

Identification de formes à partir de la transformée de Radon

Antoine Tabbone
LORIA-Equipe QGAR
tabbone@loria.fr
www.loria.fr/~tabbone

QGAR

*Querying Graphics through Analysis
and Recognition*

Avant-projet issu du projet ISA



Composition de l'équipe

- S. Collin, MC UHP
- Ph. Dosch, MC Nancy 2
- X. Hilaire, CIFRE FS2I
- B. Lamiroy, MC Mines
- G. Masini, CR CNRS
- J. Rendek, ingénieur associé INRIA
- S. Tabbone, MC Nancy 2
- K. Tombre, PR Mines
- L. Wendling, MC UHP



Fonds de commerce,
domaine scientifique

Complémentaire avec
recherches dans ce domaine,
fournisseurs d'indices

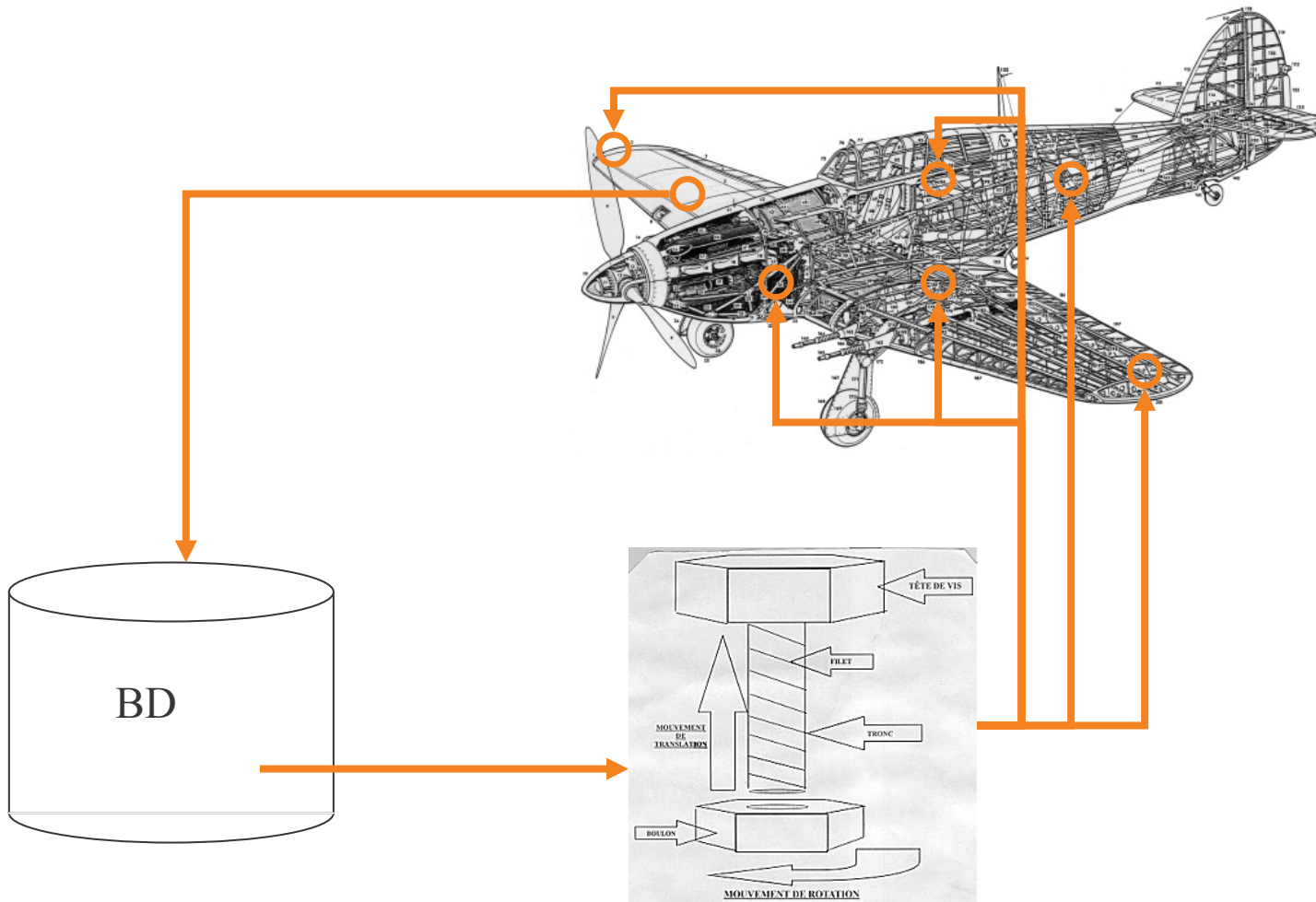
Explorer la capacité
reconnaissance graphique
à fournir à **l'innovation** et à la
recherche d'information
les **indices nécessaires**,

Texte/graphique
vecteurs,
signatures
symboles...

contexte de la
documentation

Domaine applicatif
Documents avec graphique
Papier ou électronique

Exemples





Axes de recherche

- Extraction d'indices et segmentation
- Caractérisation et reconnaissance de signatures et de symboles
- Analyse structurelle et sémantique



Plan de l'exposé

■ Introduction

- De nombreux travaux
- Les problèmes majeurs
- Les avantages de la TR

■ Définition de la TR

- Définition mathématique
- Application à la tomographie
- Concept de R-signature


■ Applications

- A l'identification d'objets binaires
- A l'identification d'objets photométriques

Introduction

- De nombreux travaux
- Les problèmes majeurs
- Avantages de la Transformée de Radon





Beaucoup de travaux en Pattern Recognition

- Applications dans divers domaines: bio-info, data mining, analyse d'images fixes ou animées, inspection industrielle, reconnaissance de la parole, analyse de documents textuels ou graphiques...
 - *A.K. Jain et al, Statistical Pattern Recognition: A Review, PAMI 00, vol 22, no 1,*
 - *L.P. Cordella et al, Symbol recognition in documents: a collection of techniques, IJDAR 2000, vol 3,*
 - *D. Doermann, The indexing and retrieval of document images: a survey, TR, Université Maryland 1998.*
 - *Content-based Image Retrieval at the End of the Early Years, PAMI 00, vol 22, no12*



En principe trois étapes:

■ Représentation:

contours, squelette, description polygonale, composantes connexes...

■ Description:

indices géométriques et topologiques, transformées, moments invariants, descriptions structurelles et syntaxique...

■ Classification:

template matching, appariement de graphes, analyseur syntaxique, réseaux de neurones, approches à base de logique floue, méthodes statistiques...



Avantages de la TR

- Complexité faible
- Transformations géométriques
- Indice de forme discriminant

Travaux précurseurs: V. Leavers, PAMI 2000

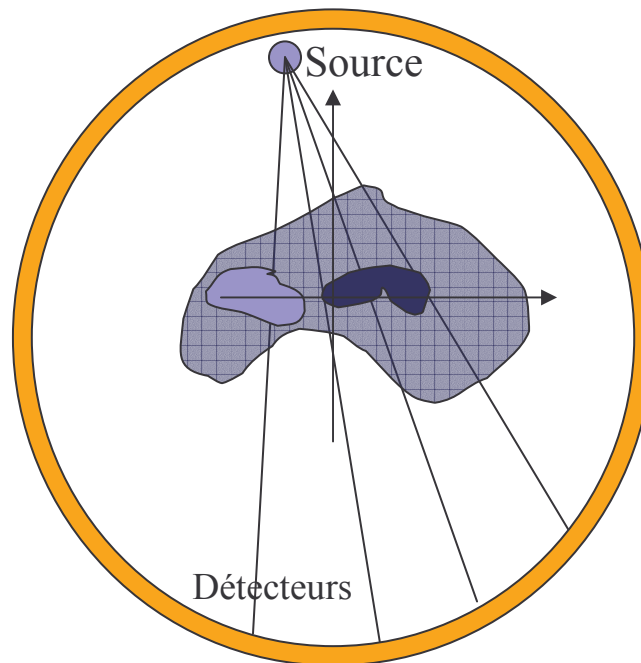
⇒ taxonomie de formes



Définition de la TR

- Définition mathématique
- Application à la tomographie
- Concept de R-signature

Application à la tomographie 2D/3D

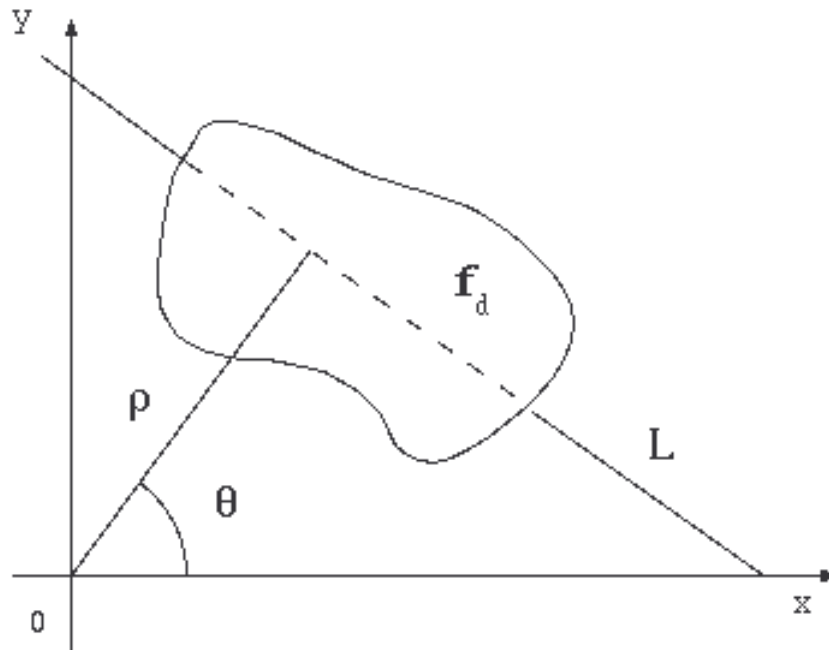


Principe du scanner

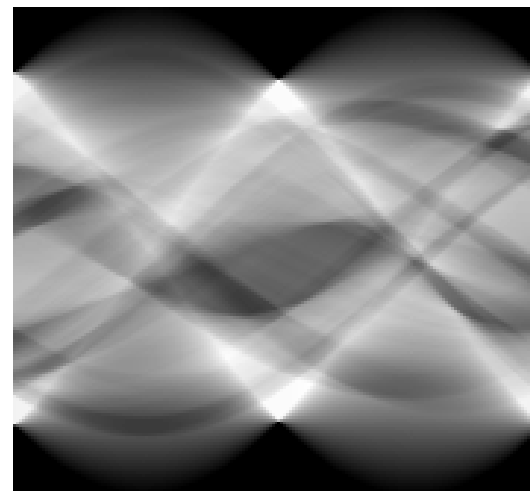
Définition mathématique

$$T_{Rf}(\rho, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(x \cos(\theta) + y \sin(\theta) - \rho) dx dy$$

$$-\infty < \rho < \infty, 0 \leq \theta < \pi$$



Une forme 2D et sa transformée de Radon





Propriétés

- Symétrie: $T_{R^f}(\rho, \theta) = T_{R^f}(-\rho, \theta \pm \pi)$
- Périodicité: $T_{R^f}(\rho, \theta) = T_{R^f}(\rho, \theta + 2k\pi)$
- Translation: $T_{R^f}(\rho - x_o \cos(\theta) - y_o \sin(\theta), \theta)$
- Rotation: $T_{R^f}(\rho, \theta + \theta_o)$
- Échelle: $\frac{1}{|\alpha|} T_{R^f}(\alpha\rho, \theta)$



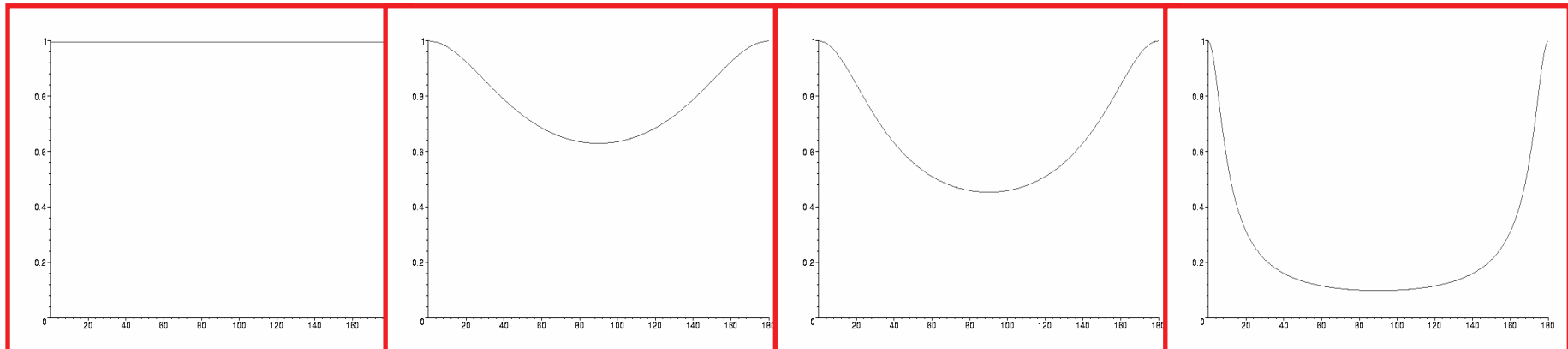
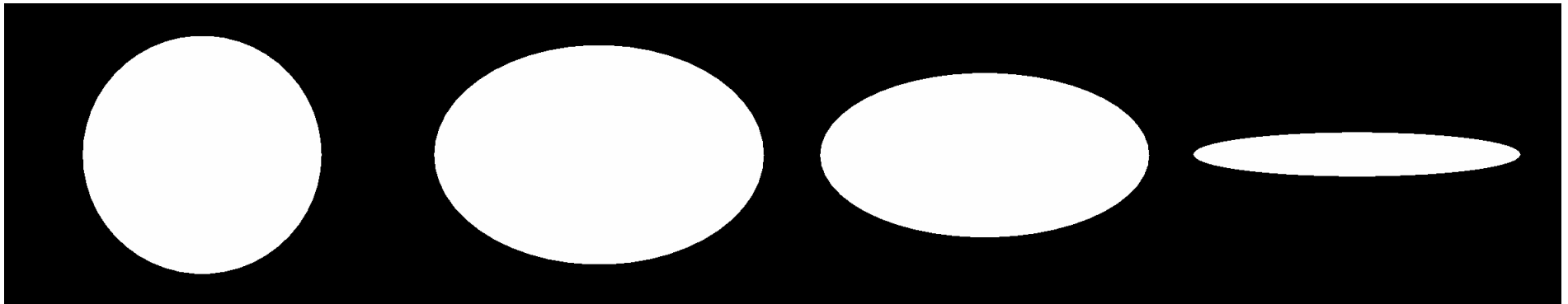
Concept de R-signature

$$R_f(\theta) = \int_{-\infty}^{\infty} T_{R^f}^2(\rho, \theta) d\rho$$

- Périodicité: $R_f(\theta) = R_f(\theta \pm \pi)$
- Rotation: $R_f(\theta + \theta_0)$
- Translation: $R_f(\theta)$
- Échelle: $\frac{1}{\alpha^2} R_f(\theta)$

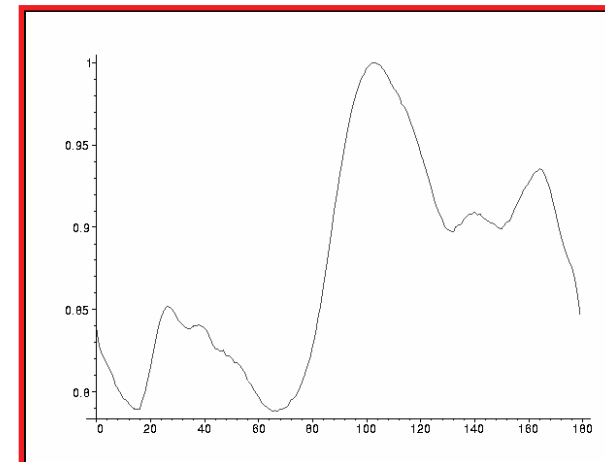
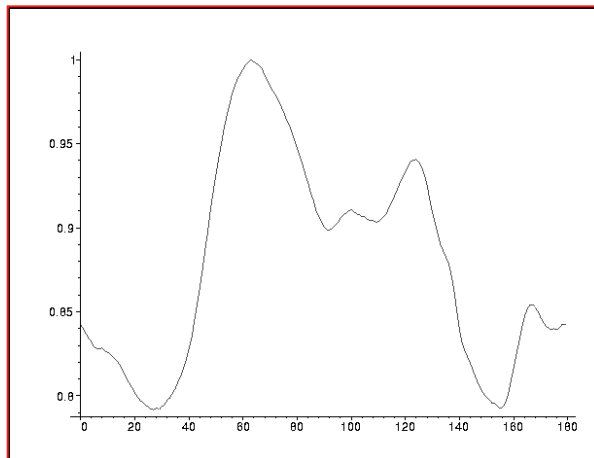
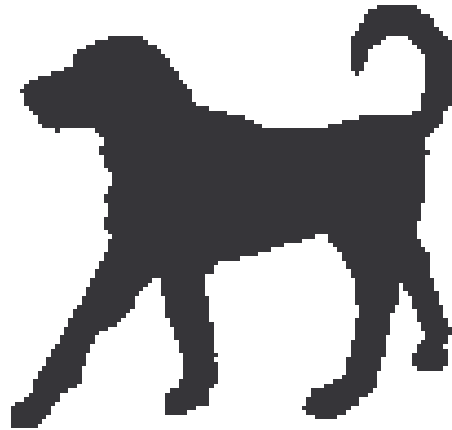


Exemple de R-signature





Exemple de R-signature





Algorithmes

- naïf :

- tous les ρ possibles, tous les θ possibles
- on recalcule l'intégrale à chaque fois

$$k_1 \cdot O(n^3) \text{ avec } k_1 \text{ grand}$$

- naïf amélioré :

- on projette chaque pixel dans le plan de Radon : $(x,y) \rightarrow (\rho=x.\cos \theta + y.\sin \theta, \theta)$

$$k_2 \cdot O(n^3) \text{ avec } k_2 \text{ petit}$$

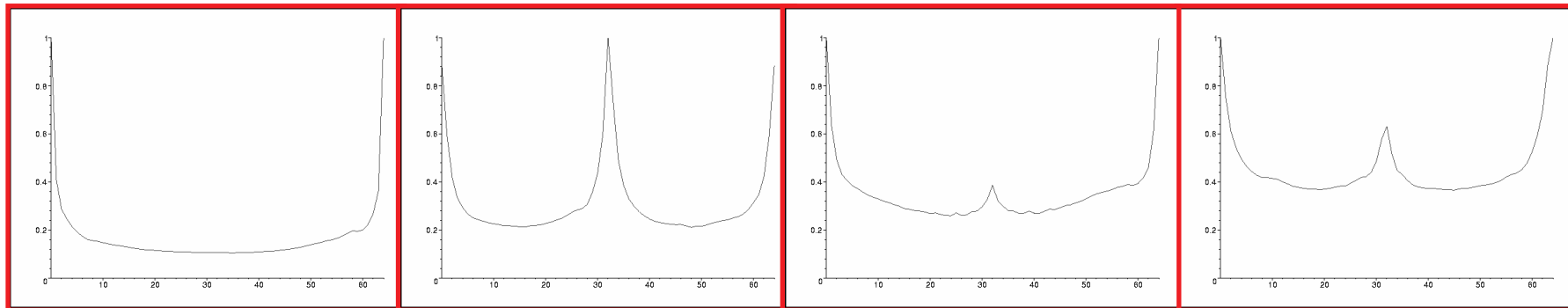
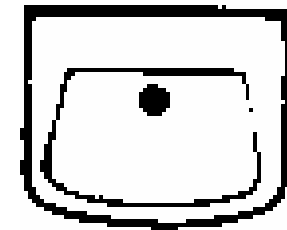
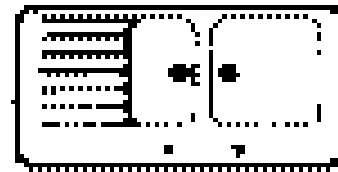
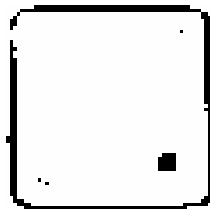
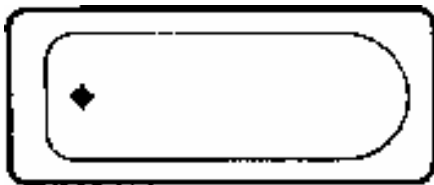
- meilleur :

- entre deux orientations successives, certaines sommes de pixels sont communes. De manière itérative, on arrive à une bonne approximation de la matrice de Radon :

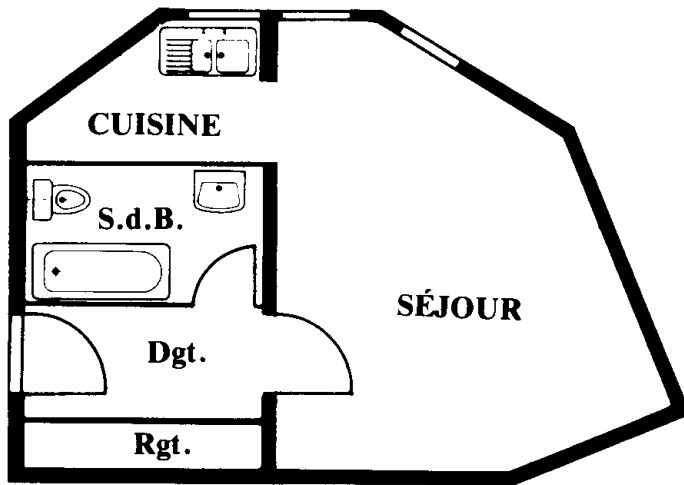
$$O(n^2.\ln(n))$$

→ R-signature discrète = somme au carré pour chaque direction

Application à des symboles de formes complexes et déconnectés

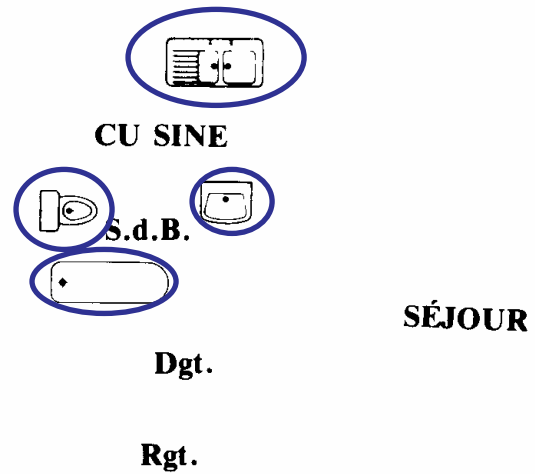


Plan d'un studio de 31,51 m²



Plan architectural

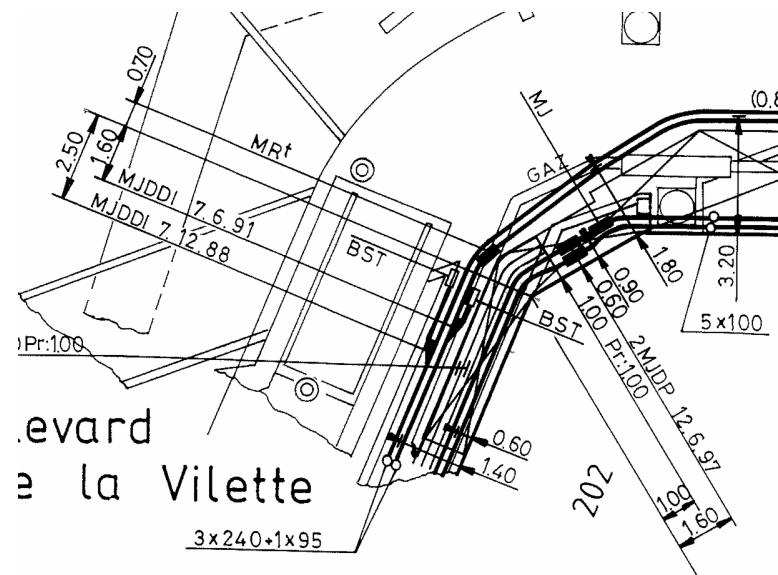
Plan d'un studio de 31,51 m²



Séparation texte/graphique

Problèmes majeurs rencontrés dans le document

- Symboles en continuités avec d'autres couches et/ou occultés



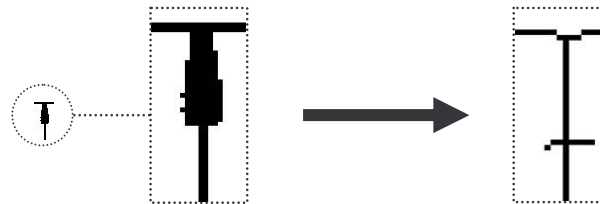


Détection de symboles quelconques

- *Application*: identification de flèches de cotation
- *Principe*: localiser rapidement les zones de l'image où des flèches peuvent se trouver, puis vérifier l'hypothèse à partir de la R-signature

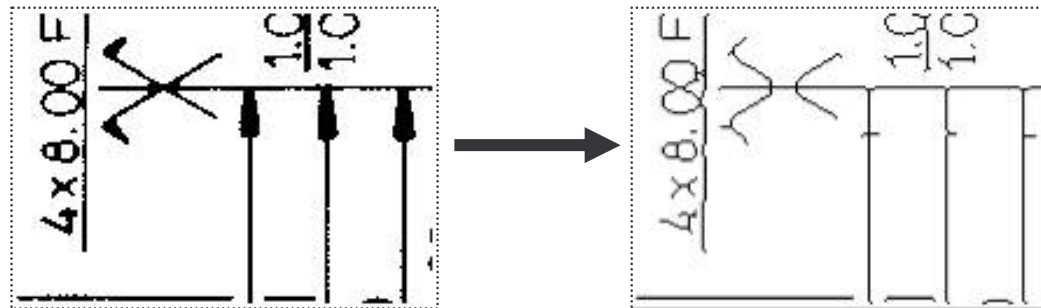
Détection de symboles (*suite*)

- Méthode : localisation des flèches à partir de points ayant une connexité multiples.



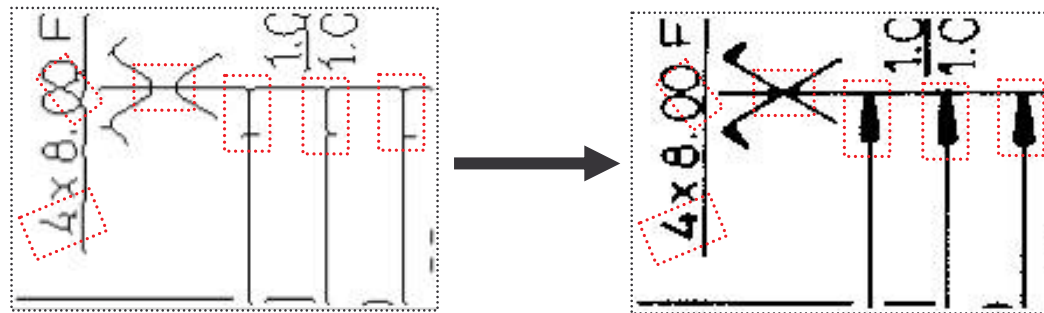
Détection de symboles (*suite*)

- Calcul du squelette du plan
 - détermination de tous les couples de points multiples appartenant à la même masse connexe.



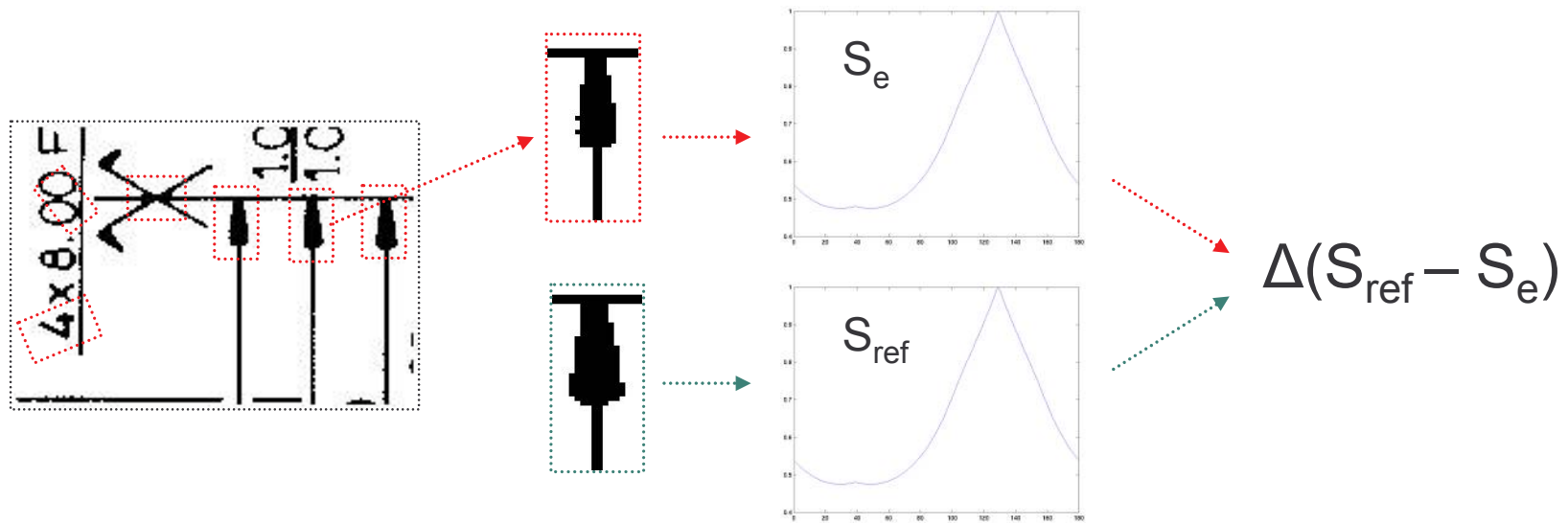
Détection de symboles (*suite*)

- Détermination d'un rectangle englobant à partir de deux points multiples



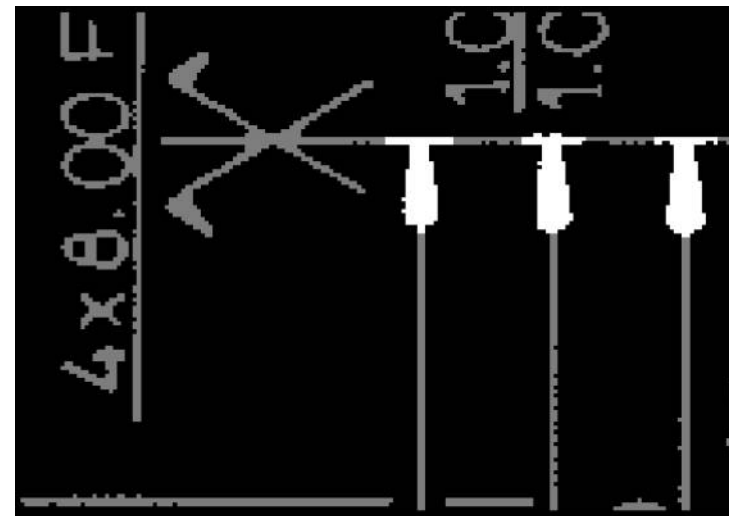
Détection de symboles (*suite*)

- Calcul d'une R-signature de la portion d'image correspondante
 - mise en correspondance avec la R-signature du symbole de référence



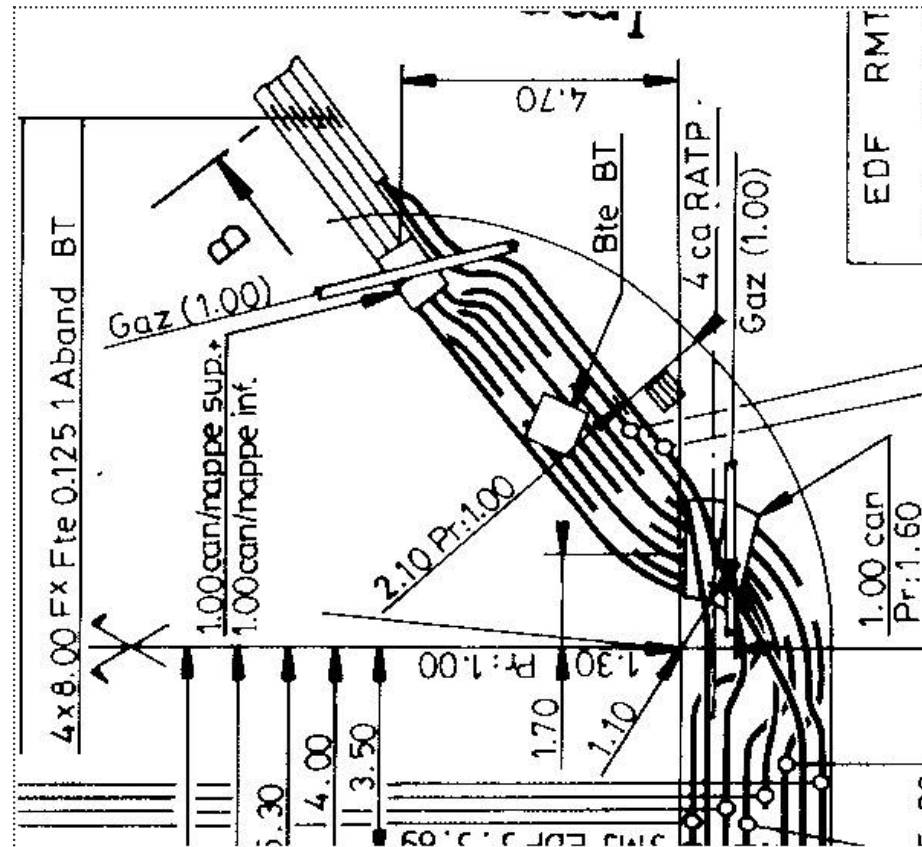
Détection de symboles (*suite*)

- A partir d'un seuil de référence, acceptation ou rejet du symbole à identifier

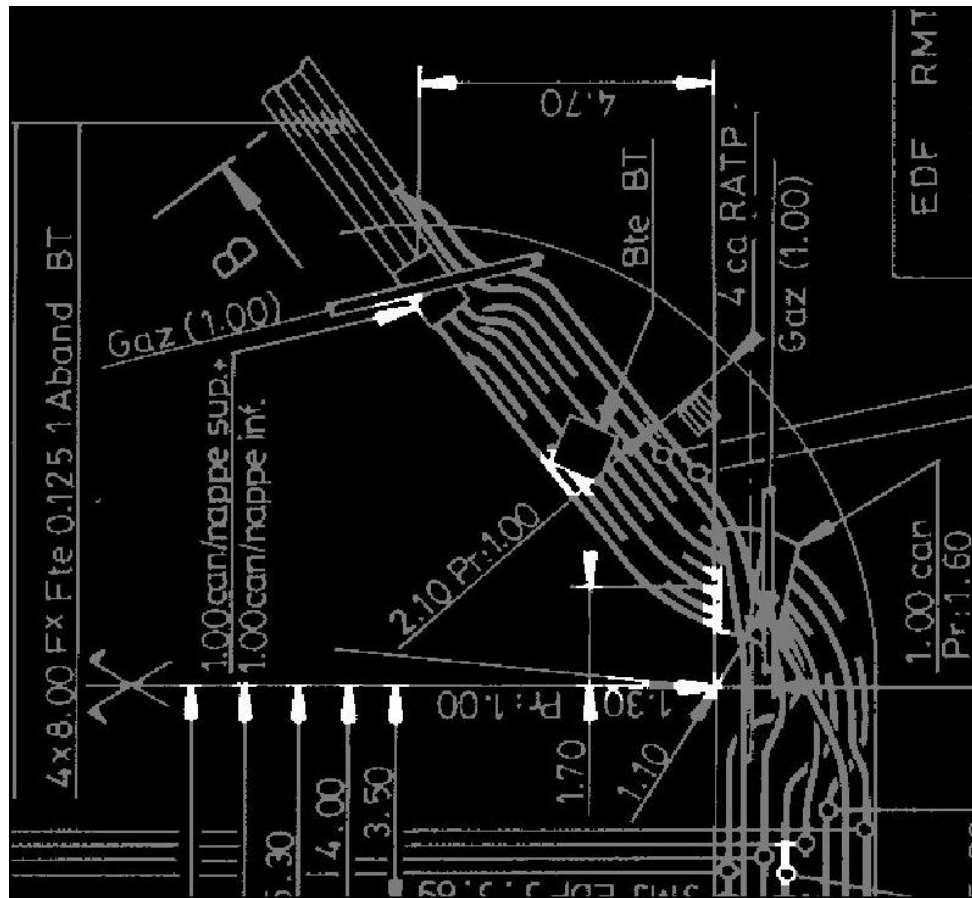


24 couples testés, 3 occurrences trouvées

Détection de symboles (*exemples*)

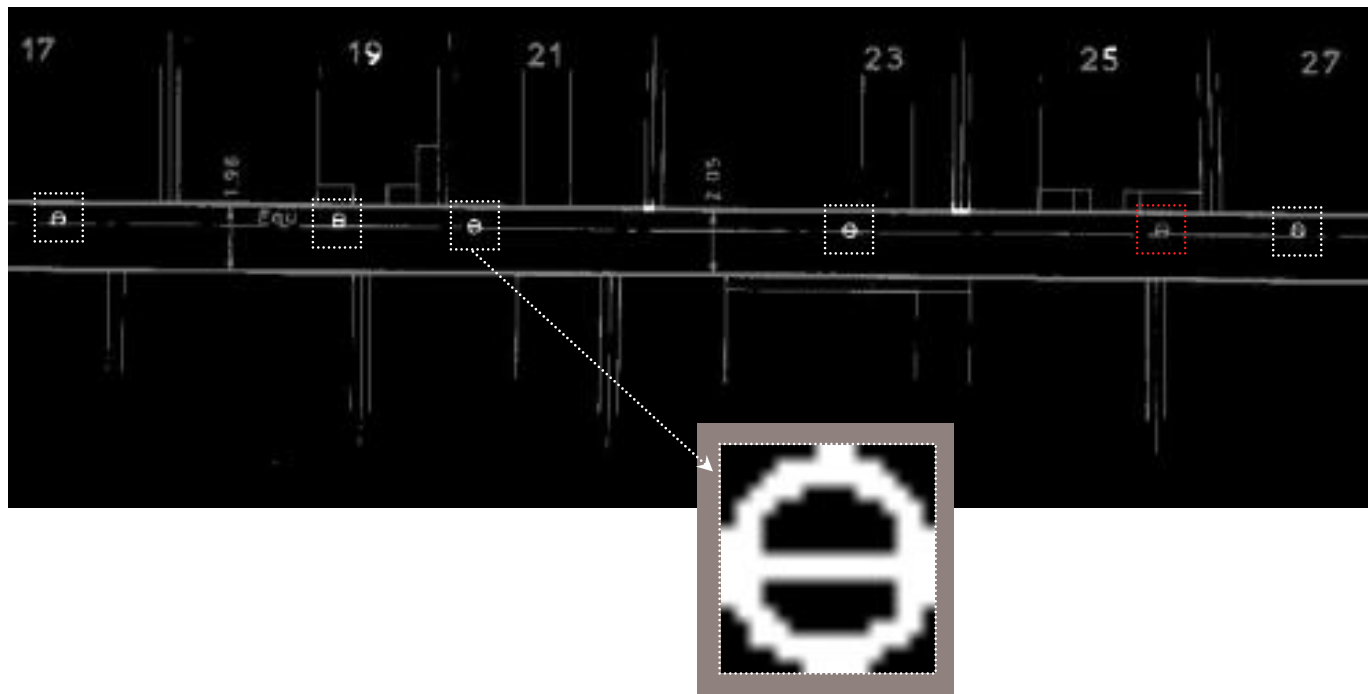


Détection de symboles (*exemples*)



Détection de symboles (*exemples*)

- Avec un autre symbole de référence



Une projection 1D est insuffisante pour extraire les redondances dans une large collection d'images

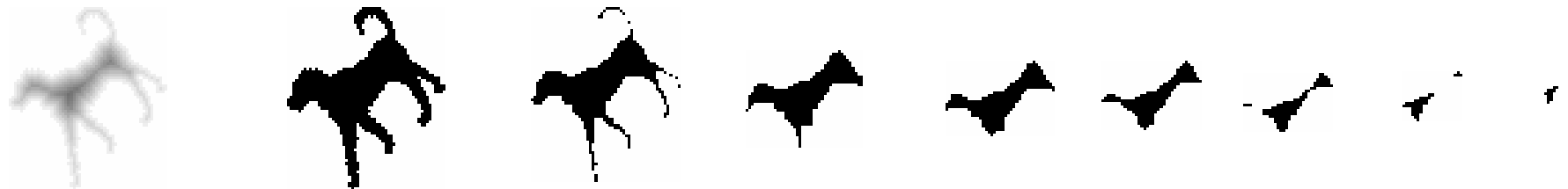
⇒ Projection dans l'espace de Radon pour différents niveaux de la transformée de distance (3-4)

4	3	4
3	0	

Avant

	0	3
4	3	4

Arrière





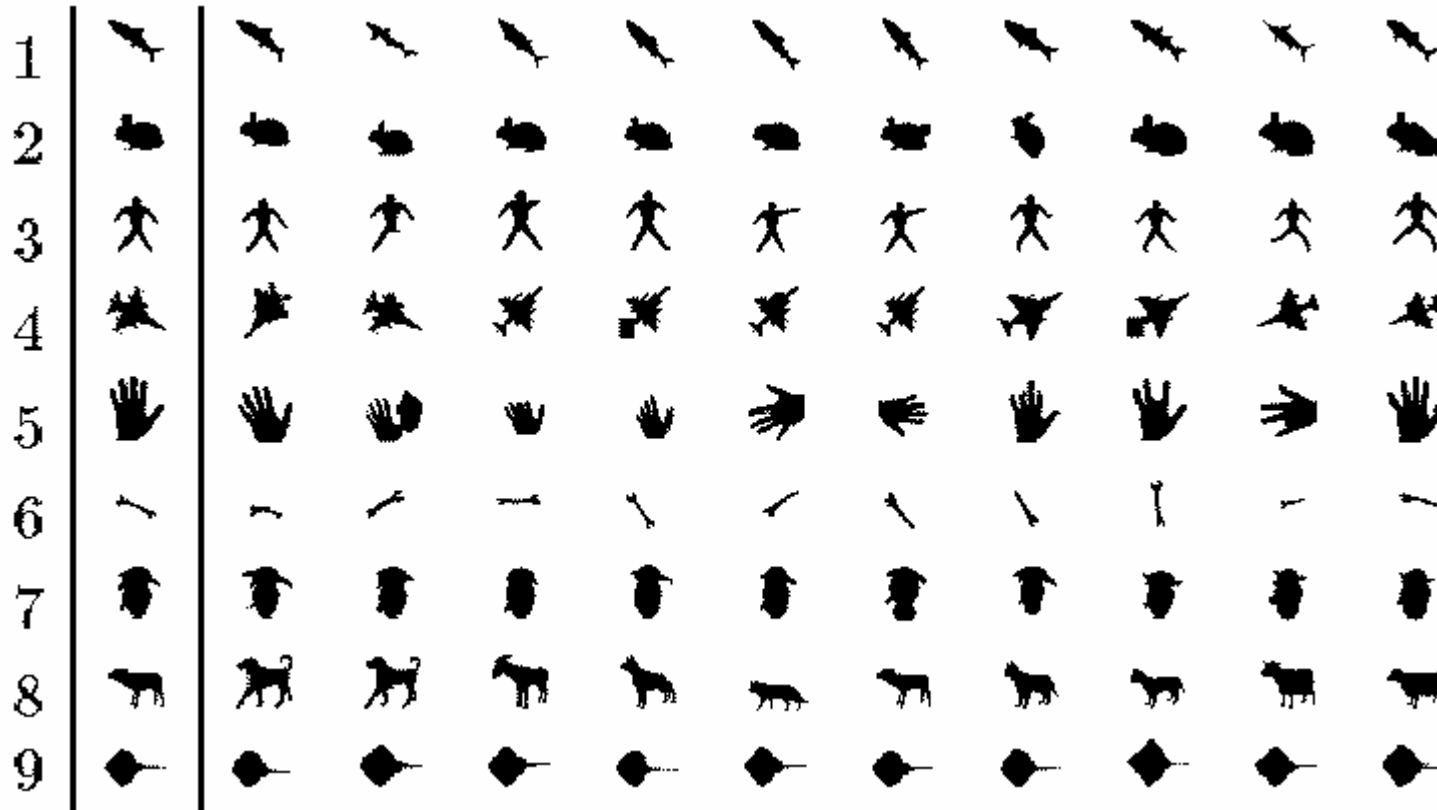
➤ Toute les transformées de Radon pour chaque image binaire sont calculées en une étape.

➤ Soit une forme sur laquelle nous avons calculé la transformée de distance

➤ Soit L_i l'image binaire contenant les pixels dont la valeur de la transformée est inférieure ou égale à i : $L_0 \subset L_1 \subset \dots \subset L_N$

où L_N correspond à la forme initiale et L_0 au niveau de segmentation le plus bas

➤ Tous les pixels d'un niveau n'apparaissant pas à un niveau inférieur sont projetés dans les matrices de Radon des niveaux supérieurs



Exemple sur une base de données de 99 objets



1										
2										
3										
4										
5										
6										
7										
8										
9										

Résultats de la mise en correspondance



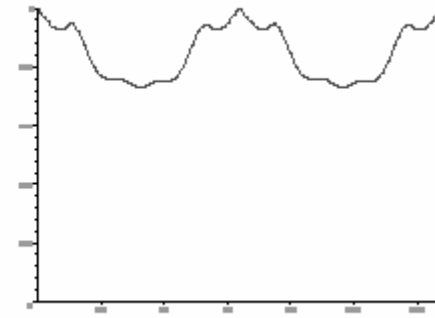
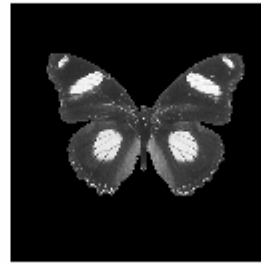
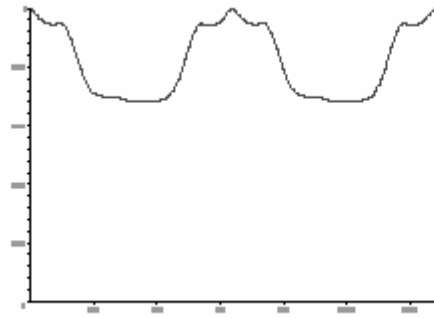
Classe	1	2	3	4	5	6	7	8	9
Taux(%)	77	76	100	68	92	60	75	61	98

- Taux moyen d'identification: 78 %
- Taux moyen d'identification sans transformée 3-4: 58%



Extension aux objets photométriques (ng et couleurs)

- ✓ **Une signature 2D n'est pas assez pertinente:**
 - ✓ **Effet de « compactage »**
 - ✓ **Sur la forme de l'objet**
 - ✓ **Sur la photométrie de l'objet**



Exemple de signatures 2D



Extension aux objets photométriques

Pour tenir compte des variations photométriques:

✓ Ajout d'une dimension supplémentaire

Soit O un objet composé de $\#ng$: $O = \prod_{i=1}^{\#ng} \lambda_i$ avec $\lambda_{\#ng} \subseteq \lambda_{\#ng-1} \subseteq \dots \subseteq \lambda_1$

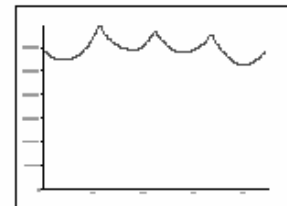
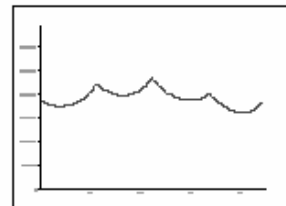
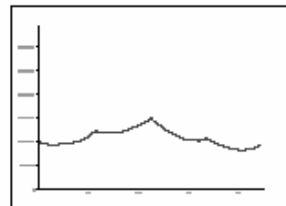
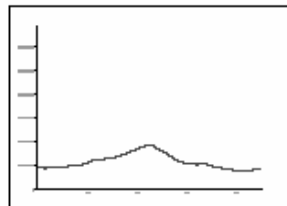
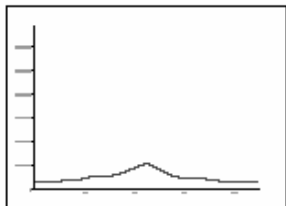
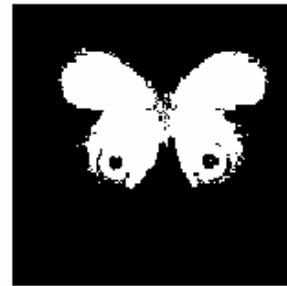
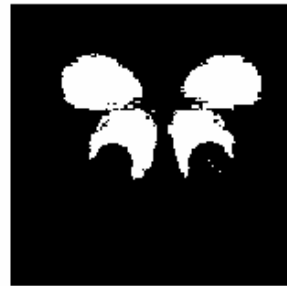
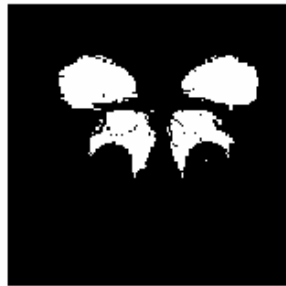
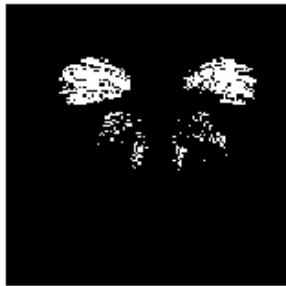
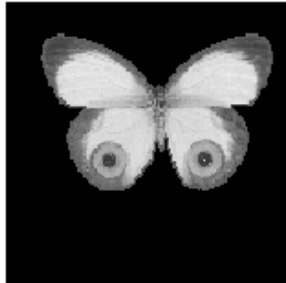
$$R_{3D} = \prod_{i=1}^{\#ng} R_{\lambda_i}$$



Extension aux objets photométriques (ng et couleurs)

Mise en œuvre:

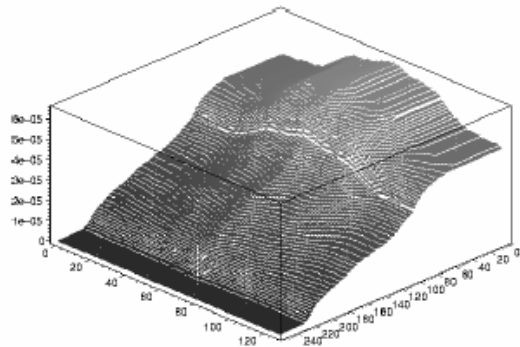
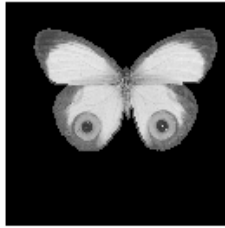
- **Seuillage successives**
- **Approche point (comme dans le binaire)**
 - Pixels => Coupes de niveaux inférieurs
 - Un seul parcours de la TR



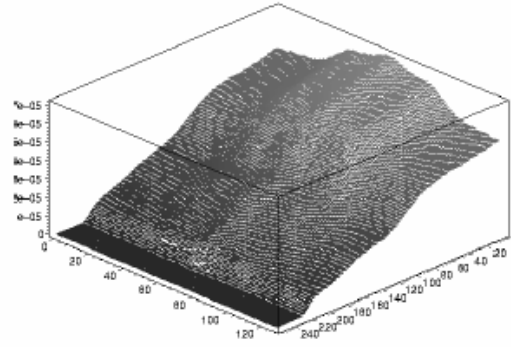
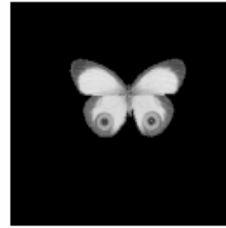
Exemple de découpage par niveaux



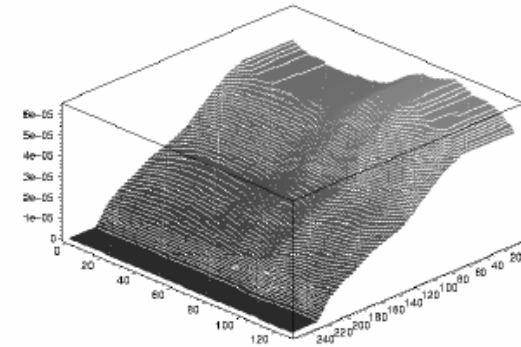
pap13_1



pap13_3

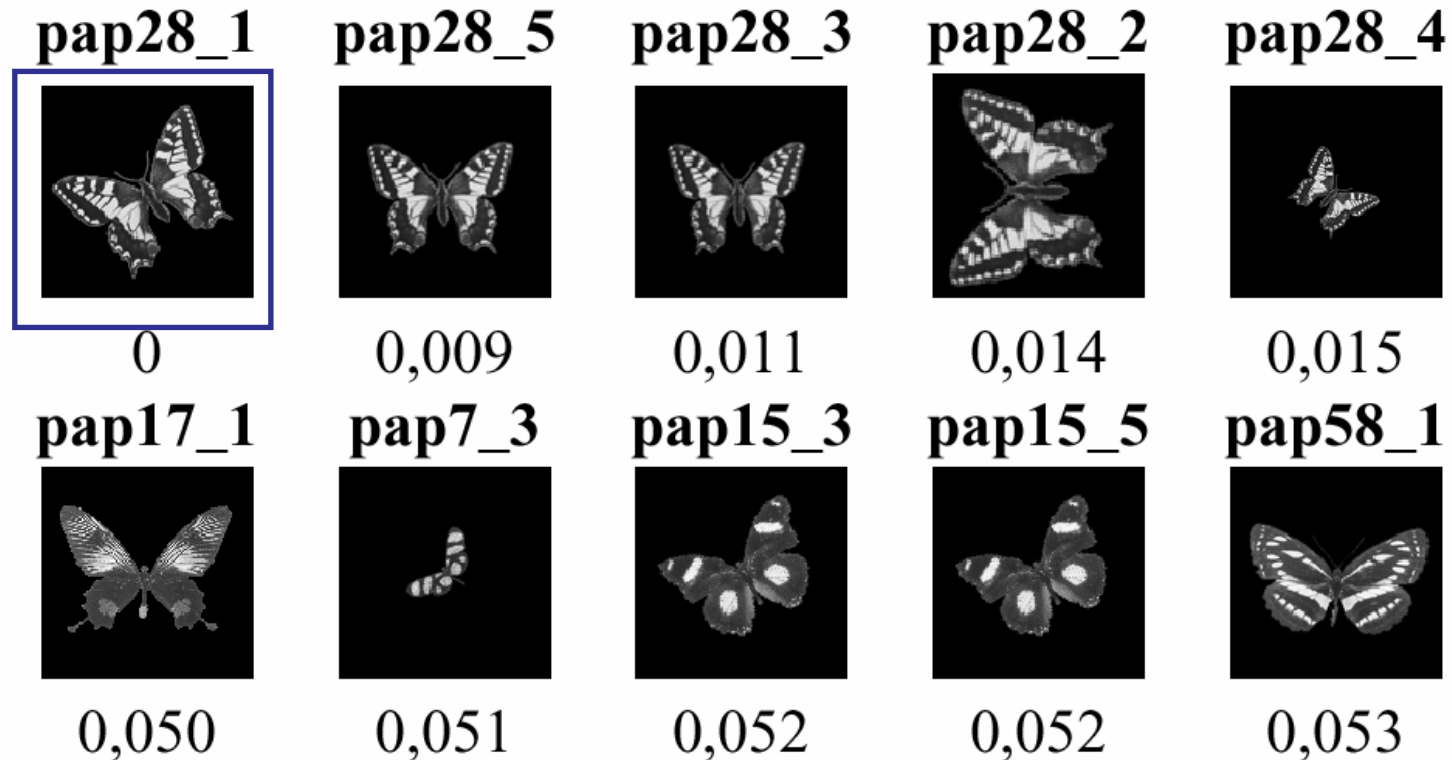


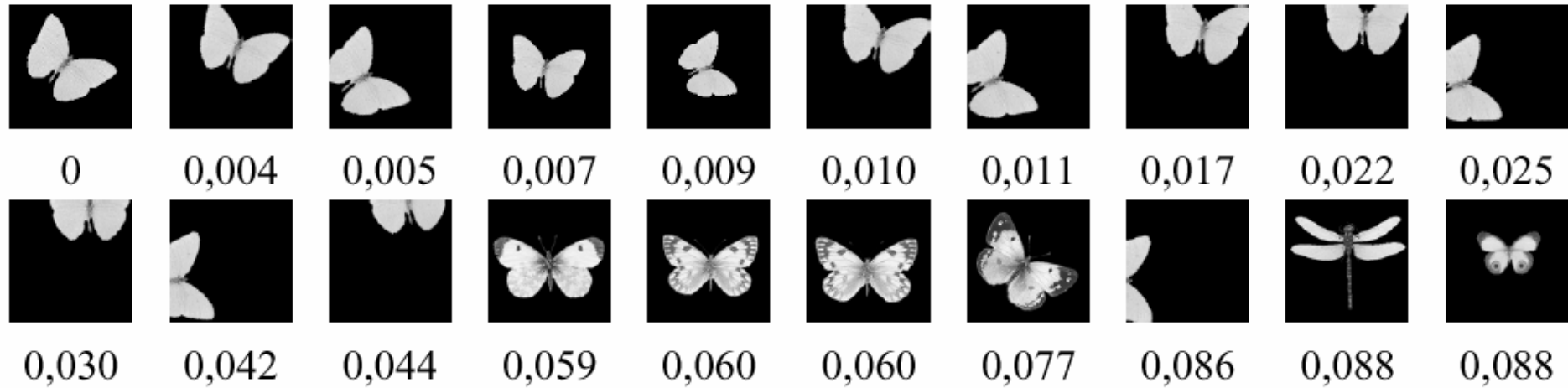
pap13_2



Propriétés géométriques conservées































Exemple sur une base de papillons (une centaine d'images)





Exemple de comportement aux occultations

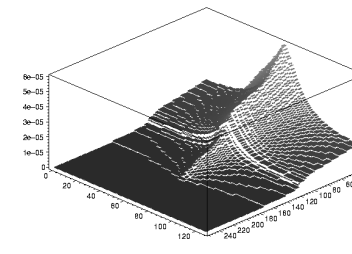
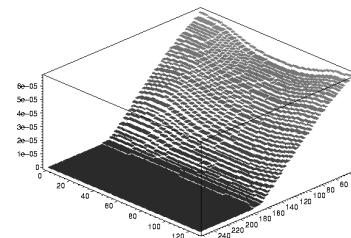
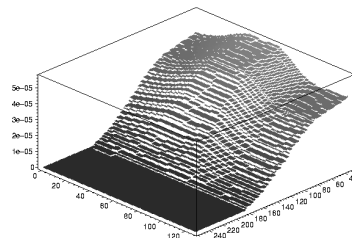
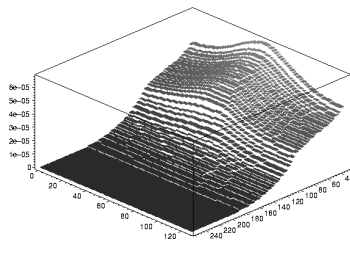
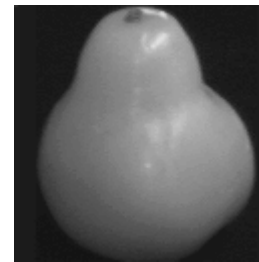
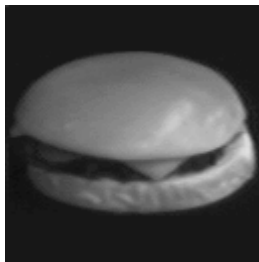


<i>HI</i>	p15_2  0	p15_3  $2E^{-15}$	p15_5  $2E^{-15}$	p12_2  $3E^{-12}$	p1_1  $4E^{-12}$	p15_1  $8E^{-12}$	p57_1  $9E^{-12}$	p12_1  $1E^{-11}$	p15_4  $3E^{-11}$	p1_2  $4E^{-11}$
<i>FM</i>	p15_2  0	p15_1  0,14	p15_4  0,42	p15_5  0,42	p15_3  0,42	p58_1  0,77	p8_1  0,83	p1_2  0,84	p59_1  1,08	p68_1  1,11
<i>R_{3D}</i>	p15_2  0	p15_6  0,007	p15_1  0,007	p15_3  0,021	p15_5  0,021	p15_4  0,023	p58_1  0,032	p17_2  0,033	p12_2  0,033	p17_1  0,036































Etudes comparatives

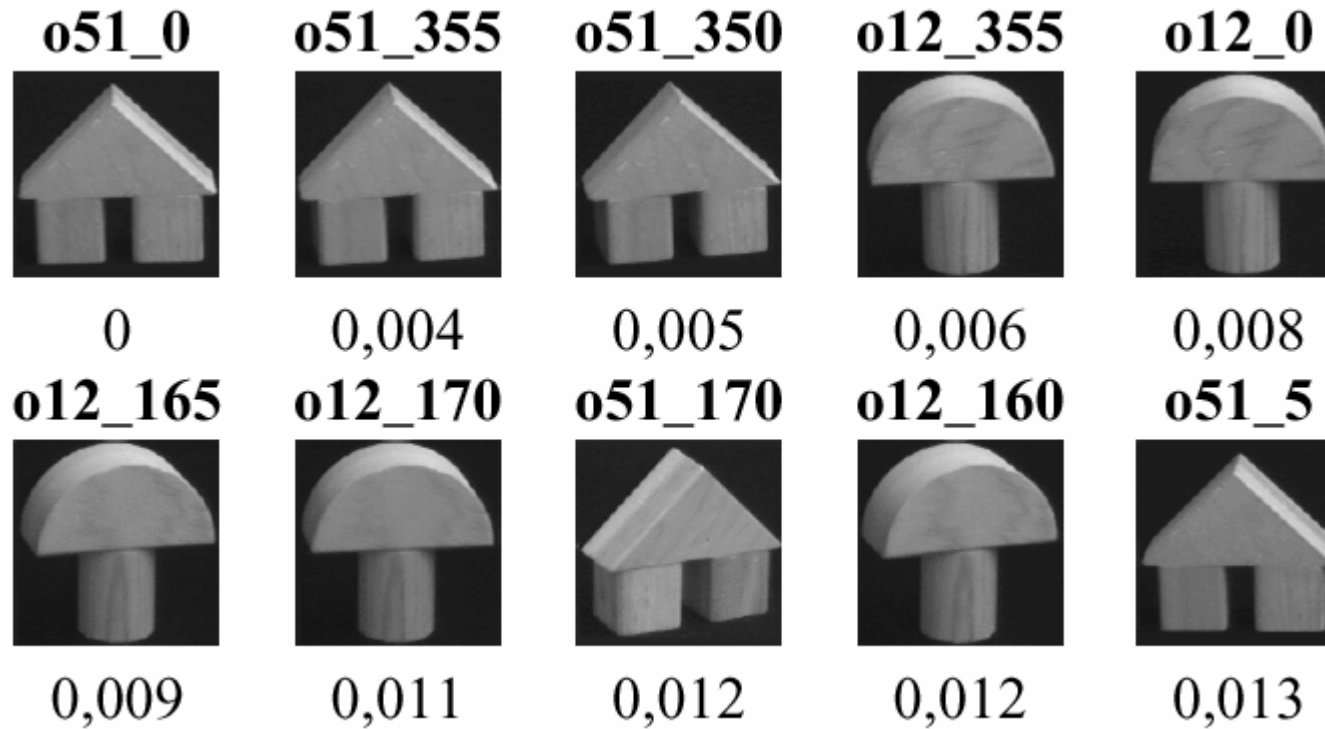
Autre exemple : base COIL (100 objets X 72 vues/objet)

■ Calcul de signatures 3D





<i>HI</i>	o5_64	o9_22	o6_6	o6_3	o10_4	o10_40	o10_41	o10_42	o10_43	o10_44
										
	0.	$2E^{-12}$	$2E^{-12}$	$2E^{-12}$	$2E^{-12}$	$3E^{-12}$	$3E^{-12}$	$3E^{-12}$	$3E^{-12}$	$3E^{-12}$
<i>FM</i>	o5_64	o5_63	o5_4	o9_2	o5_61	o5_36	o9_1	o9_61	o5_30	o5_31
										
	0	2,01	2,40	2,49	2,59	2,62	2,65	2,67	2,68	2,72
<i>R3D</i>	o5_64	o5_66	o5_65	o5_67	o5_9	o5_8	o5_63	o5_7	o5_0	o5_68
										
	0	0,005	0,007	0,012	0,012	0,016	0,017	0,018	0,022	0,022





Images en couleurs

- Matrice RGB
- Normalisation des plans (diagonales)

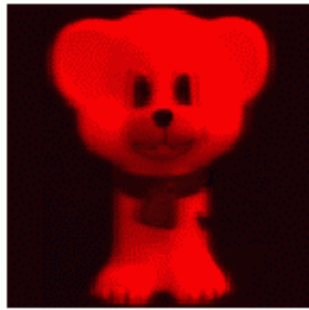
$$r(R, G, B) = \frac{R}{R + G + B}$$

- Liens entre les plans

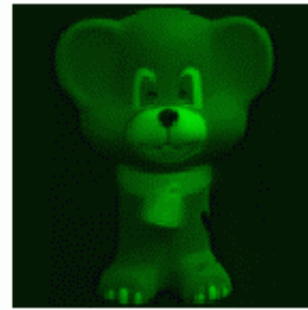
$$l_{RG} = \frac{(R - G)^2}{(R - G)^2 + (R - B)^2 + (G - B)^2}$$



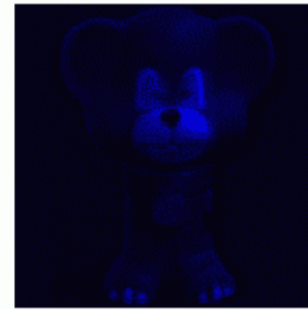
Image couleur



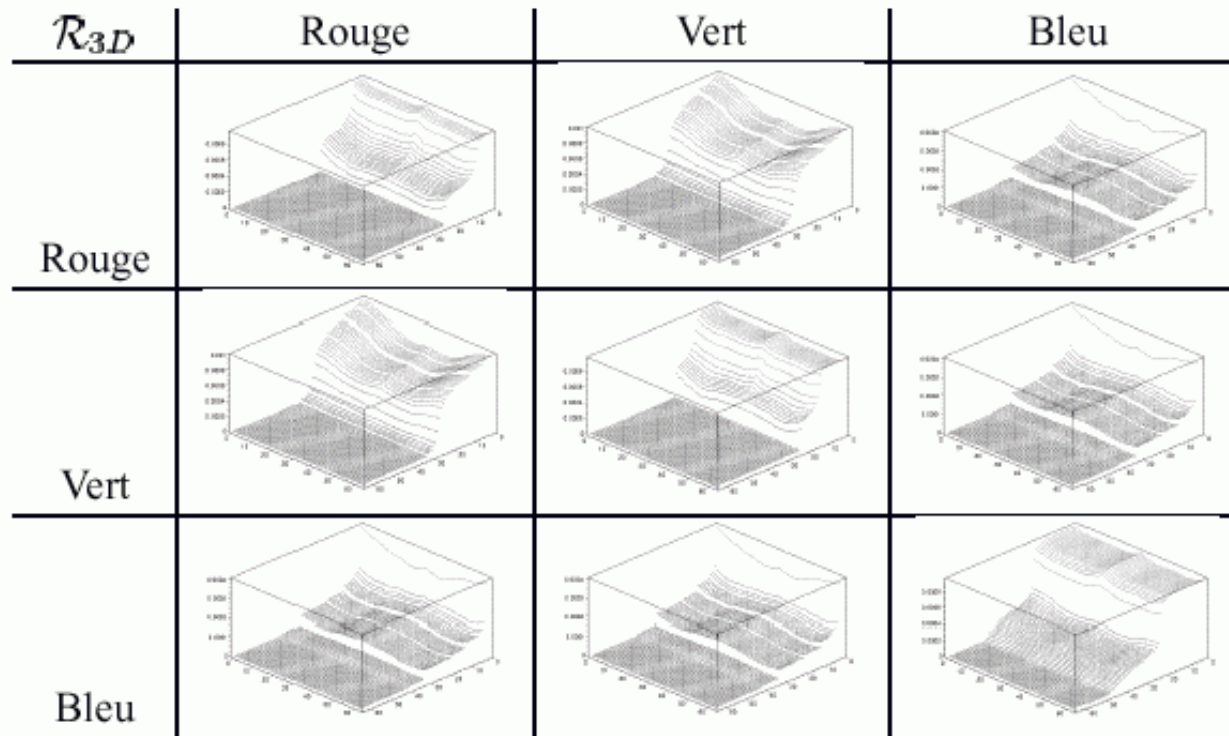
Plan rouge



Plan vert

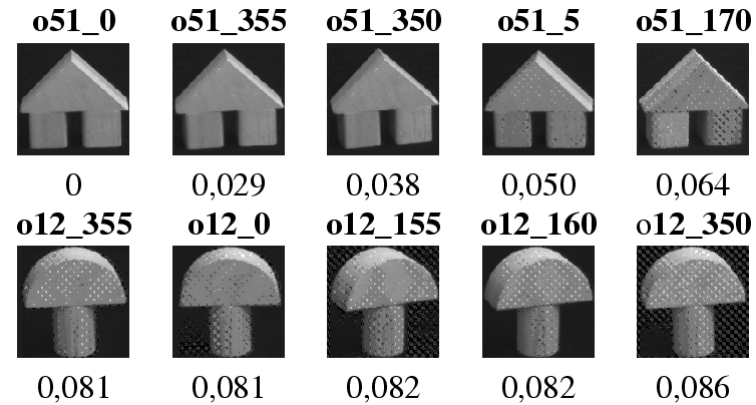


Plan bleu

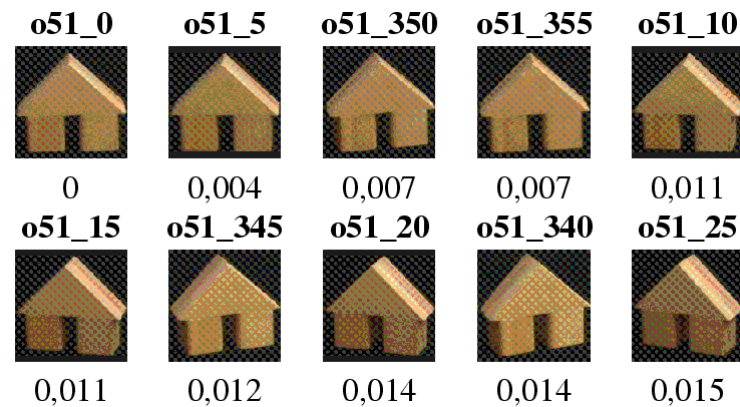




Un exemple qui donne de moins bon résultats




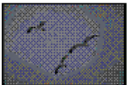
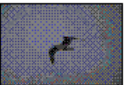
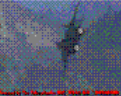
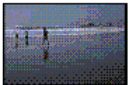
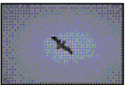

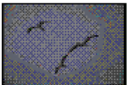
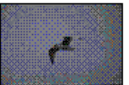
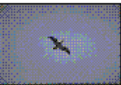

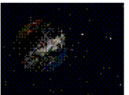





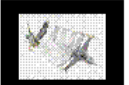


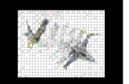








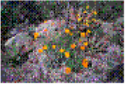






Amélioration par la couleur

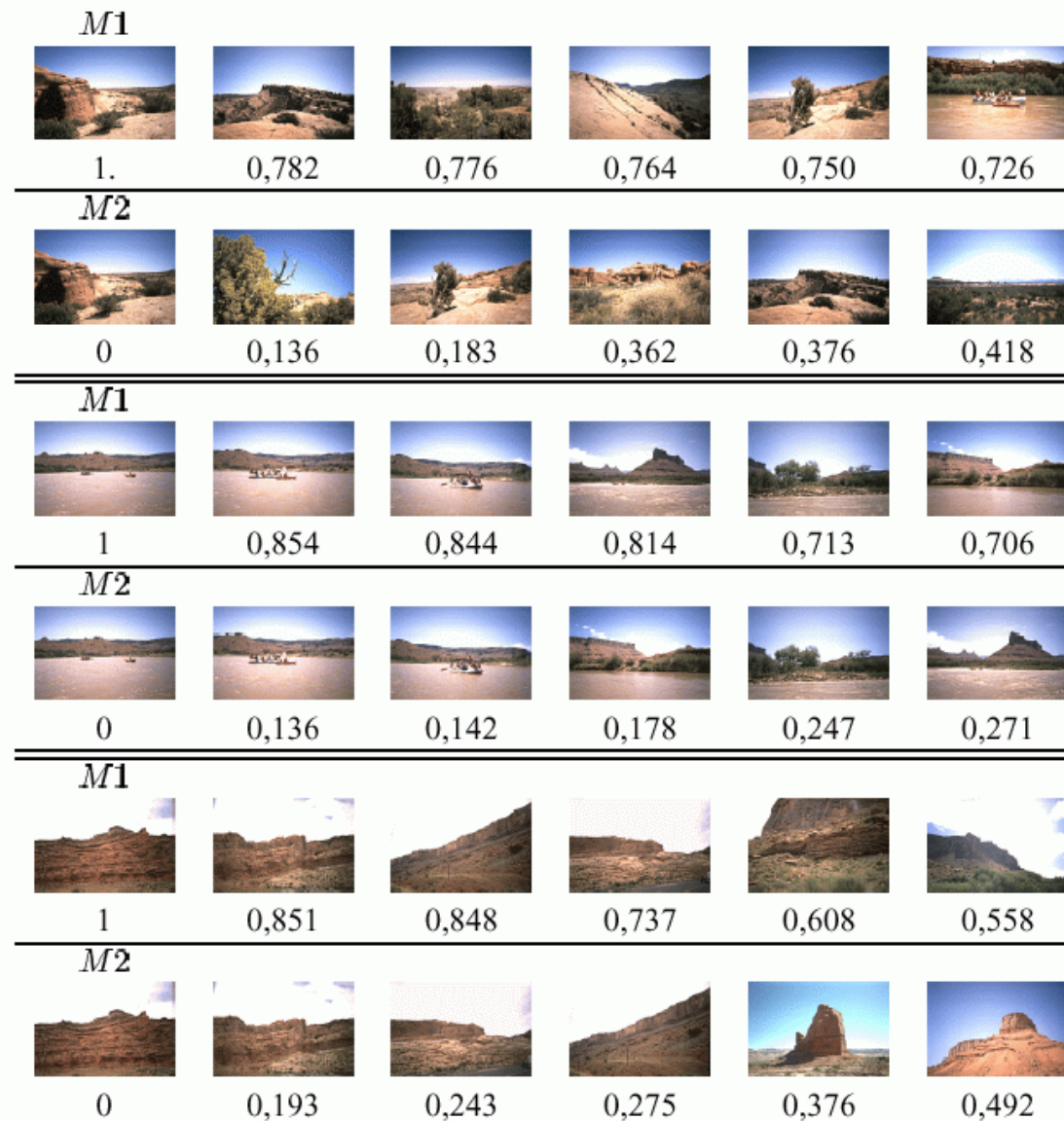


Images contenant des objets dominants

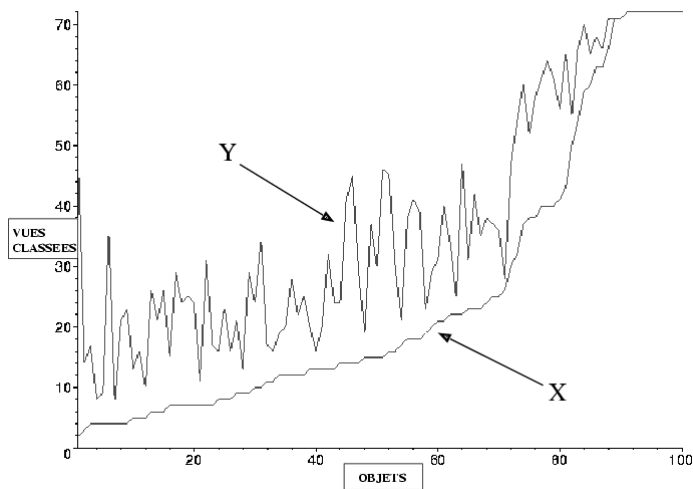
M1: méthode de Swain et Ballard

M2: notre approche

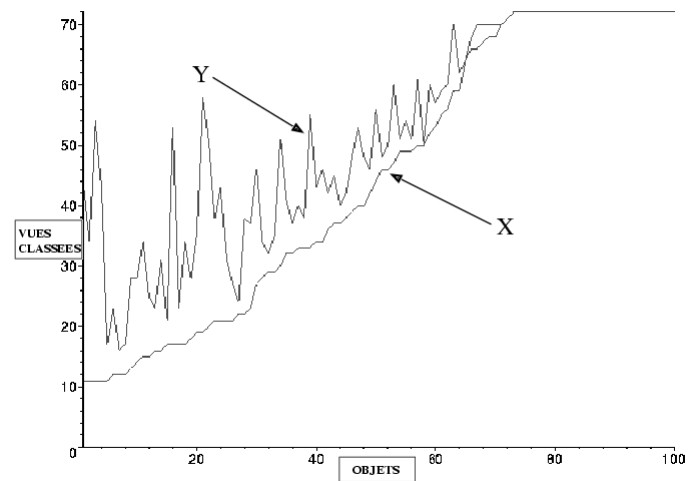
<i>M1</i>						
						
1	0,951	0,846	0,589	0,537	0,488	
<i>M2</i>						
						
0	0,213	0,253	0,304	0,354	0,376	
<i>M1</i>						
						
1.	0,849	0,643	0,610	0,587	0,574	
<i>M2</i>						
						
0	0,568	0,573	0,574	0,578	0,788	
<i>M1</i>						
						
1.	0,754	0,681	0,657	0,614	0,591	
<i>M2</i>						
						
0	0,208	0,284	0,350	0,679	0,687	



Analyse de performances en terme d'identification



Niveaux de gris



Couleur



Conclusions et perspectives

- Une adaptation de la transformée de Radon donnant des résultats prometteurs pour l'identification de formes (n&b, n.g et couleurs)
- Améliorer la pré-segmentation de l'image
 - rechercher de caractéristiques dominantes (points clés, zones dominantes)
- Passage à l'échelle pour indexer