

# Descripteur de contours pour la reconnaissance d'objets

## Contour descriptor for object recognition

S. Bourgeois<sup>1,2</sup>

S. Naudet-Collette<sup>2</sup>

M. Dhome<sup>1</sup>

<sup>1</sup> LASMEA, UMR 6602 du CNRS - Université Blaise Pascal - 63177 Aubière

<sup>2</sup> CEA Saclay - DRT/LIST/DTSI/SARC/LCEI - 91191 Gif sur Yvette cedex

steve.bourgeois@cea.fr

sylvie.naudet@cea.fr

michel.dhome@lasmea.univ-bpclermont.fr

### Résumé

*Cet article présente un descripteur local de contours s'appuyant sur les contours fermés. L'invariance de ce descripteur aux transformations affines lui permet d'être utilisé dans le cadre de reconnaissance d'objets tridimensionnels ou de mise en correspondance entre différentes vues d'une même scène.*

*Si de nombreux descripteurs locaux robustes aux transformations affines existent, ceux-ci exploitent généralement l'information de texture. Ce descripteur se différencie donc en exploitant exclusivement l'information de contours, et offre ainsi une solution pour la reconnaissance d'objets peu texturés tels que les objets métalliques rencontrés en milieu industriels.*

### Mots Clef

Reconnaissance d'objets, descripteurs locaux, contours.

### Abstract

*This paper introduces a local descriptor based on closed contour. Its affine invariance makes this descriptor adapted for 3D object recognition or wide baseline correspondances.*

*If many local descriptors with affine-transformations robustness exist, those one usually exploit texture informations. Our descriptor's originality comes from the fact that it exploits contour information exclusively. Consequently this descriptor offers a solution for textureless object recognition, case of many metallic object encountered in industrial environment.*

### Keywords

Object Recognition, local descriptor, contour descriptor, textureless.

## 1 Introduction

Les descripteurs locaux sont de plus en plus utilisés en reconnaissance d'objets. Le descripteur SIFT de Lowe [1] et le descripteur de Mikolajczyk et al.[2] en sont quelques exemples.

On remarquera que la grande majorité des descripteurs locaux exploite principalement l'information de texture. Or de telles approches ne sont pas adaptées à des objets métalliques avec aucune texture et une importante spécularité.

La texture et la couleur n'étant pas exploitable sur ce type d'objet, nous nous sommes orienté vers une approche s'appuyant sur les contours. Plus particulièrement, nous avons développé un descripteur local de contour. Celui-ci repose sur l'hypothèse que l'objet à reconnaître présente des contours fermés, ce qui est généralement le cas sur les pièces mécaniques (nombreux alésages,...).

Nous débuterons en présentant un rapide état de l'art (partie 2). Nous poursuivrons avec la présentation d'un descripteur de contours s'appuyant sur des contours fermés(partie 3). Enfin nous préciserons la mise en oeuvre de ce descripteur (partie 4) avant de présenter les premiers résultats (partie 5).

## 2 Travaux antérieurs

Nous présentons ici un rapide état de l'art sur les descripteurs de contour. Le descripteur proposé partageant à la fois des caractéristiques des approches Shape Matrix, Histogramme d'Orientation des Contours et de la Transformée de Fourier Généralisée, nous détaillerons principalement ces trois approches.

### 2.1 Approche Shape Matrix

La notion de Shape Matrix a été introduite par Goshtasby [3] en 1985 puis repris par Taza et al. [4]. Initialement, cette approche ne vise pas à décrire des contours mais des régions. Le principe est simple : on place au centre d'une région une grille radiale de taille  $M*N$  (ie. composée de  $N$

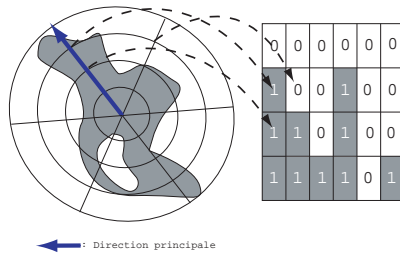


FIG. 1 – Shape Matrix

secteurs et  $M$  anneaux) et pour chaque points d'intersection rayon/cercle de cette grille, on regarde l'appartenance de celui-ci à la région (cf. Fig.(1)). On lui associe la valeur 1 si il appartient à la région, 0 dans le cas contraire. Ces résultat sont alors regroupés au sein d'une matrice binaire(cf. Fig.(1)).

Le fait de centrer la grille sur le centre de la région rend la méthode robuste aux translations. Pour obtenir l'invariance au changement d'échelle, le rayon du cercle le plus grand est défini comme le rayon du cercle circonscrit à la région, et le nombre de secteurs et cercles constituant la grille est fixe. Quant à l'invariance aux rotations, une direction principale doit être établie. Cette direction peut être définie par le point le plus éloigné du centre de gravité, bien que cette solution ne soit pas d'une grande robustesse. Pour ajouter la robustesse aux transformations affines, Zuliani et al. [5] exploite la matrice de covariance associée à la région. De plus, sa méthode ne caractérise plus une région mais les contours externe de celle-ci, ce qui fait des Shape Matrix un descripteur de contour.

Chalechale et al. proposent également une extension du Shape Matrix en appliquant celui-ci aux contours [6]. Il n'obtient pas la robustesse affine mais améliore la robustesse aux rotations en utilisant une transformée de Fourier. De plus, la Shape matrix n'est plus binaire mais pondérée par le nombre de pixel de contour par secteur, rendant ainsi la méthode moins sensible aux bruits et contours parasites. Carmichael et al. proposent une méthode appelée Edge Probe Cascade [7] s'inspirant de la technique des Shape matrix. Il présente une méthode d'échantillonnage mieux adaptée à la description de contours en introduisant la densité de contours dans le voisinage du point échantillonné. D'autre part, pour augmenter la robustesse de sa méthode, Carmichael remplace la matrice binaire ainsi que la méthode de comparaison de matrice par un arbre de probabilité [8].

Belongie et al. proposent une méthode appelée Shape Context [9, 10, 11] s'appuyant sur les Shape Matrix. Le Shape Contexte a pour principe d'échantillonner un contour en  $N$  points, et pour chacun de ces points appliquer un shape matrix centré sur celui-ci. Belongie et al. ne caractérisent alors plus un contour par une signature mais une collection de signatures. On notera que si le Shape Context présente de bons résultats pour la reconnaissance de forme isolée, ses capacités à reconnaître un objet dans un environ-

nement complexe sont pour le moment réduites en raison de la difficulté à estimer le facteur d'échelle à appliquer aux shape matrix [12].

## 2.2 Approche Histogramme des Orientations des Contours (Edge Histogram Direction)

Le premier descripteur de contours exploitant l'orientation des contours fut introduit par Jain et al. [13]. Celui-ci fut développé pour la recherche d'images dans une base de donnée. Ce descripteur consiste en un simple histogramme des orientations des contours pour l'image entière. L'histogramme représente donc une discrétisation de l'espace angulaire ( $N$  secteurs angulaire) et chaque contour incrémente le secteur angulaire comportant son orientation.

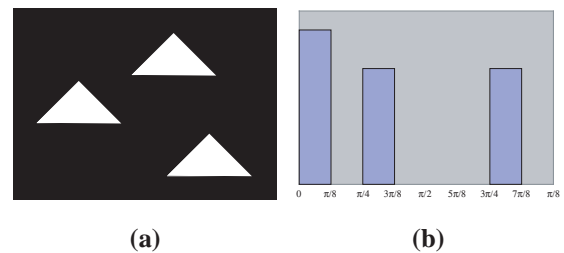


FIG. 2 – (a) image. (b) Histogramme des orientations des contours de (a)

La position des contours n'entrant pas en compte dans l'histogramme, ce descripteur est invariant aux translations. La normalisation de l'histogramme rend la signature invariante aux changement d'échelle. Il souligne cependant que cette normalisation rend impossible la détection d'un objet occulté. On notera que cet inconvénient vient avant tout de la nature globale (à opposer à la notion de descripteur local) de cette méthode.

Le point faible de ce descripteur concerne sa robustesse aux rotations. En effet, comme le souligne Jain et al. [13], une rotation de l'objet entraîne un décalage des angles dans l'histogramme, décalage d'autant plus complexe car lié à la discrétisation de l'espace angulaire de l'histogramme. Jain propose l'application d'un lissage sur l'histogramme afin d'autoriser de petites rotations.

On notera aussi que ce descripteur n'exploite pas la répartition spatiale des contours.

## 2.3 Approches mixtes

Comme nous l'avons vu plus haut, l'approche Shape Matrix s'appuie uniquement sur la distribution spatiale des contours alors que l'approche Edge Histogram Direction s'appuie sur la dispersion angulaire des contours. Evans et al. ont introduit le Pairwise Geometric Histogram (PGH) [14, 15, 16], une approche combinant à la fois l'orientation et la position des contours. Le PGH s'appuie sur l'écart angulaire entre deux contours et la distance orthogonale entre les points de ces deux mêmes contours.

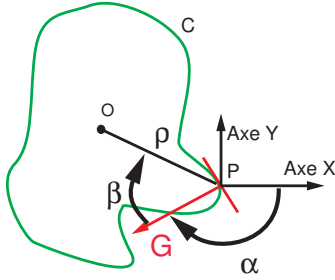


FIG. 3 – Principe de la transformée de Hough Généralisée

Mikolajczyk et al. [17] a combiné l'approche Shape Context [10] et une adaptation pour les contours du descripteur SIFT [1]. Le descripteur obtenu est robuste aux rotations et changements d'échelles uniformes.

## 2.4 Transformée de Hough Généralisée

La Transformée de Hough fut tout d'abord introduite pour la détection de droites, puis pour la détection de courbes 2D paramétrisées. Une version généralisée de la Transformée de Hough (GHT) fut introduite dans [18, 19], permettant la détection de formes quelconques. Cette méthode crée, lors d'une phase d'apprentissage, une description du contour sous la forme d'une *R-Table*.

Avant d'expliquer l'algorithme de la GHT, voici quelques définitions :

- Soit  $C$  le contour que l'on désire détecter.
- Soit  $O$  le point de référence associé à  $C$ .
- Soit  $P$  un pixel du contour  $C$ .
- Soit  $\vec{G}$  la normal à  $C$  au point  $P$
- Soit  $\alpha$  l'angle formé par l'axe des abscisses et  $\vec{G}$
- Soit  $\beta$  l'angle formé par  $\vec{G}$  et  $\vec{PO}$
- Soit  $\rho = \|\vec{PO}\|$

Lors de la phase d'apprentissage, tout point  $P \in C$ , on estime  $\alpha$ ,  $\beta$  et  $\rho$ . Ces paramètres sont alors stockés dans la *R-Table* sous la forme d'une règle :

$$\alpha \rightarrow (\rho, \beta)$$

On obtient alors une table de la forme :

$$\begin{aligned} \alpha_1 &: (\rho_1^1, \beta_1^1), (\rho_1^2, \beta_1^2), \dots, (\rho_1^{n_1}, \beta_1^{n_1}) \\ \alpha_2 &: (\rho_2^1, \beta_2^1), (\rho_2^2, \beta_2^2), \dots, (\rho_2^{n_2}, \beta_2^{n_2}) \\ \dots &: \dots \\ \alpha_m &: (\rho_m^1, \beta_m^1), (\rho_m^2, \beta_m^2), \dots, (\rho_m^{n_m}, \beta_m^{n_m}) \end{aligned}$$

Le nombre de règles associées à  $\alpha_i$  est égale au nombre  $n_i$  de pixels de contour présentant l'angle  $\alpha$  (eg. P1 et P2 sur la figure 4).

Lors de la phase de reconnaissance, une matrice d'accumulateurs de taille  $N_x, N_y$  est associée à l'image. L'angle  $\alpha$  est estimé pour chaque pixel de contour de l'image, et un

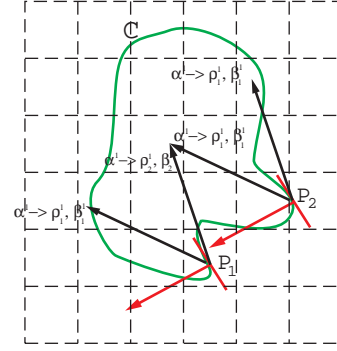


FIG. 4 – Procédure de vote du GHT

vote est ajouté à chaque accumulateur désigné par la règle  $(\rho^i, \beta^i)$  associée à  $\alpha$  (cf. Fig. (fig :HoughGeneraliseVote)). A la fin du processus, l'accumulateur associé à la position de  $O$  est sensé avoir reçu les votes de tous les points de  $C$ . On estime alors que l'accumulateur présentant le plus de votes est le meilleur candidat pour le point  $O$ .

On notera que la GHT telle que l'on vient de la présenter présente néanmoins de nombreuses faiblesses. Tout d'abord, la représentation du contour sous la forme d'une *R-Table* n'est ni robuste aux rotations et encore moins aux changements d'échelles.

## 2.5 Autres approches

De nombreuses approches existent pour décrire les contours : les Chain Codes [20], la transformée de Fourier [21], les approximations polygonales [22] ou la Curvature Scale Space [23] qui a été inclus dans la norme MPEG-7, etc. Cet article ne visant pas à présenter un état de l'art complet de ce domaine, le lecteur est invité à se référer aux articles suivants [24, 25, 26].

## 3 Un nouveau descripteur de contours : DTHGI

Le descripteur que nous proposons se classe dans les approches mixtes. Il calcule sa signature sur les contours appartenant à un voisinage d'un contour fermé. Pour obtenir une robustesse aux transformations affines, nous proposons une extension de la notion de forme d'un contour, proposée dans [5], à la notion de forme d'un voisinage d'un contour fermé. Nous proposons une nouvelle signature sur les contours robuste aux rotations. Celle-ci présente de nombreuses similarités avec le récent descripteur RIFT de Lazebnik et al. [27] utilisé pour la caractérisation de textures.

### 3.1 Forme du voisinage d'un contour fermé

La robustesse affine est un point crucial pour les descripteurs locaux [5, 28, 29]. Pour l'obtenir, nous reprenons la notion de forme d'un contour fermé proposé dans [5] pour l'étendre à la notion de forme d'un voisinage d'un contour fermé.

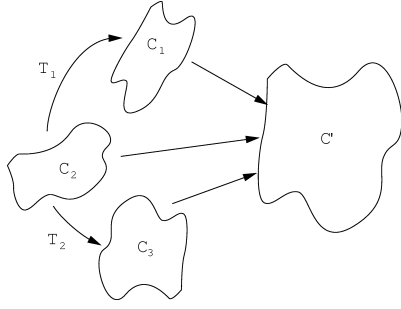


FIG. 5 –  $C_i$  : Courbes de Jordan,  $C'$  : La forme des  $C_i$ ,  $T_i$  : Transformation affine

**Forme d'un contour fermé et invariance aux transformations affines.** Pour obtenir une robustesse aux transformations affines d'une courbe, le principe consiste à déterminer un espace permettant de ramener toutes courbes fermées  $C_i$  liées par une transformation affine vers une représentation unique ou quasi unique.

Zuliani et. al ont proposé une solution dans [5] pour les courbes de Jordan<sup>1</sup>. A toute courbe de Jordan  $C$  est associée une courbe  $C'$  nommée *forme de la courbe  $C$*  (cf. Fig(5)). Cette courbe  $C'$  est définie de manière unique, modulo une Rotation/Symétrie, pour toute courbe  $C_i$  liées par une transformation affine à  $C$ .

Nous rappelons le formalisme proposé par Zuliani et. al [5] en assimilant un contour fermé à une courbe de Jordan.

*Théorème 1 (dit Théorème de Jordan) :*

*Une courbe de Jordan sépare le plan en deux parties : l'intérieur et l'extérieur de la courbe.*

Grace au Théorème 1, nous pouvons poser les définitions suivantes :

- Soit  $C$  une courbe de Jordan
- Soit  $\Omega \subset R^2$  la région intérieure de  $C$ .
- Soit  $A(\Omega) = \int_{\Omega} dx^2$  l'aire de la région  $R$ .
- Soit  $m_{pq}(\Omega) = \frac{1}{A(\Omega)} \int_{\Omega} x^p y^q dx dy$  où  $(x, y) \in \Omega^2$  le moment d'ordre  $p + q$ .
- Soit  $\Sigma(\Omega) = \frac{1}{A(\Omega)} \int_{\Omega} (X - m_{11}(\Omega))(X - m_{11}(\Omega))^T dX^2$  la matrice de covariance de  $\Omega$  avec  $X = (x, y)$

On définit alors la forme d'une courbe de Jordan  $S(C)$  de la manière suivante :

$$S(C) = \left\{ s \in R^2 / s = ((\Sigma(\Omega))^{-1})^{1/2}(x - m_{11}), x \in C \right\} \quad (1)$$

*Théorème 2 :*

<sup>1</sup>Une courbe de Jordan est une courbe simple fermée ne s'intersectant pas elle même

Soit  $C_1$  et  $C_2$  deux courbes de Jordan reliées par une transformation affine. Alors la forme de  $C_1$  et la forme de  $C_2$  sont géométriquement équivalents modulo une transformation  $T$  dont la matrice est orthonormale.

On se référera à [30, 28] pour la démonstration formelle de ce théorème.

Ce théorème montre que la forme d'une courbe de Jordan est définie de manière unique modulo une Rotation/Symétrie.

**Voisinage d'un contour fermé.** L'objectif étant d'estimer une signature sur le voisinage d'un contour fermé, il faut définir une notion de voisinage  $V(C)$  d'une courbe de Jordan  $C$  qui soit intrinsèque à  $C$  quelque soit la transformation affine  $A$  appliquée à  $C$  :

$$V(A.C) = A.V(C)$$

- Soit  $\lambda_1$  la valeur plus grande valeur propre de  $\Sigma(\Omega)$  et  $\lambda_2$  la plus petite.
- Soit  $\mu$  le vecteur propre associé à  $\lambda_1$  et  $\tau$  l'angle de  $\mu$  par rapport à l'axe des abscisses.
- Soit  $O = m_{1,1}(\Omega)$  le centre de gravité de  $\Omega$ .
- Soit  $f = \frac{\sum_{x_i \in \Omega} (\sqrt{(X_i - O)(\Sigma(\Omega))^{-1}(X_i - O)^T})}{Card(X_i)}$ .
- Soit  $E_C^1 = ellipse(O, \lambda_1, \lambda_2, \mu)$  l'ellipse associée à  $\Sigma(\Omega)$ .
- Soit  $E_C^2 = ellipse(O, f.\lambda_1, f.\lambda_2, \mu)$  l'ellipse approximant  $C$ .
- Soit *factor* une constante indiquant la taille du voisinage par rapport à la taille du contour fermé (la taille du voisinage est proportionnelle à *factor*).
- Soit  $E_C^3 = ellipse(O, f.\lambda_1.factor, f.\lambda_2.factor, \mu)$

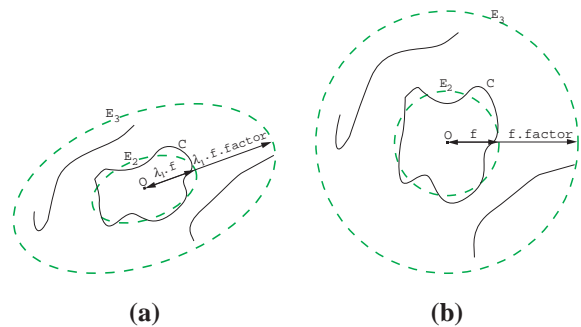


FIG. 6 – (a) Voisinage de  $C$ . (b) Forme du voisinage de  $C$

On définit alors le voisinage de  $C$  (cf. Fig(6)) par :

$$V(C, factor) = \{X_i / X_i \in E_3\} \quad (2)$$

*Démonstration :*

Montrons tout d'abord que si  $x \in E_C^1$  alors  $Ax \in E_{AC}^1$ .  
Puisque  $x \in E_C^1$ , on a

$$(x - o) \sum (\Omega)^{-1}(x - o) \leq f \quad (3)$$

Il reste à montrer que

$$(Ax - Ao)^T \sum (A\Omega)^{-1}(Ax - Ao) \leq f \quad (4)$$

Or [28] démontre que :  $\sum(A\Omega) = A \sum(\Omega)A^T$ .

D'où :

$$\begin{aligned} & (Ax - Ao)^T \sum (A\Omega)^{-1}(Ax - Ao) \\ &= (x - o)^T A^T (A \sum(\Omega)A^T)^{-1} A(x - o) \\ &= (x - o)^T A^T A^{-T} \sum(\Omega)^{-1} A^{-1} A(x - o) \\ &= (x - o)^T \sum(\Omega)^{-1}(x - o) \\ &\leq f \end{aligned}$$

On a donc montré que  $x \in E_{A\Omega}^1$ .

D'une manière similaire on démontre aisément que

$$\forall x \in E_{AC}^1, A^{-1}x \in E_C^1$$

$V(C)$  et  $V(AC)$  étant de simples changements d'échelles de  $E_C^1$  et de  $E_{AC}^1$ , on en déduit que  $AV(C) = V(AC)$

**Extension de la notion de forme d'un contour à son voisinage.** On remarquera alors que cette notion de *forme* peut aisément être étendue aux contours voisins sous l'hypothèse que ces contours appartiennent au même plan que le contour de départ. Il s'agit donc d'émettre l'hypothèse que l'objet est localement plat, ce qui est une hypothèse généralement admise dans le cadre des descripteurs locaux.

Le voisinage étant défini par une ellipse (cf.(2)), la forme d'un voisinage correspond à un disque. Elle se caractérise donc par un centre  $O$  et un rayon  $f$  factor (cf. Fig.(6)).

### 3.2 Descripteur par Transformée de Hough Généralisée Inversée (DTHGI)

A présent que nous disposons d'une notion de forme de voisinage d'un contour fermé, il nous suffit d'estimer sur celui-ci soit une signature invariante aux rotations/symétries, soit deux signatures si celle-ci n'est qu'invariante aux rotations (cf. Théorème 2).

Nous proposons un descripteur pour le voisinage d'un contour fermé que nous appelons DTHGI. Le descripteur proposé est proche du descripteur Pairwise Geometric Histogram (PGH) de Evans [14, 15, 16]. Tout comme le PGH, le DTHGI correspond à un histogramme d'orientation des contours en fonction de la distance à un centre. La

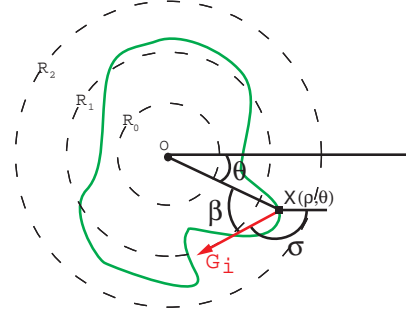


FIG. 7 – Principe du DTHGI

différence réside principalement dans le choix de l'angle et de la distance, différence qui a une influence forte sur les propriétés du descripteur.

Tout d'abord, commençons par quelques définitions :

- Soit  $C$  la forme d'un voisinage,  $O$  son centre, et  $R$  son rayon.
- Soit  $N_c$  le nombre d'anneaux concentriques partitionnant  $C$ .
- Soit  $R_i$  le  $i$ ème anneau :  $C = \bigcup_0^{N_c-1} \{R_i\}$
- Soit  $A = [0, \pi]$  le demi espace angulaire.
- Soit  $A_i$  une partition de  $A$  en  $N_a$  secteurs angulaires :  $A = \bigcup_0^{N_a-1} \left\{ A_i = \left[ \frac{i*\pi}{N_a}, \frac{(i+1)*\pi}{N_a} \right] \right\}$
- Soit  $\Omega = \{X_i = (\rho, \theta) \in [0, R] \times [0, 2\pi[ \}$  l'ensemble des points de contour contenus dans  $C$  ( $(\rho, \theta)$  étant les coordonnées polaires de  $X_i$ ).
- Soit  $\vec{G}_i$  le vecteur gradient du point de contour  $X_i \in \Omega$ .
- Soit  $\sigma \in A$  l'orientation de  $\vec{G}_i$  dans le repère global.
- Soit  $\beta = \pi - (\sigma - \theta)$

On définit alors le descripteur  $DTHGI \in R^{N_c \times N_a}$  par :

$$DTHGI_{i,j} = \frac{Card(\{X \in \Omega / \rho \in R_i, \beta \in A_j\})}{Card(\Omega)} \quad (5)$$

pour  $i = 0, 1, 2, \dots, N_c$  et  $j = 0, 1, 2, \dots, N_a$

De manière moins formelle, on peut considérer le DTHGI comme une concaténation des histogrammes des orientations  $\beta_i$  obtenu pour chaque anneau. On retrouve bien la double distribution spatiale (les anneaux) et angulaire (les  $\beta_i$ ) qui caractérise les approches mixtes.

Afin d'améliorer la robustesse du descripteur aux erreurs de positionnement des  $X_i$  (erreur provenant de

l'étape de segmentation), on associe à chaque point de contour une densité de probabilité gaussienne. Cette densité de probabilité a pour conséquence un lissage de l'histogramme selon les deux paramètres  $\rho$  et  $\beta$ . D'autre part, le découpage en anneaux est réalisé à surface constante (ie. chaque anneau couvre la même surface).

**L'Invariance aux rotations du DTHGI.** On remarque que pour un point  $X \in \Omega$  l'angle  $\beta$  est exprimé dans le repère local défini par le rayon OX. Cet angle étant local, sa valeur n'est pas affectée par les rotations. De même, la distance utilisée étant la distance euclidienne, elle non plus n'est pas sensible aux rotations. On en déduit que le descripteur DTHGI est invariant aux rotations par construction.

On remarquera par contre que l'ensemble des mesures est effectuée par rapport au point O. La qualité de la signature obtenue dépend donc de la précision avec laquelle a été obtenue la position de O. Cependant, si l'erreur reste relativement faible, la discrétisation de l'espace des paramètres ainsi que le lissage permettent de la compenser.

**Similitude avec la Transformée de Hough Généralisée.** Il suffit de comparer les figures (7) et (3) pour constater une grande similitude entre celles-ci.

En effet, l'algorithme du GHT tout comme celui du DTHGI sont basés sur la mesure de l'angle  $\beta$  et de la distance  $\rho$ . Cependant, la GHT et le DTHGI partagent des objectifs inverses. Si le GHT vise à déterminer la position du centre O à partir de points de contours, le DTHGI vise lui à caractériser les contours voisins de ce même point O. De manière symbolique on peut dire que la GHT va donc du contour vers O, alors que le DTHGI va de O vers les contours.

On notera au passage qu'il est possible de combiner le DTHGI et le GHT. En effet, il suffit d'associer une DTHGI  $D_i$  à chaque accumulateur  $A_i$  du GHT. Lors de l'exécution du GHT, pour chaque vote reçu par  $A_i$  via la règle  $\alpha \rightarrow (\rho, \beta)$ , il suffit d'incrémenter  $D_i(\rho, \beta)$ . A l'issue de cette GHT modifiée, chaque  $D_i$  représente la distribution des votes pour l'accumulateur  $A_i$  en fonction de  $\rho$  et  $\beta$ . L'objet recherché possédant une distribution des votes caractéristique, celle-ci peut être utilisée pour identifier les accumulateurs solutions. Le critère de score de l'accumulateur n'est alors plus l'unique critère de sélection, la distance entre distribution de vote venant compléter celui-ci.

## 4 Mise en oeuvre

### 4.1 Extraction des contours fermés

La première étape consiste à extraire les contours de l'image. Cette extraction est réalisée grâce à un filtre de Deriche. Afin d'aider à l'extraction de contour fermés dans les phases suivantes, un lissage par ACP est réalisé sur

l'orientation des gradients en sortie du filtre de Deriche. Un seuillage par hysteresis, une phase de "non-maximal suppression" et une phase de chaînage des pixels de contour permettent d'obtenir les contours sous la forme de listes ordonnées de pixels. Le lissage par ACP sur l'orientation a pour but d'éviter les "cassures" qui surviennent habituellement au niveau des coins lors de la phase de "non-maximal suppression", ce qui augmente le nombre de contours fermés extraits. Pour améliorer l'extraction des contours, des prétraitements de réhaussement adaptés de contraste ou des méthodes de seuillage par hysteresis adaptées peuvent être employées.

Des contours obtenus on ne retient que les contours fermés. On peut alors définir le voisinage ainsi que la forme du voisinage de ceux-ci. On calcule alors le DTHGI sur la forme du voisinage de chaque contour fermé, ce qui nous fournit deux signatures pour chacun des contours fermés.

### 4.2 Calcul de la matrice de Covariance à partir du contour

On remarque que les contours étant issus d'un filtre de Deriche, les régions associées aux contours fermés ne sont pas explicitement connues. Pourtant, (1) et (2) font intervenir la matrice de covariance  $\sum(\Omega)$ , avec  $\Omega$  la région associée au contour fermé.

En général, cette matrice est estimée en réalisant un parcours des pixels de  $\Omega$ . Cette solution n'est pas immédiatement applicable à notre problème puisque nous ne connaissons pas  $\Omega$  mais uniquement son contour C (les contours étant obtenus par un détecteur de Deriche).

On peut trouver dans [31] une solution pour estimer cette matrice à partir des points de contour de celle-ci. Pour cela, il définit deux ensembles de pixels à partir des pixels du contour (cf. Fig.(8) :

– les pixels de contours gauches définis par

$$\partial\Omega- = \{(x, y) | (x-1, y) \notin \Omega, (x, y) \in \Omega\}$$

– les pixels de contours droits définis par

$$\partial\Omega+ = \{(x, y) | (x, y) \in \Omega, (x+1, y) \notin \Omega\}$$

Notons qu'avec cette définition un pixel peut appartenir :

- à  $\partial\Omega+$ .
- à  $\partial\Omega-$ .
- à  $\partial\Omega+$  et à  $\partial\Omega-$ .
- à aucun des deux ensembles.

Ces différentes situations sont visibles dans la figure (8).

Le calcul du moment  $m_{p,q}$  est alors ramener à la formule suivante :

$$m_{pq} = \sum_{\partial\Omega+} y^q \sum_{i=1}^x i^p - \sum_{\partial\Omega-} y^q \sum_{i=1}^{x-1} i^p \quad (6)$$

On notera cependant que la formulation de  $\partial\Omega-$  et  $\partial\Omega+$  fait encore une fois intervenir la notion d'appartenance à

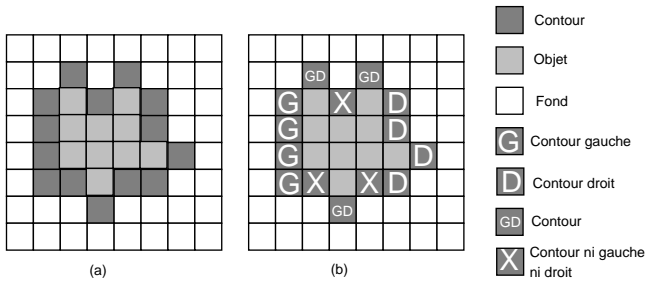


FIG. 8 – (a) : image originale, (b) décomposition du contour en contours droits et gauches

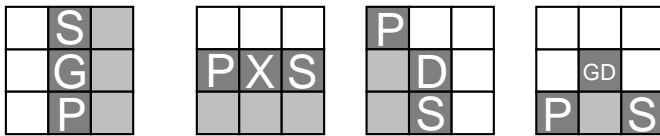


FIG. 9 – Quatre configurations parmi les 81 possibles (contour orienté dans le sens horaire). D : pixel droit, G : pixel Gauche, GD : pixel gauche et droit, X : pixel ni gauche ni droit P : pixel précédent, S : pixel suivant

$\Omega$ , impliquant du même coup la connaissance de  $\Omega$ . La solution proposée par Flusser n'est donc pas applicable immédiatement.

Pour ne pas avoir à définir  $\Omega$ , nous proposons un algorithme de calcul des contours gauches et droits basé sur l'hypothèse que le contour est orienté (ie. il existe un ordre de parcours des pixels du contour).

L'orientation du contour définissant de manière implicite la région, on peut déterminer si le pixel de contour est gauche ou droit à partir de la position relative du pixel précédent et suivant. Le problème se ramène donc à identifier pour chaque pixel du contour la configuration de celui-ci parmi les 81 configurations possibles.

La figure 9 représente 4 de ces configurations parmi les 81. Le pixel central correspond au pixel examiné et la lettre qui lui est associée indique à quelle groupe celui-ci appartient (G : groupe  $\partial(\Omega-)$ , D :  $\partial(\Omega+)$ , GD : à la fois  $\partial(\Omega-)$  et  $\partial(\Omega+)$ , X : le pixel n'appartient à aucun groupe). Les lettres P et S indiquent respectivement la position du pixel précédent et suivant. On notera que les exemples fournis sont valables pour un contour orienté dans le sens horaire.

On obtient donc  $\partial\Omega-$  et  $\partial\Omega+$  par un simple parcours du contour au cours duquel on identifie la configuration de chaque pixel du contour parmi les 81 possibles. On remarquera que le calcul de la matrice de covariance peut être, elle aussi, calculée au cours de cette même passe grâce à la formule (6). Le calcul de la matrice  $\sum(\Omega)$  présente donc une complexité en  $\mathcal{O}(n)$ , avec n le nombre de pixel du

contour.

### 4.3 Limitations

Le descripteur que nous proposons a été défini sous deux hypothèses :

- les objets sont localement plats
- les transformations supportées sont des transformations affines

Les conditions réelles ne permettant pas toujours de les vérifier, nous présentons ici les limitations induites par ces hypothèses.

**Hypothèse de transformation affine.** La notion de forme d'un contour étant définie pour des transformations affines, le descripteur proposé est robuste à ce même type de transformation. Cependant une prise de vue réalisée par une caméra induit des effets de perspectives. Dans ce cas, la transformation liant un même voisinage pris selon deux points de vues différents n'est pas affinité mais une homographie. Le calcul du descripteur sur des contours issus de prises de vues réelles n'est donc pas exact, le calcul du voisinage étant erroné ainsi que le calcul de la forme de ce dernier.

Cependant, dans le cadre de l'application visée (reconnaissance d'objets), les conditions de prises de vue (focale, distance à la pièce,...) induisent peu d'effets de perspectives. La signature est donc peu altérée par ce phénomène qui, d'autre part, est atténué par la discrétisation du descripteur. L'hypothèse d'affinité n'est donc pas limitante pour l'application visée et le descripteur peut donc être appliqué à des images réelles.

**Hypothèse de planarité du voisinage.** L'hypothèse de planarité locale est une hypothèse largement employée dans le domaine des descripteurs locaux. Cette hypothèse est justifiée par le fait que l'approximation d'une surface courbe par un plan est acceptable dès lors que le voisinage est suffisamment petit.

Cependant, ici la taille du voisinage est calculée en fonction de la taille du contour fermé étudié. Donc, dans la pratique, il peut arriver que le voisinage ne soit pas suffisamment local et que l'hypothèse de planarité ne soit pas vérifiée. La forme du voisinage calculée dans ces conditions perd sa robustesse aux changements de points de vues, comme l'illustre les figures 10.c et 10.d. On remarque que les contours (en blanc) appartenant au même plan que le contour caractérisé (en bleu) sont identiques (modulo les effets de segmentation) dans (c) et (d) alors que les contours appartenant au plan orthogonal (et donc brisant l'hypothèse de planarité) présentent des déformations (en rouge).

## 5 Résultats expérimentaux

### 5.1 Recherche dans une base de contour

La courbe de précision/rappel est une mesure généralement utilisée pour caractériser l'algorithme de recherche d'image. Cette courbe représente le rapport entre :

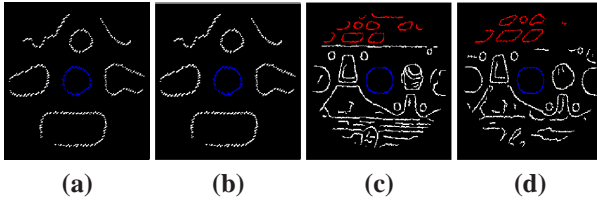


FIG. 10 – (a) et (b) : forme du voisinage avec respect de l’hypothèse de planarité. (c) et (d) : Forme de voisinage avec non respect de l’hypothèse de planarité (en rouge les contours déformés). En bleu, les contours sur lesquels sont calculés les descripteurs.

- La précision, qui correspond au rapport entre le nombre d’éléments correctement identifiés et le nombre d’éléments identifiés.
- Le rappel, qui correspond au rapport entre le nombre d’éléments correctement identifiés et le nombre d’éléments à identifier.

Le  $n^{me}$  point de la courbe indique les valeurs de précision et rappel sur les  $n$  meilleures identifications.

On peut dire qu’un descripteur est meilleur qu’un autre si sa courbe est toujours située au dessus de la courbe de l’autre descripteur.

Pour estimer les performances de notre descripteur, nous avons utilisé la base MCD (Multiview Data Curve) de Zuliani (cf. Fig.(11)). Il s’agit d’une base de silhouettes d’objets. Cette base couvre 40 objets avec 14 vues différentes pour chaque objet (les vues correspondant à des transformations affines de la silhouette initiale de l’objet).

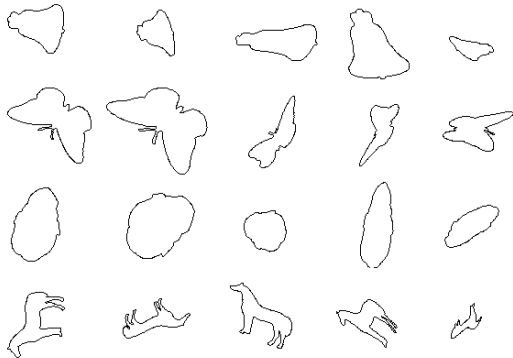


FIG. 11 – Exemples extraits de la base MCD

Nos résultats ont été obtenus avec un DTHGI de 3 anneaux, 8 intervalles angulaires et factor=1,3.

Si on compare ces résultats (Fig. (12)) à ceux présentés par Zuliani [5] sur cette même base, on constatera que ceux-ci sont très proches (léger avantage du descripteur proposé en précision). On notera que pour obtenir de tels résultats, Zuliani utilise une matrice radiale comportant 12 anneaux et 15 secteurs. Le DTHGI lui se contente de 3 anneaux et 8 intervalles angulaires pour obtenir des résultats similaires.

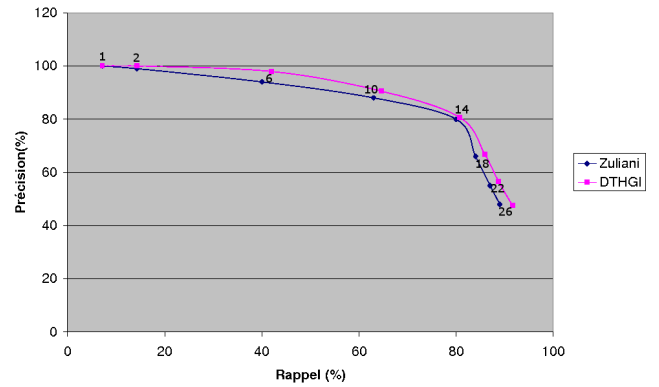


FIG. 12 – Courbe de Précision/Rappel

## 5.2 Mise en correspondance

Pour tester le descripteur sur des images réelles, nous avons choisi d’évaluer ses capacités de mise en correspondance entre deux images prises avec des points de vues différents ( Fig.(13.a) et Fig.(13.b) ) et Ce test consiste à reconnaître parmi les primitives détectées dans la seconde image celles détectées dans la première. Afin de nous placer dans des conditions réelles, les photos traitées sont celles d’une culasse. On notera que la culasse présente des motifs répétés. Nous différencions donc une mauvaise mise en correspondance avec une mise en correspondance avec une autre occurrence du même motif (traits bleus dans Fig.(13)).

Les résultats qui suivent ont été obtenus pour un descripteur de 16 secteurs, 3 anneaux et un factor de 5. Les images ont pour taille 800\*532.

Les résultats sont présentés dans la figure 13. Remarquons que sur les dix premières mises en correspondances il n’y a aucune mauvaise mise en correspondance.

## 6 Conclusions et perspectives

Dans cet article, nous avons présenté un descripteur de contours se situant dans les approches mixtes. Ce descripteur transforme le voisinage d’un contour fermé en une représentation invariante aux transformations affines modulo une rotation/Symétrie 2D : la forme du voisinage. Une signature basée sur la position des contours ainsi que leur orientation est alors calculée sur la forme de ce voisinage. L’invariance aux rotations de la signature permet finalement d’obtenir un descripteur, nommé DTHGI, invariant aux transformations affines.

Si les premiers résultats sont encourageants, une évaluation plus complète de ce descripteur reste à mener. Une fois ces actions menées, le DTHGI sera intégré dans un dispositif de reconnaissance d’objets métalliques.

## Références

[1] D. Lowe, *Distinctive image features from scale-invariant keypoints* 60 (2004), 911–110.

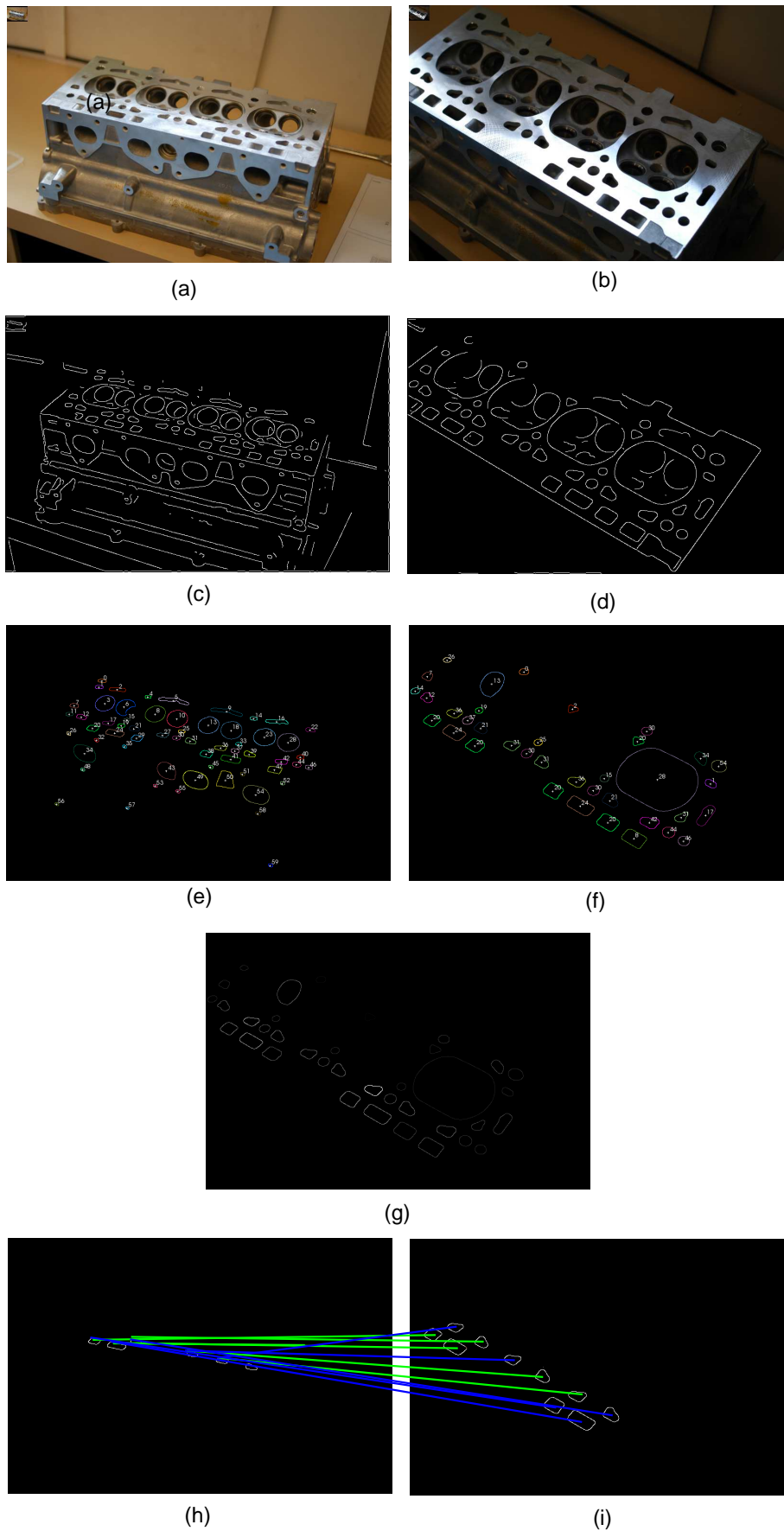


FIG. 13 – (a)Image source. (b) image destination. (c) image de contours source . (d) image de contours destinations. (e) Identités des contours sources. (f) Identités associées après mise en correspondance. (g) image de qualité (plus le contour est clair et plus la mise en correspondance est sûre) (h) et (g) mise en correspondance des 10 meilleurs résultats : un trait vert indique une mise en correspondance correcte, un trait bleu indique une mise en correspondance avec un contour équivalent (apparaît pour les motifs répétés), un trait rouges indique une mauvaise mise en correspondance.

- [2] K. Mikolajczyk and C. Schmid, *Scale and affine invariant interest point detectors*, International Journal of Computer Vision **Volume 60** (2004), 63–86.
- [3] A. Goshtasby, *Description and discrimination of planar shapes using shape matrices*, IEEE Trans. on Pattern Analysis and Machine Intelligence **7** (1985), 738–743.
- [4] A. Taza and C. Suen, *Discrimination of planar shapes using shape matrices*, IEEE Trans. on Systems, Man and Cybernetics **19** (1989), 1281–1289.
- [5] M. Zuliani, S. Bhagavathy, B. S. Manjunath, and C. S. Kenney, *Affine-invariant curve matching*, in IEEE International Conference on Image Processing, Oct 2004.
- [6] A. Chalechale, A. Mertins, and G. Naghdy, *Edge image description using angular radial partitioning*, IEE Proceedings-Vision Image and Signal Processing **151** (2004), 93–101.
- [7] O. Carmichael and M. Hebert, *Object recognition by a cascade of edge probes*, in BMVC02, 2002, p. Matching/Recognition.
- [8] O. Carmichael and M. Hebert, *Shape-based recognition of wiry objects*, IEEE Trans. on Pattern Analysis and Machine Intelligence **26** (2004), 1537–1552.
- [9] S. Belongie and J. Malik *Matching with shape context*, 2000.
- [10] S. Belongie, J. Malik, and J. Puzicha *Shape matching and object recognition using shape contexts* Technical Report UCB//CSD00, UC Berkeley, Jan. 2001.
- [11] A. Thayananthan, B. Stenger, P. Torr, and R. Cipolla, *Shape context and chamfer matching in cluttered scenes*, in IEEE Conf. on Computer Vision and Pattern Recognition, Vol. 1, 2003, pp. I–127– I–133.
- [12] P. Schwer *Shape context matching in cluttered scenes*, Dec 2003.
- [13] A. Jain and A. Vailaya, *Image retrieval using color and shape*, Pattern Recognition **29** (1996), 1233–1244.
- [14] A. Evans, N. Thacker, and J. Mayhew, *The use of geometric histogram for model-based object recognition*, in British Machine Vision Conference, 1993.
- [15] E. Huet, B. & Hancock, *Structurally gated pairwise geometric histograms for shape indexing*, in British Machine Vision Conference, 1997.
- [16] N. A. Thacker *Solving shapebased object recognition from a computational standpoint - practical and physiological constraints*, 2002.
- [17] K. Mikolajczyk, A. Zisserman, and C. Schmid, *Shape recognition with edge-based features*, in British Machine Vision Conference, 2003.
- [18] D. Ballard, *Generalizing the hough transform to detect arbitrary shapes*, Pattern Recognition **13** (1981), 111–122.
- [19] L. Davis, *Hierarchical generalized hough transforms and line-segment based generalized hough transforms*, PR **15** (1982), 277–285.
- [20] H. Freeman, *On the encoding of arbitrary geometric configuration*, IRE Transactions on Electronic Computers **10** (1961), 260–268.
- [21] D. S. Zhang and G. Lu., *A comparative study of fourier descriptors for shape representation and retrieval*, in Fifth Asian Conference on Computer Vision, Melbourne, January 2002, pp. 646–651.
- [22] A. Kaup and J. Heuer, *Polygonal shape descriptors - an efficient solution for image retrieval and object localization*, in 34th Asilomar Conference on Signals, Systems, and Computers, 2000.
- [23] S. Abbasi, F. Mokhtarian, and J. Kittler, *Curvature scale space image in shape similarity retrieval*, Springer Journal of MultiMedia Systems **7** (1999), 467–476.
- [24] I. Levner *Shape detection, analysis and recognition*, Aout 2002.
- [25] T. Pavlidis, *A review of algorithms for shape analysis*, Computer Graphics Image Processing **7** (1978), 243–258.
- [26] D. Zhang and G. Lu, *Review of shape representation and description techniques*, Pattern Recognition **37** (2004), 1–19.
- [27] S. Lazebnik, C. Schmid, and J. Ponce *A sparse texture representation using local affine regions* Technical Report CVR-TR-2004-01, Beckman Institute, 2004.
- [28] S. Obdrzalek and J. Matas, *Local affine frames for image retrieval*, in International Conference on Image and Video Retrieval, 2002, pp. 318–327.
- [29] S. Obdrzalek and J. Matas, *Object recognition using local affine frames on distinguished regions*, in BMVC2002, 2002.
- [30] M. Zuliani, S. Bhagavathy, B. S. Manjunath, and C. S. Kenney *A definition of shape for jordan manifolds* Technical Report TR-VRL-03-2004, VRL, ECE Department, University of California Santa Barbara, 2004.
- [31] J. Flusser and S. Tomas *On the calculation of image moments*, 1999.