

NEUVIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

611



NICE du 16 au 20 MAI 1983

RECONNAISSANCE RECURSIVE ET LOCALISATION DE FORMES PLANES PARTIELLEMENT VISIBLES DANS UNE IMAGE

N. AYACHE, C. DARMON

INRIA - Domaine de Voluceau - Rocquencourt - 78153 Le Chesnay Cedex

RESUME

Sommaire

. La méthode présentée a pour but d'identifier et localiser dans l'image d'une scène, une forme plane dont on possède un modèle et qui a subi une similitude positive arbitraire (combinaison d'une rotation, translation et homothétie planes). Cette méthode a l'originalité de pouvoir s'appliquer à des formes partiellement cachées et pouvant présenter des excroissances.

. Les formes considérées sont réduites par une technique appropriée à une approximation polygonale de leurs contours. Le principe de la méthode est alors de générer et d'évaluer un certain nombre d'hypothèses.

. Une hypothèse est générée en appariant un segment du modèle avec un segment de la scène observée. Cette hypothèse est ensuite évaluée par une méthode d'estimation récursive (filtrage de Kalman) associée à un opérateur de décision qui en teste la validité (en termes de longueur de contour identifié et de la norme d'erreur entre contours homologues) et fournit une estimation a posteriori de la similitude recherchée. L'hypothèse retenue est celle donnant le meilleur score.

. La méthode a été appliquée avec succès à la reconnaissance et à la localisation de pièces de fonderie partiellement cachées et présentant des excroissances (masselottes, attaques de coulée) et s'est avérée être robuste, précise, rapide.

SUMMARY

Summary

. We introduce a new method to identify and locate a model planar form in a planar scene up to an arbitrary planar similarity transform (combination of rotation, translation and scale). The method is robust to partial occlusions.

. Planar forms are represented by polygonal approximations of their contours. Then hypotheses are successively generated and evaluated.

. Generating an hypothesis consists in matching a model-segment to a scene-segment. Each hypothesis is then evaluated by a recursive estimation method (Kalman Filtering) coupled with a decision-operator, providing a score value (in terms of the total length of identified segments and the overall error norm between matched segments) and an a posteriori estimate of the sought similarity transform. The best scored hypothesis is retained as being a solution.

. This method has been successfully applied to identify and locate severely occluded castings, and proved itself to be robust, accurate and fast.



RECONNAISSANCE RECURSIVE ET LOCALISATION DE FORMES PLANES PARTIELLEMENT VISIBLES
DANS UNE IMAGE

A - INTRODUCTION

La vision par ordinateur est un outil de perception puissant au service des robots "intelligents" de demain [1-2]. La quasi-totalité des tâches d'inspection et de manipulation automatique requièrent une connaissance a priori précise du nom et de la position des objets à traiter. Dans beaucoup de situations, cette connaissance a priori est insuffisante : c'est le cas lorsque les objets sont posés de façon arbitraire sur un support plan. Le rôle du système de vision est alors de préciser le nom des objets visibles et d'indiquer leurs positions respectives. Des programmes de décision au niveau supérieur peuvent alors tirer parti de ces informations pour contrôler des tâches d'assemblage ou/et d'inspection automatiques.

Des systèmes efficaces ont été conçus pour reconnaître et localiser des objets parfaitement isolés placés sur un fond fortement contrasté [3-5], mais les techniques employées par ces systèmes sont inadaptées dès lors que les objets sont partiellement visibles ou mal séparés. Cela se produit dans de nombreuses situations où les objets peuvent se toucher, se superposer, être partiellement hors du champ de vision, ou présenter des parties additionnelles comme les masselottes de fonderie par exemple.

Pour traiter le problème de la reconnaissance des objets partiellement cachés, différentes approches ont été explorées par de nombreux chercheurs. Elles diffèrent par le type de représentation choisi pour décrire les images et aussi par les algorithmes d'identification mis en oeuvre.

Dessimoz [6] et Perkins [7,8] calculent une corrélation croisée de la fonction intrinsèque des contours. Leurs méthodes, appliquées avec succès à des exemples concrets, requièrent néanmoins un prétraitement coûteux et ne sont pas indépendante des homothéties de l'image.

Bolles [9], Ambler [10], Hattich [11] et Rummel [12] appliquent des techniques de recherche arborescente qui nécessitent la détection d'attributs locaux tels que les trous circulaires, les coins, les segments de droite. Ces techniques sont donc limitées aux formes qui présentent des attributs de cette nature.

Ballard [16] calcule une transformation de Hough généralisée qui, bien qu'applicable à une forme quelconque, nécessite un grand volume mémoire autant que le stockage des modèles que pour la tâche de reconnaissance.

Enfin les techniques de relaxation développées par Faugeras et al. [13,14,15], bien qu'appliquées avec succès à de nombreux problèmes, semblent perdre de leur robustesse lorsque le nombre des attributs à identifier est très grand.

Dans notre approche, les silhouettes des objets sont décrites par une approximation polygonale de leurs contours. Cette approximation s'obtient facilement et peut être calculée sur toutes les formes planes. De plus elle est très compacte, et permet donc le stockage d'un grand nombre de modèles.

La technique de reconnaissance combine la génération d'un certain nombre d'hypothèses quant à la position des modèles dans la scène, et une évaluation quantitative de ces hypothèses. Un ajustement séquentiel de la position des modèles (filtrage de KALMAN) est calculé pendant le processus d'évaluation des hypothèses, et rend la reconnaissance précise et robuste [cf. 17].

Cette méthode présente l'avantage d'être peu coûteuse en capacité mémoire requise et en temps calcul. (Des structures de contrôle très simples, incluant des techniques de "Branch and Bound", sont utilisées pour accroître encore l'efficacité des calculs sur un ordinateur séquentiel). De plus la reconnaissance est in-

dépendante des occlusions partielles et des variations du facteur d'échelle (homothéties de l'image).

B - DESCRIPTION DU SYSTEME D'ANALYSE DE SCENE

I - Généralités :

L'image d'une scène quelconque est acquise par la caméra. Aucun, un ou plusieurs objets sont présents dans une position d'équilibre stable, et peuvent se toucher, se recouvrir partiellement, être partiellement hors du champ de vision. De plus les objets peuvent présenter des parties additionnelles comme les masselottes de fonderie par exemple. Le résultat de l'analyse de l'image est le nom des objets identifiés, ainsi que leurs positions dans le plan de la scène. Le schéma global de l'algorithme est le suivant.

Tout d'abord, l'image est segmentée en une approximation polygonale de ses contours. Puis chaque modèle est successivement sélectionné (les modèles sont les images segmentées de chaque objet dans chacune de ses positions d'équilibre, et sont calculés dans une phase préliminaire indépendante). La présence dans la scène d'un modèle sélectionné est testée en générant puis en évaluant un certain nombre d'hypothèses. L'hypothèse donnant le meilleur score est retenue si son score est suffisamment élevé. Dans ce cas l'objet est identifié et sa position dans l'image de la scène est donnée. Dans le cas contraire, l'objet est supposé absent (ou trop occulté), et un nouveau modèle est sélectionné.

II - Segmentation des images :

La silhouette des objets est extraite par un simple seuillage de luminance sur l'image, et un opérateur de suivi de contours est utilisé pour produire des chaînes de contours. A ce stade la représentation des objets est déjà réduite à seulement quelques milliers de points (comparés au quelque 250 000 points de l'image initiale).

Une réduction supplémentaire de la taille de la description est obtenue en représentant les contours de l'objet sous forme d'un ensemble de segments de droite (approximation polygonale). La taille de la représentation est alors réduite à quelques centaines de points (voire quelques dizaines) seulement (cf. Figure 1). L'approximation polygonale d'un contour peut être effectuée par de nombreuses techniques. L'algorithme classique de "segmentation par divisions successives" (cf. [18], pp. 176, pour une description de l'algorithme) est simple et efficace, et a été choisi pour cette application. De plus, la forme segmentée des contours tend à demeurer relativement robuste à l'occlusion et aux déplacements, c'est-à-dire que les portions visibles du contour tendent à être segmentées d'une façon similaire, indépendamment de la position de l'objet et de ses parties cachées. Enfin cette représentation est très générale puisqu'elle peut s'appliquer à toutes les formes planes (dont les silhouettes des pièces de fonderie qui ne présentent que rarement de "bons" attributs locaux tels que les coins marqués, les petits trous, les longs segments de contours linéaires, etc...)

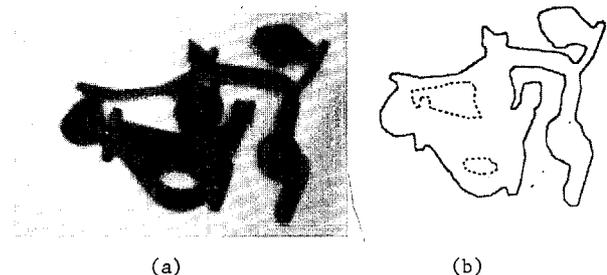


Figure 1 : (a) Scène originale (512 x 512 points) (b) Approximation polygonale des contours (150 points).

RECONNAISSANCE RECURSIVE ET LOCALISATION DE FORMES PLANES PARTIELLEMENT VISIBLES
DANS UNE IMAGE

III - Génération d'hypothèses :

Réaliser une estimation récursive des paramètres de la similitude T qui fait passer du modèle à la scène nécessite l'utilisation d'un minimum de connaissance a priori sur cette similitude (ainsi qu'on le verra au IV). En pratique, au début du traitement de la scène, on ne dispose pas de cette connaissance a priori ; par conséquent on procède par génération d'hypothèses a priori sur T, chaque hypothèse servant à l'initialisation de la phase d'EVALUATION-ESTIMATION (paragraphe IV). La phase d'évaluation-estimation calcule pour chaque hypothèse générée :

- (1) L'estimée récursive de la similitude T correspondante
- (2) Le score de qualité associé.

En fin de traitement, on ne conserve que l'hypothèse du meilleur score et la similitude estimée \hat{T} correspondante.

On pourrait croire qu'il existe une infinité d'hypothèses possibles sur T. Ce n'est pas le cas, de par la nature segmentée du modèle et de la scène. En effet, expérimentalement, à condition que le modèle ne soit pas complètement caché dans la scène, il existe au moins un segment M_0 du modèle et un segment S_{j_0} de la scène qui se correspondent (cf. figure 2.a), c'est-à-dire tels que $S_{j_0} = T(M_0)$, aux bruits près.

Donc parmi toutes les hypothèses consistant à appairier un quelconque des segments du modèle avec un quelconque des segments de la scène, il en existe une qui est la bonne : celle qui apparie M_0 et S_{j_0} . Les hypothèses à évaluer sont donc en nombre fini. Néanmoins ce nombre reste élevé et on le réduit en tenant compte d'attributs locaux mesurés sur les segments du modèle et de la scène. Dans la solution actuelle, on sélectionne comme segments privilégiés (M_0) les segments les plus longs du modèle, et comme segments (S_{j_0}) à leur appairier ceux dont la longueur est telle que $||S_{j_0}||/||M_0||$ soit compatible avec la connaissance a priori du facteur d'homothétie k (cf. figure 2-a).

IV - Evaluation des hypothèses :

Une hypothèse est caractérisée par un segment privilégié du modèle M_0 , et un segment associé de la scène S_{j_0} . Chaque hypothèse implique une position hypothétique de l'objet dans la scène. Afin d'évaluer la plausibilité de ces positions hypothétiques, le programme tire parti de l'information apportée par l'ensemble des segments du modèle autres que M_0 . Le processus d'évaluation fait appel à deux techniques combinées : une technique d'appariement des segments du modèle M_1 avec des segments de la scène S_j , et une technique de recalage récursif de la position du modèle dans la scène (Filtrage de Kalman). A sa sortie, le processus d'évaluation rend pour chaque hypothèse, un score de qualité, une liste de segments appariés (M_1, S_j) et l'estimée récursive \hat{T}_N de la similitude T définissant la position du modèle dans la scène.

IV -1- Appariement des segments du modèle avec les segments de la scène :

Pour chaque hypothèse générée, après initialisation (§ IV-3) et appariement du segment privilégié du modèle M_0 avec son correspondant S_{j_0} , le programme sélectionne parmi les segments du modèle non encore appariés, celui qui est le plus proche de M_0 . Ce segment M_1 est appliqué sur la scène par la transformation courante estimée \hat{T} . Puis une mesure de similarité est calculée entre l'image de ce segment modèle et chacun des

segments de la scène :

Ainsi si M_1 est le segment du modèle à appairier et M_1^* son image par la transformation estimée $\hat{T}(i-1)$, et si S_j est le segment de la scène que l'on désire comparer,

la mesure de similarité s s'écrit :

$$s = a \frac{\Delta M_{\max} - \Delta M}{\Delta M_{\max}} + b \frac{\Delta \Theta_{\max} - \Delta \Theta}{\Delta \Theta_{\max}} + c \frac{\Delta \ell_{\max} - \Delta \ell}{\Delta \ell_{\max}} \quad \text{si } \begin{cases} \Delta M < \Delta M_{\max} \\ \Delta \Theta < \Delta \Theta_{\max} \\ \Delta \ell < \Delta \ell_{\max} \end{cases}$$

$$s = 0 \quad \text{sinon}$$

a, b, c sont trois coefficients de pondération positifs et tels que $a+b+c=1$, ΔM , $\Delta \Theta$, $\Delta \ell$ sont respectivement une norme de l'écart entre les milieux des segments $m(M_1^*)$ et $m(S_j)$, une norme de l'écart entre les orientations des segments $\Theta(M_1^*)$ et $\Theta(S_j)$, enfin une norme de l'écart entre les longueurs des segments $||M_1^*||$ et $||S_j||$.

$$\Delta M = ||m(M_1^*) - m(S_j)||$$

$$\Delta \Theta = |\Theta(M_1^*) - \Theta(S_j)|$$

$$\Delta \ell = |||M_1^*|| - ||S_j|||$$

La norme de similarité s prend la valeur 1 maximum lorsque les segments M_1^* et S_j sont exactement superposés, et décroît lorsque la norme de l'écart entre les milieux des segments, entre leurs orientations, ou leurs longueurs, augmente. La mesure de similarité s prend une valeur nulle dès que l'un de ces écarts (en position, orientation ou longueur) dépasse en norme les valeurs de saturations prédéfinies ΔM_{\max} , $\Delta \Theta_{\max}$, $\Delta \ell_{\max}$.

Si aucune des mesures de similarité calculées entre M_1^* et chacun des segments S_j de la scène n'est strictement positive, le segment M_1 est apparié avec un segment fictif "NIL". Dans ce cas le segment est considéré comme étant occulté.

Dans le cas contraire, M_1 est apparié avec le segment S_j le plus similaire (S_j tel que $s(M_1^*, S_j)$ maximum). Ce processus est répété jusqu'à ce que tous les segments M_1 du modèle aient été examinés.

IV -2- Recalage récursif de la transformation estimée :
Description :

Chaque fois qu'un nouveau segment du modèle M_1 est identifié, (apparié avec un segment S_j non fictif de la scène), un filtre de Kalman récursif est utilisé pour optimiser les paramètres \underline{a} de la transformation estimée T, ainsi que leur matrice de covariance associée S.

Le résultat est une transformation qui tend à minimiser la distance entre les centres des segments appariés. (cf. figure 2-b et 2-c).

Mise en équations :

Après chaque appariement (M_1, S_j), on déduit de l'écart entre les milieux des segments $m(M_1^*)$ et $m(S_j)$ une information sur les quatre paramètres k, θ, t_x, t_y .

Soit \underline{a} un jeu de paramètres $[k \cos \theta, k \sin \theta, t_x, t_y]^T$. La similitude associée T transforme le point générique $\begin{bmatrix} x \\ y \end{bmatrix}$ du modèle en un point $\begin{bmatrix} x^* \\ y^* \end{bmatrix}$ dans la scène tel que :

$$\begin{bmatrix} x^* \\ y^* \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \end{bmatrix} + k R(-\theta) \begin{bmatrix} x \\ y \end{bmatrix}$$

où $R(-\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$



soit $\begin{bmatrix} x_i \\ y_i \end{bmatrix}$ les coordonnées de $m(M_i)$, et $\begin{bmatrix} x_j \\ y_j \end{bmatrix}$ celles de $m(S_j)$; dire que S_j est le segment de la scène homologue du segment M_i du modèle par la transformation T , c'est dire en particulier que :

$$\begin{bmatrix} x_j \\ y_j \end{bmatrix} = \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} \varepsilon_{x_i} \\ \varepsilon_{y_i} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} + \begin{bmatrix} a_1 & -a_2 \\ +a_2 & a_1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} \varepsilon_{x_i} \\ \varepsilon_{y_i} \end{bmatrix}$$

où $\begin{bmatrix} \varepsilon_{x_i} \\ \varepsilon_{y_i} \end{bmatrix}$ sont les bruits d'image dus à différentes cau-

ses (non-positionnement à plat de l'objet, bruits de quantification de l'image, etc...) et sont raisonnablement approximables par des bruits noncorrélés d'un point à un autre de la scène. Les coefficients a_1 et a_2 sont respectivement $a_1 = k \cos \theta$, $a_2 = k \sin \theta$.

En notant $\begin{bmatrix} x_i \\ y_i \end{bmatrix} = Y_i$, et :

$$\begin{bmatrix} x_i & -y_i & 1 & 0 \\ y_i & x_i & 0 & 1 \end{bmatrix} = C_i$$

enfin $\varepsilon_i = \begin{bmatrix} \varepsilon_{x_i} \\ \varepsilon_{y_i} \end{bmatrix}$, il vient l'équation :

$$Y_i = C_i \underline{a} + \varepsilon_i \quad \text{où} \quad E(\varepsilon_i \varepsilon_i^T) = W$$

Ainsi les deux équations en x_j' et y_j' sont des équations de mesure -au sens du filtrage de Kalman- linéaires sur le vecteur paramètre \underline{a} représentatif de la similitude T .

Egalement le vecteur-paramètre \underline{a} est solution d'une équation markovienne $\underline{a}(i) = \underline{a}(i-1)$ [c'est-à-dire que sur tous les segments M_i du modèle, on effectue le calcul d'un même et unique vecteur \underline{a} , à savoir celui représentatif de la similitude T permettant de passer du modèle à la scène].

Résolution :

A présent, grâce à l'équation d'état $\underline{a}(i) = \underline{a}(i-1)$ et à l'équation de mesure $Y_i = C_i \underline{a}(i) + \varepsilon_i$ il devient possible d'appliquer les équations d'estimation récursive par filtrage de Kalman de la similitude T sous forme d'une estimée récursive optimale $\hat{T}(i)$ raffinée à chaque itération. Les équations sont (cf. [17] par exemple) :

$$\begin{aligned} \hat{\underline{a}}(i) &= \hat{\underline{a}}(i-1) + K(i) \cdot [Y_i - C_i \hat{\underline{a}}(i-1)] \\ K(i) &= S(i-1) C_i^T [W + C_i S(i-1) C_i^T]^{-1} \\ S(i) &= [I - K(i) \cdot C_i] \cdot S(i-1) \end{aligned}$$

Ces équations expriment les valeurs nouvelles estimées $\hat{\underline{a}}(i)$ du vecteur-paramètre \underline{a} après traitement de la $i^{\text{ème}}$ mesure, ainsi que sa variance d'erreur associée $S(i) = E[(\underline{a} - \hat{\underline{a}}(i))(\underline{a} - \hat{\underline{a}}(i))^T]$ à partir de $\hat{\underline{a}}(i-1)$ et $S(i-1)$. $K(i)$ est un intermédiaire de calcul qui porte le nom de gain de Kalman.

IV -3- Initialisation :

L'initialisation de l'algorithme de recalage, sous l'hypothèse générée que M_0 et S_{j_0} sont homologues, consiste à calculer un $\hat{\underline{a}}(0)$ estimée initiale de \underline{a} et un $S(0)$ variance d'erreur associée, avant traitement de tous les segments-modèle autres que M_0 .

Pour calculer $\hat{\underline{a}}(0)$ et $S(0)$, dans la version actuelle, on considère de la façon la plus générale que l'estimée a priori de \underline{a} est nulle $\hat{\underline{a}}(-2) = 0$ et sa variance d'erreur d'estimation associée

$S(-2) = E\{[a - \hat{a}(-2)][a - \hat{a}(-2)]^T\} \approx \infty \cdot I$. Si toutefois on dispose d'une certaine connaissance a priori sur \underline{a} , on peut régler en conséquence $\hat{\underline{a}}(-2)$ et $S(-2)$.

On traite ensuite les deux extrémités du segment M_0 par l'algorithme de recalage (cf. paragraphe IV.2) en les associant avec les extrémités correspondantes du segment S_{j_0} . Ceci constitue les deux premiers pas de calcul de l'algorithme et permet de passer de $\hat{\underline{a}}(-2)$, $S(-2)$ à $\hat{\underline{a}}(0)$ et $S(0)$.

IV -4- Calcul itératif du score de qualité :

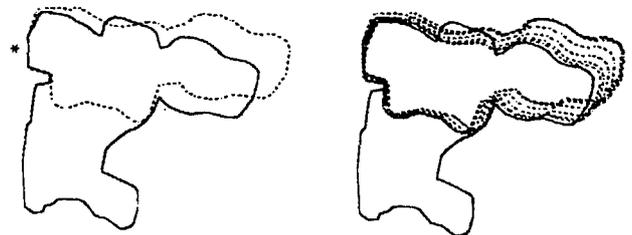
Après chaque appariement de segments homologues M_i et S_j et avant recalage, le score de qualité Q_i de l'hypothèse est incrémenté sous la forme :

$$Q_i = Q_{i-1} + s(M_i^*, S_j) \cdot \frac{|M_i|}{\sum_{i=1}^N |M_i|}$$

où $M_i^* = \hat{T}_{i-1}(M_i)$, et où \hat{T}_{i-1} est l'estimée de T avant traitement du segment M_i .

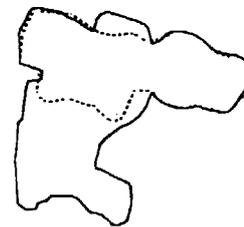
La mesure de similarité s est comprise entre 0 et 1 donc le score de qualité est borné supérieurement par le périmètre du modèle, et est atteint quand le modèle est parfaitement reconnu. D'autre part il décroît lorsque certains segments du modèle sont occultés et/ou quand l'image segmentée de l'objet présente des distortions (bruits, objet incliné, manque de robustesse de l'algorithme de segmentation...). Un cas pratique est représenté figures 2.d et 2.e.

Figure 2 : Evaluation d'une hypothèse



(a)

(b)



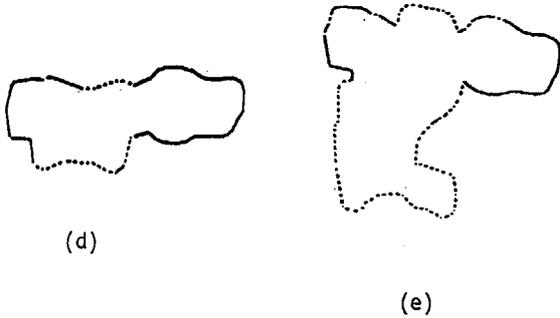
(c)

(a) Une hypothèse est générée en associant deux segments compatibles M_0 et S_{j_0} (*) entre le modèle (représenté en pointillés) et la représentation segmentée de la scène (traits pleins). Après initialisation, le modèle est superposé dans la scène par \hat{T}_0 .

(b) Les segments M_i du modèle sont appariés avec les segments S_j de la scène, en s'éloignant progressivement de M_0 . La similitude \hat{T}_i est affinée après chaque nouvel appariement.

(c) Le modèle est superposé sur la scène par la similitude \hat{T}_N estimée après traitement de tous les segments du modèle.

Figure 2 (suite)



(d,e) En traits pleins sont représentés les segments appariés entre le modèle (d) et la scène (e) à la fin du processus d'évaluation. Le score de qualité est $Q_N = 49 \%$.

V - Tests d'arrêt :

Le processus d'évaluation des hypothèses s'achève lorsque toutes les hypothèses générées ont été évaluées. L'hypothèse de meilleur score est alors conservée, ainsi que la liste des segments identifiés.

Avant de décider si le modèle est reconnu dans la scène, le programme compare ce score à un "seuil de qualité". Lorsque le seuil est dépassé, le programme décide que le modèle est identifié. Dans le cas contraire, l'hypothèse est rejetée.

Plus le seuil de qualité est élevé, plus fiables sont les détections de modèles, puisqu'elles requièrent alors la détection de longues portions des contours avec de faibles distorsions. Mais aussi les objets occultés de façon importante ou/et fortement bruités sont alors plus rarement détectés.

Par contre, un seuil de qualité plus faible autorise la détection d'objets moins visibles et/ou plus bruités mais tout en augmentant le risque des détections erronées. C'est donc en évaluant l'importance relative des fausses-détections et des faux rejets que l'utilisateur ajuste la valeur du seuil de qualité. Ce compromis entre fausses-détections et faux-rejets était prévisible puisque par nature même, le processus de la vision est d'autant moins fiable que les objets sont moins visibles.

C - REDUCTION DU TEMPS DE CALCUL

I - Structures de contrôle logicielles :

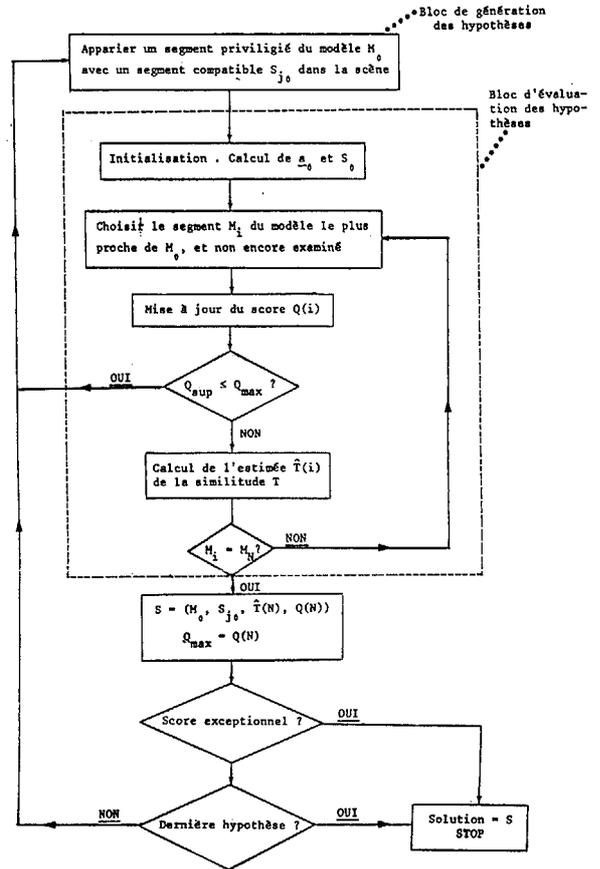
(Celles-ci sont décrites en détail dans une autre article [20].)

Le temps de calcul nécessaire à l'identification de chaque modèle est égal au produit du temps d'évaluation de chaque hypothèse, par le nombre total d'hypothèses générées. Des réductions considérables de ce temps de calcul ont été obtenues en réduisant d'une part le nombre des hypothèses prises en compte, et d'autre part en rejetant de façon précoce les hypothèses peu vraisemblables et en accélérant le processus d'appariement des segments.

Le nombre d'hypothèses générées est réduit en appariant les segments privilégiés M_0 du modèle avec des segments S_{j_0} compatibles seulement. Actuellement la compatibilité est vérifiée en termes de longueurs relatives seulement (cf. B.III). D'autres tests, comparant par exemple les orientations relatives des segments M_0 et M_{j_0} avec leurs voisins respectifs pourraient également être utilisés pour réduire le nombre des hypothèses générées. D'autre part, lorsqu'une hypothèse, obtient un score exceptionnel comme résultat

de son évaluation, le programme considère celle-ci comme une hypothèse solution. Cela permet de détecter très rapidement les objets isolés ou très peu occultés.

Figure 3 : Organigramme de l'algorithme de reconnaissance



Le temps moyen d'évaluation de chaque hypothèse est réduit principalement par deux techniques. Tout d'abord le calcul itératif du score de qualité permet de calculer à chaque itération une borne supérieure Q_{sup} sur le score de qualité final ($Q_N \leq Q_{sup}$). Lorsque cette borne supérieure Q_{sup} est inférieure au meilleur score Q_{max} déjà obtenu par une autre hypothèse, il est inutile de poursuivre l'évaluation, et une nouvelle hypothèse est générée. Enfin, le calcul des orientations des segments de la scène et leur tri préalable, permet d'apparier par un algorithme dichotomique très rapide les segments $M_i^* (= \hat{T}_{i-1}(M_i))$ avec leurs homologues S_j .

La figure 3 présente un organigramme complet de l'algorithme de reconnaissance.

II - Solution matérielle :

Comme chaque hypothèse peut être générée et évaluée de façon indépendante, l'algorithme peut être implémenté directement sur des unités de calcul parallèles. Dans ce cas les temps de calcul seraient divisés (approximativement) par le nombre des unités de calcul utilisées.

D - RESULTATS EXPERIMENTAUX

Ce système a été appliqué avec succès à de nombreux exemples concrets. On ne présente ici que deux exemples typiques.



Le premier exemple concerne la reconnaissance dans l'image de la figure 1 d'une pièce de fonderie dont le modèle est représenté par la figure 4.a.

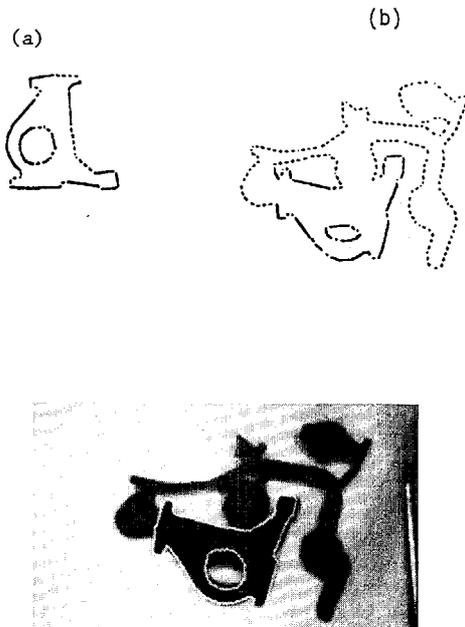
Le résultat de la reconnaissance est une hypothèse de meilleur score $Q_{\max} = 47\%$ suffisamment élevé pour être retenue.

Les segments identifiés entre les représentations segmentées du modèle et de la scène sont représentés en trait pleins dans les figures 4.a et 4.b.

Enfin le modèle (trait blanc) est superposé dans la scène par l'estimée a posteriori \hat{T}_N de la similitude de T (figure 4.c).

Le temps de calcul requis pour le processus d'analyse est de moins de deux secondes. (Les programmes sont écrits en Fortran et fonctionnent sur un mini ordinateur Perkin Elmer 32-44.

Figure 4 : Reconnaissance d'une pièce de fonderie présentant des masselottes



Le second exemple illustre la capacité du système à reconnaître et localiser plusieurs modèles dans une scène complexe (figure 5.a) comportant plusieurs pièces en vrac planaire.

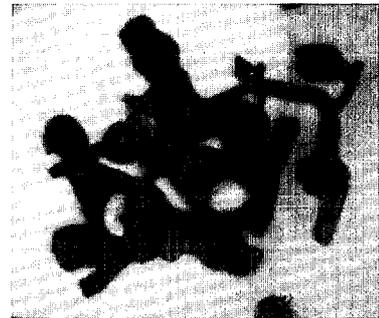
Le programme identifie et localise correctement chacun des quatre modèles testés. Chacun des modèles est alors superposé dans l'image de la scène par l'estimée a posteriori \hat{T}_N de la similitude T lui correspondant (figure 5.b). Les scores de qualité obtenus par chaque modèle sont respectivement (dans le sens trigonométrique indirect en partant du haut) 42 %, 34 %, 36 % et 29 %. La représentation segmentée de la scène contient près de 300 segments, et les temps de calcul varient entre 3,5 et 12 secondes par modèle en fonction du modèle testé.

Figure 5 : Analyse d'une scène comportant quatre pièces automobiles en vrac planaire

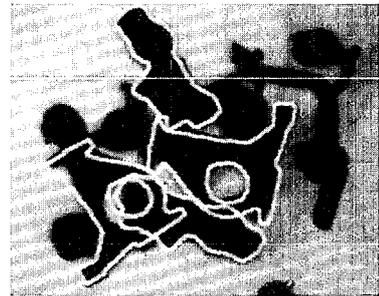
(a) image originale

(b) résultat du processus d'analyse

(a)



(b)



E - CONCLUSION

Nous avons décrit une nouvelle méthode pour identifier et localiser des objets de forme connue dans des scènes planaires. L'originalité de la méthode réside dans son large champ d'application (toutes les formes planes) et dans sa robustesse aux occlusions partielles, aux changements d'échelle, et aux distorsions géométriques. Enfin l'algorithme de reconnaissance est rapide.

Le système de reconnaissance utilise une représentation segmentée des objets à identifier et de la scène à analyser. Un certain nombre d'hypothèses sont générées puis évaluées. Une hypothèse est générée en associant un segment du modèle avec un segment de la scène observée. Cette hypothèse est ensuite évaluée par une méthode d'estimation récursive (filtrage de Kalman) associé à un opérateur de décision qui en teste la validité (en termes de longueur de contour identifié et de la norme d'erreur entre contours homologues) et fournit l'estimée a posteriori de la similitude recherchée. L'hypothèse retenue est celle donnant le meilleur score.

La méthode a été appliquée avec succès à la reconnaissance et à la localisation de pièces de fonderie partiellement cachées et présentant des excroissances (masselottes, attaques de coulée) et s'est avérée être robuste, précise, rapide. Nos travaux s'orientent maintenant vers le couplage du système de vision avec un manipulateur industriel programmable, afin d'étendre les résultats de l'analyse de la scène à la saisie automatique des objets. [21]

F - REMERCIEMENTS

Les auteurs remercient Mlle Nathalie ROCHER pour l'excellente frappe de cet article.



RECONNAISSANCE RECURSIVE ET LOCALISATION DE FORMES PLANES PARTIELLEMENT VISIBLES
DANS UNE IMAGE

G - REFERENCES

- (1) O.D. FAUGERAS et al. "Toward a flexible vision System", in ROBOT VISION, Ed. A. PUGH, IFS Publ., Springer Verlag 1983
- (1a) C.A. ROSEN "Machine Vision and Robotics : Industrial Requirements", International Symposium on Computer Vision and Sensor Based Robots, General Motors Research Laboratories, Warren, Michigan, Oct 1978.
- (2) R.P. KRUGER and W.B. THOMPSON "A Technical and Economic Assessment of Computer Vision for Industrial Inspection and Robotic Assembly", proc. of the IEEE, vol. 69, number 12, Dec 1981.
- (3) G.J. GLEASON and J. AGIN "A Modular Vision System for Sensor-controlled Manipulation and Inspection", Proc. of the 9th Int. Symp. On Indust. Robots, Washington, D.C., 1978, pp 57-70.
- (4) S.W. HOLLAND et al. "Consight 1 : A Vision-Controlled Robot System for Transferring Parts from Belt Conveyors, International Symposium on Computer Vision and Sensor Based Robots, General Motors Research Laboratories, Warren, Michigan, Oct 1978.
- (5) B. CARLISLE, S. ROTH, J. GLEASON, D. McGHIE "The Puma/VS-100 Robot Vision System", Robot Vision, Alan Pugh editor, IFS Publications, 1983.
- (6) J.D. DESSIMOZ et al. "Recognition and Handling of Overlapping Industrial Parts", International Symposium on Computer Vision and Sensor Based Robots, General Motors Research Laboratories, Warren, Michigan, Oct 1978.
- (7) W.A. PERKINS "A Model-Based Vision System for Industrial Parts", IEEE Transactions on Computer, Vol C27, No 2, Feb 1978.
- (8) W.A. PERKINS "Simplified Model-Based Part Locator", 5th International Conf. on Pattern Recognition, Vol 1, Miami Beach, Florida, Dec 1980.
- (9) R.C. BOLLES and R.A. CAIN "Recognizing and Locating Partially Visible Workpieces", Proceedings of the Conference on Pattern Recognition and Image Processing, June 1982, Las Vegas, Nevada.
- (10) A.P. AMBLER et al. "A versatile Computer Controlled Assembly System", proc. I.J.C.A.I, 1973, Stanford, California, pp 298-307.
- (11) W. HATTICH et al. "Experience with two Hybrid Systems for the Recognition of Overlapping Workpieces", proc. of the 6th Int. Conf. on Pattern Recognition, Munich, Sept 1982.
- (12) P. RUMMEL and W. BEUTEL "A Model-Based Image Analysis System for Workpiece Recognition", Proc. of the 6th Int. Conf. on Pattern Recognition, Munich, Sept 1982
- (13) O.D. FAUGERAS and B. BHANU "Recognition of Occluded Two Dimensional Objects", the Second Scandinavian Conf. on Image Analysis, Helsinki, Finland, June 15-17, 1981, pp 72-77.
- (14) N.J. AYACHE and O.D. FAUGERAS "Recognition of Partially Visible Planar Shapes", proc. of the 6th Int. Conf. on Pattern Recognition, Munich, Sept 1982.
- (15) O.D. FAUGERAS and M. BERTHOD "Improving Consistency and Reducing Ambiguity in Stochastic Labeling : an Optimization Approach", IEEE Trans. on Pattern Analysis and Machine Intelligence, Volume 3, No 4, July 1981.
- (16) D.H. BALLARD "Generalizing the Hough Transform to Detect Arbitrary Shapes", Pattern Recognition, 13, Vol. 2, pp 111-122, 1981.
- (17) C.A. DARMON "A New Recursive Method to Detect Moving Objects in a Sequence of Images", Proceedings of the Conference on Pattern Recognition and Image Processing, June 1982, Las Vegas, Nevada.
- (18) T. PAVLIDIS "Structural Pattern Recognition", New York : Springer-Verlag, 1977.
- (19) J.SERRA "Image Analysis and Mathematical Morphology, Academic Press, New York, 1982.
- (20) N.J. AYACHE "A Model Based Vision System to identify and locate partially visible planar shapes", to appear.
- (21) J.D. BOISSONNAT "Stable Matching Between a Hand Structure and an Object Silhouette", IEEE Trans. on PAMI, vol 4, No 6, November 1982.

