



**HAL**  
open science

## Reconnaissance d'objets 3D par réseau d'automates

Gilles Burel

► **To cite this version:**

Gilles Burel. Reconnaissance d'objets 3D par réseau d'automates. RFIA, Nov 1991, Lyon-Villeurbanne, France. hal-03222644

**HAL Id: hal-03222644**

**<https://hal.univ-brest.fr/hal-03222644v1>**

Submitted on 19 Mar 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

**RECONNAISSANCE D'OBJETS 3D  
PAR RESEAU D'AUTOMATES**

Gilles BUREL <sup>(\*)</sup><sup>(\*\*)</sup>

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

tél : 99.25.42.88 fax : 99.25.43.34

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

THEME : Vision (V)

RESUME :

*De nombreuses applications de vision, notamment en robotique, nécessitent l'identification d'objets tri-dimensionnels. L'emploi de méthodes complexes, à base de mise en correspondance avec des modèles, n'est pas toujours justifié, et peut être jugé trop coûteux en temps de calcul. Nous proposons une méthode rapide, et de mise en œuvre aisée, pour la reconnaissance d'objets 3D.*

*La méthode exploite les capacités d'apprentissage et de généralisation d'un réseau d'automates. L'idée consiste à faire apprendre à un tel réseau un certain nombre de vues de chaque objet possible. Pour réduire la quantité de données à traiter, l'objet est caractérisé par sa silhouette. En fin d'apprentissage, les capacités de généralisation du réseau d'automates lui permettent une reconnaissance sur des vues non apprises.*

*Après une description de la méthode proposée, nous présenterons les résultats obtenus sur une base de données de 216 images.*

ABSTRACT :

*A lot of vision applications, for instance in robotics, require identification of 3D objects. Use of complex methods, based on model matching, is not always necessary, and may be too computationally expensive. We propose a fast and simple method for recognition of 3D objects.*

*The method takes profit of the learning capabilities of a neural network. The idea is to train a neural network on some views of each object. In order to reduce the amount of data, the object is characterized by its silhouette. At the end of the learning phase, the generalization capabilities of the network allow it to recognize non-learned views.*

*After a description of the proposed method, we will present experimental results obtained on a data base of 216 images.*

MOTS CLES : Apprentissage, Descripteurs de Fourier, Réseaux d'Automates, Silhouette, Vision 3D.

KEYWORDS : Machine Learning, Fourier Descriptors, Neural Networks, Silhouette, 3D Vision.

## 1 Introduction

La reconnaissance automatique d'objets tri-dimensionnels est considérée comme une tâche difficile du fait de l'importance des données à traiter et de la complexité de l'information sous-jacente. En raison de l'essor de la robotique, ce domaine est appelé à se développer dans l'avenir.

Dans la littérature, plusieurs approches ont été proposées pour la reconnaissance d'objets 3D. La plupart des 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 :

- Méthodes exploitant plusieurs vues de l'objet.
- Méthodes exploitant une seule vue de l'objet.

Le premier type de méthode consiste à reconstruire en 3 dimensions l'objet présent dans la scène (en mettant en œuvre des techniques de type stéréoscopie), puis à réaliser la mise en correspondance avec des modèles 3D. Ces méthodes supposent la mise en place d'un dispositif d'acquisition important (au moins 2 caméras, parfaitement calibrées), et nécessitent un traitement informatique complexe (pour résoudre le problème de la mise en correspondance). Mais elles présentent l'intérêt de fournir la position précise de l'objet en plus de la reconnaissance.

Le second type de méthode consiste à tenter d'interpréter l'image comme la projection de l'un des modèles 3D dont on dispose. La focalisation sur des structures caractéristiques (par exemple des groupes de 3 segments adjacents) permet de réduire la complexité du problème [3].

Les 2 types de méthodes nécessitent la création préalable de modèles 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. De plus, les segments sont des primitives relativement facile à détecter dans une image.

Du fait de leur capacité d'apprentissage, les réseaux d'automates pourraient permettre d'éviter la

phase de modélisation. Même si les performances obtenues (taux de reconnaissance) sont finalement plus faibles que celles des systèmes à base de modèle, 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 d'automates 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]. Cette description dans le domaine de Fourier présente les avantages suivants :

- Les descripteurs de Fourier portent la même information que le contour original (i.e. il est possible de reconstruire le contour à partir des descripteurs de Fourier).
- Ils constituent une représentation multi-résolution de la silhouette.
- 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.

Le paragraphe suivant est une rapide description de la base d'images constituée pour les expérimentations et du traitement bas niveau mis en œuvre. Ensuite, nous détaillerons la méthode adoptée pour utiliser les descripteurs de Fourier en association avec un réseau d'automates pour l'identification d'objets 3D. La méthode inclut une normalisation par rapport à une translation de l'objet dans le plan image, à une rotation de la caméra autour de son axe et à une modification de la distance caméra-objet. Nous étudierons également diverses stratégies pour fournir les données au réseau d'automates dans une représentation adéquate. Les résultats expérimentaux en apprentissage et généralisation seront ensuite présentés.

## 2 La base d'images et le traitement bas niveau

### 2.1 La base d'images

Les expérimentations ont été conduites sur une base de 216 images dont la taille est d'environ 300x200 pixels. Il y a 3 objets différents : un distributeur de ruban adhésif (SCOTCH), une agrafeuse (STAPLER) et une pièce industrielle (WORKPIECE). Ces images font partie de l'ensemble d'images test utilisé dans le projet Esprit n°2059 "Pygmalion". Pour chaque objet, il y a deux séries de 36 images. Elles ont été obtenues en plaçant l'objet sur un support que l'on fait tourner par pas de 10° autour de l'axe vertical. Une série a été réalisée avec la caméra sous une élévation de 0°, et l'autre série avec la caméra sous une élévation de 30°. Dans les 2 cas, la caméra vise approximativement le centre de gravité de l'objet.

Ces 216 images ont été divisées en 2 groupes de 108 images : l'ensemble d'apprentissage, qui permet d'entraîner notre système, et l'ensemble d'évaluation qui permet de tester ses performances. L'ensemble d'apprentissage contient les images prises sous un azimut multiple de 20° (0°, 20°, 40°, ..., 340°). L'ensemble d'évaluation contient les autres images (azimut = 10°, 30°, ..., 350°).

Quelques images sources sont visibles sur les figures 3 à 5. On remarquera des problèmes d'ombres et de reflets. Ceci est particulièrement vrai pour la pièce industrielle, dont la surface métallique produit d'importants reflets, de sorte que les contours extraits seront très bruités. Mais comme nous le verrons plus loin, les performances obtenues en reconnaissance sont satisfaisantes, ce qui confirme la robustesse de la méthode.

### 2.2 Extraction de la silhouette

L'objet est en moyenne de brillance supérieure au fond. Par conséquent, un simple seuillage permet d'isoler la plus grande partie des pixels correspondant à l'objet.

On a utilisé ici un seuil fixe, réglé légèrement au dessus du niveau moyen du fond. Cependant, du fait des reflets sur le fond, certaines zones du fond passent le seuil. Afin de les éliminer on effectue une détection de zones connexes. Seule la zone connexe de plus forte surface est conservée, ce qui permet d'éliminer les reflets sur le fond, et de ne conserver que l'objet. Puis, on effectue un suivi de contour.

Pour information, quelques contours d'objets sont présentés sur les figures 6 et 7. On remarquera les perturbations importantes sur les contours (principalement pour la pièce industrielle), dues aux problèmes d'ombres et de reflets.

## 3 Les descripteurs de Fourier

### 3.1 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. De nombreux auteurs ont étendu la liste des propriétés connues des descripteurs et les méthodes de normalisation [7] [6].

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. Un re-échantillonnage plus élaboré, utilisant par exemple une fine approximation polygonale pour la définition de l'abscisse curviligne, aurait pu être envisagée pour des images de plus faible résolution. Il n'est pas justifié ici.

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 comme :

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

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 :

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

### 3.2 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 :

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

En conséquence, pour être invariant par rapport à ces transformations, nous réalisons les normalisations suivantes :

1.  $C'_0$  n'est pas pris en compte, donc l'invariance en TRANSLATION est assurée.
2. Tous les coefficients sont divisés par  $C'_1$ , donc les invariances en ROTATION et HOMOTHETIE sont assurées. Ceci peut aisément être vérifié :

$$\begin{aligned} K'_n &= \frac{C'_n}{C'_1} \\ &= \frac{\rho C_n e^{j\phi} e^{2j\pi \frac{kn}{N}}}{\rho C_1 e^{j\phi} e^{2j\pi \frac{k}{N}}} \\ &= \frac{C_n e^{2j\pi \frac{k(n-1)}{N}}}{C_1} \\ &= K_n e^{2j\pi \frac{k(n-1)}{N}} \end{aligned}$$

3. Tous les coefficients  $K'_n$  sont ensuite multipliés par  $e^{j(1-n)\phi'_2}$ , où  $\phi'_2$  est la phase de  $K'_2$ . Comme

$\phi'_2 = \phi_2 + 2\pi \frac{k}{N}$ , nous avons

$$\begin{aligned} P'_n &= K'_n e^{j(1-n)\phi'_2} \\ &= K_n \left( e^{2j\pi \frac{k(n-1)}{N}} \right) \left( e^{j(1-n)(\phi_2 + 2\pi \frac{k}{N})} \right) \\ &= K_n e^{j(1-n)\phi_2} \\ &= P_n \end{aligned}$$

Donc l'invariance par rapport au point de départ sur le contour est assurée. Ici, il est important de noter que la phase  $\phi'_2$  doit être arrondie à la valeur la plus proche de la forme  $\frac{2\pi l}{N}$ , où  $l$  est un entier. Ne pas le faire causerait l'incohérence des descripteurs normalisés, car il ne serait plus possible de les interpréter comme étant les descripteurs d'un objet réel, obtenu en faisant subir à l'objet source des transformations géométriques de base, et un décalage du point de départ sur le contour. Ceci peut facilement être vérifié en reconstruisant le contour à partir des descripteurs de Fourier.

Cette normalisation n'est efficace que sous réserve que le module des coefficients  $C_1$  et  $C_2$  ne soit pas trop faible. En effet, dans le cas contraire, l'imprécision sur leur phase est importante. On peut montrer que le module du coefficient  $C_1$  est le rayon du cercle qui approxime au mieux la forme. Ce module est donc assez élevé pour les objets réels. Quant au module de  $C_2$ , la figure 9 (sur laquelle nous reviendrons), montre quelle est généralement assez importante. Lorsque ceci n'est pas le cas, on peut envisager une méthode de normalisation plus élaborée.

Il est important de noter qu'il n'y a pas de perte d'information concernant la forme des objets dans notre prétraitement. L'application de la transformation inverse aux coefficients normalisés  $P_n$  fournit une silhouette identique à l'originale, à une transformation géométrique près (composition d'une translation, d'une rotation, et d'une homothétie).

La figure 8 montre la reconstruction de l'agrafeuse (vue sous  $0^\circ$  d'élévation et  $90^\circ$  d'azimut) en fonction du nombre de coefficients conservés. On ne conserve que les coefficients d'indice  $n$  tel que  $-\Delta \leq n \leq \Delta$ , pour les valeurs  $\Delta = 1, 2, 4, 6, 8, 10, 16, 32$ . On constatera que les coefficients d'ordre inférieur à 32 sont très largement suffisants pour reconstruire fidèlement la forme. Les coefficients d'ordre élevé correspondent à des hautes fréquences qui ne sont pas utiles pour la représentation de la forme (ces coefficients représentent

plutôt du bruit de quantification).

Compte tenu du nombre relativement faible d'exemples à notre disposition pour l'apprentissage (108 exemples), nous conserverons seulement les coefficients d'ordre inférieur à 16 (ces coefficients décrivent encore relativement bien la forme). Ceci permet de limiter la dimension du réseau de automates. En effet, un réseau sur-dimensionné risquerait de faire de l'apprentissage "par cœur", ce qui produirait une mauvaise généralisation.

## 4 Le réseau d'automates

### 4.1 Structure du réseau

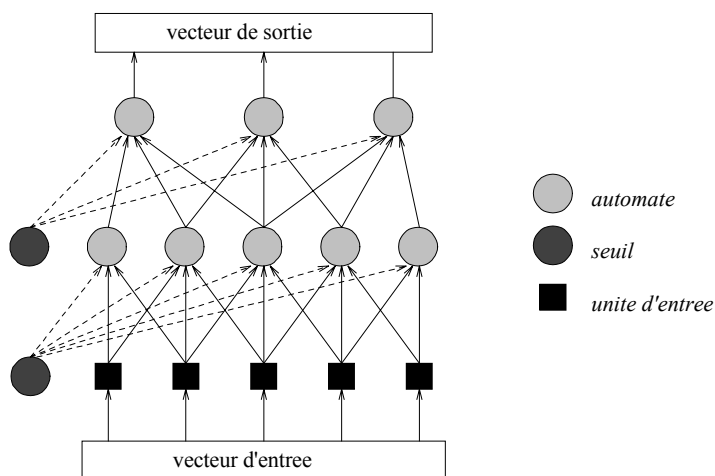


FIG. 1: Perceptron multi-couches

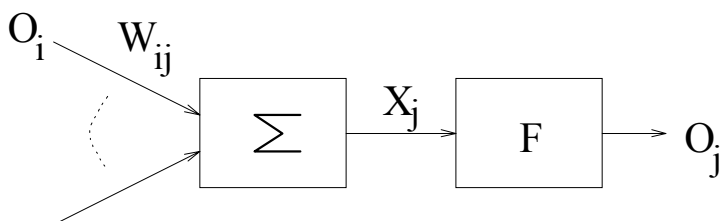


FIG. 2: modèle de l'automate

Nous avons utilisé un perceptron multi-couches (fig 1), entraîné par l'algorithme de rétropropagation [8], auquel nous avons apporté diverses améliorations [1] (réglage automatique des paramètres de l'algorithme et contrôle de la saturation). L'automate (fig 2) est un sommateur pondéré suivi 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 automate sans entrées, et dont la sortie est toujours égale à 1. Cet automate, nommé "automate seuil", permet d'améliorer les performances du réseau. La couche d'entrée est alimentée avec les valeurs des descripteurs de Fourier normalisés d'ordre inférieur à 16.

### 4.2 Apprentissage et Classification

La base d'images a été divisée en 6 classes, et la couche de sortie du réseau d'automates contient 6 automates (un par classe). Le tableau suivant indique le vecteur de sortie souhaité pour chacune des 6 classes :

| Classe | Dénomination       | Sorties souhaitées       |
|--------|--------------------|--------------------------|
| 1      | Scotch site=0°     | (+1, -1, -1, -1, -1, -1) |
| 2      | Scotch site=30°    | (-1, +1, -1, -1, -1, -1) |
| 3      | Stapler site=0°    | (-1, -1, +1, -1, -1, -1) |
| 4      | Stapler site=30°   | (-1, -1, -1, +1, -1, -1) |
| 5      | Workpiece site=0°  | (-1, -1, -1, -1, +1, -1) |
| 6      | Workpiece site=30° | (-1, -1, -1, -1, -1, +1) |

L'apprentissage consiste à adapter les coefficients de pondération  $W_{ij}$  afin d'obtenir la réponse souhaitée en sortie. Il se fait conformément à l'algorithme de rétropropagation. Lorsque l'apprentissage est terminé, le réseau peut traiter de nouvelles données.

Pour la classification, l'automate dont la sortie est la plus forte détermine la classe. Une mesure de confiance peut également être définie comme la moitié de l'écart entre la plus forte sortie et la sortie immédiatement inférieure. Ainsi, si les sorties obtenues sont  $(-0.9, -0.3, 0.7, -0.8, 0.2, -0.8)$ , la

classe sera 3, et la confiance sera  $(0.7 - 0.2)/2$ , soit 0.25

### 4.3 Représentation des entrées

Trois possibilités pour représenter les descripteurs de Fourier en entrée du réseau d'automates ont été envisagées.

**Représentation 1 :** Les descripteurs normalisés sont présentés sous forme cartésienne :

$$\begin{aligned} &Re\{P_n\} \\ &Im\{P_n\} \end{aligned}$$

**Représentation 2 :** Seulement les modules des descripteurs normalisés sont présentés :

$$Module\{P_n\}$$

**Représentation 3 :** Les descripteurs normalisés sont présentés sous forme polaire :

$$\begin{aligned} &100 * Module\{P_n\} \\ &cos(phase\{P_n\}) \\ &sin(phase\{P_n\}) \end{aligned}$$

Le sinus et le cosinus de la phase sont présentés à la place de la phase elle-même, de façon à éviter le problème de discontinuité lors du passage de  $359^\circ$  à  $0^\circ$ . Le module est également multiplié par 100 afin de l'amener à l'ordre de grandeur des sinus et cosinus (cette multiplication ne change pas les performances obtenues, mais permet d'accélérer l'apprentissage).

La silhouette est décrite par les coefficients  $P_{-16}$  à  $P_{-1}$  et  $P_2$  à  $P_{16}$ , soit un total de 31 coefficients. La taille de la couche d'entrée du réseau d'automates sera donc 62, 31 ou 93 en fonction de la représentation choisie.

Pour information, les modules des descripteurs de Fourier pour les images de l'ensemble d'évaluation sont représentées sur la figure 9. Les modules sont représentées en luminance (la brillance est proportionnelle à la racine carrée du module). En abscisse, on a l'indice du descripteur (2 à 16, puis  $-16$  à  $-1$ ), et en ordonnée on a le numéro de l'image. De gauche à droite et de haut en bas, on a les 18 exemples de la classe 0,

puis les 18 exemples de la classe 1, etc. Cette figure permet de mieux apprécier la difficulté du problème. Aucune classe ne se distingue en effet clairement des autres, si ce n'est l'agrafeuse sous  $0^\circ$  d'élévation (classe 3) par l'importance des hautes fréquences.

## 5 Résultats expérimentaux

Les tableaux ci-dessous comparent les résultats obtenus avec différents réseaux d'automates : un réseau à 2 couches, un réseau à 3 couches (9 automates sur la couche intermédiaire, automate seuil non compris) et un réseau à 4 couches (9 et 6 automates sur les couches intermédiaires). A titre de comparaison, les résultats obtenus en remplaçant le réseau d'automates par un classifieur aux k-Plus-Proches-Voisins (k-ppv) sont également indiqués. La dernière colonne indique le nombre de multiplications. Les classifieurs sont rangés par ordre décroissant des taux de généralisation :

Représentation cartésienne :

| Classifieur | Apprentissage | Généralisation | mult. |
|-------------|---------------|----------------|-------|
| 4 couches   | 93%           | 70%            | 648   |
| ppv         | 100%          | 65%            | 6696  |
| 3 couches   | 91%           | 63%            | 612   |
| 3-ppv       | 67%           | 59%            | 6696  |
| 2 couches   | 90%           | 57%            | 372   |
| 5-ppv       | 55%           | 52%            | 6696  |

Module :

| Classifieur | Apprentissage | Généralisation | mult. |
|-------------|---------------|----------------|-------|
| 4 couches   | 94%           | 90%            | 369   |
| 3 couches   | 92%           | 86%            | 333   |
| 3-ppv       | 95%           | 84%            | 3348  |
| 2 couches   | 87%           | 83%            | 3348  |
| ppv         | 100%          | 80%            | 186   |
| 5-ppv       | 90%           | 75%            | 3348  |

Représentation polaire :

| Classifieur | Apprentissage | Généralisation | mult. |
|-------------|---------------|----------------|-------|
| 4 couches   | 96%           | 86%            | 927   |
| 3-ppv       | 94%           | 84%            | 10044 |
| 3 couches   | 95%           | 83%            | 891   |
| ppv         | 100%          | 81%            | 10044 |
| 5-ppv       | 91%           | 77%            | 10044 |
| 2 couches   | 93%           | 72%            | 558   |

## 6 Conclusion

Nous avons spécifié et expérimenté une méthode de reconnaissance d'objets 3D qui a l'avantage d'être simple à mettre en œuvre. En dépit de problèmes d'éclairage (ombres et reflets), les résultats obtenus sont satisfaisants, ce qui prouve la robustesse de la méthode face à un prétraitement ou une acquisition déficients. Le système proposé est facile à implémenter et à utiliser car il n'y a ni modélisation, ni stéréoscopie, ni reconstruction 3D (une seule vue est suffisante). De plus la reconnaissance est très rapide.

La matrice de confusion du réseau à 4 couches avec une représentation "module" est la suivante :

| Classe                                    | Matrice de confusion |     |      |     |     |     |
|---|----------------------|-----|------|-----|-----|-----|
| 0   | 83%                  | 17% | 0%   | 0%  | 0%  | 0%  |
| 1   | 0%                   | 94% | 0%   | 0%  | 6%  | 0%  |
| 2   | 0%                   | 0%  | 100% | 0%  | 0%  | 0%  |
| 3   | 0%                   | 0%  | 6%   | 83% | 0%  | 11% |
| 4   | 0%                   | 5%  | 0%   | 6%  | 89% | 0%  |
| 5   | 0%                   | 0%  | 0%   | 0%  | 11% | 89% |
| Taux moyen de généralisation : <b>90%</b> |                      |     |      |     |     |     |

En concaténant les lignes et colonnes correspondant à un même objet, on obtient la matrice de confusion suivante :

| Classe                                    | Effectif   | Matrice de confusion |     |     |
|---|------------|----------------------|-----|-----|
| 0-1                                       | 36 (33.3%) | 97%                  | 0%  | 3%  |
| 2-3                                       | 36 (33.3%) | 0%                   | 94% | 6%  |
| 4-5                                       | 36 (33.3%) | 3%                   | 3%  | 94% |
| Taux moyen de généralisation : <b>95%</b> |            |                      |     |     |

Enfin, dans le cas à 6 classes, en rejetant les formes pour lesquelles la confiance est inférieure à 0.33, on obtient les résultats suivants sur l'ensemble d'évaluation :

correct : 81%  
 rejet : 17%  
 erreur : 2%

Il est donc possible de faire passer le taux d'erreur de 10% à 2%, au prix d'une non-décision dans 17% des cas. Si la caméra est fixée sur le bras d'un robot, une bonne stratégie consisterait à déplacer le bras et acquérir une nouvelle vue lorsqu'il y a rejet.

Des expérimentations futures seront menées sur un plus grand nombre d'images. En particulier, la représentation polaire nous semble préférable à la représentation sous forme de module uniquement, car elle n'induit pas de perte d'information. Le fait que cette représentation n'ait pas conduit aux meilleurs résultats dans nos expérimentations nous semble plutôt dû au faible nombre d'exemples disponibles (ce qui favorise les réseaux de faible taille).

La représentation cartésienne semble être difficile à exploiter par le réseau. Ceci confirme l'idée intuitive selon laquelle il est plus facile d'interpréter un module et une phase qu'une partie réelle et une partie imaginaire. Les faibles performances du 3-ppv et du 5-ppv en apprentissage donnent d'ailleurs une indication quant à la complexité de cette représentation. En effet, ces performances signifient que les plus proches voisins d'un exemple appartiennent souvent à des classes différentes de la sienne.



## Références

- [1] G. BUREL, D. CAREL, J.Y. CATROS  
“A connectionist system for recognition of 2D workpieces”  
Revue Technique THOMSON-CSF, vol.22, n<sup>o</sup>4, Décembre 1990
  
- [2] Bart DE GREEF  
“Classical approaches to 3 dimensional object recognition”  
Esprit II project n<sup>o</sup>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<sup>o</sup>2059 “Pygmalion”, report R36-4, June 1990
  
- [4] C. GUERRA, E.N. HOUSTIS  
“Definition of algorithms for 3D pattern recognition”  
Esprit II project n<sup>o</sup>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<sup>o</sup>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

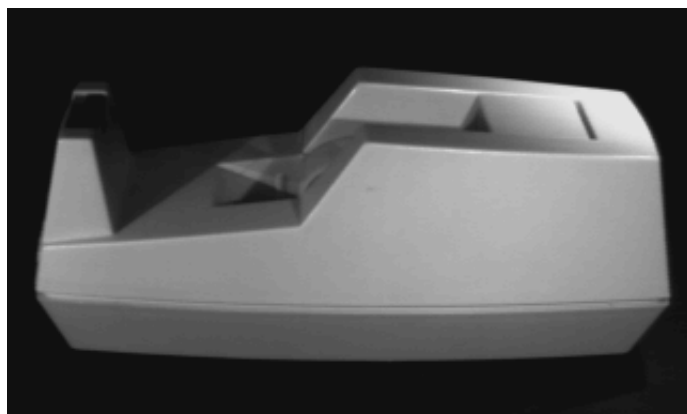


FIG. 3: SCOTCH, site=30°, azimuth=270°

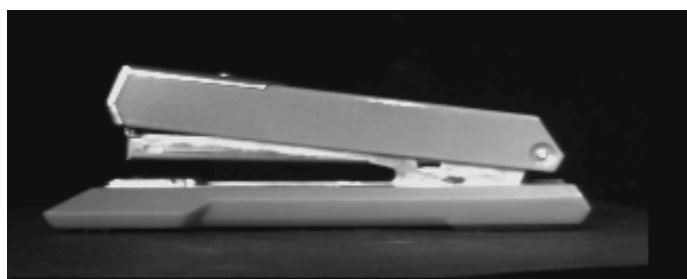


FIG. 4: STAPLER, site=0°, azimuth=90°

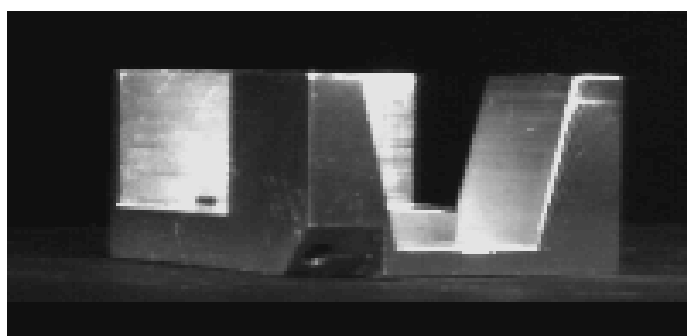


FIG. 5: WORKPIECE, site=0°, azimuth=150°

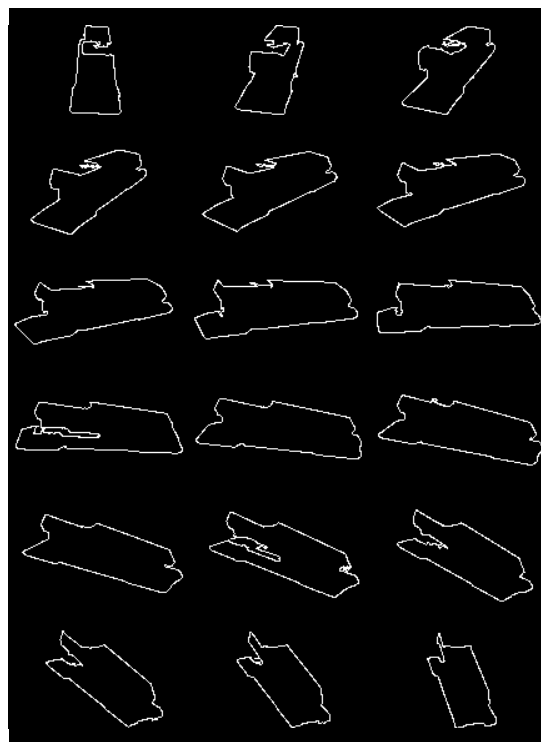


FIG. 6: Contours de STAPLER, site=30°, azimuth=0° à 170°

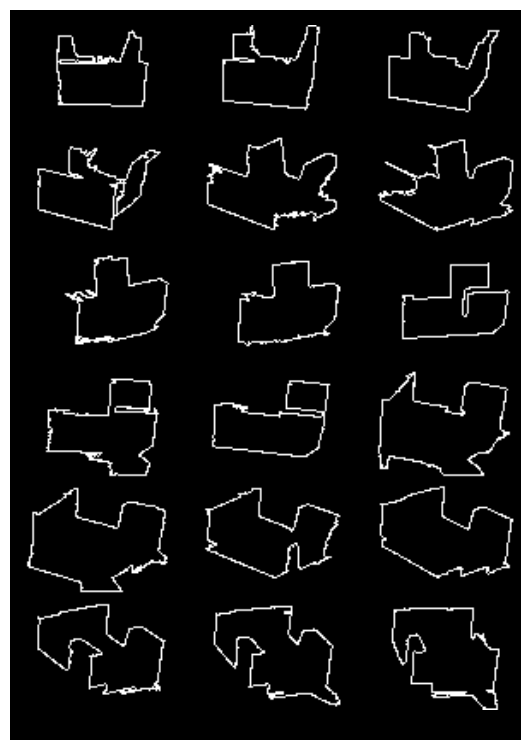


FIG. 7: Contours de WORKPIECE, site=30°, azimuth=0° à 170°

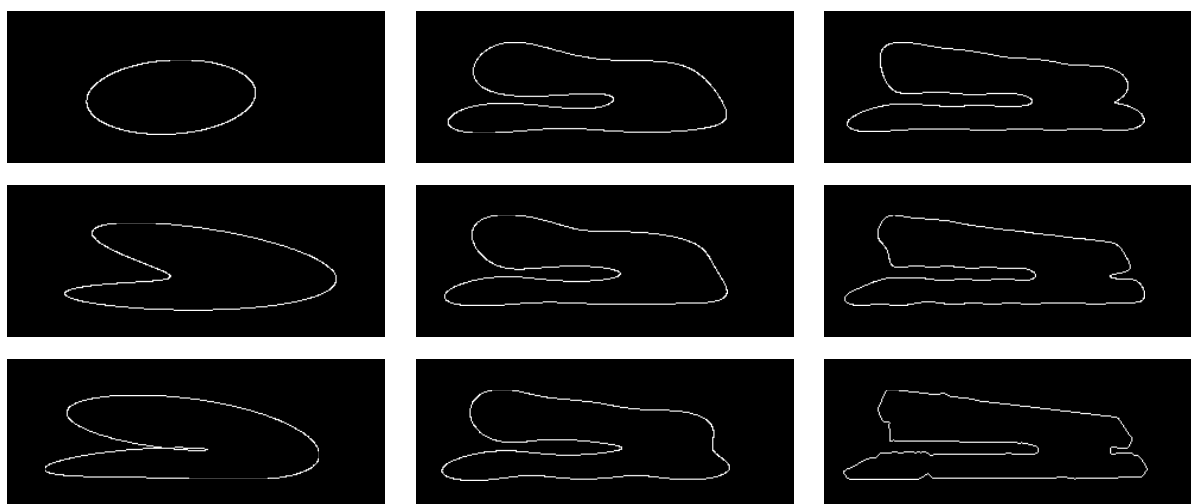


FIG. 8: Reconstruction de l'agrafeuse ( $\Delta = 1, 2, 4, 6, 8, 10, 16, 32$ ) et original

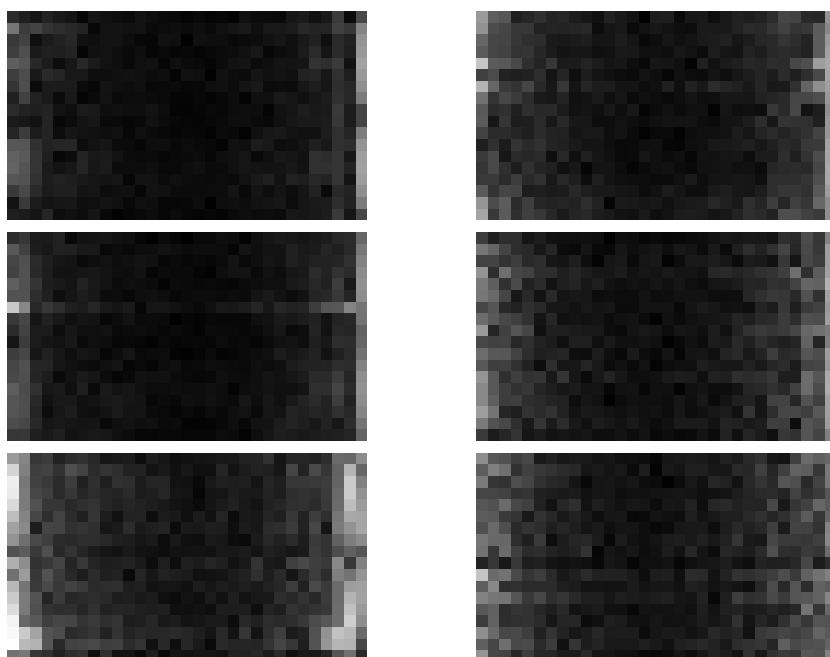


FIG. 9: Module des descripteurs de Fourier (ensemble d'évaluation, racine carrée). En abscisse : l'ordre du coefficient (2 à 16 puis -16 à -1). En ordonnée : le numéro de l'exemple (de gauche à droite et de haut en bas : classes 1 à 6).