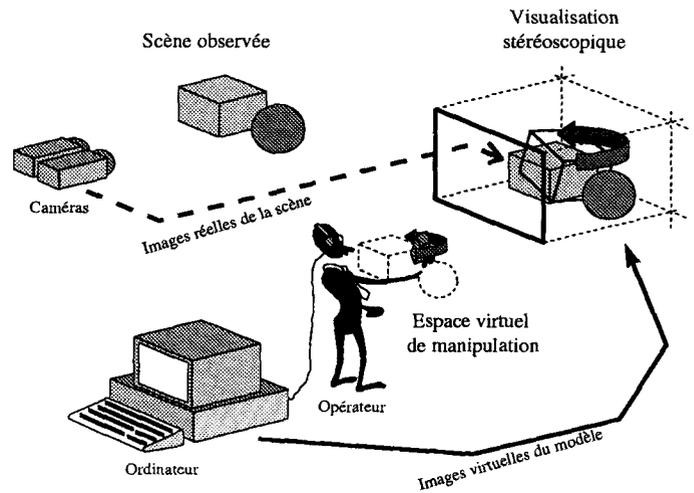


Figure 5. - Principe de fonctionnement de la modélisation interactive.

la distance mesurée. C'est celle que nous utiliserons lors de la minimisation.

La véritable distance euclidienne pourrait également être calculée, moyennant une double élimination en plus d'une recherche de racine comme pour la pseudo-distance présentée précédemment. En raison des problèmes évidents de performance que cela impliquerait et du faible intérêt pratique de cette démarche, nous laisserons de côté cette possibilité. En effet, la précision obtenue ne serait pas meilleure puisque la pseudo distance utilisée converge vers cette véritable distance euclidienne quand le modèle se rapproche de l'objet réel au cours de la minimisation.

6.4. minimisation de l'écart

Maintenant que nous disposons d'une fonction permettant d'évaluer la distance entre un point de l'écran et une courbe du modèle, nous allons pouvoir définir une mesure globale de l'écart entre l'objet réel et son modèle virtuel. Cette mesure sera simplement la somme des carrés des distances entre les sommets des contours image et les contours du modèle auxquels ils ont été appariés. Le calcul de cette mesure et de ses dérivées partielles permet alors, avec les mêmes outils que ceux employés par le recalage polyédrique, de recalcr le modèle courbe sur les images.

La différence fondamentale avec le recalage basé sur l'approximation polyédrique provient du fait que le calcul des distances et de ses dérivées partielles ne peut pas être fait numériquement car il ne dépend pas directement et simplement des matrices de transformations apparaissant dans l'arbre de construction du modèle global. Une bibliothèque de fonctions C, PARI-GP [Bat95], permettant de représenter des polynômes formels et de réaliser certaines opérations dessus (multiplication, addition, dérivation, élimination) est nécessaire. Ceci a une incidence sur les performances en temps de cette phase du recalage. Des résultats quantitatifs seront présentés dans la section 7.

7. résultats et discussion

Les algorithmes présentés ont été implémentés en C++ sur une station O2 (R10000) de Silicon Graphics.

Le test que nous donnons en exemple met en scène une pièce industrielle complexe. Il s'agit d'une tige de bol de générateur de vapeur, objet couramment impliqué dans des opérations de maintenance à EDF. Cette pièce est en partie composée de deux demis « disques » articulés (cf. figure 6). Sa position n'est pas connue (6 degrés de liberté) et l'angle entre ses deux parties non plus (1 degré de liberté). Dans cet exemple, la mise en correspondance a été rendue difficile en raison de la complexité de la scène. Les détails dans la forme, les ombres et les réflexions spéculaires produisent de nombreux contours parallèles et très rapprochés de ceux que l'algorithme tente de retrouver. Le coût en temps d'exécution pour la phase de mise en correspondance est donc élevé car certaines fausses associations peuvent facilement être considérées et testées avant les bonnes.

Les performances obtenues sont données dans les tableaux 1 et 2. La première indique les temps d'exécution des différentes étapes du recalage et la seconde présente les précisions supposées pour l'estimation des différents paramètres. La solution retenue semble s'être présentée relativement tôt dans l'exploration et le recalage n'a pas été trop long (par rapport au nombre de contours et à la complexité de la pièce).

De nombreux tests, portant sur des images réelles et sur des images de synthèse variées ont été réalisés. Les expérimentations effectuées sur les images réelles ont permis de démontrer la robustesse de l'approche dans des conditions d'utilisation normales. Dans ces tests, les positions et configurations effectives des

objets n'étaient pas connues et les précisions obtenues ont été estimées grâce à la matrice de covariance retournée par l'algorithme d'optimisation non linéaire utilisé. La validité de ces estimations dépend donc de la fidélité du modèle d'erreur considéré pour les données extraites des images. Les précisions présentées ici sont donc discutables. Cependant, les tests réalisés sur les images de synthèse, pour lesquels la précision obtenue peut directement être mesurée, ont conforté la validité de ces estimations. L'ensemble des tests réalisés confirme la généralité de l'approche vis-à-vis du nombre et de la nature des paramètres à recaler. D'autre part, les temps d'exécution sont en général compatibles avec l'application visée.

Les résultats obtenus, moyennés sur l'ensemble des tests réels, indiquent une erreur supposée d'estimation de la profondeur (erreur la plus significative) de l'ordre de 3.9 mm pour des objets placés à une profondeur d'environ 1.5 m. Cette dernière correspond donc à une erreur relative de 0.26%. Pour les translations suivant les axes du capteur, l'erreur vaut à peu près 0.6 mm. La précision obtenue pour l'estimation des angles vaut en moyenne 5.8 mrad. Dans le cas des tests réalisés sur les images de synthèse, les erreurs mesurées (et non plus supposées) sur l'estimation de la profondeur valent en moyenne 3.7 mm pour des objets placés à une profondeur de 3 m. Cette erreur correspond à 0.12% de la profondeur effective des objets. L'erreur en translation selon les axes du capteur est de 0.9 mm. L'estimation des angles vaut en moyenne 1.25 mrad.

Dans leur très grande majorité, les publications dont nous disposons ne présentent pas de résultats expérimentaux chiffrés. Seules deux d'entre elles nous permettent de comparer les résultats que nous avons obtenu avec des résultats existants. La première [Kim96] propose un système de modélisation interactive et de recalage des modèles sur les images ayant les mêmes objectifs que ceux qui ont motivés notre étude. Cependant, la modélisation interactive ne se fait pas en superposant les modèles aux images des objets correspondants mais nécessite une intervention plus concrète de l'opérateur qui doit spécifier un ensemble de correspondances point à point entre des sommets 3D du modèle virtuel et des couples de sommets 2D pris dans les images. Le problème de la mise en correspondance ne se pose donc pas et le problème du recalage s'en trouve simplifié. D'autre part, la méthode ne gère que des objets polyédriques et la calibration des caméras ne prend pas en compte la distorsion radiale. Enfin, la méthode a été développée pour la détermination des 6 paramètres de positionnement spatial et ne semble pas gérer de façon générique d'éventuelles contraintes géométriques entre les objets, ni d'éventuels paramètres de taille ou de forme pour les composants de la scène. Avec deux caméras et dans des conditions d'expérimentation similaires, les erreurs moyennes de positionnement obtenues sont de 10.7 mm pour l'estimation de la profondeur (contre 3.9 mm pour nos tests) et de 5.6 mm en translation selon les axes du capteur (contre 0.6 mm dans notre cas). Les angles sont estimés avec une précision de 10.8 mrad contre une moyenne de 5.8 mrad pour notre approche.

Tableau 1. - Recalage d'une pièce industrielle complexe.

traitements d'images (pour 158772 pixels)	1.49
mise en correspondance (avec 40 appariements)	15.2
minimisation (polyédrique ou courbe)	0.68 ou 7.53
temps d'exécution total (polyédrique ou courbe)	17.37 ou 24.22

Tableau 2. - Précision supposée dans les paramètres estimés (recalage polyédrique ou courbe).

	écart moyen (polyédrique)	écart moyen (courbe)
tx (mm)	0.69	0.66
ty (mm)	0.52	0.50
tz (mm)	5.93	3.79
α (10^{-3} rad)	2.32	2.34
β (10^{-3} rad)	2.68	2.73
γ (10^{-3} rad)	3.68	3.60
ϵ (10^{-3} rad)	10.73	10.64

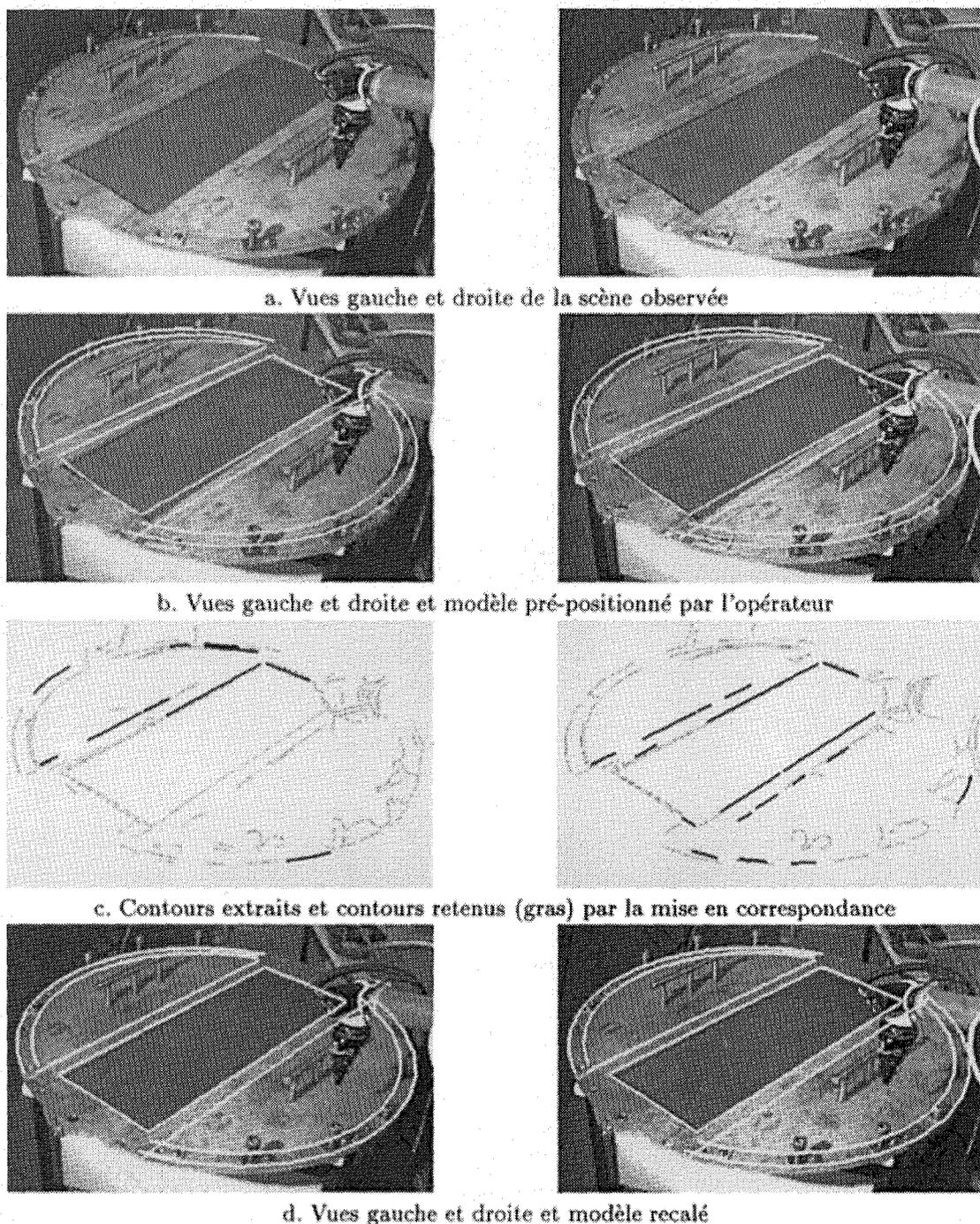


Figure 6. – Recalage d'une pièce industrielle complexe.

Le deuxième article [Coo94] présentant des résultats chiffrés propose un système de modélisation basé sur la spécification d'une boîte englobant l'objet par pointages stéréoscopiques. Des traitements d'images sont alors effectués pour extraire l'information pertinente dans la zone désignée. Dans la version présentée, seuls les cylindres et les parallélépipèdes rectangles sont gérés. D'autre part, aucune recherche effective n'est réalisée dans le cas où plusieurs indices concurrents seraient trouvés dans les images.

L'approche semble donc très peu robuste vis-à-vis des occultations éventuelles et vis-à-vis de la présence d'autres objets dans la région désignée. Il ne permet pas de spécifier d'éventuelles contraintes de positionnement entre les modèles et se contente de recalibrer les objets selon leurs 6 degrés de liberté. Toutefois, des paramètres représentant les dimensions des objets peuvent également être déterminés. Le système utilise aussi deux caméras et les conditions d'expérimentation sont du même ordre que les

nôtres. L'erreur la plus significative a été observée pour l'estimation de la profondeur avec un écart (maximal?) de 38.1 mm à comparer à une erreur moyenne supposée de 3.9 mm pour nos tests. L'estimation des dimensions des objets a d'autre part été réalisée avec une erreur (maximale?) de 6.3 mm contre 0.6 mm dans nos expérimentations.

8. conclusion et perspectives

Le travail présenté a concerné l'amélioration et la recherche de nouvelles fonctionnalités pour les postes de téléopération et plus particulièrement au niveau de leur partie logicielle.

Le principal outil du système est un algorithme de recalage automatique des modèles géométriques sur les images vidéo stéréoscopiques provenant du site distant. Il permet de localiser avec précision des objets de la scène dont les positions sont mal estimées. Cette assistance peut être utilisée lors d'une session de modélisation interactive en la rendant particulièrement conviviale et précise. Elle permet également de corriger les erreurs d'estimation des déplacements du robot ou d'objets du site tout au long de la mission. Les modèles utilisés peuvent présenter des surfaces courbes et appartenir à des classes d'objets paramétrables. Le recalage fonctionne selon deux niveaux de précision, lui permettant d'être soit approximatif et rapide, soit plus long mais précis. Enfin, le recalage produit une modélisation compatible avec d'éventuelles spécifications de contraintes géométriques entre les modèles.

Les résultats obtenus démontrent l'intérêt de l'approche qui est à la fois plus complète (mise en correspondance automatique), plus conviviale (modélisation interactive par superposition des modèles par des techniques de réalité augmentée), plus générique (classe de modèles représentables et gestion unifiée des paramètres de positionnement, de taille, de forme, de liaisons entre objets et de calibration des caméras) et plus précise que les méthodes avec lesquelles une comparaison a été possible. La méthode est d'autre part robuste vis-à-vis des éventuelles occultations et de la complexité de la scène observée.

Les perspectives de développement concernent principalement diverses optimisations dans l'algorithme de recalage, qui viseront à réduire les temps d'exécution de la phase de mise en correspondance [Mam98b]. En particulier, l'utilisation d'un algorithme de recherche des parties cachées plus complet que celui qui est utilisé actuellement permettrait de réduire l'espace de recherche et d'éviter certaines ambiguïtés inutiles dans les associations potentielles. Une amélioration non négligeable au niveau de l'exploration de l'arbre de recherche est également possible. Elle permettrait de tester les hypothèses au fur et à mesure de leur génération plutôt qu'après la construction complète (souvent inutile) de chaque descendance dans l'arbre. Enfin, étant donné la

structure arborescente de l'exploration de l'espace des correspondances, une parallélisation est d'autre part possible sans changement dans la méthode. Au niveau de l'établissement des équations implicites des contours courbes en projection, la technique d'élimination, qui constitue la partie la plus sensible du recalage basé sur les contours courbes, peut être améliorée en utilisant une autre méthode s'appuyant sur les bases de Gröbner [Hof89].

Un des intérêts d'atteindre des performances temps réel est de permettre d'appliquer l'algorithme de recalage au suivi d'objet. En effet, si un objet est en mouvement, une succession de recalages appliqués à chaque image de la séquence vidéo peut permettre de suivre ses déplacements et ses changements d'orientation, à condition que son déplacement d'une image à l'autre soit petit. La position estimée à un instant donné est simplement la position recalée, obtenue grâce à l'image précédente.

Remerciements

Ce travail est le fruit d'un partenariat entre le Centre de Robotique de l'École des Mines de Paris et la Direction des Études et Recherches du Groupe Téléopération et Robotique d'EDF. Nous tenons à remercier particulièrement Jean-Claude Bordas et Valérie Moreau.

BIBLIOGRAPHIE

- [Abr68] M. Abramovitz, «Handbook of Mathematical Functions» *Dover Publications*, Vol. 55, 1968, New York, USA.
- [Bat95] C. Batut, D. Bernardi, H. Cohen, M. Olivier, «User's Guide to PARI-GP», *Laboratoire A2X, U.M.R. 9936 du CNRS*, Université de Bordeaux I, 1995.
- [Chi92] E.W. Chionh, R.N. Goldman, «Using Multivariate Resultants to Find the Implicit Equation of a Rational Surface», *The Visual Computer*, Vol. 8, n°3, 1992, p. 171-180.
- [Coo94] C. Cooke, S. Stansfield, «Interactive Graphical Model Building using Telepresence and Virtual Reality», *IEEE?*, 1994, p. 1436-1440.
- [Der87] R. Deriche, «Using Canny's Criteria to Derive a Recursively Implemented Optimal Edge Detector», *International Journal of Computer Vision*, 1(2), 1987, p. 167-187.
- [Dho89] M. Dhome, M. Richetin, J-T. Lapresté, G. Rives, «Determination of the Attitude of 3D Objects from a Single Perspective View», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, 1989, p. 1265-1278.
- [Eve93] P. Even, R. Fournier, «Telerobotics Tasks Execution Based on 3D Geometric Modelling and Graphical Programming», *IEEE / SMC Conference*, 1993, Le Touquet, France.
- [Fuc96] P. Fuchs, «Les Interfaces de la Réalité Virtuelle», *Le Corum*, 1996, Montpellier.
- [Gri89] W.E.L. Grimson, «On the Recognition of Curved Objects», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, n°6, 1989, p. 632-643.
- [Hof89] C.M. Hoffmann, «Geometric and Solid Modeling, an Introduction», *Morgan Kaufmann Publishers*, 1989.
- [Hor93] R. Horaud, O. Monga, «Vision par Ordinateur, Outils Fondamentaux», *Editions Hermès*, 1993.
- [Kim96] W.S. Kim, «Virtual Reality Calibration and Preview/Predictive Displays for Telerobotics», *Presence*, Vol. 5, n°2, 1996, p. 173-190.

- [Kri90] D.J. Kriegman, J. Ponce, «On Recognizing and Positioning Curved 3D Objects from Image Contours», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, n°12, 1990, p. 1127-1137.
- [Low91] D.G. Lowe, «Fitting Parameterized Three-Dimensional Models to Images», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, n°5, 1991, p. 441-450.
- [Low92] D.G. Lowe, «Robust Model-Based Motion Tracking Through the Integration of Search and Estimation», *University of British Columbia (Canada), Computer Science Department, Technical Report*, 1992.
- [Mam96] D. Maman, P. Fuchs, F. Nashashibi, C. Lurgeau, «Utilisation des Techniques de la Réalité Virtuelle pour la Modélisation Interactive en Téléopération», *Congrès Informatique 96 : l'Interface des Mondes Réels et Virtuels*, 1995, Montpellier.
- [Mam98a] D. Maman, F. Nashashibi, P. Fuchs, J.-C. Bordas, «Stereo Vision Based Mixed Reality : an Assistance for Teleoperation», *SPIE International Symposium on Intelligent Systems and Advanced Manufacturing*, 1998, Boston, USA.
- [Mam98b] D. Maman, «Recalage de Modèles Tridimensionnels sur des Images Réelles : Application à la Modélisation Interactive pour la Téléopération», *Thèse de Doctorat, Ecole des Mines de Paris*, 1998.
- [Pon92] J. Ponce, S. Petitjean, D.J. Kriegman, «Computing Exact Aspect Graphs of Curved Objects : Algebraic Surfaces», *Second European Conference on Computer Vision*, p. 599-614, 1992, Santa Margherita Ligure, Italy.
- [Pre92] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery «Numerical Recipes in C. The Art of Scientific Computing», *Cambridge University Press*, 2nd edition, 1992.
- [Pyr95] «Pyramide : Modélisation Interactive d'Environnement», CEA - Aleph Technologies, 1995, Saint-Martin d'Hères, France.
- [Ric91] M. Richetin, M. Dhome, J.-T. Lapresté, G. Rives, «Inverse Perspective Transform using Zero-Curvature Contour Points : Application to the Localization of Some Generalized Cylinders from a Single View», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, 1991.
- [Zha93] Z. Zhang, «Recalage de Deux Nuages de Points 3D», *Traitement du Signal*, Vol. 10, n°4, 1993, p. 263-281.

Manuscrit reçu le 8 juin 1998.

LES AUTEURS

Didier MAMAN



Didier Maman a obtenu un doctorat en Informatique Temps Réel, Automatique et Robotique au centre de Robotique de l'École Nationale Supérieure des Mines de Paris en 1998. Les travaux réalisés au cours de sa thèse ont porté sur la modélisation interactive d'environnements et sur le recalage temps réel d'objets 3D sur des images stéréoscopiques. Ses autres centres d'intérêt concernent les techniques de la réalité virtuelle et augmentée ainsi que le suivi de mouvement en temps réel basé sur des modèles géométriques.

Philippe FUCHS



Philippe Fuchs, Professeur agrégé au Centre de Robotique de l'École des Mines de Paris, est le responsable de l'équipe de recherche «Réalité Virtuelle et Réalité Augmentée», orientée vers les applications industrielles. Il est l'auteur du livre «Les Interfaces de la Réalité Virtuelle».

Fawzi NASHASHIBI



Fawzi Nashashibi a obtenu un Doctorat de Robotique de l'Université Paul Sabatier en 1993, sur la «Perception et modélisation d'environnements tridimensionnels pour la navigation autonome d'un robot mobile». De 1990 à 1994, il a travaillé dans le Groupe Robotique et Intelligence Artificielle du LAAS-CNRS à Toulouse où il a été impliqué dans les divers projets robotiques nationaux et internationaux (EDEN, AMR, VAP and IARES). Depuis, il est chercheur au Centre de robotique à l'École Nationale Supérieure des

Mines de Paris. Ses principaux domaines de recherche sont centrés autour des Applications des technologies de la Réalité Virtuelle/Réalité Augmentée et leurs Interfaces.

Jean-Claude BORDAS

Jean-Claude Bordas est chercheur à la Direction des Études et Recherches d'Électricité de France. Il travaille dans un groupe chargé de téléopération et de robotique sur les problèmes de téléprésence et de communication entre le site opératoire et la supervision.