



HAL
open science

Localisation et reconnaissance d'objets 3D en robotique

Hugues Henocq, Gilles Burel, Christophe Venaille

► **To cite this version:**

Hugues Henocq, Gilles Burel, Christophe Venaille. Localisation et reconnaissance d'objets 3D en robotique. Journée d'étude SEE, Société des électriciens et des électroniciens, Apr 1993, Gif-sur-Yvette, France. hal-04053855

HAL Id: hal-04053855

<https://hal.univ-brest.fr/hal-04053855>

Submitted on 1 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

LOCALISATION ET RECONNAISSANCE D'OBJETS 3D EN ROBOTIQUE

Journée d'étude de la SEE (Société des Electriciens et des
Electroniciens), Gif-Sur-Yvette, France, 2 avril 1993

Hugues HENOCQ (*)(**)
Gilles BUREL (*)
Christophe VENAILLE (*)

(*) Thomson CSF-LER, Avenue de Belle Fontaine,
F-35510 CESSON-SEVIGNE

tél: 99.25.42.00 fax: 99.25.43.34

(**) L.E.S.T., URA CNRS 1329, 6 Avenue Le Gorgeu,
F-29200 BREST

RESUME:

Dans cet article, nous proposons une solution au problème de la reconnaissance d'objets 3D placés aléatoirement sous un robot 6 axes, par des techniques de vision. Les sous-tâches à résoudre sont la calibration du robot et de la caméra, la localisation des objets situés dans l'espace de travail du robot, la navigation du robot vers ces objets pour un examen approfondi et enfin la reconnaissance et le recalage de ces pièces par rapport à un modèle connu. Dans cette communication, nous aborderons plus particulièrement le problème de la calibration dynamique de la caméra ainsi que les aspects localisation / reconnaissance des objets.

1 INTRODUCTION

Dans le cadre d'une étude générale sur les problèmes liés à la vision associée à la robotique, nous avons été amené à développer des méthodes de calibration du robot et de la caméra, de localisation d'objets placés dans l'espace de travail, de navigation du robot et enfin de reconnaissance et de recalage de ces pièces par rapport à un modèle connu. Les différentes solutions adoptées sont validées à l'aide de démonstrations utilisant comme support une caméra montée sur le dernier axe d'un bras manipulateur. Actuellement, une démonstration met en œuvre les méthodes de calibration dynamique, localisation et reconnaissance d'objets 3D exposées dans cet article. Cette démonstration consiste en une première phase de détection et de localisation d'objets 3D posés aléatoirement sur une table, utilisant les systèmes de calibration dynamique et

de localisation. Dans une seconde phase, le robot s'approche et observe, les uns après les autres, les objets localisés, sous éventuellement plusieurs angles de vues, afin de les identifier en utilisant la méthode de reconnaissance d'objets 3D.

La calibration dynamique consiste à estimer, quel que soit le mouvement du robot, la loi de projection du monde 3D sur le plan image 2D de la caméra. De nombreuses méthodes ont été proposées pour résoudre ce problème lorsque la caméra est fixe (calibration statique). Elles utilisent l'observation de la projection 2D dans l'image de points 3D de coordonnées connues. Si la caméra est déplacée en une nouvelle position, la calibration précédemment calculée n'est plus valide. Nous développons ici une solution, que nous appellerons méthode de calibration par transformée inverse. L'originalité de la méthode est de calibrer la caméra par rapport à une position pour laquelle on dispose d'une matrice de calibration statique. Cette approche, par rapport à la méthode habituelle (cf [10], [11]), permet de s'affranchir du calcul explicite des paramètres intrinsèques et remplace l'estimation de la matrice de passage entre le repère de la caméra et le repère de son support par l'estimation de la matrice de passage entre le repère absolu et le repère dans lequel est défini le mouvement du robot (cf figures 1, 2).

Disposant d'une matrice de calibration pour toutes positions du robot, la localisation et l'approche des objets devient possible. La solution adoptée consiste à calculer l'équation de la droite joignant le barycentre de l'objet au centre optique de la caméra à l'aide de la matrice de calibration et connaissant l'équation du plan de la table sur lequel sont posés les objets, il est alors possible d'estimer leur position 3D.

Pour l'identification d'objets 3D, nous avons mis en œuvre une méthode simple et rapide ne nécessitant pas l'emploi de techniques à base de mise en correspondances de modèles. Les capacités d'apprentissage des réseaux de neurones nous ont amené à penser qu'un tel réseau pourrait, par apprentissage d'un certain nombre de vues de chaque objet possible, reconnaître ces objets pour des vues différentes. Afin de diminuer le nombre des données à traiter, l'objet sera caractérisé par sa silhouette.

Après description des différentes étapes de notre système de localisation et reconnaissance, des résultats obtenus sur les performances de la méthode de calibration et de reconnaissance sont présentés.

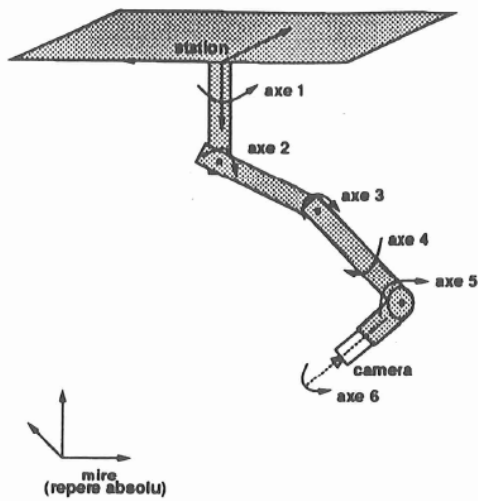


Figure 1: La caméra associée au robot 6 axes

2 CALIBRATION DYNAMIQUE

La plupart des études menées en calibration dynamique se placent dans une configuration où la caméra est fixe par rapport à la scène et permettent ce que nous appellerons une "calibration initiale" de la caméra. Les deux catégories de méthodes connues pour résoudre ce problème sont la méthode multi-plans [12] [13] [14] avec interpolation linéaire, quadratique ou spline linéaire et la méthode basée sur une modélisation de la caméra, soit linéaire (modèle sténopé), soit non linéaire avec prise en compte d'un coefficient de distorsion radiale [10] [11] [15] ou quadratique [16]. Dans le cas du modèle sténopé pour la caméra, on montre par exemple que la relation entre un point 3D P et sa projection 2D U , exprimés en coordonnées homogènes, s'écrit dans le repère image:

$$U = M.P$$

où M est la matrice de calibration. Cette matrice dépend des paramètres intrinsèques de la caméra et des paramètres extrinsèques qui sont la rotation et la translation reliant le repère associé à la caméra au repère de la scène.

Le problème suivant concerne la calibration d'une caméra en mouvement montée sur le dernier axe d'un bras manipulateur. Les méthodes existantes cherchent à estimer les paramètres intrinsèques de la caméra après calibration de celle-ci dans une position fixe, puis à évaluer la matrice de passage entre le repère de la caméra et le repère lié à son support (M_{pc}). Nous proposons une solution différente qui permet de s'affranchir de la connaissance explicite des paramètres intrinsèques et de remplacer l'estimation de la transformation 3D M_{pc} précédente par l'estimation directe de la transformation 3D en-

tre le repère mire (absolu) et le repère station dans lequel est défini le mouvement du poignet (M_{sm}) (cf figure 2). La méthode suppose que le mouvement du support de la caméra est connu (fourni par le transformateur de coordonnées du robot) et que l'on dispose d'une calibration de la caméra dans une position initiale.

Le principe de la transformée inverse est basée sur l'idée que le mouvement d'une caméra autour d'un point fixe est équivalent au mouvement inverse du point autour d'une caméra fixe.

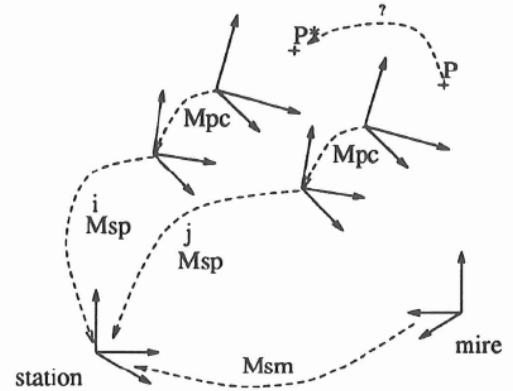


Figure 2: Représentation des repères pour 2 positions du poignet

2.1 Calibration initiale de la caméra

Un point 3D de coordonnées $P_m(X_m, Y_m, Z_m)$ dans le repère mire se projette dans le plan image de la caméra en $U(u, v)$ par l'intermédiaire de la matrice de calibration notée M_c tel que $U = M_c.P_m$.

Dans le cas du modèle sténopé (cf figure 3), on peut supposer que la matrice de calibration se décompose en une matrice de projection perspective (3x4) K référencée dans le repère caméra et d'une matrice (4x4) M_{cm} de rotation (R_{cm}) -translation (T_{cm}) reliant le repère mire au repère caméra, tel que $M_c = K.M_{cm}$ avec:

$$K = \begin{pmatrix} \frac{f}{\Delta_{xim}} & 0 & u_0 & 0 \\ 0 & \frac{f}{\Delta_{yim}} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$M_{cm} = \begin{pmatrix} R_{cm(3,3)} & T_{cm(3,1)} \\ 0_{(1,3)} & 1 \end{pmatrix}$$

$\Delta_{xim}, \Delta_{yim}$ étant respectivement la hauteur et la largeur d'un pixel et f la distance focale.

La matrice est estimée à partir de l'observation de la projection dans l'image de points 3D de coordonnées connus dans le repère mire. Pour le détail des calculs, se référer à (cf [10]).

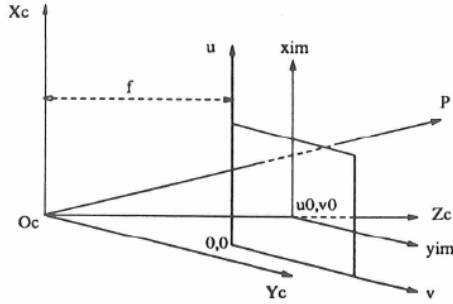


Figure 3: Modèle sténopé

2.2 Calibration et mouvement

2.2.1 Formulation du problème:

Connaissant la matrice de calibration M_c^i pour une position i de la caméra, comment en déduire la matrice de calibration M_c^j pour une position j différente et la relation \mathcal{F} telle que (cf figure 2):

$$M_c^j = \mathcal{F}(M_c^i)$$

En notant M_{sm} la matrice de passage entre le repère mire et le repère station, M_{cm}^i la matrice de passage entre le repère mire et le repère caméra pour une position i et M_{sp}^i la matrice de passage entre le repère poignet et le repère station pour une position i , on peut écrire pour 2 positions i et j :

$$M_{sm} = M_{sp}^i \cdot M_{pc} \cdot M_{cm}^i \quad (1)$$

$$M_{sm} = M_{sp}^j \cdot M_{pc} \cdot M_{cm}^j \quad (2)$$

$$(1) \rightarrow M_{pc} = M_{sp}^{i-1} \cdot M_{sm} \cdot M_{cm}^{i-1} \quad (3)$$

$$(2).(3) \rightarrow M_{cm}^j = M_{cm}^i \cdot M_{sm}^{-1} \cdot M_{sp}^i \cdot M_{sp}^{j-1} \cdot M_{sm} \quad (4)$$

$$M_c = K \cdot M_{cm} \quad (5)$$

$$(4).(5) \rightarrow M_c^j = M_c^i \cdot M_{sm}^{-1} \cdot M_{sp}^i \cdot M_{sp}^{j-1} \cdot M_{sm} \quad (6)$$

Cette dernière équation permet, à l'aide de plusieurs matrices de calibration obtenues en sta-

tique, d'estimer la matrice de transformation 3D M_{sm} [18].

3 LOCALISATION

3.1 Introduction

Pour la phase de localisation, l'environnement du robot est supposé assez bien connu. On sait, par exemple, où se trouve la zone d'intérêt (zone de travail, d'inspection ou de surveillance selon les applications). Dans un premier temps, le bras manipulateur est suffisamment éloigné de la scène pour pouvoir la scruter entièrement en nombre faible de déplacements et d'acquisitions (dans la démonstration décrite ce nombre est égal à 6). Les objets sont détectés par segmentation de chaque image (il peut y en avoir plusieurs par image).

La localisation consiste à calculer une estimation de la position 3D de tous les objets présents dans la zone de travail et détectés dans au moins une des images acquises en position éloignée. Le but de la localisation est de pouvoir s'approcher individuellement vers chaque objet, de positionner celui-ci au centre de l'image, éventuellement de tourner autour de celui-ci pour acquérir des vues multiples; tout ceci afin de faciliter la tâche de reconnaissance.

Deux méthodes ont été proposées pour cette phase de localisation. Toutes deux s'appuient sur une détermination préliminaire de l'équation de la droite 3D reliant le centre optique de la caméra à l'objet (cf figure 4).

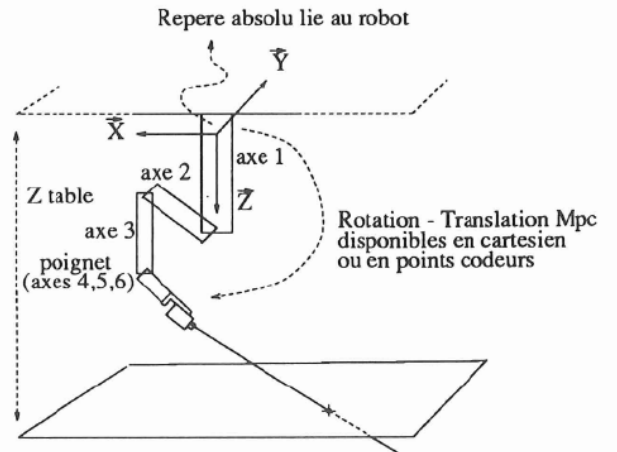


Figure 4: Localisation

3.2 Calcul de la droite 3D de visée

Connaissant la matrice de calibration M_c^i de la caméra lorsque le robot est en position i , et les coordonnées $g(u_i, v_i)$ du centre de gravité de l'un des objets segmenté dans l'image acquise dans cette position, on montre aisément qu'une équation

paramétrique de la droite 3D passant par le centre optique C_i et issue de g s'écrit dans le repère mire:

$$P = C_i + \lambda \vec{N}_i$$

$$\text{avec } \vec{N}_i = u_i \vec{r}_i + v_i \vec{s}_i + \vec{t}_i \quad (8)$$

où $\vec{r}_i, \vec{s}_i, \vec{t}_i$ sont définis à partir de M_c^i comme étant les produits vectoriels des vecteurs lignes $\vec{i}, \vec{j}, \vec{k}$ de la sous-matrice (3x3) issue de M_c^i par suppression de la dernière colonne:

$$\begin{aligned} \vec{r}_i &= (\vec{j} \wedge \vec{k}) \\ \vec{s}_i &= (\vec{k} \wedge \vec{i}) \\ \vec{t}_i &= (\vec{i} \wedge \vec{j}) \end{aligned}$$

L'expression (8) peut donc se réécrire:

$$\vec{N}_i = E_i \cdot \begin{pmatrix} u_i \\ v_i \\ 1 \end{pmatrix}$$

où la matrice (3x3) E_i a pour vecteurs colonnes, les vecteurs $\vec{r}_i, \vec{s}_i, \vec{t}_i$. Le calcul des coordonnées (X_c, Y_c, Z_c) du centre optique C_i s'effectue aisément en résolvant l'équation :

$$M_c^i \cdot \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Pour un objet détecté dans le plan image lorsque le robot est en position i , on peut donc lui associer une droite 3D le long de laquelle se trouve l'objet.

Si on connaît a priori une équation du plan sur lequel se trouve l'objet, on obtient une estimée de sa position 3D en calculant l'intersection de la droite avec ce plan.

Si aucune hypothèse ne peut être faite sur la disposition des objets dans la scène, le problème peut être résolu par stéréovision en choisissant de déplacer le robot dans une autre position proche.

4 RECONNAISSANCE D'OBJETS 3D

Dans la littérature, plusieurs approches ont été proposées pour la reconnaissance d'objets 3D. Toutes les méthodes sont basées sur la comparaison avec des modèles des objets à reconnaître [2]. On peut en première approximation les classer en 2 grandes catégories, les méthodes nécessitant une reconstruction 3D et les méthodes basées sur l'analyse de projections 2D.

Les 2 types de méthodes nécessitent la création préalable de modèles explicites des objets. Ceci explique probablement le fait que les solutions proposées sont généralement limitées aux objets polygonaux [2] [4]. Les modèles polygonaux conduisent

à manipuler des segments, alors que des modèles plus généraux nécessiteraient la gestion de primitives complexes.

Les capacités d'apprentissage des réseaux de neurones pourraient permettre d'éviter la phase de modélisation explicite. En effet, l'apprentissage permet de réaliser automatiquement une modélisation implicite des objets à reconnaître. Ceci est une propriété intéressante qui permettrait de réduire le coût de mise au point d'un système. L'idée consiste à apprendre à un réseau de neurones un certain nombre de vues de chaque objet. En fin d'apprentissage, il devrait être capable de reconnaître des vues non apprises.

Nous avons spécifié et expérimenté une approche qui est basée sur la reconnaissance de l'objet à partir de sa silhouette (contour externe). La silhouette peut être entièrement décrite par un ensemble de coefficients complexes connus sous le nom de "descripteurs de Fourier" [5]. Les descripteurs de Fourier portent la même information que le contour original et ont pour avantages de constituer une représentation multi-résolution de la silhouette et il est possible de les normaliser pour les rendre invariants par rapport aux transformations de translation, rotation, et homothétie dans le plan, ainsi que par rapport à un décalage du point de départ sur le contour.

5 LES DESCRIPTEURS DE FOURIER

5.1 Introduction

La théorie des descripteurs de Fourier est une méthode de décomposition d'une forme en série de Fourier [5] [7]. Il s'agit d'une transformation qui présente des propriétés intéressantes: les opérations de translation, rotation, dilatation et permutation circulaire des indices s'expriment de façon très simple dans le domaine transformé.

Etant donné le contour (dans le sens trigonométrique) d'un objet sous forme d'une liste de coordonnées de pixels, nous construisons un contour de vélocité uniforme de N points équidistants en re-échantillonnant le contour original. Sur le contour original, la distance entre deux pixels 4-connexes est 1.0 et la distance entre deux pixels 8-connexes est $\sqrt{2}$. Après re-échantillonnage, toutes les distances entre points successifs sont identiques.

Notons U_m les points du contour re-échantillonné (ils peuvent être vus comme des vecteurs dans le plan complexe). Les descripteurs de Fourier sont définis

par:

$$C_n = \frac{1}{N} \sum_{m=0}^{N-1} U_m e^{-j2\pi \frac{nm}{N}}$$

Il y a N descripteurs de Fourier, et les indices peuvent être vus modulo N (car $C_n = C_{n \bmod N}$). La transformation inverse est donnée par :

$$U_m = \sum_{n=0}^{N-1} C_n e^{j2\pi \frac{nm}{N}}$$

5.2 Normalisation

Les transformations géométriques de base (translation, rotation, homothétie) et le décalage du point de départ sur le contour agissent sur les descripteurs de Fourier comme suit:

Translation	$U'_m = U_m + T$	$C'_0 = C_0 + T$
Rotation	$U'_m = U_m e^{j\phi}$	$C'_n = C_n e^{j\phi}$
Homothétie	$U'_m = \rho U_m$	$C'_n = \rho C_n$
Décalage	$U'_m = U_{(m+k) \bmod N}$	$C'_n = C_n e^{2j\pi \frac{kn}{N}}$

En conséquence, pour être invariant par rapport à ces transformations, il suffit de ne pas prendre en compte C_0 (invariance en translation) et de modifier les descripteurs de la manière suivante, en appelant ϕ_2 la phase de $(\frac{C_2}{C_1})$ [19]:

$$P_n = \frac{C_n}{C_1} e^{j(1-n)\phi_2}$$

La figure 5 montre la reconstruction du contour d'un avion en fonction du nombre de coefficients conservés. On constatera que les coefficients d'ordre inférieur à 16 sont très largement suffisants pour reconstruire la forme.

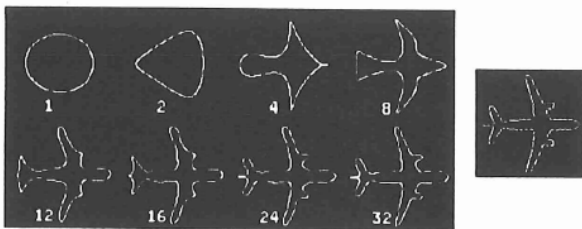


Figure 5: Reconstruction d'un avion (a320) en fonction du nombre de coefficients conservés et contour initial

6 LE RÉSEAU DE NEURONES

6.1 Structure du réseau

Nous avons utilisé un perceptron multi-couches (fig 6), entraîné par l'algorithme de rétropropagation

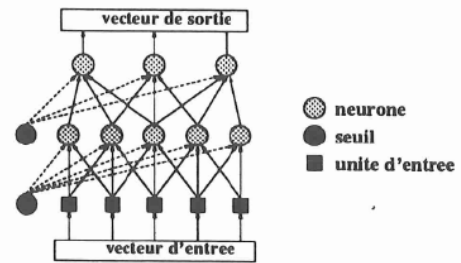


Figure 6: Perceptron multi-couches

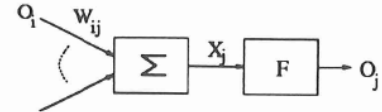


Figure 7: modèle du neurone

[8], auquel nous avons apporté diverses améliorations [1]. Un neurone (fig 7) est un automate réalisant une sommation pondérée suivie d'une fonction non-linéaire en tangente hyperbolique: $O_j = F(X_j) = th(X_j)$.

Le réseau est entièrement connecté entre 2 couches successives. Sur chaque couche, sauf sur la couche de sortie, on a disposé un neurone sans entrées, et dont la sortie est toujours égale à 1. Ce neurone, nommé "neurone seuil", est indispensable aux bonnes performances du réseau. La couche d'entrée est alimentée avec les valeurs des descripteurs de Fourier normalisés d'ordre inférieur à 16. Par convention, dans la suite de cet article un réseau à n couches désigne un réseau comprenant une couche d'entrée, une couche de sortie et $(n-2)$ couches cachées.

6.2 Représentation des entrées

Trois possibilités pour représenter les descripteurs de Fourier en entrée du réseau de neurones ont été envisagées, la représentation sous forme cartésienne $(Re\{P_n\}, Im\{P_n\})$, sous forme polaire $(Module\{P_n\}, cos(phase\{P_n\}), sin(phase\{P_n\}))$ ou uniquement par leur module $(Module\{P_n\})$. Des tests présentés dans un article précédent (cf [9]) nous ont permis de faire ressortir que la représentation en module pour des réseaux 3 et 4 couches donne les meilleurs résultats.

7.1 Mise en œuvre

Nous disposons d'un robot manipulateur 6 axes GT6A, distribué par GT productique, équipé d'une caméra montée sur le dernier axe du robot. Ce banc robotique a pour objectif à terme d'être capable d'analyser des scènes tri-dimensionnelles complexes. L'intégration sur ce robot de la méthode présentée est un premier pas vers cet objectif.

7.1.1 Constitution d'une base d'exemples

La constitution d'une base d'apprentissage est réalisée par l'acquisition préalable d'un grand nombre de vues 2D de chaque objet. Chacune de ces vues représente la projection d'un objet pour une élévation et un azimut différents.

Cette base d'images, après extraction des silhouettes, permettra la réalisation d'exemples en vue d'un apprentissage de réseau. L'extraction de la silhouette de l'objet est réalisée par l'intermédiaire d'un seuillage adapté complété par un suivi de contour. Le calcul des descripteurs de Fourier sur ce contour nous renvoie enfin le vecteur descriptif de cet objet pour un angle de vue donné (cf figure 8).

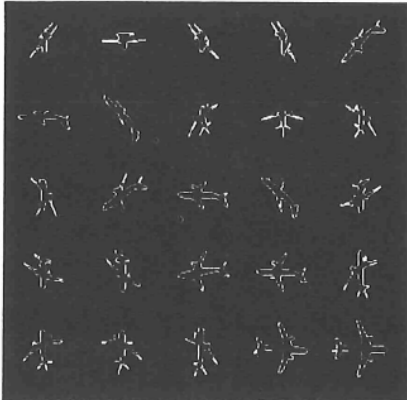


Figure 8: Contours extrait d'un avion (a320)

7.1.2 Analyse en composantes principales

En considérant les exemples comme vecteurs d'attributs, on peut alors envisager de réduire la taille de ces vecteurs en réalisant une analyse en composantes principales (ACP). Ceci permet de réduire la complexité du réseau et donc d'accélérer l'apprentissage.

L'ACP consiste en la recherche des directions privilégiées d'énergie, ce qui revient à décorréler les différents attributs au moyen d'équations linéaires. Ceci est équivalent à la recherche d'un changement de base pour les attributs, conduisant à une matrice

de covariance de type diagonale.

Le changement de base conduit à prendre pour axes de la nouvelle base les vecteurs propres de la matrice de covariance des attributs. Les valeurs propres correspondent à l'énergie distribuée dans la direction du vecteur propre associé.

En se fixant une certaine énergie à conserver, par exemple 98% de l'énergie totale, ce qui revient à ne conserver pour chaque vecteur d'attributs dans cette nouvelle base que les composantes des projections sur les vecteurs propres dont la somme des valeurs propres associées est juste supérieure à 98% de l'énergie totale, on peut espérer une nette réduction de la taille de ce vecteur.

7.2 Résultats expérimentaux

7.2.1 Résultats en calibration dynamique

On note $\hat{(\)}$ les paramètres estimés à distinguer des paramètres exacts.

La solution développée dans l'article (méthode inverse) est comparée à une méthode classique (méthode directe) dans les cas où les matrices de calibrations estimées sont utilisées directement et lorsqu'elles sont optimisées par la méthode développée par P. Puget et T. Skordas (cf [15]). Rappelons les expressions des matrices de calibration pour une position j du poignet du robot pour les 2 méthodes:

méthode inverse:

$$\widehat{M}_c^j = \widehat{M}_c^i \cdot \widehat{M}_{sm}^{-1} \cdot M_{sp}^i \cdot M_{sp}^{j-1} \cdot \widehat{M}_{sm}$$

méthode directe:

$$\widehat{M}_c^j = \widehat{K} \cdot \widehat{M}_{pc}^{-1} \cdot M_{sp}^{j-1} \cdot M_{sm}$$

Nous utilisons une mire constituée de 3 plans parallèles 400mm x 400mm séparés de 200mm et contenant chacun 100 points régulièrement espacés de 40mm. Après s'être fixé les matrices de transformation 3D M_{sm} et M_{pc} on génère, pour 10 positions M_{sp}^j différentes, les matrices de calibration exactes correspondantes. La caméra matricielle simulée est une caméra de type CCD 16mm. Les paramètres intrinsèques sont $u_0=250$, $v_0=275$, $\frac{f}{\Delta_{xim}} = 1000$ et $\frac{f}{\Delta_{yim}} = 1500$. Les estimations des matrices de calibrations initiales sont perturbées en ajoutant un bruit gaussien de moyenne nulle et d'écart type σ aux coordonnées des projections image $U_i(u_i, v_i)$ des points 3D de la mire. Ces calibrations sont utilisées, ensuite, pour estimer les paramètres \widehat{M}_{sm} et \widehat{M}_{pc} et

\widehat{K} résultants. Les paramètres intrinsèques retenus, résumés par la matrice \widehat{K} , sont leur moyenne sur tous les paramètres intrinsèques obtenus pour chaque calibration. La matrice \widehat{M}_{pc} est estimée par la méthode développée par F. Chaumette et P. Rives (cf [11]). La qualité des matrices de calibration est mesurée en projetant les points 3D d'une mire virtuelle liée artificiellement au repère caméra et placée devant celle-ci. La somme des différences entre les projections par la matrice exacte (u_i, v_i) et les projections par la matrice estimée ($\widehat{u}_i, \widehat{v}_i$) constitue l'évaluation de l'erreur ($Err = \sum_i [(\widehat{u}_i - u_i)^2 + (\widehat{v}_i - v_i)^2]$). Le tableau ci-dessous résume l'erreur Err en pixels obtenue pour une position proche et une position éloignée de la position de référence pour différents niveaux de bruit. Nous précisons que l'influence de la distance à une position de référence n'intervient pas dans le cas de la méthode directe.

(*): Déviation moyenne provoquée sur les projections $U_i(u_i, v_i)$

position (*)	Err (position proche)				Err (position éloignée)			
	0.5	1.0	1.5	2.0	0.5	1.0	1.5	2.0
méthode 1	0.527	1.16	1.89	2.73	0.527	1.16	1.89	2.73
méthode 2	0.113	0.224	0.33	0.436	0.523	1.03	1.53	2.0
méthode 1b	0.105	0.214	0.327	0.443	0.105	0.214	0.327	0.443
méthode 2b	0.083	0.171	0.261	0.356	0.158	0.323	0.49	0.673

- méthode 1 : Méthode directe
- méthode 2 : Méthode inverse
- méthode 1b : Méthode directe + optimisation des matrices de calibration
- méthode 2b : Méthode inverse + optimisation des matrices de calibration

7.2.2 Résultats en reconnaissance d'objets 3D

La base d'images et conditions de tests: Nous allons appliquer la méthode des descripteurs de Fourier à la reconnaissance de 3 types d'objets, un type AVION, un type VOITURE, et un type CHAR. Le réseau réalisé devra reconnaître tout objet appartenant à ces 3 catégories.

Chacune des classes est composée de 3 modèles lui appartenant:

- AVION: A320, Dc10, Dc8
- VOITURE: Ferrari, Mercedes, Fiat tipo
- CHAR: Gal-Lee, R35, AML H90 F1

La base d'apprentissage contient 143 exemples par modèle, ce qui fait un total de 429 exemples par classe. La généralisation est effectuée, d'une part sur des vues différentes des modèles appris (99 exemples par modèles) et d'autre part sur des modèles inconnus pour les classes AVION et VOITURE (52 exemples par modèles).

Les modèles inconnus sont un Concorde et un B707 pour la classe AVION, une Mercedes cabriolet (300SL) et une Chevrolet corvette pour la classe VOITURE.

Nous exposons ici les résultats obtenus dans les cas de réseaux de neurones 3 et 4 couches alimentés par une représentation des données en module avec une étude des performances après réduction par analyse en composantes principales en conservant 98% de l'énergie provoquant un passage de 31 composantes à 11 pour les vecteurs d'entrées.

Nous avons également étudié la possibilité de discriminer, à l'intérieur de chaque classe, les modèles entre eux.

Nous rappelons que dans ce dernier cas, la base d'exemples contient 143 exemples par modèle pour l'apprentissage et 99 exemples par modèle pour la généralisation. L'étude a été faite à l'aide d'un réseau 3 couches avec 9 neurones sur la couche cachée, avec ACP (98% de l'énergie conservée) et sans ACP. La réduction à 98% de l'énergie fait passer le nombre d'entrée de 31 à 13 dans les cas de l'identification des avions et chars et de 31 à 9 pour les voitures.

Résultats sans ACP:

a- En généralisant sur des vues différentes des modèles appris:

	Généralisation
4 couches	96.2%
3 couches	96.5%

modèles	4 couches	3 couches
Concorde	75%	80.6%
B707	97.2%	94.4%
Mercedes cabriolet	88.9%	88.9%
Chevrolet corvette	94.4%	94.4%

c- Discrimination des modèles entre eux:

	Généralisation
AVIONS	98.7%
VOITURES	68.7%
CHARS	92.3%

Résultats avec ACP:

On ne conserve que 98% de l'énergie.

a- En généralisant sur des vues différentes des modèles appris:

	Généralisation
4 couches	96.6%
3 couches	97.3%

b- En généralisant sur des modèles inconnus:

modèles	4 couches	3 couches
Concorde	94.4%	97.2%
B707	97.2%	97.2%
Mercedes cabriolet	94.4%	92%
Chevrolet corvette	94.4%	94.4%

c- Discrimination des modèles entre eux:

	Généralisation
AVIONS	98.7%
VOITURES	61.3%
CHARS	89.6%

Les performances en généralisation d'un réseau de neurone sont nettement validées par les résultats précédents. En effet, on remarque que les objets inconnus (le réseau ne les a jamais appris) sont très correctement reconnus. Leurs taux de généralisation sont comparables au taux de généralisation obtenu pour des modèles appris. L'exemple le plus net est le Concorde qui, malgré sa différence d'aspect par rapport aux modèles appris (A320, Dc10, Dc8), est reconnu sur 52 vues différentes à 97.2% après analyse en composantes principales et avec un réseau 3 couches.

La méthode de calibration par transformée inverse, que nous venons de développer, est actuellement utilisée pour calibrer un banc robotique constitué d'une caméra montée sur un bras manipulateur 6 axes (figure 1). Plusieurs applications mettant en œuvre des algorithmes de stéréovision et de reconstruction d'objets 3D [17], nous ont permis de valider qualitativement la méthode de calibration.

De nombreuses démonstrations du système de localisation / reconnaissance d'objets 3D implanté sur le robot ont permis de confirmer la robustesse de la méthode précédemment développée. Les erreurs parfois constatées ont été essentiellement provoquées par une mauvaise segmentation de l'image et de la silhouette.

Des expérimentations futures seront menées afin d'étudier les performances d'un tel système pour des conditions de prises de vues plus réalistes. L'utilisation d'une classe de rejet sera également étudiée afin de détecter les cas où l'objet présenté n'appartient à aucune des classes définies.

REFERENCES

- [1] G. BUREL
"Réseaux de neurones en traitement d'images - des modèles théoriques aux applications industrielles"
Thèse de Doctorat, Université de Brest, 6 Décembre 1991
- [2] Bart DE GREEF
"Classical approaches to 3 dimensional object recognition"
Esprit II project n°2059 "Pygmalion", report R36-2, September 1989
- [3] Bart DE GREEF
"Prototype implementation of a classical approach to 3D object recognition"
Esprit II project n°2059 "Pygmalion", report R36-4, June 1990
- [4] C. GUERRA, E.N. HOUSTIS
"Definition of algorithms for 3D pattern recognition"
Esprit II project n°2059 "Pygmalion", report R35-1, January 1990
- [5] G.H. GRANLUND
"Fourier processing for hand print character recognition"
IEEE Trans. Computers, vol 21, pp 195-201, 1972
- [6] C.S. LIN, C.L. HWANG
"New forms of shape invariants from elliptic Fourier Descriptors"
Pattern Recognition, vol 20, n°5, pp 535-545, 1987
- [7] O.R. MITCHELL, T.A. GROGEN
"Evaluation of Fourier Descriptors for target recognition in digital imagery"
Purdue University, report RADC-TR-83-33, February 1983
- [8] D.E. RUMELHART, G.E. HINTON, R.J. WILLIAMS
"Learning internal representations by error backpropagation"
Parallel Distributed Processing, D.E. RUMELHART and J.L. Mc CLELLAND
Chap 8, Bradford Book - MIT Press - 1986
- [9] Gilles BUREL
"Reconnaissance d'objets 3D par réseau d'automates"
Congrès AFCET-RFIA, Lyon-Villeurbanne, 25-29 Novembre 1991
- [10] O.D. FAUGERAS, G. TOSCANI
"The calibration problem for stereo"
Proceedings of CVPR'86
Miami Beach, Florida, pp 15-20, 1986
- [11] F. CHAUMETTE, P. RIVES
"Réalisation et calibration d'un système expérimental de vision
composé d'une caméra mobile embarquée sur un robot-manipulateur"
IRISA de Rennes, Campus universitaire de Beaulieu
Rapport de Recherche n° 454, Février 1989
- [12] H.A. MARTINS, J.R. BIRK and R.B. KELLEY
"Camera models based on data from two calibration planes"
Computer Graphics and Image Processing
vol 17 n° 2 p 173-180, 1981
- [13] K.D. GREMBAN, C.E. THORPE, T. KANADE
"Geometric camera calibration using system of linear equations"
Proceedings of Image Understanding Workshop - DARPA, April 1988, Massachusetts
- [14] A. IZAGUIRRE, P. PU and J. SUMMERS
"A new development in camera calibration calibrating a pair of mobile cameras"
Proceedings of IEEE, International Conference on robotics and automation, 74-79, 1985.
- [15] P. PUGET, T. SKORDAS
"An optimal solution for mobile camera calibration"
Proceedings of First European Conference on Computer Vision, Antibes, France, April 1990
- [16] M.E. BOWMAN and A.K. FORREST
"Transformation Calibration of a camera mounted on a robot"
Image and Vision Computing
vol 5 n°4, Novembre 1987
- [17] C. VENAILLE, D. MISCHLER, J.L. COATRIEUX
"Un algorithme peu contraint d'appariement de primitives courbes par stéréovision trinoculaire"
Revue technique THOMSON-CSF, numéro spécial analyse d'images
Tome 1, Décembre 1992
- [18] H. HENOCQ, C. VENAILLE
"Calibration Dynamique d'une caméra CCD par Transformée Inverse"
Revue technique THOMSON-CSF, vol. 1, Mars 1993
- [19] H. HENOCQ, G. BUREL
"Reconnaissance d'objets 3D par analyse de Fourier de la silhouette"
Revue technique THOMSON-CSF, vol. 1, Mars 1993