



HAL
open science

Reconnaissance de formes basée géodésiques et déformations locales de formes

Mayss'Aa Merhy

► **To cite this version:**

Mayss'Aa Merhy. Reconnaissance de formes basée géodésiques et déformations locales de formes. Traitement du signal et de l'image [eess.SP]. Université de Bretagne occidentale - Brest; Université Libanaise, 2017. Français. NNT: 2017BRES0051 . tel-01617137

HAL Id: tel-01617137

<https://theses.hal.science/tel-01617137>

Submitted on 9 Feb 2018

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.

UBO

université de bretagne
occidentale



الجامعة اللبنانية
UNIVERSITÉ LIBANAISE

UNIVERSITÉ
BRETAGNE
LOIRE

THÈSE EN COTUTELLE
UNIVERSITÉ DE BRETAGNE OCCIDENTALE

sous le sceau de l'Université Bretagne Loire

UNIVERSITÉ LIBANAISE

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE

Spécialité : Traitement des Signaux et des Images

École Doctorale SICMA

DOCTEUR DE L'UNIVERSITÉ LIBANAISE

Spécialité : Mathématiques et Images

École Doctorale EDST

présentée par

Mayss'aa MERHY

Préparée au LabSTICC, UMR CNRS 6285, ENIB
et au centre de recherche AZM, UL

Reconnaissance de formes basée géodésiques et déformations locales de formes

Thèse soutenue le 29 juin 2017

devant le jury composé de :

Denis HAMAD

Professeur des Universités, ULCO / *Rapporteur*

Eric MOREAU

Professeur des Universités, Université de Toulon / *Rapporteur*

Denis FRIBOULET

Professeur des Universités, INSA de Lyon / *Président*

Sandy RIHANA

Maître de Conférences, USEK-Kaslik Liban / *Examinatrice*

Abdesslam BENZINO

Maître de Conférences HDR, ENI de Brest / *Directeur de thèse*

Kamal NASREDDINE

Maître de Conférences, ENI de Brest / *Examineur*

Mohamad KHALIL

Professeur des Universités, Univ. Libanaise / *Co-Directeur de thèse*

Ghaleb FAOUR

Professeur des Universités, CNRS-Libanais / *Examineur*

Remerciements

A l'issue de la rédaction de ce mémoire, je suis convaincue que la thèse est loin d'être dénuée de vie, de créativité et d'interactions humaines. En effet, je n'aurais jamais pu réaliser ce travail doctoral sans le soutien d'un grand nombre de personnes dont la générosité, la bonne humeur et l'intérêt manifestés à l'égard de ma recherche m'ont permis de progresser dans cette phase délicate de ma vie.

En premier lieu, je tiens à remercier M. Denis FRIBOULET, Professeur des Universités à l'INSA de Lyon, qui m'a fait l'honneur de présider le jury de cette thèse. J'exprime ma sincère gratitude à M. Denis HAMAD, Professeur des Universités à l'ULCO et à M. Eric MOREAU, Professeur des Universités à l'Université de Toulon, pour avoir accepté d'être rapporteurs de cette thèse. Je les remercie bien pour le temps passé à la lecture de mes travaux et à la rédaction des rapports, pour l'intérêt porté à mon travail et pour leurs précieuses remarques. Je tiens aussi à remercier Mme Sandy RIHANA, Maître de Conférences à l'USEK-Kaslik Liban et M. Ghaleb FAOUR, Professeur des Universités au CNRS-Libanais, pour avoir accepté d'examiner ce travail.

Je souhaite exprimer ma reconnaissance à mon directeur de thèse, M. Abdesslam BENZINO, Maître de Conférences HDR à l'ENIB, pour m'avoir accueilli dans le laboratoire et pour m'avoir donné l'occasion de travailler sur un sujet passionnant. Je tiens à le remercier également pour l'encadrement et pour la sympathie qu'il m'a témoignée au cours des années de thèse. Ses compétences, sa rigueur scientifique et sa clairvoyance m'ont beaucoup appris. Elles ont été et resteront des moteurs de mon travail de chercheur.

Mes plus vifs remerciements s'adressent à M. Kamal NASREDDINE, Maître de Conférences à l'ENIB, co-encadrant, sans qui cette thèse ne serait pas ce qu'elle est. Merci pour m'avoir orienté, soutenu et encouragé tout au long de cette thèse. Face à mon exigence, il s'est toujours montré disponible et patient. Son encadrement, ses remarques et sa grande qualité humaine m'ont permis de mener à bout cette thèse.

J'adresse mes profonds remerciements à M. Mohamad KHALIL, Professeur des Universités à l'Université Libanaise, pour avoir investi de son temps à diriger ce travail durant mon séjour de recherche au Liban. J'aimerais également lui dire à quel point j'ai apprécié sa grande disponibilité, ses qualités humaines d'écoute et de compréhension tout au long de ce travail doctoral.

Je remercie l'Agence Universitaire de la Francophonie (AUF) et le CNRS libanais

pour leurs soutiens financiers pendant mes trois années de thèse.

Mes remerciements vont particulièrement à mon père, Merhy MERHY, et à ma mère, Maha EL BAKKAR, qui m'ont constamment encouragé et soutenu tout au long de ces années. Je ne saurai passer sous silence l'apport inestimable de mes frères et ma soeur qui m'ont soutenu de près durant mes études doctorales.

Les mots me manquent pour remercier chaleureusement mon mari Hassan EL HAYEK, pour son soutien moral et psychologique indispensables pour maintenir ce projet à flot au travers des aléas de la vie et pour avoir cru en mes capacités et à mon sens de l'organisation pour le réaliser.

Je remercie sincèrement tous les collègues du laboratoire LabSTICC-ENIB pour leur soutien et leur encouragement.

Je ne saurais terminer sans souligner le soutien amical et chaleureux de mes copines et copains de tous les jours qui m'ont soutenu durant ce parcours doctoral.

Table de matières

Table des matières	1
Table des figures	5
Liste des tableaux	9
Glossaire	13
Introduction générale	15
I Généralités	23
1 État de l’art en reconnaissance de parties de formes	25
1.1 Introduction	27
1.2 Reconnaissance de formes entières	27
1.2.1 Représentation et description de formes planaires	28
1.2.2 Invariances aux transformations géométriques de formes $2D$	29
1.2.3 Mesures de similarité	32
1.3 Reconnaissance de formes basée-contour	33
1.3.1 Méthodes primaires	34
1.3.2 Méthodes basées recalage	34
1.3.3 Méthodes basées géodésiques de formes	35
1.4 Normalisation de parties de formes	38
1.4.1 Représentation invariante dans un repère canonique	38
1.4.2 Normalisation par alignement angulaire	39
1.4.3 Alignement Procuste	39
1.4.4 Alignement optimal par estimation du point initial et de pose	40
1.5 Représentation et description de parties de formes	42
1.5.1 Fonction d’angle	42
1.5.2 Fonction de courbure	43
1.5.3 Contexte de forme	43

1.5.4	Descripteurs de Fourier	45
1.5.5	Représentations élastiques	48
1.6	Mise en correspondance de parties de formes	49
1.6.1	Distance de Minkowski	51
1.6.2	Distance du ‘Khi-deux’	51
1.6.3	Distance de Hausdorff	52
1.6.4	Distance de Procuste	52
1.6.5	Distance basée sur le contexte de forme	53
1.7	Conclusion	53
2	État de l’art sur la décomposition des formes	55
2.1	Introduction	56
2.2	Motivations	56
2.3	Règles générales de décomposition	58
2.4	Travaux antérieurs	58
2.5	Évolution de la courbe discrète (DCE)	62
2.5.1	Description de la DCE	63
2.5.1.1	Objectifs	63
2.5.1.2	Caractéristiques	63
2.5.1.3	Mesure de relevance	64
2.5.1.4	Critère d’arrêt de l’évolution	65
2.5.2	Décomposition par DCE	70
2.5.2.1	Étape d’évolution optimale	70
2.5.2.2	Méthode des arcs convexes/concaves	72
2.5.2.3	Méthode des segments aux sommets	74
2.6	Conclusion	75
II	Méthodologies et validation	77
3	Reconnaissance de parties de formes basée sur les géodésiques de formes	79
3.1	Introduction	80
3.2	Recalage de parties de contours $2D$	81
3.2.1	Mesure de similarité	82
3.2.2	Recalage robuste	83
3.2.3	Métrie proposée	85
3.2.4	Minimisation par programmation dynamique	86
3.3	Système de reconnaissance de parties de formes	88
3.3.1	Construction de la base des parties de formes étiquetées	88
3.3.2	Stratégie de sous-sélection	91
3.3.3	Schéma proposé pour la reconnaissance	92

3.3.4	Évaluation des performances	92
3.3.4.1	Depuis la base de formes MPEG-7	92
3.3.4.2	Sur des images réelles	98
3.4	Reconnaissance de formes partielles	103
3.4.1	Stratégie de reconnaissance proposée	103
3.4.2	Évaluation depuis la base de formes MPEG-7	104
3.4.3	Évaluation sur des images réelles	106
3.5	Conclusion	116
4	Invariance à la rotation des parties de formes	119
4.1	Introduction	120
4.2	Méthode proposée	121
4.3	Critères proposés pour la comparaison des méthodes d'orientation	122
4.3.1	Critère basé sur la correspondance	122
4.3.2	Critère basé sur la distance géodésique	125
4.4	Comparaison des méthodes d'orientation	126
4.4.1	Évaluation sur des paires de parties de formes identiques	127
4.4.2	Évaluation sur des parties de formes synthétiques	131
4.4.3	Évaluation sur des parties de formes de la base des formes MPEG-7	142
4.5	Conclusion	145
5	Reconnaissance des formes entières à partir des parties de formes	147
5.1	Introduction	148
5.2	Travaux antérieurs	149
5.3	Système de correspondance basé sur la métrique quadratique	150
5.3.1	Décomposition des formes 2D fermées	150
5.3.2	Mesure de similarité quadratique	151
5.4	Approche proposée	153
5.4.1	Méthode de décomposition proposée	153
5.4.2	Distance utilisée	154
5.5	Applications sur la base de formes MPEG-7	155
5.5.1	Description des bases de test	155
5.5.2	Résultats et discussion	156
5.6	Conclusion	164
	Conclusion	165
	Bibliographie	169
	Résumés	193

Liste des figures

1	Schéma typique de la reconnaissance d'objet dans des images.	16
2	Illustration de la différence entre la comparaison globale et locale en présence d'une certaine déformation locale.	17
3	Un exemple de deux formes de différentes classes qui ne se différencient qu'en petite partie.	18
4	Illustration de la représentation invariante des parties de formes obtenue par la méthode du repère canonique et la méthode proposée.	19
5	Schéma typique de la reconnaissance de parties de formes et de formes partielles.	19
6	Schéma typique de la reconnaissance de formes entières à partir des parties de formes.	20
1.1	Illustration de la normalisation de formes par rapport à l'axe principal. . .	30
1.2	Alignement par analyse Procuste.	31
1.3	Alignement en utilisant la boîte rectangulaire englobante.	32
1.4	Alignement par représentation dans le repère référentiel canonique.	33
1.5	Représentation invariante dans un repère canonique.	39
1.6	Normalisation par alignement angulaire.	40
1.7	Alignement Procuste des parties de formes.	41
1.8	Alignement des parties de formes en utilisant la méthode proposée.	42
1.9	Descripteur du contexte de forme.	44
1.10	Illustration de la méthode de déformation et d'extension du segment. . . .	47
1.11	Correspondance obtenue en utilisant la fonction Log-vitesse et la fonction d'angle.	48
1.12	Comparaison de deux courbes ouvertes par correspondance directe des échantillons et par DTW.	50
1.13	Distance de Hausdorff.	52
2.1	Similarité globale et locale des formes.	57
2.2	Décomposition des contours de formes en utilisant les points minima de courbures négatives consécutives.	60
2.3	Points maxima locaux de courbure positive d'un contour.	60

2.4	Détermination des points critiques des formes.	60
2.5	Illustration de la simplification DCE.	61
2.6	Différents types de coupure.	62
2.7	Décomposition correspondant au squelette de la forme.	63
2.8	Quelques étapes de l'évolution par DCE d'un contour.	64
2.9	La méthode DCE lors du passage de l'étape d'évolution e à l'étape d'évolution $e+1$	65
2.10	Quelques étapes d'évolution par DCE d'un contour d'une forme la base des images MPEG-7.	65
2.11	Simplification DCE en utilisant 16 sommets comme critère d'arrêt.	67
2.12	Simplification DCE en utilisant 22 sommets comme critère d'arrêt.	68
2.13	Simplification DCE en utilisant 25 sommets comme critère d'arrêt.	68
2.14	Simplification DCE en utilisant la convexité comme critère d'arrêt.	68
2.15	Simplification DCE avec ou sans rétention des points extrémités.	71
2.16	Le sens de la recherche de l'étape d'évolution optimale.	72
2.17	Schéma descriptif de la recherche de l'étape de l'évolution optimale en utilisant la méthode DCE.	73
2.18	Définition des arcs convexes/concaves maximaux des formes.	73
2.19	Décomposition des formes en utilisant la méthode DCE.	74
3.1	Schéma proposé pour la reconnaissance de parties de formes et de formes partielles.	80
3.2	Représentation des déformations continues entre toute paire de parties de formes par des chemins géodésiques dans l'espace de formes.	82
3.3	Illustration de la signature angulaire et la fonction de correspondance entre deux parties de formes obtenue par l'approche de recalage proposée.	84
3.4	Comparaison entre un estimateur robuste et un estimateur non robuste.	85
3.5	Illustration du schéma de minimisation par programmation dynamique.	86
3.6	Décomposition d'une forme d'un cerf en utilisant la méthode DCE.	89
3.7	Évolution optimale de trois formes de <i>Cerf</i> obtenue par la méthode de retour en arrière en utilisant la mesure de similarité partielle ms avec un seuil fixe ($S = 0.4$).	89
3.8	Évolution optimale de trois formes de <i>Cerf</i> obtenue par la méthode de retour en arrière en utilisant la mesure de similarité partielle proposée ms' avec un seuil fixe ($S = 0.0006$).	90
3.9	Illustration des bases de test.	93
3.10	Illustration de la correspondance entre deux parties de formes obtenue avec notre méthode géodésique, avec la méthode du contexte de forme et avec la méthode Procuste.	94
3.11	Taux de recherche correcte de chaque requête obtenus par la méthode géodésique, la méthode du contexte de forme et la méthode Procuste.	96

LISTE DES FIGURES

3.12	Comparaison des performances de la méthode géodésique, la méthode du contexte de forme et la méthode Procuste en recherche.	97
3.13	Exemples de parties de formes extraites démontrant que notre approche est robuste aux déformations élastiques et aux articulations.	97
3.14	Illustration d'une base de test.	105
3.15	Les 20 formes les plus similaires à la paire des parties requêtes recherchées par notre schéma de combinaison basée sur les géodésiques. .	116
4.1	Exemple de perturbations d'extrémités des parties de formes visuellement similaires extraites par l'utilisation de la technique de la décomposition DCE.	120
4.2	Représentation invariante de deux parties de formes par la méthode du repère canonique et par notre méthode.	122
4.3	Illustration de la fonction de transformation utilisée qui permet d'avoir des parties de formes différentes dépourvues des problèmes aux extrémités.	123
4.4	Comparaison de la correspondance optimale des extrémités des parties de formes basée sur l'utilisation de la fonction angulaire avec celle qui correspond à la vérité terrain.	124
4.5	Un exemple de l'une des paires des points caractéristiques bornes possibles.	125
4.6	Perturbations d'extrémité des parties de formes dans cinq différentes directions.	126
4.7	Paires de parties de formes identiques.	127
4.8	La différence entre la fonction de correspondance initialement imposée et son équivalent vérité-terrain en présence d'une perturbation d'extrémité. .	128
4.9	Illustration des fonctions de correspondance trouvées $\phi_{estimée}$ et de la correspondance entre des points caractéristiques de deux parties de formes, obtenue respectivement par chacune des méthodes d'orientation (1-3).	129
4.10	L'étape de la réinitialisation des points extrémités, adoptée dans notre méthode d'orientation.	130
4.11	Parties de formes références utilisées pour l'évaluation de la performance des 4 méthodes d'orientation en présence des perturbations d'extrémité. .	131
4.12	Représentation de la fonction de correspondance initialement imposée dans le nouveau repère de la vérité-terrain	133
4.13	Illustration de la fonction de correspondance $\phi_{estimée}$ et la correspondance entre des points caractéristiques de deux parties de formes, obtenues respectivement par chacune des méthodes d'orientation (1-3).	134
4.14	Illustration de la fonction de correspondance vérité-terrain et la correspondance entre des points caractéristiques de deux parties de formes obtenues suite à l'utilisation de notre méthode d'orientation. . . .	135

4.15	Les 70 paires de parties de formes utilisées dans l'étude comparative des 4 méthodes d'orientation.	142
4.16	Les valeurs moyennes de la distance géodésique calculées entre des parties de formes obtenues en présence des perturbations d'extrémité à la suite d'un alignement par les 4 méthodes d'orientation.	143
4.17	Illustration de la variation des distances D calculées en présence de différentes perturbations a été attribué pour chacune des 4 méthodes d'orientation.	144
5.1	Similarité globale des formes.	148
5.2	Arcs convexes/concaves maximaux d'une forme d'une voiture de la base de données MPEG-7.	151
5.3	Trois sous-bases de la base d'images MPEG-7 dont chacune contient deux classes différentes de formes qui sont similaires du point de vue globale. .	156
5.4	Visualisation de la correspondance optimale entre les parties de différentes formes obtenue par la méthode géodésique, la méthode angulaire, la méthode du contexte de forme et la méthode Procruste. . . .	158
5.5	Visualisation de la correspondance optimale entre les parties de différentes formes obtenue par la méthode géodésique, la méthode angulaire, la méthode du contexte de forme et la méthode Procruste. . . .	159

Liste des tableaux

2.1	Évolution de trois formes différentes de la base MPEG-7 avec l'utilisation de 5 différentes valeurs du K_{max}	69
3.1	Taux de classification correcte obtenus avec les classifieurs $1-NN$, $5-NN$ et $10-NN$ en utilisant les approches géodésique, du contexte de forme et Procuste sur les ensembles des requêtes $E1$ et $E1_{\text{ét}}$	98
3.2	Exemples d'images réelles et leurs cartes de contours. Résultats de la classification en utilisant chacune des trois méthodes utilisées (géodésique, contexte de forme et Procuste) sur les requêtes retenues d'images réelles utilisées.	99
3.3	Résultats de la recherche en utilisant notre méthode sur les requêtes retenues d'images réelles.	100
3.4	Résultats de la recherche en utilisant l'approche du contexte de forme sur les requêtes retenues d'images réelles.	101
3.5	Résultats de la recherche en utilisant la distance de Procuste sur les requêtes retenues d'images réelles.	102
3.6	Taux de précision de recherche des formes partielles en utilisant notre méthode géodésique, la méthode basée sur le contexte de forme et la méthode Procuste.	106
3.7	Taux de classification des formes partielles en utilisant notre méthode géodésique, la méthode basée sur le contexte de forme et la méthode Procuste avec le classifieur $K-NN$	106
3.8	Exemples d'images réelles, leurs cartes de contours et les paires de parties de contours les plus significatives retenues pour la reconnaissance des formes dans les images segmentées.	108
3.9	Résultats de la recherche sur des images réelles en utilisant notre méthode géodésique.	109
3.10	Résultats de la recherche sur des images réelles en utilisant l'approche du contexte de forme.	110
3.11	Résultats de la recherche sur des images réelles en utilisant l'approche Procuste.	111

3.12	Résultats de classification sur des paires de requêtes extraites des images réelles en utilisant les trois méthodes : géodésique, contexte de forme et Procuste.	112
3.13	Comparaison des performances de recherche de formes partielles entre notre approche géodésique et l'approche basée sur les fonctions d'angle MTA.	114
4.1	L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues par l'application de chacune des 4 méthodes d'orientation pour l'alignement des courbes données.	131
4.2	L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues par l'application de chacune des 4 méthodes d'orientation pour l'alignement des courbes données.	132
4.3	L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre deux courbes synthétiques alignées avec les 4 méthodes d'orientation.	135
4.4	L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre deux courbes synthétiques alignées avec les 4 méthodes d'orientation.	136
4.5	L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre deux courbes synthétiques alignées avec les 4 méthodes d'orientation.	137
4.6	L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre deux courbes synthétiques alignées avec les 4 méthodes d'orientation.	138
4.7	L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre deux courbes synthétiques alignées avec les 4 méthodes d'orientation.	139
4.8	L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre deux courbes synthétiques alignées avec les 4 méthodes d'orientation.	140
5.1	Taux de recherche obtenus en utilisant les approches : géodésique, angulaire, contexte de forme et Procuste sur les formes de chaque classe des trois sous-bases étudiées.	156
5.2	Comparaison entre les correspondances des parties de formes de la classe de <i>Chiens</i> obtenues avec l'approche géodésique et avec la méthode angulaire.	160
5.3	Comparaison entre les correspondances des parties de formes de la classe de <i>Chiens</i> obtenues avec l'approche géodésique et avec la méthode Procuste.	161

5.4	Comparaison entre les correspondances des parties de formes de la classe de <i>Chevaux</i> obtenues avec l'approche géodésique et avec la méthode angulaire.	162
5.5	Comparaison entre les correspondances des parties de formes de la classe de <i>Chevaux</i> obtenues avec l'approche géodésique et avec la méthode du contexte de forme.	162
5.6	Résultats de classification obtenus avec le classifieur <i>K-NN</i> en utilisant les approches : géodésique, angulaire, contexte de forme et Procuste sur chaque classe des trois sous-bases étudiées.	163
5.7	Comparaison des taux de recherche sur les sous-bases étudiées obtenues par les stratégies de la reconnaissance basée-parties de formes et la reconnaissance en utilisant la forme entière.	163
5.8	Taux de recherche et de classification correcte obtenus en utilisant l'approche géodésique proposée pour la reconnaissance des parties de formes et des formes partielles.	166

Glossaire

$1D$: Mono-dimensionnel
 $2D$: Bi-dimensionnel
 $3D$: Tri-dimensionnel
AP : Analyse procustéenne
BCF : Bag of contour fragments (Sac de fragments de contour)
BoW : Classical bag of words (Sac classique de mots)
CPF : Chance Probability Function (Fonction de probabilité de chance)
CSS : Curvature Scale Space (Courbure multi-échelle)
 d_{Eu} : Distance euclidienne
 d_{Geod} : Distance géodésique entre les parties de formes
 d_H : Distance de Hausdorff
 d_p : Distance de Minkowski
 d_{Pr} : Distance de Procuste
 d_q : Distance angulaire quaratique
 d_{SC} : Distance du contexte de forme
 D : Distance géodésique entre les formes partielles
DCE : Discrete Curve Evolution (Évolution de la courbe discrète)
DFT : Discrete Fourier Transform (Transformée de Fourier discrète)
DFTC : Descripteur de Fourier par tangente complexe
dis : Mesure de dissimilarité
DTW : Dynamic Time Warping (Déformation temporelle dynamique)
EDP : Équations aux dérivations partielles
EQM : Erreur Quadratique Moyenne
 $Id_{[0,1]}$: Transformation identique de $[0, 1]$ dans $[0, 1]$
 k -NN : k -Nearest Neighbors (k -Plus Proches Voisins)
LCSS : The Longest Common Subsequence (La plus longue sous-séquence commune)
 \mathbb{L}^p : Espace de Lebesgue
MAR : Minimum Area Rectangle (Boîte rectangulaire englobante minimale de la surface)
MTA : Multiple turning angle (Fonction angulaire multiple)
MS : Mesure de similarité
 MS_R : Mesure de similarité robuste

MVM : The Minimum Variance Matching (La correspondance à variance minimale)

ODE : Ordinary differential equation (Équation différentielle ordinaire)

OSB : The Optimal Subsequence Bijection (La bijection optimale entre les sous-séquences)

PF : Filtre à particules

PR : Taux de précision de la recherche

SC : Shape Context (Contexte de forme)

SIFT : Scale-invariant feature transform (Transformation de caractéristiques visuelles invariante à l'échelle)

SRV : Square root velocity (Fonction de la racine carrée de la vitesse)

VACF : Vector of aggregated contour fragments (Vecteur de fragments de contour agrégés)

χ^2 : Distance du 'Khi-deux'

Introduction générale

Contexte et motivations

La simulation des fonctions humaines par la machine est un domaine de recherche très actif depuis l'avènement des ordinateurs. Une de ces fonctions est l'analyse de l'information visuelle, dite tout simplement l'analyse de la vision. La vision, plus que tout autre sens, nous fournit des informations sur l'environnement et guide la plupart de nos activités et nos interactions. En vision par ordinateur, l'un des principaux enjeux consiste à prendre une image (ou séquence d'images) de la scène qui nous entoure pour la traiter et en dégager les informations pertinentes (en lien avec la tâche à accomplir). Parmi les différents aspects de l'information visuelle, la forme joue un rôle crucial dans différentes tâches de la vie quotidienne. La forme représente la caractéristique la plus pertinente de l'objet et elle peut être complétée par des informations de couleur ou de texture. En effet, la forme est l'un des plus importants attributs visuels d'un objet. La caractérisation de la forme fournit un indice puissant de l'identité et de la fonctionnalité de l'objet, et conduit à la reconnaissance de formes.

La reconnaissance de formes reste à nos jours l'un des domaines ouverts de recherche et d'expérimentation. En vision par ordinateur, la reconnaissance de formes a été largement étudiée au cours des dernières années dans la littérature scientifique. Elle est motivée par un large éventail d'applications importantes telles que la robotique [77, 195], la biologie [7], l'Internet [11], l'analyse de documents [53, 203], les arts visuels [87], l'industrie [252], la sécurité [129], etc. Cependant, malgré toutes les recherches intensives, les techniques de reconnaissance disponibles restent en deçà des capacités visuelles humaines. Nous pouvons dire que les techniques développées pour des applications particulières ont rencontré un succès limité sur des domaines restreints [19, 231].

Dans une application typique, la reconnaissance de formes commence par l'acquisition de l'image, ce qui donne une image numérique qui doit être pré-traitée pour accentuer certaines informations visuelles en rapport avec le domaine du problème posé. Les objets d'intérêt sont ensuite détectés et représentés pour un niveau haut d'analyse, pour des problèmes de reconnaissance (classification et recherche). Les étapes de base dans une application de reconnaissance de formes sont données dans la figure 1

ci-dessous. Essentiellement, l'étape de l'extraction des caractéristiques contenues dans l'image fait appel à la segmentation, en particulier la détection de contours, qui est l'une des tâches les plus difficiles dans le domaine de la vision par ordinateur.

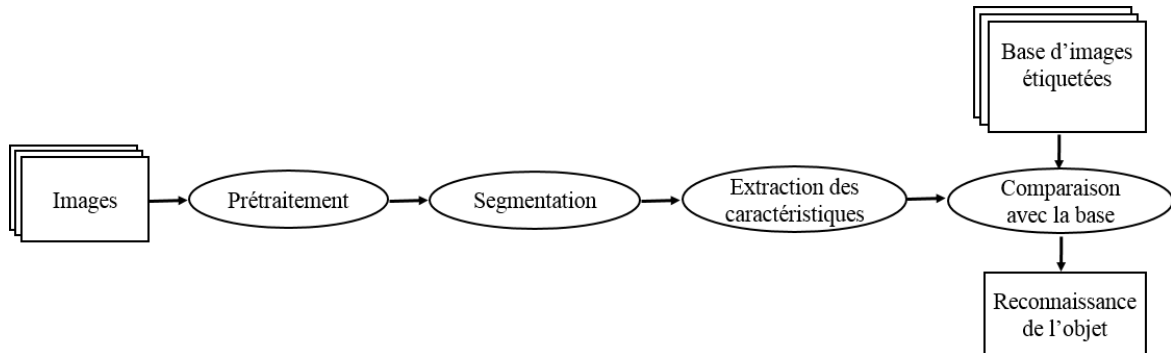


Figure 1 – Schéma typique de la reconnaissance d'objet dans des images.

En général, les performances d'un système de reconnaissance de formes dépendent en bonne partie de la qualité de l'image segmentée. Le traitement des images a connu ces dernières années un essor considérable et un nombre conséquent de travaux de recherche ont été proposés au niveau de la détection de contours. Malgré les progrès effectués ces dernières années visant la méthode optimale [207], les chercheurs s'accordent que la méthode de segmentation idéale n'existe pas. Autrement dit, une segmentation complète (c'est-à-dire avec des contours entiers) ne peut pas être toujours atteinte et reste un problème difficile à résoudre. Les opérateurs classiques de détection de contours ont en effet beaucoup de limitations en présence de bruit, de texture, de changement d'éclairage, d'occlusion ou d'objets en contact dans les images à segmenter. En raison de ces différents facteurs, il est souvent impossible de segmenter l'objet en entier et seulement certaines parties peuvent être détectées. Ceci justifie pleinement la nécessité d'un système de reconnaissance de parties de formes (ou de formes partielles) qui reste très peu abordé dans la littérature.

Pour plus de précision, une partie de forme est définie comme une partie significative de son contour représentatif, tandis qu'une forme partielle désigne un ensemble de parties de forme non nécessairement consécutives.

La représentation basée sur le contour, comparée à la représentation basée sur la région, peut fournir une description plus précise des formes ou des parties de formes. Elle est plus invariante aux conditions d'éclairage, variations de la couleur et de la texture de l'objet. Elle permet aussi une description plus stable des structures locales (parties de formes).

Ainsi, nous nous intéressons ici aux approches de reconnaissance de forme basées sur le contour. La plupart des méthodes de reconnaissance de formes basées-contour utilisent des techniques de correspondance de formes $2D$ où les formes de courbes fermées sont entièrement considérées dans la mise en correspondance [25, 81, 206, 229].

Dans un travail antérieur [178], une méthode de reconnaissance de formes 2D a été proposée en se basant sur la mise en correspondance de courbes fermées et sur les géodésiques de formes. Il est montré dans [178] que cette approche surpasse toutes les approches de l'état d'art. Malgré sa grande efficacité, les performances de cette approche diminuent sur certaines bases d'images du fait que la comparaison entre les formes se fait au niveau des formes entières. Bien que la description soit locale, la comparaison globale mène à des résultats erronés avec cette méthode comme avec n'importe quelle méthode basée sur la comparaison des formes entières. D'une part, il existe des formes qui sont visuellement jugées comme différentes des autres membres de leur propre classe dues aux différentes poses, aux déformations et aux auto-occlusions. Ces différences ne sont que locales sur les formes mais peuvent influencer fortement sur la comparaison globale. Par exemple, dans la figure 2 nous montrons un exemple de deux images de chien où la queue est vers le bas dans une image, et est vers le haut dans l'autre image. Cette déformation locale au niveau de la queue aura une influence importante sur la similarité entre les formes si la comparaison est globale. Bien au contraire, une comparaison par parties de formes ne mènera pas à une différence aussi importante.

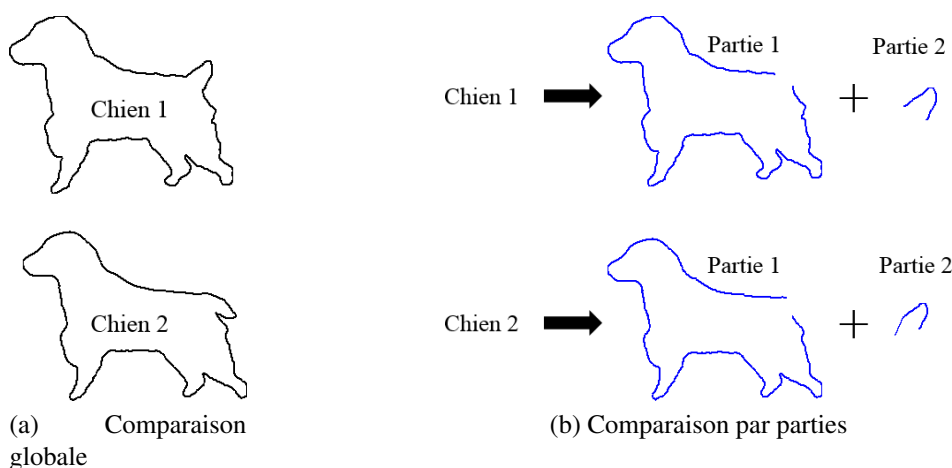


Figure 2 – En (a), exemple de deux formes de chiens globalement jugées comme différentes à cause d'une différence locale au niveau de la queue. Cependant en (b), avec une comparaison par parties ces mêmes formes sont jugées comme similaires.

D'autre part, certaines formes sont très similaires à des exemples d'autres classes où elles ne se différencient qu'en petite partie de la forme. Par exemple, les formes de la figure 3 ne se différencient qu'en petite partie de la forme. Cette différence peut ne pas se révéler dans la mesure de similarité globale si la comparaison est faite d'une façon globale. Cela justifie encore la nécessité d'un système de reconnaissance de formes basé sur la correspondance des parties de formes.

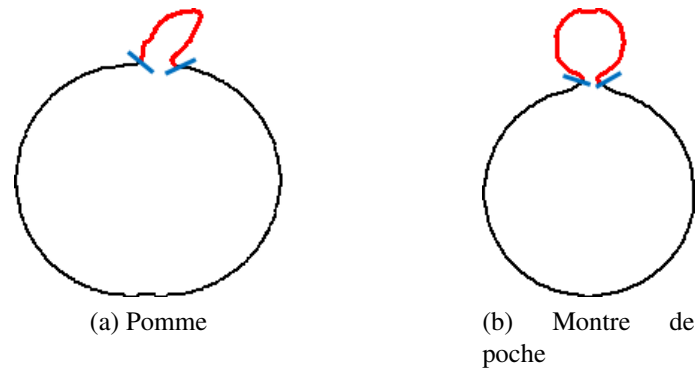


Figure 3 – Un exemple d’une pomme (a) et d’une montre de poche (b) qui ne se différencient qu’en petite partie (parties en rouge) de la forme.

Contributions de la thèse

Les contributions de cette thèse sont de trois ordres. En effet, nous nous sommes focalisés sur les trois tâches principales pour la reconnaissance de formes et nous avons proposé une nouvelle méthode pour chacune de ces tâches :

- **Représentation invariante des parties de formes**

Nous développons une nouvelle méthode de représentation invariante des parties de formes. Afin d’assurer l’invariance aux transformations géométriques (translation, rotation et facteur d’échelle), toutes les méthodes de la littérature exigent une correspondance fixée des points extrémités des parties de formes à comparer (voir par exemple la méthode la plus utilisée, dite du repère canonique en figure 4c). Une telle fixation n’est pas toujours optimale. Nous proposons ici une méthode d’optimisation de l’analyse procustéenne avec une fenêtre englobante en chaque point extrémité (figure 4d). Les nouveaux points extrémités sont retrouvés en leur permettant de glisser le long du contour à proximité des points extrémités initiaux afin de minimiser la distance de Procuste du type moindre carré. Une étude comparative a montré une grande stabilité aux perturbations d’extrémité et une supériorité par rapport aux méthodes de l’état de l’art.

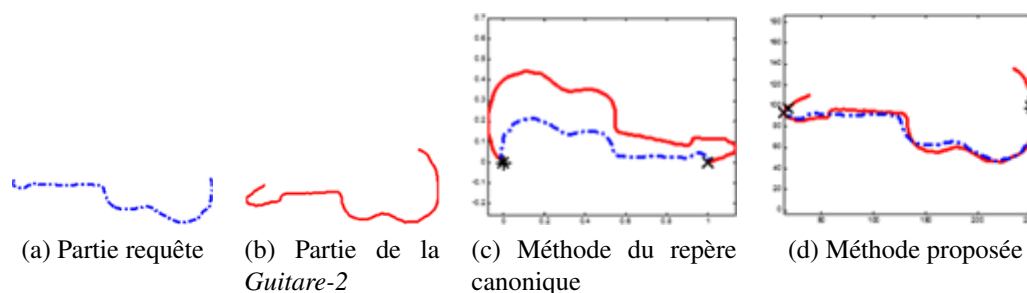


Figure 4 – La représentation invariante d’une partie requête (partie de la guitare donnée en (a)) et un élément de la base MPEG-7 (partie de la *Guitare-2*) par la méthode du repère canonique et par la méthode que nous proposons. Les points extrémités sont marqués par des (x) et des (*).

• **Reconnaissance de parties de formes ou de formes partielles**

Nous proposons des schémas de reconnaissance, basés-contour, de parties de formes et de formes partielles. Les approches de reconnaissance que nous développons sont fondées sur un processus de recalage robuste de courbes 2D ouvertes. Cette méthode de recalage, autrement dit de mise en correspondance, consiste à optimiser une mesure de similarité élastique basée sur les géodésiques dans l’espace de formes. Le résidu de recalage est ensuite exploité pour la reconnaissance de parties de formes et de formes partielles (figure 5 suivante).

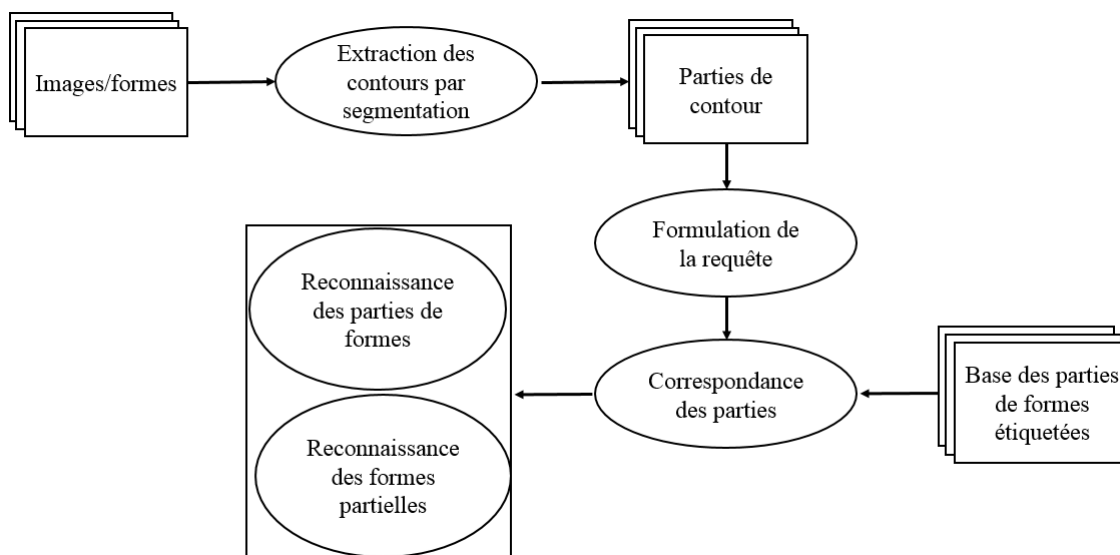


Figure 5 – Schéma typique de la reconnaissance de parties de formes et de formes partielles.

• Reconnaissance de formes 2D fermées à partir des parties de formes

Nous proposons un nouveau système de reconnaissance de formes 2D fermées, basé sur la correspondance de leurs parties significatives. Au niveau de ce système, nous exploitons les similarités locales estimées à partir de notre approche géodésique, proposée pour la comparaison de parties de formes, pour calculer la similarité globale entre les formes entières (figure 6 suivante).

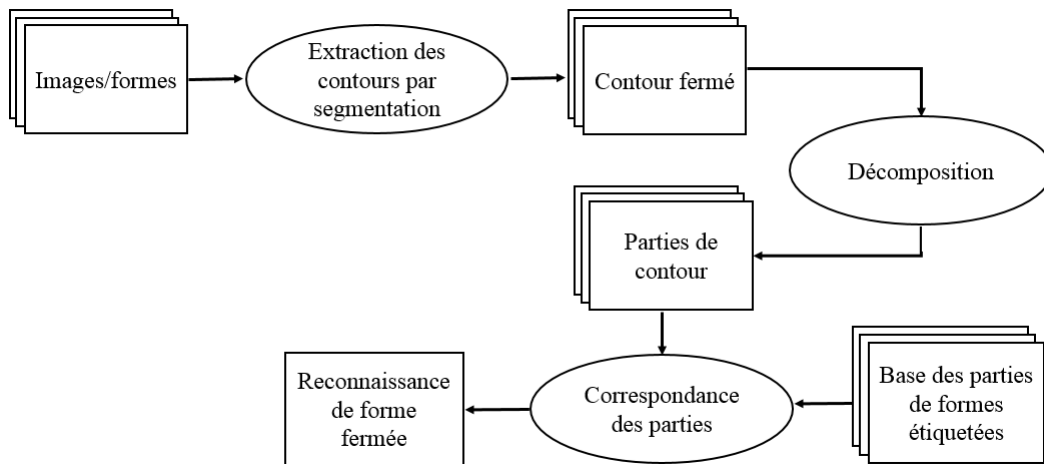


Figure 6 – Schéma typique de la reconnaissance de formes entières à partir des parties de formes.

Organisation du document

Cette thèse est composée principalement de deux parties, chacune formée de plusieurs chapitres.

La première partie, ‘Généralités’, est consacrée à la présentation du cadre général de cette thèse. Le chapitre 1 donne un aperçu sur les techniques de représentation, de normalisation et de reconnaissance de formes et parties de formes par ordinateur. Le chapitre 2 a pour objectif de donner au lecteur un état de l’art sur les méthodes de décomposition de formes entières en des parties de formes significatives.

La deuxième partie, ‘Méthodologies et validation’, présente les contributions développées dans ce travail de thèse. Dans le chapitre 3 intitulé “Reconnaissance de parties de formes basée sur les géodésiques de formes”, nous proposons deux approches de reconnaissance (classification et recherche) basées-contour ; toutes les deux sont fondées sur un processus de comparaison entre les parties de formes par mise en correspondance (ou recalage). La première approche concerne la reconnaissance de parties de formes 2D et la deuxième concerne la reconnaissance de formes partielles 2D. Ces parties de formes 2D étant codées par leurs fonctions tangentes aux contours, leur

recalage revient à un recalage de signaux 1D. Ces deux approches sont comparées à des méthodes employées dans l'état de l'art, à partir des parties de formes de la base de test MPEG-7, et à partir des parties de formes issues d'une segmentation incomplète d'images réelles. Dans le chapitre 4 intitulé "Invariance à la rotation des parties de formes", nous proposons un nouvel algorithme basé sur l'alignement Procuste et la redétermination des points extrémités pour résoudre le problème de l'orientation des parties de formes. Son efficacité est démontrée en utilisant des parties de formes synthétiques et des parties de formes de référence (base de test MPEG-7). Nous proposons ensuite, dans le chapitre 5, "Reconnaissance des formes entières à partir des parties de formes", une nouvelle technique de classification et de recherche des formes fermées utilisant les géodésiques de formes et la mise en correspondance des parties de formes significatives. Comme expérimentations, nous réalisons des tests sur les formes de différentes classes de la base de test MPEG-7 qui ont des formes globales très similaires.

Dans la conclusion, nous résumons les réalisations effectuées et nous rappelons les principaux résultats obtenus. Nous explicitons également l'apport scientifique du point de vue expérimental. Enfin, nous proposons de nouvelles voies d'investigation et nous évoquons les perspectives qui pourraient orienter nos recherches futures.

Partie I

Généralités

Chapitre 1

État de l'art en reconnaissance de parties de formes

Sommaire

1.1	Introduction	27
1.2	Reconnaissance de formes entières	27
1.2.1	Représentation et description de formes planaires	28
1.2.2	Invariances aux transformations géométriques de formes 2D	29
1.2.3	Mesures de similarité	32
1.3	Reconnaissance de formes basée-contour	33
1.3.1	Méthodes primaires	34
1.3.2	Méthodes basées recalage	34
1.3.3	Méthodes basées géodésiques de formes	35
1.4	Normalisation de parties de formes	38
1.4.1	Représentation invariante dans un repère canonique	38
1.4.2	Normalisation par alignement angulaire	39
1.4.3	Alignement Procuste	39
1.4.4	Alignement optimal par estimation du point initial et de pose	40
1.5	Représentation et description de parties de formes	42
1.5.1	Fonction d'angle	42
1.5.2	Fonction de courbure	43
1.5.3	Contexte de forme	43
1.5.4	Descripteurs de Fourier	45
1.5.5	Représentations élastiques	48

*CHAPITRE 1. ÉTAT DE L'ART EN RECONNAISSANCE DE PARTIES DE
FORMES*

1.6	Mise en correspondance de parties de formes	49
1.6.1	Distance de Minkowski	51
1.6.2	Distance du 'Khi-deux'	51
1.6.3	Distance de Hausdorff	52
1.6.4	Distance de Procuste	52
1.6.5	Distance basée sur le contexte de forme	53
1.7	Conclusion	53

1.1 Introduction

Ce chapitre est destiné à présenter les outils de la reconnaissance de parties de formes par ordinateur. En général, la reconnaissance automatique de parties de formes est le processus visant à trouver, pour une partie de forme donnée, l'ensemble de parties de formes auquel elle appartient (classification), ou à dire quelles sont les parties de formes de la base qui ont les mêmes propriétés que cette partie de forme (recherche).

La reconnaissance de formes dans une image segmentée est un problème largement ouvert en vision par ordinateur [211]. Dans les images réelles, la segmentation complète d'objet en contour entier est souvent impossible [19]. À cause du phénomène de sur-ou sous-segmentation, la segmentation d'image ne donne qu'un ensemble de parties de contour qui peuvent être reconstituées et regroupées ensuite en des structures locales de formes : les parties de formes et les formes partielles. Pour plus de précision, une partie de forme est définie comme une partie significative de son contour représentatif, tandis qu'une forme partielle désigne un ensemble de parties de forme non nécessairement consécutives. En dépit de la facilité de la reconnaissance visuelle des formes à partir des différentes parties de formes, cette tâche impose de nombreux défis et reste difficile à automatiser.

Ce chapitre commence par introduire en section 1.2 les méthodes les plus utilisées pour assurer l'invariance aux transformations usuelles ainsi que les techniques classiquement utilisées en reconnaissance de formes *2D* entières. La section 1.3 présente les grandes tendances en matière de reconnaissance de formes basée sur le contour. Les méthodes de normalisation de parties de formes sont décrites dans la section 1.4. En section 1.5, nous présentons ensuite les méthodes de représentation et de description de parties de formes. Finalement, nous décrivons en section 1.6 les mesures de similarité proposées dans la littérature pour établir la meilleure correspondance entre les parties de formes.

1.2 Reconnaissance de formes entières

La reconnaissance de formes est un domaine très important de la vision par ordinateur. Elle consiste à reconnaître les objets de l'image en fonction des informations liées à leurs formes [63]. En effet, la forme est considérée comme l'une des principales caractéristiques décrivant le contenu d'une image. Cependant, la représentation et la description de la forme constituent une tâche difficile à entreprendre [45, 261].

Les principales qualités d'un descripteur sont d'une part sa grande capacité à discriminer des formes [18, 107, 261] et d'autre part, la faible redondance de ses composantes de sorte qu'un nombre minimal de celles-ci suffit à caractériser efficacement la forme. Il est évident qu'un descripteur fiable doit d'abord être invariant aux transformations géométriques usuelles : translation, rotation et changement

d'échelle. Ensuite, il doit être robuste au bruit, aux distorsions et aux occlusions car de telles perturbations peuvent changer la forme d'une manière subtilement complexe.

1.2.1 Représentation et description de formes planaires

Les techniques de représentation et de description de la forme peuvent être généralement scindées en deux classes principales de méthodes : les méthodes basées-*contour* et les méthodes basées-*région*. Ce classement dépend de la manière dont les caractéristiques de forme sont extraites : à partir du contour seulement ou à partir de toute la région de la forme. Sous chaque catégorie, les différentes approches sont divisées en approches *globales* et approches *locales* (structurelles). Ce sous-classement est basé sur la représentation de la forme qui soit dépend de la forme toute entière, soit des parties de la forme (primitives). Ces approches peuvent aussi être distinguées en fonction du domaine *spatial* ou de *transformation*, dans lequel les caractéristiques de forme sont calculées. Les méthodes globales ne sont pas toujours robustes contre les occlusions et le bruit de l'image. De plus, elles nécessitent une segmentation entière et correcte des objets dans les images, ce qui est en soi un problème mal posé. En général, le processus de segmentation aboutit à partitionner les objets en des régions ou en des parties de contour qui ne correspondent pas nécessairement aux objets entiers.

Les approches basées-contour n'exploitent que la frontière de l'objet pour la caractérisation de la forme en ignorant son contenu intérieur. La représentation la plus utilisée dans les méthodes de reconnaissance basées sur le contour est la signature de la forme [70, 262]. Pour une forme donnée, la signature est essentiellement une représentation basée sur une paramétrisation 1D du contour de la dite forme. Ceci peut être réalisé en utilisant une valeur scalaire de la fonction de distance radiale, d'angle, de courbure ou de vitesse. Notons ici que la signature d'une forme entière (courbe fermée) est souvent une fonction périodique, ce ne sera pas le cas d'une partie de forme (courbe ouverte) pour laquelle les deux extrémités ne sont pas jointives. Parmi les descripteurs basés sur le contour, on peut citer les descripteurs de Fourier [98, 118, 184, 261], les descripteurs d'ondelettes [20, 48, 175], la courbure multi-échelle [32, 168, 261], le contexte de forme [25], les moments du contour [218], la chaîne de symboles [74, 92]... Comme ces descripteurs sont calculés en utilisant uniquement les pixels du contour, la complexité de calcul est faible et leurs vecteurs caractéristiques sont généralement compacts.

Dans les approches basées-région, tous les pixels de l'objet sont pris en compte pour la caractérisation de la forme. Ce type de méthodes vise à exploiter non seulement les informations de la frontière de la forme, mais aussi celles de la région interne de la forme. La majorité des méthodes basées sur la région utilisent des descripteurs de moments pour décrire les formes tels que les moments de Zernike [105], les moments de Legendre [113] ou les moments géométriques invariants [91]. D'autres méthodes incluent les descripteurs de grille [142], la matrice de forme [73], les descripteurs de

Fourier génériques [260], l'axe médian (ou le squelette) [213] ou la transformation radiale angulaire [194]. Comme le descripteur basé-région fait usage de tous les pixels constituant la forme, il peut décrire efficacement diverses formes dans un seul descripteur. Toutefois, la taille des caractéristiques basées sur la région est généralement grande. Ce descripteur conduit à un temps de calcul qui reste considérable.

Reste à souligner que la description des formes basée sur le contour est considérée comme plus pertinente que celle basée sur la région car la forme d'un objet se distingue essentiellement par la frontière. Dans la plupart des cas, la partie centrale de l'objet ne contribue guère à la reconnaissance de la forme [264].

1.2.2 Invariances aux transformations géométriques de formes 2D

Un système de reconnaissance de formes doit identifier un objet comme tel en dépit de sa position, son orientation et sa taille. Une forme d'objet dans l'image peut être différente de la forme standard car elle est souvent affectée par certains facteurs naturels ou issus d'algorithmes de traitement. Ces facteurs comprennent les paramètres régissant le processus de capture par camera (angle et distance de prise de vue), l'occlusion, le bruit, etc. En effet, la même forme peut apparaître différemment lorsqu'elle est vue de différentes perspectives.

Afin d'associer la forme à un modèle utile pour les applications de reconnaissance, toute représentation et toute description de forme doivent être invariantes par rapport aux transformations géométriques usuelles telles que la translation, la rotation et le changement d'échelle. Dans la littérature, le problème de l'invariance est traité de deux façons différentes. Certaines méthodes choisissent de représenter et décrire les formes par des descripteurs invariants. Tandis que d'autres abordent ce sujet d'invariance en utilisant des techniques de transformation qui s'appliquent directement aux formes en s'appuyant sur quelques caractéristiques avant toute technique de représentation et de description des formes.

Avec des descripteurs basés-région, le premier type de méthodes consiste à construire et à associer à chaque forme des invariants aux transformations géométriques. Hu s'est inspiré par des invariants algébriques [40] pour calculer des moments invariants à la translation, la similitude et à la rotation qui sont basés sur les moments géométriques centrés de la forme [91]. Ultérieurement, plusieurs méthodes dérivées ont été proposées comme les moments de Zernike [105, 234], les moments de Legendre [234] et la transformation Fourier-Mellin [133]. Les moments complexes ont été également utilisés pour dériver des moments invariants aux transformations géométriques [3]. Les descripteurs SIFT [141] sont aussi largement utilisés pour assurer l'invariance des formes à comparer, mais ces descripteurs sont généralement de grande dimension (128 dimensions).

Avec des descripteurs basés-contour, les techniques consistent à extraire des caractéristiques invariantes aux transformations géométriques des courbes fermées.

Parmi ces méthodes, nous citons les descripteurs de Fourier [82, 114, 258] appliqués à un contour fermé ou à plusieurs contours fermés d'objets déconnectés [61]. Les descripteurs de Fourier elliptiques ont été également beaucoup utilisés pour la description invariante des contours et la reconnaissance de formes [112, 180, 236]. Les ondelettes [48], les méthodes basées sur la courbure [39, 171] ou sur des intégrales invariantes [150], et les descripteurs de contexte de forme [25] ont été aussi très utilisés en tant que descripteurs invariants des contours fermés.

D'un autre côté, la normalisation *a priori* des formes a été largement étudiée pour assurer l'invariance des formes entières à la translation, rotation et au changement d'échelle. Par exemple, l'axe principal a été utilisé dans différents travaux antérieurs pour rendre des formes invariantes quelques soient leurs positions et orientations initiales [8, 34, 49, 54, 117, 239]. Cette technique impose de trouver d'abord le centre de masse de chaque forme et sa ligne optimale qui passe par ce centre de masse et qui minimise la distance totale à tous les points de la forme. Ensuite, on vient confondre les deux centres de masse et mettre les axes principaux des deux formes à comparer à la même direction pour assurer l'invariance à la translation et à la rotation respectivement. Cependant, cette méthode n'est pas invariante au changement d'échelle. La figure 1.1 montre l'efficacité de cette méthode à déplacer et à tourner la guitare noire pour qu'elle ait une position et une orientation identiques à celles de la guitare bleue.

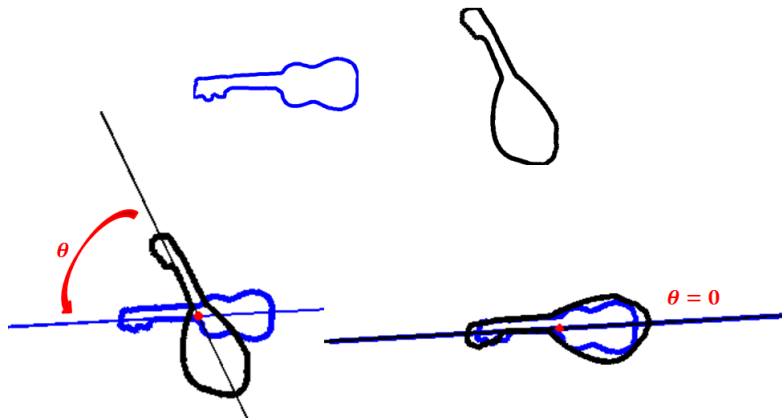


Figure 1.1 – En première ligne, deux formes de guitares positionnées et orientées différemment. En deuxième ligne, à gauche, les points de masse (en rouge) et les directions des axes principaux correspondant aux deux formes; à droite, les formes avec la même position et la même orientation une fois leurs points de masse sont confondus (en rouge) et leurs axes principaux sont placés à la même direction.

D'autres approches de normalisation [81, 145, 173] utilisent l'alignement Procuste pour aligner une forme à une autre. L'analyse de Procuste cherche à maximiser l'accord entre deux formes à travers des transformations géométriques. En premier lieu, l'effet du facteur d'échelle et l'effet de la position sont annulés. Ensuite, les deux formes sont

1.2. RECONNAISSANCE DE FORMES ENTIÈRES

ajustées à la même direction par la recherche de l'angle optimal θ duquel il faut tourner la forme à aligner pour la faire correspondre au mieux avec l'autre forme référence. La correspondance optimale est atteinte par une minimisation de la distance entre la forme à aligner et celle de référence. Cette distance peut être calculée par la méthode des moindres carrés. Par exemple dans la figure 1.2, le quadrilatère vert (Fig. 1.2a) est translaté, dilaté et tourné de l'angle optimal θ pour mieux correspondre au quadrilatère bleu.

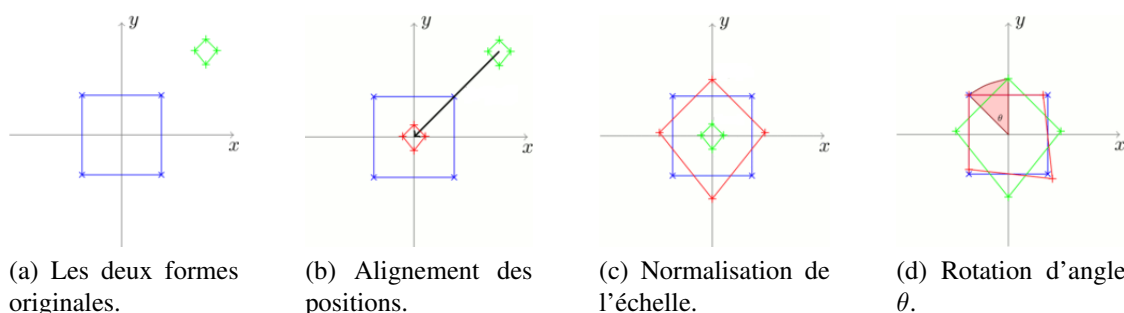


Figure 1.2 – Aligned par analyse Procuste. Étant donné les deux quadrilatères vert et bleu (a), le quadrilatère vert est translaté (b), dilaté (c) et tourné d'un angle θ (d) afin de l'aligner avec le quadrilatère bleu.

Par ailleurs dans [115], les auteurs ont utilisé la boîte rectangulaire englobante de surface minimale¹ pour obtenir des formes ayant des orientations relativement similaires. La boîte associée à une forme est définie comme le plus petit rectangle minimisant la zone avec la forme [205]. Autrement dit, la boîte englobante d'une forme donnée tend à ajuster l'orientation de cette forme de sorte à faire coïncider l'axe des abscisses avec la longueur du rectangle et l'axe des ordonnées avec la largeur du rectangle (voir la figure 1.3 suivante).

La méthode de normalisation proposée par Super [228] est basée sur des points caractéristiques extraits du contour représentatif de la forme. Ces points caractéristiques (figure 1.4(a)) correspondent aux points de maxima et de minima locaux de la courbure. À chaque point caractéristique x est associé le point du contour y le plus éloigné de x et le point du contour z le plus éloigné de la ligne droite définie par la paire des points (x, y) . Par suite, les points x, y et z sont utilisés pour trouver la similitude T_x permettant de représenter le contour dans un repère référentiel canonique tel que $T_x(x) = (0, 0)$, $T_x(y) = (1, 0)$ et $T_x(z)$ est au dessus de l'axe des abscisses (figure 1.4(b)). Ainsi, chaque forme a un ensemble de poses/représentations invariantes dont le nombre est égal au nombre des points caractéristiques. La comparaison de deux formes données nécessite de traiter toutes les paires de représentations invariantes de ces

¹Minimum Area Rectangle (MAR) dans la littérature anglo-saxonne.

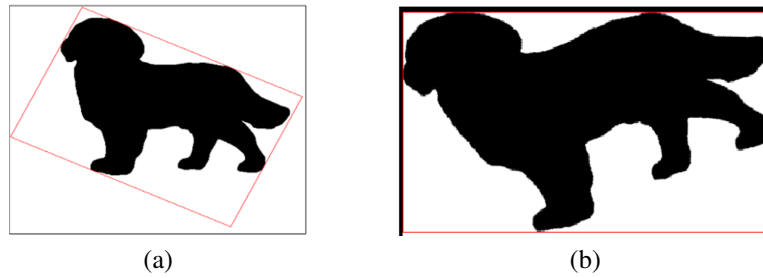


Figure 1.3 – Le rectangle minimum (en rouge) englobant la forme d'un chien (a) est utilisé pour ajuster l'orientation de cette forme (b).

deux formes dans le repère référentiel canonique et de trouver la paire de poses qui se correspondent le mieux (figure 1.4(c)).

1.2.3 Mesures de similarité

En général, une personne peut juger que deux images observées reviennent à un même objet grâce à des caractéristiques visuelles similaires des deux formes correspondantes. Cependant en vision par ordinateur, la similarité d'objets est jugée par le calcul d'une distance entre les caractéristiques extraites des deux formes. Pratiquement, toutes les méthodes de reconnaissance de formes ou de parties de formes nécessitent une telle étape de calcul pour comparer les formes selon une certaine mesure de similarité. Les mesures de similarité peuvent elles aussi être définies à partir des informations extraites de la totalité des pixels de l'objet (techniques basées-région) [107], ou à partir de certaines caractéristiques qui décrivent seulement la frontière de l'objet (techniques basées-contour) [128]. Formellement, une mesure de similarité assigne une distance de grande valeur aux formes non similaires. Autrement dit, plus la mesure de similarité est faible plus les formes sont similaires. Toutefois les propriétés que doit vérifier une mesure de similarité dépendent directement du problème étudié. Un espace E est dit métrique si pour tous éléments X, Y et Z de E , il existe une distance, autrement dit une application continue, $d : E \times E \rightarrow \mathbb{R}^+$ qui satisfait les propriétés suivantes :

- (unicité) : $d(X, Y) = 0$ si et seulement si $X = Y$,
- (symétrie) : $d(X, Y) = d(Y, X)$,
- (inégalité triangulaire) : $d(X, Z) \leq d(X, Y) + d(Y, Z)$.

Les choix simples d'une distance élémentaire telle que la distance euclidienne ne sont pas toujours optimaux. Généralement, la similarité entre deux formes données est calculée à l'aide d'une métrique bien adaptée à la représentation choisie de ces formes et à l'espace de formes construit. De plus, il est souvent souhaitable que la mesure de

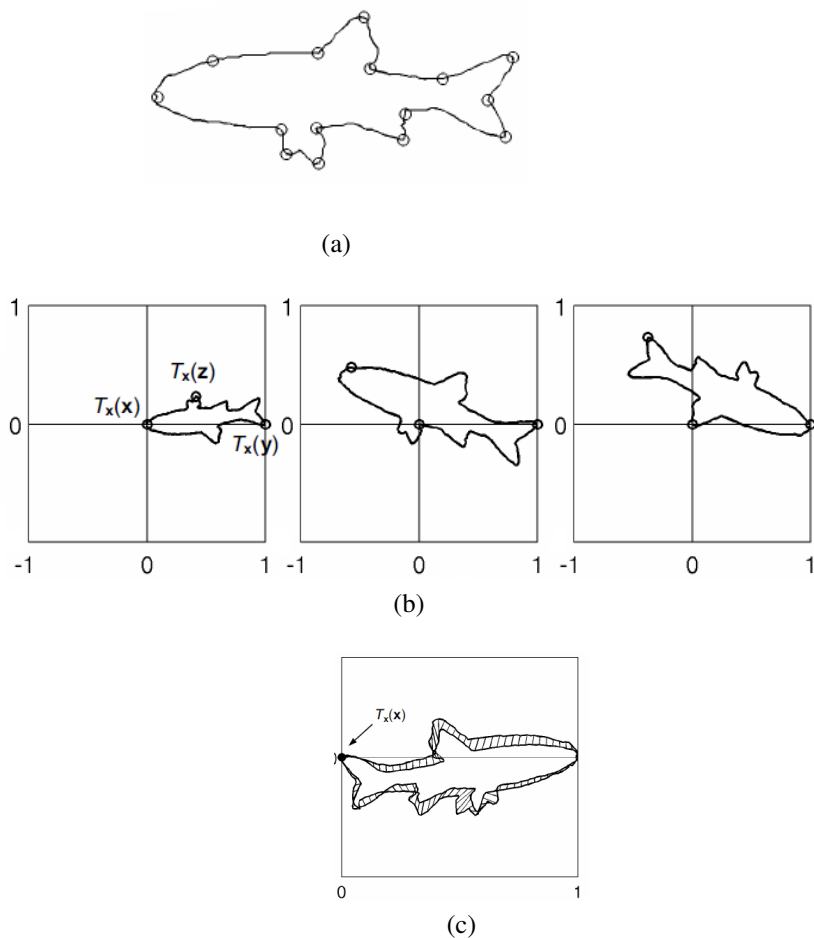


Figure 1.4 – Exemple d’une forme d’un poisson (a) et ses 15 points caractéristiques (petits ronds). Les points caractéristiques sont des points de maxima et de minima locaux de la courbure. En (b), trois poses ou contours normalisés, parmi les 15 poses possibles, de la forme représentée en (a). En (c) la mise en correspondance de la forme en (a) avec une autre forme de poisson dans le repère référentiel canonique (relatif à une des poses).

similarité soit invariante aux transformations géométriques (translation, rotation et facteur d’échelle) et aux reparamétrisations. Elle doit aussi être robuste au bruit et aux occlusions, etc.

1.3 Reconnaissance de formes basée-contour

Nous allons présenter dans cette section les grandes tendances de la reconnaissance de formes basée-contour. Nous ne pouvons pas fournir ici une liste exhaustive les

concernant, nous allons tout de même tenter de montrer la diversité des techniques employées dans ce domaine. Les méthodes basées recalage sont les techniques dont la reconnaissance est établie sur un processus de comparaison entre les formes par mise en correspondance ou recalage. Les méthodes primaires regroupent celles qui ne sont pas basées recalage.

1.3.1 Méthodes primaires

Les méthodes primaires de reconnaissance de formes sont les méthodes non-basées recalage, c'est-à-dire celles qui ne cherchent pas une correspondance optimale entre les points des courbes à comparer. Ce type de techniques ne produit pas toujours les résultats intuitifs car ces techniques ne prennent pas bien en compte les caractéristiques géométriques et topologiques des formes. Souvent, une approche naïve de correspondance est adoptée pour faire correspondre les vecteurs caractéristiques de deux formes données. Cette approche de correspondance autorise ainsi une comparaison directe des points de même indice sur les deux formes considérées.

Parmi ces méthodes de reconnaissance, on trouve les méthodes qui comparent les formes localement. Par exemple, la méthode proposée en [48] utilise les ondelettes tandis que d'autres [169, 170] utilisent la technique basée sur la courbure multi-échelle². On trouve également les méthodes qui font une comparaison spatialement globale. Par exemple, celles qui utilisent le graphe orienté acyclique [138], les vecteurs propres multicouches [229], les descripteurs de Fourier elliptiques [180] ou les descripteurs de Fourier 2D multi-échelle [67].

1.3.2 Méthodes basées recalage

Pour faire correspondre les vecteurs caractéristiques de deux formes données, ces techniques s'appuient sur un recalage préalable des représentations correspondantes. Le recalage consiste à trouver la transformation qui fait correspondre au mieux les points de caractéristiques similaires des deux formes considérées. L'objectif principal du recalage est de mettre deux formes *a priori* différentes dans un même système de référence.

Le recalage de formes 2D nécessite donc une méthode de correspondance entre leurs courbes représentatives. Pour réaliser une telle correspondance, chaque courbe est échantillonnée pour générer des séquences de points à recaler en fonction des caractéristiques géométriques de ces points. Ensuite, le recalage fait intervenir des transformations mathématiques dans le but de déformer une forme pour l'adapter à une forme de référence. Ainsi, c'est le type de déformation appliquée à la forme qui va déterminer le type de recalage. On entend par recalage rigide [6, 36, 144, 181, 250] et recalage non-rigide (élastique) [2, 55, 56, 183, 209] deux types de recalage de dimensions

²Curvature Scale Space (CSS) dans la littérature anglo-saxonne.

différentes. Malgré la simplicité du recalage rigide qui ne nécessite qu'un petit nombre de paramètres, ce recalage est peu utilisé en pratique car il confère peu de liberté aux transformations effectuées, appelées transformations géométriques de rotation, translation et d'homothétie. Par exemple, il est impossible de trouver une approximation rigide de la transformation entre deux formes d'une même classe qui présente une forte variabilité inter-individuelle. En outre, la principale caractéristique du recalage rigide est qu'il représente un recalage global, c'est-à-dire de petite dimension. Le recalage élastique est le type de recalage le plus utilisé car il donne un grand nombre de degrés de liberté aux transformations effectuées (appelées transformations non linéaires ou déformables); ces transformations ont un caractère local. En général, dans le cas où les formes sont représentées de manière à être invariantes par l'action des transformations colinéaires, le recalage rigide n'est pas nécessaire. En revanche, dans le cas de formes déformées, il faut trouver les transformations non rigides permettant à faire correspondre les formes entre elles. Ce type de recalage est utilisé dans nombreuses tâches du monde réel, y compris la reconnaissance de caractères écrits à la main [38], la reconnaissance de formes [178], le suivi de mouvement d'objets déformables [88, 143] et le recalage d'images médicales [56]. La plupart des méthodes de recalage fondées sur les transformations non rigides sont itératives et s'appuient sur la minimisation d'un critère d'appariement défini par une fonction de coût ou d'énergie associée à une mesure de similarité. Ces choix dépendent fortement du type d'application traitée et des points caractéristiques retenus sur la forme.

La diversité des domaines où le recalage est impliqué a mené à un développement très vaste dans les techniques employées et les descripteurs utilisés. Tout descripteur de forme invariant et assez discriminant peut être employé dans la mise en correspondance des courbes fermées [268]. Dans [79, 233], un descripteur général en matrice binaire a été proposé. Goshtasby *et al.* [80] utilisent les matrices de formes comme des descripteurs simples et rapides pour le recalage de formes dans des images satellitaires. Les matrices de formes sont utilisées ensuite sous différentes variantes dans plusieurs applications [268]. Dans [131] les auteurs proposent d'utiliser une représentation des contours par chaîne de codes et d'opérer une mise en correspondance moyennant une mesure de corrélation. Suk *et al.* [222] ont proposé d'utiliser une description invariante des polygones.

1.3.3 Méthodes basées géodésiques de formes

L'analyse de formes est un problème fondamental en vision par ordinateur. Au cours des dernières années, les géodésiques dans l'espace des formes ont été largement utilisées pour la comparaison des formes [21, 22, 109, 110, 160, 163, 256]. En outre, les espaces de formes ont été récemment étudiés pour l'analyse des variations de formes dans de nombreuses applications de vision par ordinateur [42, 51, 110, 219].

Espace de formes : L'étude quantitative d'une forme commence par la formulation d'un espace de formes \mathcal{C} , dans lequel chaque forme est représentée par un point. Étant donné un contour fermé planaire, le groupe des translations, rotations et changements d'échelle crée une famille de contours planaires. La seule propriété commune de tous les éléments de cette famille est leur forme. L'ensemble de toutes ces familles est appelé l'espace de formes. En général, l'étude de la géométrie de l'espace de forme est basée sur l'exploration des espaces non-linéaires, les variétés, qui sont généralement de grande dimension.

Variété de formes : L'espace de formes muni de la structure d'une variété différentielle, et en particulier de celle d'une variété Riemannienne permet de résoudre des problèmes variationnels en utilisant un calcul différentiel intrinsèque et des représentations locales. Une variété différentielle est un élément base de la topologie et de la géométrie différentielle. C'est l'espace sur lequel on peut avoir un espace tangent en tout point, ce qui nous permet également d'effectuer des opérations de calcul différentiel et intégral. Autrement dit, une variété différentielle, est un espace topologique sur lequel des systèmes de coordonnées locales, cartes locales, peuvent être définies au voisinage de n'importe quel point de cet espace. Par exemple en définissant l'espace de formes, Grenander [83] a considéré les formes comme des points sur une variété différentielle de dimension infinie. En plus, une variété Riemannienne est une variété différentielle munie d'une métrique riemannienne (c'est-à-dire un produit scalaire défini positif) sur l'espace tangent. Une variété Riemannienne de dimension finie est isométriquement incorporée dans un espace euclidien de dimension plus grande [166]. Dans la théorie de Kendall [102], les formes sont représentées par un nombre fini de points repères, et l'espace de formes est considéré comme une variété Riemannienne de dimension finie, souvent nommée variété de formes. Cependant, nombreuses sont les approches qui ont étudié les formes comme des éléments d'une variété de dimension infinie [110, 159, 204, 256].

L'étude mathématique des espaces de courbes a été largement abordée ces dernières années [37, 255]. L'ensemble des courbes est naturellement modélisé sur une variété riemannienne [158]. Plusieurs travaux récents [44, 158, 253, 254] soulignent que le choix de la métrique affecte notamment les résultats des algorithmes numériques de minimisation des fonctionnelles d'énergie. D'où, la conception d'une métrique avec des comportements spécifiques est très cruciale pour des applications dans le traitement des images (par exemple, la segmentation d'image) et la vision par ordinateur (par exemple, recalage des formes). Des exemples typiques de ces métriques riemanniennes sont des mesures de type Sobolev [224, 226, 227], qui conduisent à l'évolution des courbes lisses sans limiter leurs déformations.

Géodésiques de formes : Pour calculer la distance entre deux formes de l'espace de formes, il suffit simplement de prendre la longueur minimale parmi les chemins sur la variété qui relie ces deux formes [158, 225, 237, 241]. Pour deux formes quelconques

de la variété, le chemin réalisant ce minimum est appelé géodésique. Ces chemins géodésiques sont trouvés en minimisant une fonction de coût, qui mesure la similarité locale entre les formes. Autrement dit, leur détermination se fait en utilisant la géométrie locale de la variété.

Analyse de formes basée géodésiques : Les travaux sur les espaces de formes se distinguent principalement par la représentation des courbes qui peut être discrète (par l'utilisation des points repères) [33, 102] ou continue (par des fonctions continues) [83, 84, 162, 256]. En outre, les espaces de formes peuvent se différencier par les méthodes du calcul des géodésiques et les métriques développées pour l'analyse des formes sur l'espace de formes.

Étant donné un espace de formes muni de la structure d'une variété Riemannienne, le calcul des géodésiques dans l'espace de formes revient à résoudre une équation aux dérivées partielles (EDP) [158]. À cause du cadre mathématique rigoureux dans cette approche, il est difficile de trouver la plus courte géodésique entre deux formes données. En plus, Klassen *et al.* [110] ont également utilisé les variétés pour présenter les espaces métriques. Leur approche est axée sur l'étude des courbes planes fermées paramétrées par la longueur d'arc. Cette simplification a conduit à une équation différentielle ordinaire (ODE) au lieu d'un EDP pour le calcul des géodésiques. Une limitation clé de cette approche est que les distances calculées sont en général non symétriques, leur calcul est aussi intrinsèquement instable. Comparée à cette méthode, la méthode variationnelle est par définition plus stable, symétrique et permet un calcul des géodésiques plus rapide [204]. Par des approches variationnelles, [75, 161, 204, 256] ont proposé d'utiliser une métrique qui mesure le coût de déformation entre les éléments (les courbes) de l'espace de formes. Cette métrique utilise une combinaison de flexion et d'étirement/compression pour trouver les déformations optimales d'une forme à une autre. Grâce à cette métrique élastique, Mio *et al.* [164] ont développé un algorithme pour calculer les géodésiques entre des formes arbitraires. Plusieurs autres travaux ont mis en évidence les avantages de la métrique élastique [95, 158, 201, 221].

Récemment, de nombreuses approches d'analyse de formes fondées sur l'utilisation des géodésiques dans l'espace de formes ont été développées. Différentes techniques algorithmiques ont été proposées pour le calcul numérique des géodésiques entre les courbes fermées dans des espaces appropriés [85, 95, 204, 221]. Ces approches sont utilisées par exemple pour interpoler entre les formes [165, 219], pour extrapoler un changement de forme [110], ou pour étudier les statistiques de formes [110, 220]. Elles sont également utilisées pour l'analyse des variations de formes dans de nombreuses applications de la surveillance [42, 219], de l'imagerie médicale [43, 85, 152, 165, 219], de la biologie [177] et de la météorologie [51].

L'avantage principal de ce type d'approches est qu'il permet de calculer la similarité entre des paires de formes arbitraires représentées sur l'espace de formes. Étant données deux formes, la longueur minimale parmi les chemins géodésiques qui relient ces deux

formes fournit une métrique qui mesure la similarité locale entre elles. Nombreux sont les travaux qui ont développé des métriques basées sur les géodésiques de formes, mais peu sont les méthodes qui ont exploité de telles métriques pour la reconnaissance de formes. Par exemple, par des approches variationnelles, certains travaux [75, 161, 204, 256] ont proposé d'utiliser une métrique géodésique pour mesurer le coût de déformation entre les éléments (les courbes) de l'espace de formes. Dans les travaux de thèse de Nasreddine [176], la distance géodésique a été exploitée comme une métrique pour la classification et la recherche des formes fermées [178].

1.4 Normalisation de parties de formes

Le problème de la normalisation de formes entières (c'est-à-dire de courbes fermées) a été largement abordé de différentes manières dans de nombreux travaux antérieurs. Cependant, la normalisation de parties de formes a été peu abordée dans l'état de l'art. Ce sujet constitue un problème non-trivial à étudier car pas toutes les méthodes proposées pour les formes fermées peuvent être directement appliquées sur les courbes ouvertes. Les invariants ou les descripteurs de Fourier par exemple supposent une périodicité sur le signal, ce qui est le cas des courbes fermées seulement. Dans cette section, nous décrivons brièvement les différentes méthodes utilisées dans la littérature pour normaliser les parties de formes (c'est-à-dire les courbes ouvertes). La plupart de ces méthodes sont des techniques de prétraitement permettant d'éliminer tout effet de translation, rotation et de facteur d'échelle avant de passer à la phase d'extraction des caractéristiques.

1.4.1 Représentation invariante dans un repère canonique

L'approche de Sun et Super [223] s'appuie sur le fait que le segment d'un contour est un simple modèle générateur dans un espace invariant de segments. Chaque segment, autrement dit partie de forme, est normalisé par rapport aux transformations géométriques par une représentation dans un repère référentiel canonique (figure 1.5). La même représentation invariante a été utilisée par Super [228] pour estimer la pose optimale d'une forme fermée de contour fermé dans le but de l'utiliser dans le système de correspondances fixées avec mesure de probabilité de chance³. Dans un premier temps, chaque partie de forme P est échantillonnée sur N points à intervalle régulier de longueur d'arc, c'est-à-dire $P = \{p_1, \dots, p_N\}$. Ensuite, la partie de forme P est transformée dans un repère référentiel canonique (x', y') en une autre partie de forme $P' = \{p'_1, \dots, p'_N\}$. Les deux points extrémités (p_1, p_N) de P sont utilisés pour calculer la transformation de P à P' : p_1 est mappé au point $(0, 0)$ et p_N est mappé au point $(1, 0)$ comme illustré en figure 1.5 ci-dessous. Ainsi, les coordonnées des points du segment P

³Chance Probability Function (CPF) dans la littérature anglo-saxonne.

dans la nouvelle référence $\{p'_1, \dots, p'_N\}$ sont invariantes à la translation, à la rotation et au changement d'échelle. Cette méthode a l'avantage d'être rapide et facile à implémenter mais elle est sensible aux extrémités des courbes dans le cas de fortes distorsions.

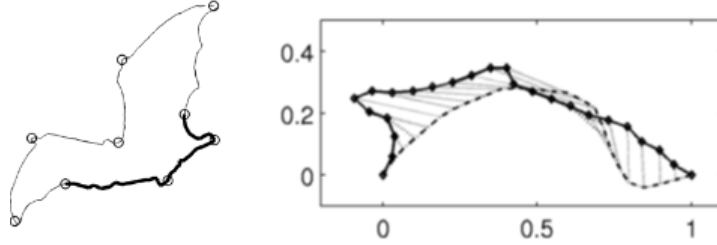


Figure 1.5 – La partie de forme d'une chauve-souris (en gras à gauche) est représentée dans un repère invariant où les extrémités sont mappées aux points $(0, 0)$ et $(1, 0)$.

1.4.2 Normalisation par alignement angulaire

Étant données deux parties de formes P et \tilde{P} échantillonnées à N points, Latecki *et al.* [122] ont représenté chaque partie de forme par sa fonction d'angle (θ et $\tilde{\theta}$) paramétrée par l'abscisse curviligne s normalisée à l'unité. L'utilisation de la signature angulaire $\theta(s)$ réalise naturellement l'invariance à la translation et au changement d'échelle. En effet, toute translation de la courbe n'a aucune incidence sur θ , et toute homothétie n'a aucun effet sur le paramètre normalisé s . En outre, un alignement angulaire a été proposé pour gérer la différence d'orientation entre chaque paire de parties de formes. La méthode utilisée cherche à trouver l'angle optimale θ_0 qui minimise la distance \mathbb{L}^2 entre les fonctions d'angle θ et $\tilde{\theta}$ des parties de formes P et \tilde{P} respectivement. Autrement dit, θ_0 est la translation de l'une des deux fonctions d'angle (θ si \tilde{P} est choisie comme référence) qui minimise l'intégrale de la différence des fonctions d'angle en tout point des deux parties de formes (voir figure 1.6). Cette intégrale est donnée par l'équation suivante :

$$\theta_0^* = \underset{\theta_0 \in [0, 2\pi[}{\operatorname{argmin}} \int_{s=0}^1 \left(\theta(s) - \tilde{\theta}(s) + \theta_0 \right)^2 ds. \quad (1.1)$$

Cette méthode est simple à implémenter, mais elle présente une complexité au niveau du temps de calcul et une grande sensibilité aux valeurs aberrantes en présence du bruit.

1.4.3 Alignement Procruste

Dans le cadre d'une application biomédicale, Mani [152] a utilisé l'analyse Procrustéenne pour aligner des courbes ouvertes. Pour ce faire, il a cherché à optimiser la différence entre les courbes données en utilisant une approximation du type moindre carrés qui

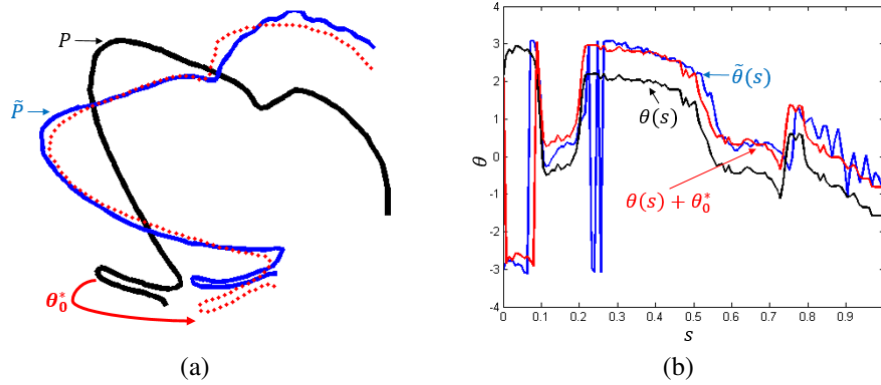


Figure 1.6 – Normalisation par alignement angulaire de deux parties de papillons P et \tilde{P} données en (a). En (b), la fonction $\theta(s) + \theta_0^*$ représente la translation de la fonction d'angle $\theta(s)$ de la partie de forme P de l'angle optimale θ_0^* qui minimise l'intégrale de la différence des fonctions d'angle $\theta(s)$ et $\tilde{\theta}(s)$ des deux parties de formes P et \tilde{P} .

exige une correspondance point-à-point des deux courbes. Ce type d'alignement consiste à suivre les mêmes étapes utilisées pour l'alignement des courbes fermées représentant des formes entières. Au niveau de l'alignement Procuste, on trouve tout d'abord le barycentre de chaque forme, puis on change les échelles des formes pour qu'elles aient une taille égale. L'alignement des formes se fait en position par rapport aux coordonnées des centroïdes et les effets rotationnels sont filtrés par le calcul de l'angle d'orientation optimale qui minimise la différence entre les courbes étudiées (figure 1.7). Étant données les deux parties de formes P et \tilde{P} représentées par N points, $P = \{p_1, \dots, p_N\}$ et $\tilde{P} = \{\tilde{p}_1, \dots, \tilde{p}_N\}$, les paramètres optimaux $(\alpha, t, \gamma)^*$ sont respectivement les paramètres d'échelle, de position et d'orientation qui minimisent une somme quadratique suivant l'équation suivante :

$$(\alpha, t, \gamma)^* = \operatorname{argmin}_{\alpha, t, \gamma} \frac{1}{N} \sum_{i=1}^N \|\tilde{p}_i - (\alpha \Gamma p_i + t)\|^2 \quad (1.2)$$

où α est un paramètre d'échelle, $t = \begin{pmatrix} t_x \\ t_y \end{pmatrix} \in \mathbb{R}^2$ est un vecteur de translation et $\Gamma = \begin{pmatrix} \cos \gamma & \pm \sin \gamma \\ \sin \gamma & \mp \cos \gamma \end{pmatrix}$ est une matrice de rotation $2D$ avec $\gamma \in [0, 2\pi[$.

1.4.4 Alignement optimal par estimation du point initial et de pose

Marques et Arbrantes [153] proposent un algorithme pour l'estimation optimale des paramètres de la pose, l'échelle et du point initial. Étant données deux parties de formes $P = \{p_1, \dots, p_N\}$ et $\tilde{P} = \{\tilde{p}_1, \dots, \tilde{p}_N\}$, elles sont d'abord centrées à l'origine du système

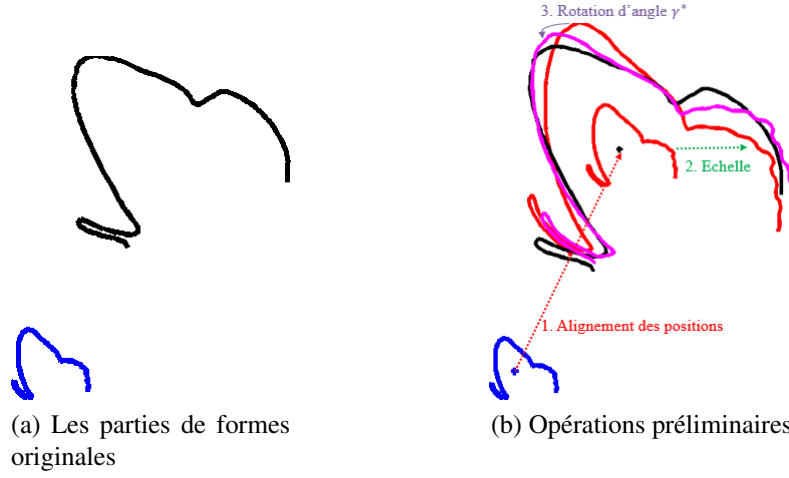


Figure 1.7 – L’alignement Procuste de deux parties de papillons (a) qui consiste à les aligner en position, en échelle et en rotation (b).

de coordonnées afin de les rendre invariantes à la translation. Cette méthode est basée sur la supposition que l’une des deux parties, P par exemple, est une version déformée, tournée et multipliée par un facteur d’échelle α de l’autre partie \tilde{P} , dite partie modèle. Comme le point initial est souvent inconnu, l’algorithme proposé autorise un décalage circulaire du point initial sur la partie de forme modèle pour établir la meilleure correspondance des points entre P et \tilde{P} (figure 1.8). Par conséquent, la formulation adoptée dans cette méthode est donnée par :

$$P = D_m \Gamma \alpha (\tilde{P} + d) \quad (1.3)$$

où d est une matrice de déformation de formes de dimension $2 \times N$, Γ est la matrice de rotation et D_m est l’opérateur de décalage circulaire d’amplitude m définie par :

$$D_m \{\tilde{p}_1, \dots, \tilde{p}_N\} = \{\tilde{p}_m, \dots, \tilde{p}_N, \tilde{p}_1, \dots, \tilde{p}_{m-1}\}. \quad (1.4)$$

Ensuite, les paramètres de la pose, l’échelle et du point initial m sont estimés en minimisant une mesure de déformation de forme S :

$$S = \|P - D_m \Gamma \alpha \tilde{P}\|^2 \quad (1.5)$$

L’algorithme proposé est simple à implémenter. Des essais expérimentaux montrent que cette technique est capable d’aligner des formes complexes et elle est robuste contre les déformations de parties de formes et les changements de point initial. Malgré toutes ces avantages, l’implémentation de cette méthode nécessite beaucoup de temps à cause du décalage circulaire du point initial.

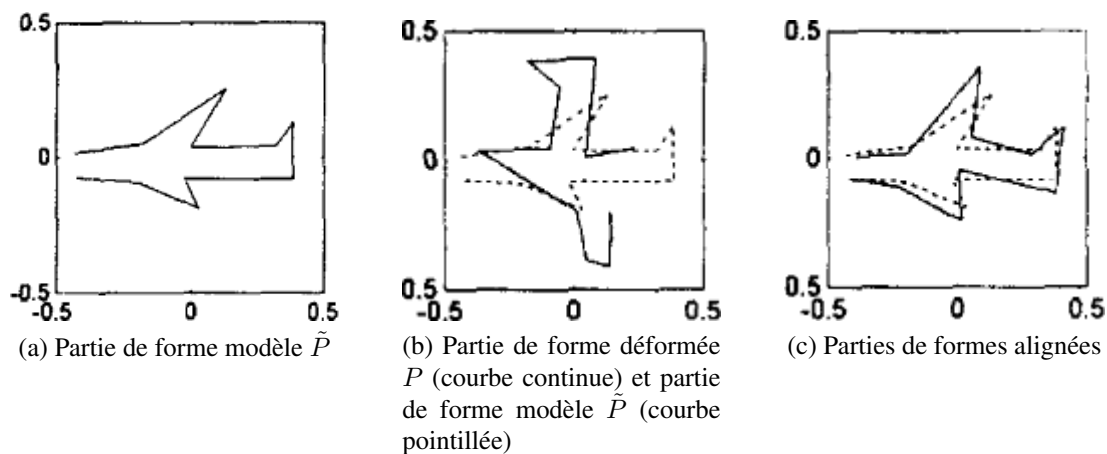


Figure 1.8 – Alignement optimal des parties de formes par estimation du point initial, la rotation, facteur d'échelle et de la déformation appliqués [153].

1.5 Représentation et description de parties de formes

Dans la littérature, de nombreuses techniques de description ont été utilisées en reconnaissance de formes planaires [41, 71, 114, 139, 171, 229, 261]. Cependant, seulement quelques-unes ont été adaptées pour la description de parties de formes. Dans ce qui suit, chaque partie de forme est une courbe ouverte C échantillonnée sur N points équidistants. La représentation usuelle des courbes est donnée par leurs coordonnées x et y . Une courbe paramétrique C est une courbe dont les coordonnées cartésiennes (x, y) de ses points sont exprimées en fonction d'un paramètre intermédiaire t tel que : $C(t) = (x(t), y(t))$. Différentes représentations autres que les fonctions coordonnées sont également utilisées pour la correspondance des courbes. La représentation des caractéristiques extraites est généralement faite par un descripteur de forme constitué par un ensemble de vecteurs. Dans cette section, nous présentons plusieurs représentations et descripteurs de parties de formes proposés dans la littérature.

1.5.1 Fonction d'angle

La fonction d'angle connue par la fonction tangente, en un point p d'une courbe C , est définie par l'angle formé entre la tangente à la courbe au point p et l'axe des abscisses. Elle s'exprime en fonction de la longueur d'arc, c'est-à-dire de l'abscisse curviligne s . Le changement des directions angulaires est une caractéristique importante de la forme. Par conséquent, toute partie de forme peut être représentée par sa fonction d'angle $\theta(s)$ donnée par la formule suivante :

$$\theta(s) = \arctan \frac{y(s) - y(s - w)}{x(s) - x(s - w)} \quad (1.6)$$

où $w \in \mathbb{N}$ représente le pas utilisé pour le calcul. La fonction d'angle $\theta(s)$ ne peut prendre des valeurs que dans un intervalle de longueur 2π , habituellement dans l'intervalle $[-\pi, \pi]$ ou $[0, 2\pi]$. Cette représentation a été utilisée en reconnaissance de formes partielles [126, 231, 238]. Il est important de noter que la fonction d'angle représentative d'une polyligne issue d'une simplification de la courbe C est une fonction constante par morceaux; elle augmente ou diminue au niveau des sommets et elle reste constante entre deux sommets consécutifs.

Cette signature se caractérise par l'invariance à la translation. En outre, le problème de la longueur est facilement résolu en représentant chaque partie de forme par sa fonction d'angle $\theta(s)$ paramétrée par l'abscisse curviligne s normalisée à l'unité [119, 123, 126]. Cette signature n'est par contre pas invariante à la rotation. Une rotation d'angle c transforme la fonction $\theta(s)$ en $\theta(s) + c$ modulo 2π .

1.5.2 Fonction de courbure

La courbure est une caractéristique très utilisée en reconnaissance de formes pour la représentation du contour [248]. Pour un contour paramétré par la longueur d'arc s , la courbure est définie par :

$$k(s) = \frac{d\theta(s)}{ds} \quad (1.7)$$

où θ est la fonction d'angle donnée par l'équation 1.6. On peut aussi utiliser une autre définition équivalente de la courbure donnée par :

$$k(s) = \frac{\dot{x}(s)\ddot{y}(s) - \ddot{x}(s)\dot{y}(s)}{(\dot{x}(s)^2 + \dot{y}(s)^2)^{\frac{3}{2}}}, \quad (1.8)$$

où $\dot{x}(s)$ et $\ddot{x}(s)$ représentent respectivement la dérivée première et la dérivée seconde de x au point s . De même $\dot{y}(s)$ et $\ddot{y}(s)$ représentent respectivement la première et la seconde dérivée de y au point s . En utilisant une abscisse curviligne s normalisée à l'unité, la signature $k(s)$ est invariante en rotation, en translation et en changement d'échelle. Cui et al. [57] en ont développé une méthode de correspondance des courbes ouvertes basée sur une nouvelle signature invariante au changement d'échelle. Cette signature est dérivée du concept de l'intégrale de la courbure non signée $|k(s)|$.

1.5.3 Contexte de forme

Dans [24, 25], Belongie *et al.* ont introduit une nouvelle méthode de représentation de formes fermées basée sur le contexte de forme⁴. Ce descripteur a été récemment proposé comme un outil puissant pour la description de fragments de contours [14, 19, 247]. Le principe est de décrire localement une partie de forme en codant la distribution relative

⁴Shape Context (SC) dans la littérature anglo-saxonne.

(distance et orientation) des points du plan par rapport à chaque point de la forme pris comme référence. Pour chaque point de référence p_i , la répartition de la silhouette par rapport au point p_i est exprimée par les vecteurs originaires de ce point à tous les autres points échantillonnés sur la partie de forme considérée. Le contexte de forme d'un point p_i est l'histogramme log-polaire h_i (de tous les autres points) défini par l'équation suivante :

$$h_i(k) = \#\{u \neq p_i : (u - p_i) \in \text{classe}(k)\}, \quad (1.9)$$

où les classes $\text{classe}(k)$ sont définies en utilisant une partition uniforme des directions et des distances de l'histogramme Log-polaire centré au point de référence p_i . L'histogramme $h_i(k)$ est le nombre de points du contour de l'objet qui se trouvent dans la $k^{\text{ième}}$ classe ($\text{classe}(k)$).

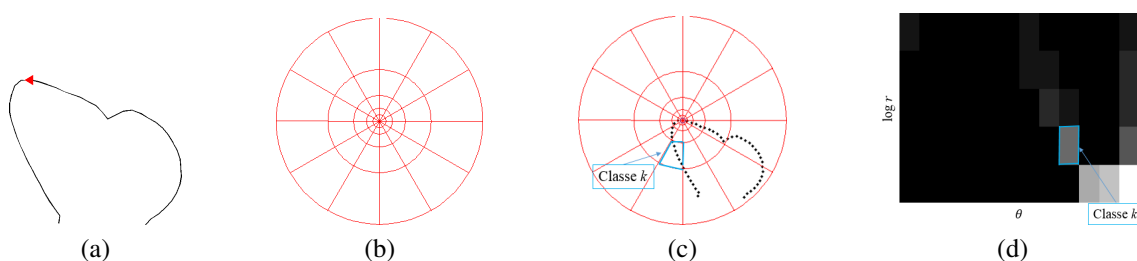


Figure 1.9 – Calcul du descripteur contexte de forme. (a) : Partie de forme et point de référence considérés. (b) : Le diagramme de l'histogramme log-polaire utilisé pour le calcul du contexte de forme. Cinq points sur la direction radiale $\log r$ et douze points sur la direction angulaire θ sont considérés. (c) : Le contexte de forme par rapport au point indiqué \triangleleft sur la partie de forme en (a). (d) : Le contexte de forme est un histogramme log-polaire des coordonnées des autres points de la forme mesurées par rapport au point de référence. Sur les figures (c) et (d) est repérée une même classe k (contour bleu).

La figure 1.9 représente un exemple de calcul du contexte de forme d'une partie d'un papillon. En figure 1.9b est illustré le diagramme log-polaire utilisé pour le calcul du contexte de forme. Ce diagramme considère cinq points sur la direction radiale et douze points sur la direction angulaire. L'histogramme est formé en plaçant le centre de ce diagramme sur le point de référence considéré (marqué par \triangleleft sur la partie de forme de la figure 1.9a).

Grâce à ses propriétés d'invariance à la translation, rotation et au changement d'échelle, les contextes de forme ont été utilisés dans une variété de problèmes de reconnaissance d'objets [25, 174, 235, 263].

Dans [147], Ma et Latecki ont proposé deux matrices, des distances et des orientations, comme un nouveau descripteur de fragments de contours. Ces distances et ces orientations sont calculées entre toutes les paires de points des deux parties de formes à comparer, formant ainsi une description de la relation géométrique relative aux

parties. La distance $D^{(P,\tilde{P})}(i, j)$ entre le point p_i et le point \tilde{p}_j sur les parties de formes P et \tilde{P} respectivement est définie à partir de la distance euclidienne $\|\cdot\|_2$ dans l'espace log comme suit :

$$D^{(P,\tilde{P})}(i, j) = \log(1 + \|p_i - \tilde{p}_j\|_2). \quad (1.10)$$

L'orientation $\theta^{(P,\tilde{P})}(i, j)$ entre le point p_i et le point \tilde{p}_j est l'angle formé par le vecteur $\overrightarrow{p_i\tilde{p}_j}$ et l'axe horizontal :

$$\theta^{(P,\tilde{P})}(i, j) = \angle(\overrightarrow{p_i\tilde{p}_j}) \in [-\pi, \pi]. \quad (1.11)$$

Géométriquement, ce descripteur est basé sur le même principe du contexte de forme. Cependant, il ne nécessite pas une construction des histogrammes.

1.5.4 Descripteurs de Fourier

Chaque point de la courbe considérée C est représenté par une valeur numérique associée à l'information géométrique correspondante. Le descripteur de Fourier de la forme est défini comme la transformée de Fourier de la séquence ordonnée des valeurs obtenues, du point de départ jusqu'au point final. Ce descripteur représente l'information globale de la forme dans l'espace de fréquences.

Le descripteur de Fourier est un descripteur compact des formes utilisé pour la correspondance et la reconnaissance de formes. Il est ainsi connu comme un moyen efficace de reconstruction, d'enregistrement et de compression d'une courbe : quelques coefficients basses fréquences sont suffisants pour reconstruire le contour d'origine. Cependant, lorsque les deux points d'extrémité ne sont pas adjacents, il en résulte un signal discontinu. Le segment reconstruit après avoir éliminé la partie à fréquences hautes présente une distorsion importante autour des deux extrémités. Dans ce paragraphe, nous présentons les différents schémas de descripteurs de Fourier proposés pour traiter néanmoins le cas des courbes ouvertes.

Descripteur de Fourier par Tangente Complexe (DFTC) : Le descripteur de Fourier par tangente complexe⁵, connu encore sous le nom de descripteur de Fourier de type-P, est le premier descripteur de Fourier proposé pour caractériser les courbes ouvertes [240] à partir des informations de pente. La courbe ouverte considérée est ré-échantillonnée sous la forme d'une séquence ordonnée de N points également espacés :

$$\left\{ (x_0, y_0), \dots, (x_N, y_N) \mid \sqrt{(x_{n+1} - x_n)^2 + (y_{n+1} - y_n)^2} = \delta, \forall n \in \{0, \dots, N-1\} \right\},$$

où δ est une constante. Uesaka [240] a défini ensuite des fonctions complexes w_n comme suit :

$$w_n = \frac{x_{n+1} - x_n}{\delta} + j \frac{y_{n+1} - y_n}{\delta}, \forall n \in \{0, \dots, N-1\} \text{ où } j = \sqrt{-1}.$$

⁵Complex Tangent Fourier Descriptor (CTFD) dans la littérature anglo-saxonne.

Enfin, la transformée de Fourier discrète⁶ est appliquée à la séquence $\{w_0, \dots, w_{N-1}\}$:

$$c_k = \frac{1}{N} \sum_{n=0}^{N-1} w_n \exp\left(-j \frac{2\pi kn}{N}\right), \forall k \in \{0, \dots, N-1\}. \quad (1.12)$$

L'ensemble de ces nombres complexes c_0, \dots, c_{N-1} sont les descripteurs DFTC, et l'ensemble des $2N$ nombres réels $\Re(c_0), \Im(c_0), \dots, \Re(c_{N-1}), \Im(c_{N-1})$ constitue le vecteur caractéristique de la courbe ouverte. Depuis, les DFTC ont été utilisés dans différentes applications comme la reconnaissance de profils de visage humain [4], la génération des trajectoires pour la conception automatique de mécanismes [257], l'analyse statistique sur le pétale de lotus [265] ou sur la feuille de riz [266].

Cette transformée de Fourier présente de nombreux avantages par rapport au descripteur de Fourier traditionnel qui utilise les coordonnées complexes, car la pente relative est invariante à la translation, à la rotation et à l'échelle et fonctionne bien pour les deux courbes, fermées et ouvertes.

Descripteur de Fourier Anti-symétrique : Récemment, Ding *et al.* [66] ont construit un nouveau schéma pour la description des segments non fermés. Dans cette approche, un segment de contour est d'abord représenté par une séquence $s(n)$ de N coordonnées complexes comme suit :

$$s(n) = x_n + jy_n, \forall n \in \{0, \dots, N-1\}.$$

La séquence $s(n)$ est évidemment discontinue aux deux extrémités (voir figure 1.10a) et sa transformée de Fourier discrète contiendra plus de coefficients de haute fréquence. Afin de réduire ces composantes de fréquences hautes dérivées des points extrémités, une idée intuitive consiste à déformer le segment original ouvert en un segment fermé. Pour ce faire, Ding *et al.* ont proposé d'utiliser l'algorithme de décalage qui consiste à décaler linéairement tout point (x_n, y_n) du segment original au (x'_n, y'_n) en fonction de son indice et de la différence entre les points extrémités. La règle est la suivante :

$$\begin{aligned} x'_n &= x_n - x_0 - \{(x_{N-1} - x_0) \times n / (N - 1)\}, \forall n \in \{0, \dots, N - 1\} \\ y'_n &= y_n - y_0 - \{(y_{N-1} - y_0) \times n / (N - 1)\}, \forall n \in \{0, \dots, N - 1\} \end{aligned}$$

avec $x'_0 = x'_{N-1} = 0$ et $y'_0 = y'_{N-1} = 0$. Le segment fermé $s_D(n)$ obtenu par une telle déformation est illustré dans la figure 1.10b :

$$s_D(n) = \{x'_n + jy'_n, \forall n \in \{0, \dots, N - 1\}\}.$$

Cependant, même après ce décalage linéaire, les deux extrémités reliées peuvent conduire à un angle aigu, ce qui peut augmenter à son tour la possibilité d'obtenir des

⁶Discrete Fourier Transform (DFT) dans la littérature anglo-saxonne.

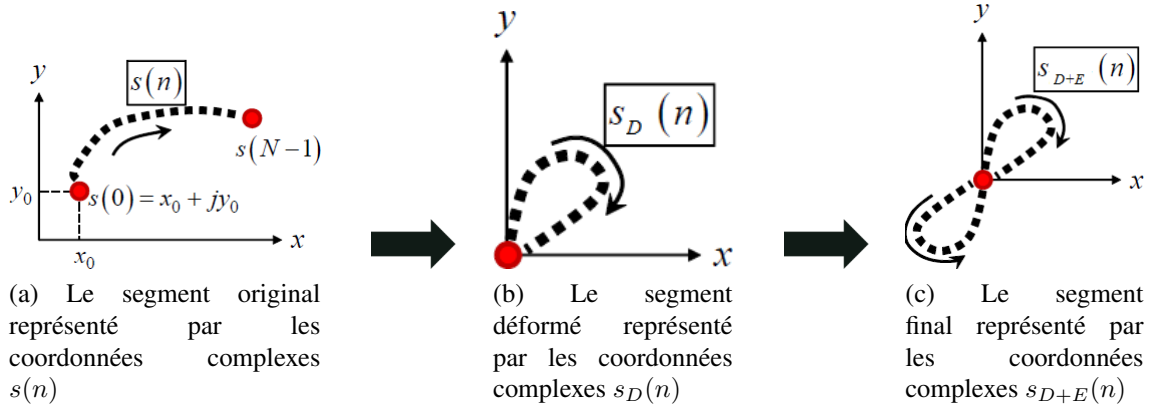


Figure 1.10 – Illustration de la méthode de déformation et d'extension du segment. (a) : Segment ouvert d'origine. (b) : Le décalage linéaire du segment ouvert en (a). (c) : L'extension anti-symétrique du segment résultant du décalage linéaire en (b). Le segment final en (c) est une courbe fermée, continue et différentielle d'ordre 1 au niveau des deux points d'extrémité.

composantes de fréquences hautes. Pour lisser l'angle aigu généré, Ding *et al.* [66] ont proposé un système d'extension anti-symétrique qui consiste à ajouter un segment anti-symétrique de $s_D(n)$ au segment $s_D(n)$ comme illustré en figure 1.10c. Ce nouveau segment s_{D+E} à $2N - 2$ points, obtenu par la déformation et l'extension appliquées, est défini comme suit :

$$s_{D+E}(n) = \begin{cases} s_D(n), \forall n \in \{0, \dots, N-1\}, \\ -s_D(2N-2-n), \forall n \in \{N, \dots, 2N-3\} \end{cases}$$

En appliquant la transformée de Fourier classique sur la séquence $s_{D+E}(n)$ du segment fermé lisse, obtenu par le décalage linéaire et l'extension anti-symétrique du segment ouvert $s(n)$, les coefficients de la série de Fourier $S(k)$ sont donnés par l'équation suivante :

$$S(k) = \frac{1}{2N-2} \sum_{n=0}^{2N-3} s_{D+E}(n) e^{-j2\pi nk/(2N-2)}. \quad (1.13)$$

La valeur moyenne de $s_{D+E}(n)$ est nulle et $S(k)$ est anti-symétrique avec $S(k) = -S(2N-2-k)$, les deux coefficients $S(0)$ et $S(N-1)$ sont égaux à 0. Ainsi, en pratique seulement les coefficients $S(k)$ pour $k \in \{1, 2, \dots, N-2\}$ sont requis.

Pour appliquer aux courbes ouvertes les descripteurs de courbes fermées, il existe d'autres approches basées sur la fermeture des courbes ouvertes. Lestrel *et al.* [130] ont utilisé la courbe symétrique par rapport à la ligne droite qui passe par les deux points

d'extrémité de la courbe ouverte initiale. Tandis que Kawamura et Yokota [100] ont proposé de relier les extrémités du segment ouvert original avec celles de sa courbe demi-tournée. Ces techniques ne fonctionnent pas toujours bien. En effet, dans le cas où la forme géométrique obtenue au voisinage des points de connexion affecte fortement la description d'une courbe, le descripteur peut ne pas saisir la subtilité de la forme. En plus, ces opérations peuvent générer des croisements de courbes.

1.5.5 Représentations élastiques

Fonction d'angle et fonction de Log-vitesse : Dans [164,165], Mio *et al.* ont développé une théorie de formes planes en utilisant des modèles élastiques des courbes ouvertes ou fermées pour la représentation et l'analyse de formes. Les courbes planes codées par leurs fonctions de vitesse et d'angle (φ, θ) sont représentées en coordonnées Log-polaires. Dans cette approche, toute courbe élastique paramétrée $C(t)$ est représentée par une paire de fonctions : la fonction d'angle θ et la fonction de Log-vitesse φ . Afin de calculer ces deux fonctions, le vecteur vitesse $\dot{C}(t)$ de la courbe $C(t)$ est calculé et posé sous la forme suivante :

$$\dot{C}(t) = e^{\varphi(t)} e^{j\theta(t)}. \quad (1.14)$$

La fonction Log-vitesse φ code le taux d'étirement ou de compression le long de la courbe $C(t)$ tandis que la fonction d'angle θ reflète sa flexion.

$$\varphi(t) = \log |\dot{C}(t)|, \quad (1.15)$$

et θ est l'angle que fait le vecteur vitesse \dot{C} avec l'axe horizontal. La figure 1.11 montre que l'utilisation des fonctions de Log-vitesse φ et d'angle θ pour la correspondance des formes est indépendante des déformations élastiques. La paire (φ, θ) associée à une courbe paramétrique donnée est une représentation invariante par translation, mais sensible à la rotation et à l'échelle. Pour rendre cette représentation invariante à l'échelle, les fonctions φ et θ sont utilisées avec normalisation de la longueur des courbes à l'unité. Pour obtenir l'invariance à la rotation, on fixe la moyenne des valeurs d'angle θ à une valeur $\bar{\theta}$ donnée.

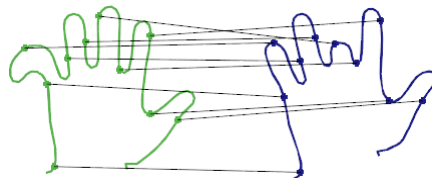


Figure 1.11 – Correspondance optimale entre deux formes de geste manuel obtenue avec une représentation élastique de la forme basée sur l'utilisation de la fonction Log-vitesse et la fonction d'angle [165].

Fonction de la racine carrée de la vitesse⁷ : En vertu de l'utilisation de la représentation par la fonction d'angle θ et la fonction Log-vitesse φ , la métrique varie d'un point à l'autre sur la variété et elle devient donc difficile à implémenter. Pour cette raison, Joshi *et al.* [95, 96] et Srivastava *et al.* [221] ont proposé une nouvelle représentation d'une forme élastique en utilisant la racine carrée de la vitesse (SRV). Pour une courbe paramétrique $C(t)$ donnée, la fonction de la racine carrée de la vitesse $q(t)$ est définie comme suit :

$$q(t) = \dot{C}(t) / \sqrt{\|\dot{C}(t)\|}. \quad (1.16)$$

Un des avantages de cette représentation est qu'elle est valable pour l'espace euclidien \mathbb{R}^n , $\forall n \in \mathbb{N}$. En effet, toute courbe simple et non nécessairement fermée peut être représentée par une seule fonction $q(t)$ au lieu d'une paire de fonctions (φ, θ) [164, 165]. De plus, cette représentation est plus facile à implémenter.

1.6 Mise en correspondance de parties de formes

En reconnaissance de formes, la plupart des méthodes utilisent des techniques de correspondance de formes ou de correspondance partielle de formes, mais la correspondance entre les parties de formes reste très peu abordée dans la littérature. Dans les techniques de correspondance de formes $2D$, les courbes fermées sont entièrement considérées dans la mise en correspondance [176]. En revanche, les méthodes de correspondance partielle de formes sont basées sur la recherche de la meilleure correspondance entre une forme partielle et une partie, voire des parties, d'une forme fermée donnée. Ces méthodes admettent que le contour ouvert (requête) est inclus dans l'autre forme fermée considérée. Selon le contexte applicatif, la forme partielle pourrait bien être une partie de forme ou un ensemble de parties de formes non adjacentes. Par exemple, pour reconnaître des formes partiellement occultées/déformées, les méthodes locales de correspondance basées sur les algorithmes génétiques [182] ou sur l'algorithme de correspondance sous-matricielle [198], l'algorithme de Smith-Waterman [46] ou encore la technique de déformation temporelle dynamique DTW⁸ [104] ont été utilisées pour trouver les parties similaires entre la requête et d'autres formes fermées. Dans le cas où la requête est une partie de forme, plusieurs travaux ont proposé différentes signatures pour étudier le problème de la correspondance entre contour ouvert et contour fermé. Par exemple, les algorithmes DTW avec un descripteur d'échelle [157, 167], OSB⁹ [127], LCSS¹⁰ [127] et MVM¹¹ [126] traitent

⁷Square Root Velocity (SRV) dans la littérature anglo-saxonne.

⁸Dynamic Time Warping.

⁹Optimal Subsequence Bijection.

¹⁰Longest Common Subsequence.

¹¹Minimum Variance Matching.

différemment les problèmes de correspondance et des valeurs aberrantes. Dans [116], les filtres à particules PF¹² et l'alignement Procuste ont été utilisés pour construire des correspondances entre les formes partielles.

La correspondance de parties de formes et la correspondance partielle de formes sont deux mécanismes qui se réfèrent à deux tâches différentes. Étant données deux parties de formes représentées par des séquences de points, leur mise en correspondance consiste à trouver la correspondance optimale entre ces deux parties de formes en tant que deux courbes ouvertes à comparer en tout point [14,19,71,223]. Avec l'utilisation de différentes mesures de similarité, cette correspondance de séquences de points peut être achevée par :

- une correspondance directe des points [116], ou
- une correspondance élastique des points [104,126].

En général, les algorithmes de correspondance directe consistent à faire correspondre directement les points un à un par le même ordre que l'échantillonnage. Néanmoins, ils ne peuvent pas faire correspondre correctement les échantillons homologues comme illustré en figure 1.12 ci-dessous. De fait, les résultats intuitifs de correspondance ne peuvent pas être toujours atteints avec l'utilisation de la distance euclidienne [90,155,187,188]. Par conséquent, la similarité calculée à partir de cette correspondance n'est pas toujours la distance optimale car elle est très sensible aux valeurs aberrantes. Pour résoudre ce problème de correspondance sous-optimale de points, différents

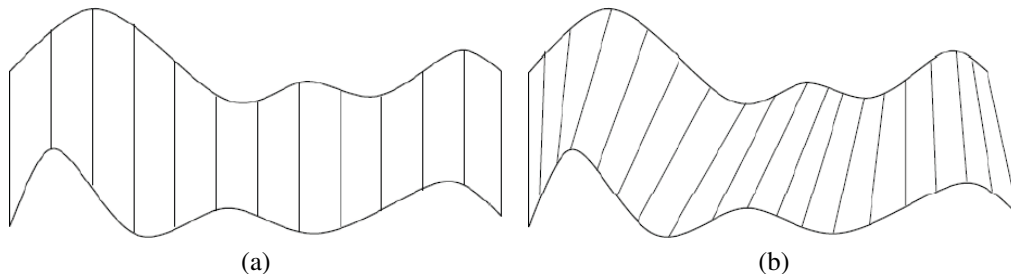


Figure 1.12 – Comparaison de deux courbes ouvertes par correspondance directe des échantillons (a) et par DTW (b) [188]. Comme on peut le voir, la comparaison par DTW est plus intuitive que la comparaison directe des échantillons.

algorithmes ont été utilisés pour établir une correspondance élastique des points [26,172,199]. Certains de ces algorithmes permettent d'ignorer les points aberrants [125,127,243] ou d'ignorer l'ordre des points échantillonnés [127]. À l'encontre, le DTW consiste à établir une correspondance élastique entre tous les points

¹²Particle Filters.

de séquences sans perdre l'ordre d'échantillonnage. L'efficacité de ce type de correspondance réside dans la capacité de garder les petits détails aidant à discriminer des parties de formes similaires mais appartenant à des classes différentes. La figure 1.12 montre la supériorité de la correspondance élastique optimale des points, obtenue en minimisant la distance DTW définie comme une fonction de coût.

Bai *et al.* [19] ont utilisé le contexte de forme à des fins de classification et de recherche dans une grande base de parties étiquetées. D'autres auteurs [14, 223] ont exploité la correspondance entre les parties de formes pour la reconnaissance de formes fermées à partir de leurs parties de formes.

Dans la suite de cette section, nous décrivons les mesures de similarité les plus utilisées pour comparer des parties de formes. Étant données deux parties de formes P et \tilde{P} représentées par leurs vecteurs caractéristiques $A = (a_1, a_2, \dots, a_N)$ et $\tilde{A} = (\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_N)$, la distance entre ces deux parties de formes est essentiellement la distance entre les deux vecteurs A et \tilde{A} .

1.6.1 Distance de Minkowski

La distance de Minkowski est utilisée pour mesurer la similarité entre les éléments de l'espace de Lebesgue \mathbb{L}^p en se basant sur la norme $\|\cdot\|_p$. Cette distance est définie par :

$$d_p(P, \tilde{P}) = \left[\sum_{i=1}^N |a_i - \tilde{a}_i|^p \right]^{1/p} \quad (1.17)$$

La distance de Minkowski est plus utilisée avec $p = 1, 2$ ou ∞ . Par exemple, Latecki *et al.* [119, 123] ont utilisé la distance Euclidienne $d_{Eu} = d_2$ ($p = 2$) pour comparer les parties de formes.

1.6.2 Distance du 'Khi-deux'

La distance euclidienne d_{Eu} augmente avec le nombre de variables, et sa valeur dépend également de l'échelle des variables. Ce problème peut être contourné en standardisant les variables. Dans ce contexte, pour pallier à ces inconvénients liés à l'utilisation de la distance euclidienne, il est possible d'utiliser la distance du χ^2 définie par :

$$d_{\chi^2}(P, \tilde{P}) = \frac{1}{N} \sum_{i=1}^N \frac{(a_i - \tilde{a}_i)^2}{(a_i + \tilde{a}_i)} \quad (1.18)$$

Cette distance est souvent utilisée pour évaluer la proximité de variables quantitatives; elle mesure l'indépendance des éléments des deux vecteurs A et \tilde{A} . Une valeur importante de cette distance signifie que les écarts entre les éléments sont dus au hasard et que la situation observée est compatible avec l'hypothèse d'indépendance. Si par contre la distance calculée est faible, on peut conclure à une liaison significative.

1.6.3 Distance de Hausdorff

La distance de Hausdorff dirigée de $P = \{p_1, \dots, p_N\}$ vers $\tilde{P} = \{\tilde{p}_1, \dots, \tilde{p}_N\}$ est définie comme la limite supérieure des distances à \tilde{P} sur tous les points de P (voir figure 1.13). Elle est donnée par l'équation suivante :

$$d_{P\tilde{P}} = \max_{p_i \in P} \{ \min_{\tilde{p}_j \in \tilde{P}} d_{Eu}(p_i, \tilde{p}_j) \}, \quad (1.19)$$

avec d_{Eu} est la distance euclidienne. Il s'agit de la plus grande distance euclidienne obtenue de n'importe quel point de P au point le plus proche correspondant de \tilde{P} . Par suite la distance de Hausdorff est définie par :

$$d_H(P, \tilde{P}) = \max\{d_{P\tilde{P}}, d_{\tilde{P}P}\} \quad (1.20)$$

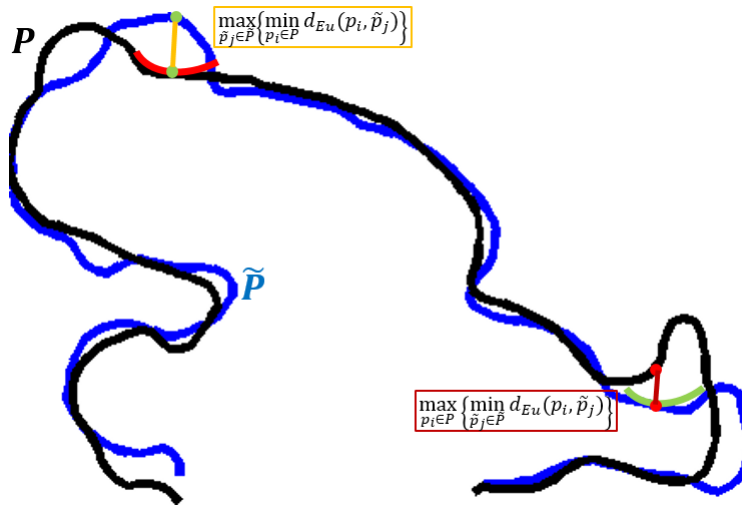


Figure 1.13 – Exemple de calcul de la distance de Hausdorff entre deux parties de formes de chameaux P et \tilde{P} .

La distance de Hausdorff d_H est sensible aux valeurs aberrantes et n'est pas robuste au bruit, car elle est définie à partir des distances euclidiennes maximales. Cette distance est clairement invariante par reparamétrisation. Cependant, elle n'est pas invariante aux transformations géométriques. Saber *et al.* l'ont utilisée pour aborder la correspondance partielle de formes en vue d'étiqueter des formes qui sont partiellement occultées [198]. Leur approche reste tout de même très sensible aux distorsions.

1.6.4 Distance de Procruste

La distance de Procruste [81] est une métrique du type moindre carrés qui exige une correspondance des points des deux parties de formes. Pour déterminer la distance de

1.7. CONCLUSION

Procuste entre deux parties de formes $P = \{p_1, \dots, p_N\}$ et $\tilde{P} = \{\tilde{p}_1, \dots, \tilde{p}_N\}$, on établit d'abord l'alignement Procuste décrit en section 1.4.3. La distance de Procuste entre P et \tilde{P} est la somme des distances carrées entre les N paires de points $p_i = (x_i, y_i)$ et $\tilde{p}_i = (\tilde{x}_i, \tilde{y}_i)$ après alignement :

$$d_{Pr}(P, \tilde{P}) = \frac{1}{N} \sum_{i=1}^N \|\tilde{p}_i - (\alpha^* \Gamma^* p_i + t^*)\|^2 = \frac{1}{N} \sum_{i=1}^N [(\tilde{x}_i - x'_i)^2 + (\tilde{y}_i - y'_i)^2] \quad (1.21)$$

Cette distance a été utilisée dans [116] pour la comparaison et la correspondance partielle des formes pour la classification et la détection des objets/formes déformées dans des images bruitées [71].

1.6.5 Distance basée sur le contexte de forme

Le principe de la distance basée sur le contexte de forme¹³ d_{SC} consiste à calculer le coût $SC_{i,j}$ attribué à la correspondance entre tout point $p_i \in P$ et tout point $\tilde{p}_j \in \tilde{P}$. Étant donnés les contextes de forme h_i et \tilde{h}_j associés aux points $p_i \in P$ et $\tilde{p}_j \in \tilde{P}$ respectivement, ce coût est donné par la distance \mathcal{X}^2 comme :

$$SC_{i,j} = SC(p_i, \tilde{p}_j) = \frac{1}{2} \sum_{k=1}^T \frac{[h_i(k) - \tilde{h}_j(k)]^2}{h_i(k) + \tilde{h}_j(k)}$$

où T est le nombre total des classes du diagramme de l'histogramme log-polaire utilisé (voir section 1.5.3). Par suite, la distance basée sur le contexte de forme d_{SC} est donnée par l'équation suivante :

$$d_{SC}(P, \tilde{P}) = \frac{1}{N} \sum_{p_i \in P} \min_{\tilde{p}_j \in \tilde{P}} SC_{i,j} + \frac{1}{N} \sum_{\tilde{p}_j \in \tilde{P}} \min_{p_i \in P} SC_{i,j} \quad (1.22)$$

Dans [19], Bai *et al.* ont utilisé cette distance basée sur le contexte de forme pour la détection et la reconnaissance des parties de formes. Cette distance se caractérise par sa robustesse aux petits désalignements des parties de formes.

1.7 Conclusion

La reconnaissance de formes 2D entières a été largement étudiée en vision par ordinateur. Ces méthodes ne peuvent pas être appliquées directement aux parties de formes. Dans ce chapitre, nous avons présenté le concept de base de la reconnaissance de parties de formes. Parce que l'intérieur d'une partie de forme n'est pas aussi significatif, la partie de forme se

¹³Shape Context (SC)

CHAPITRE 1. ÉTAT DE L'ART EN RECONNAISSANCE DE PARTIES DE FORMES

distingue essentiellement par la frontière et fait intervenir des techniques basées-contour, techniques moins compliquées et moins sensibles au bruit. Au travers de ce chapitre, nous avons décrit différentes techniques de représentation et de description de parties de formes. Nous avons également présenté les mesures de similarité les plus utilisées en reconnaissance de parties formes. Dans le chapitre suivant, nous avons présenté les méthodes de l'état de l'art les plus utilisées pour la décomposition de formes 2D en des parties de formes significatives.

Chapitre 2

État de l'art sur la décomposition des formes

Sommaire

2.1	Introduction	56
2.2	Motivations	56
2.3	Règles générales de décomposition	58
2.4	Travaux antérieurs	58
2.5	Évolution de la courbe discrète (DCE)	62
2.5.1	Description de la DCE	63
2.5.1.1	Objectifs	63
2.5.1.2	Caractéristiques	63
2.5.1.3	Mesure de relevance	64
2.5.1.4	Critère d'arrêt de l'évolution	65
2.5.2	Décomposition par DCE	70
2.5.2.1	Étape d'évolution optimale	70
2.5.2.2	Méthode des arcs convexes/concaves	72
2.5.2.3	Méthode des segments aux sommets	74
2.6	Conclusion	75

2.1 Introduction

La décomposition de formes $2D$ et $3D$ est considérée comme un problème crucial et une étape fondamentale vers une analyse et une interprétation approfondie des formes des objets [29, 89, 120, 148, 216]. En vision par ordinateur et en infographie, on trouve de nombreuses applications qui utilisent une méthode de décomposition efficace et fiable dans le prétraitement des formes [267]. Une telle méthode de décomposition est utilisée pour :

- la classification de formes [149, 191, 192],
- la recherche de formes [215, 269],
- la correspondance de formes [64, 76, 151, 210],
- l'extraction de squelette [13, 97, 136, 230],
- la simplification de formes [52],
- la détection de collision [132],
- la planification de mouvement [99, 134, 137], et
- la cartographie de texture [202].

Ce chapitre est consacré à la présentation générale du problème de la décomposition de formes $2D$ et ses solutions. En particulier, nous étudions le problème de décomposer le contour externe d'une forme donnée en des parties visuelles, robustes au mouvement articulé et aux occlusions partielles.

Pour comprendre les motivations de la décomposition de formes présentées en section 2.2, nous commençons à introduire en section 2.3 les règles de la perception humaine adoptées pour décomposer une forme en vision par ordinateur. Les méthodes les plus utilisées pour la décomposition de formes $2D$ en des parties de formes significatives sont décrites dans la section 2.4. En section 2.5, nous décrivons la méthode de l'évolution de la courbe discrète (DCE) utilisée dans la littérature pour la décomposition de formes $2D$.

2.2 Motivations

Généralement, le système visuel humain n'est pas capable de voir et d'analyser une forme entière en une seule fois. Tout comme l'analyse syntaxique qui décompose une phrase en ses constituants (nom, verbe, adverbe, etc.) de structure, le système visuel humain interprète et représente une forme par une hiérarchie de parties perceptives [89, 216]. De

2.2. MOTIVATIONS

fait, différentes approches en vision par ordinateur [120, 196, 212] ont proposé d'analyser les formes après décomposition en se basant sur leurs structures locales (c'est-à-dire sur leurs parties de formes).

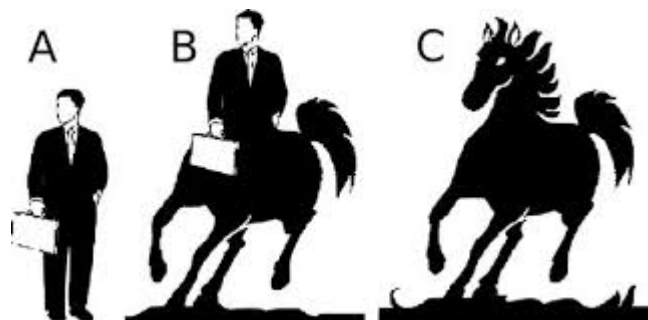


Figure 2.1 – A, B et C se ressemblent respectivement à un homme, un centaure et un cheval. Du point de vue global, les trois formes sont différentes, mais les paires de formes (A,B) et (B,C) ont localement des parties très similaires [242].

En réalité, il est souvent difficile d'établir une correspondance significative entre les formes car cette étude nécessite une compréhension de la structure globale et locale des formes. En comparant des formes naturelles ou artificielles, telles qu'elles existent, il est rarement possible de trouver une correspondance optimale entre elles en absence d'une étude approfondie de la fonctionnalité des parties de formes. Par exemple, il est impossible d'arriver à calculer la similarité globale entre un centaure, un homme et un cheval (figure 2.1). Dans cette situation de similarité, il est préférable de négliger les parties différentes et de se concentrer à comparer seulement les parties qui se ressemblent dans les deux formes. Ceci justifie pleinement l'intérêt de chercher une technique convenable pour la décomposition de formes. De plus, plusieurs facteurs comme la distorsion, le bruit, les erreurs de segmentation, le chevauchement, l'occultation, le mouvement et l'articulation démontrent l'intérêt d'avoir un processus de décomposition conduisant à une représentation partielle d'une forme donnée qui la rend plus simple à manipuler. Par exemple, en présence d'une déformation locale provoquée par une occultation ou un mouvement partiel de la forme, la reconnaissance peut avoir lieu en utilisant les parties de la région stable non affectée.

Par conséquent, l'identification des composants d'une forme est essentielle pour la reconnaissance partielle et globale de forme, pour la correspondance partielle, et pour d'autres tâches en vision par ordinateur [267]. En fait, la décomposition de forme en des parties significatives aide à effectuer la reconnaissance de forme à partir des parties au lieu de l'effectuer à partir de la forme globale. Une telle reconnaissance basée sur les parties de formes est voulue d'être robuste à l'occlusion et au mouvement des régions.

2.3 Règles générales de décomposition

La décomposition d'une forme (2D ou 3D) donnée est l'opération qui consiste à représenter cette forme en des entités plus élémentaires, les parties de forme. Cette représentation doit être compatible avec les règles générales de la perception visuelle humaine qui utilise souvent une représentation basée sur les parties de forme pour accomplir les tâches de reconnaissance de forme. Les trois grandes règles de la perception humaine sont généralement adoptées pour décomposer une forme en vision par ordinateur. Ces règles sont :

- la règle des minima [89],
- la règle du raccourci [217], et
- la règle de la convexité [120, 244].

La règle des minima suggère qu'une coupe est habituellement localisée aux endroits où la courbure est un minimum local. La règle du raccourci tend à minimiser la longueur totale de la coupe, et la règle de la convexité contraint les parties d'être convexes.

D'après ces règles, on peut poser le problème de la décomposition d'une forme par la génération d'un ensemble des parties significatives couvrant la forme entière. Chaque coupe est déterminée par une paire d'extrémités se trouvant sur le contour extérieur de la forme et à des positions critiques. Notant que de nombreuses études psychologiques ont montré le rôle important que jouent les parties dans la perception et la reconnaissance d'objets [28, 72, 89, 214].

2.4 Travaux antérieurs

Comme nous l'avons noté précédemment, le système visuel humain utilise une représentation à base de parties de formes pour la reconnaissance de formes. Pour résoudre le problème de décomposition de formes en vision par ordinateur, de nombreuses approches anciennes et récentes s'inspirent alors des théories de la perception humaine [120, 189, 190, 196, 212]. Les parties de formes significatives sont très utiles dans de nombreuses tâches de vision par ordinateur comme la détection d'objet, parce que ces parties sont généralement plus stables dans différents environnements. Plusieurs méthodes ont été proposées pour décomposer une forme en des parties significatives en appliquant les règles de la perception humaine citées en section 2.3 [78, 94, 140, 189]. Les règles de perception sont généralement génériques et faciles à calculer. Les méthodes de décomposition approximativement convexe de la forme polygonale ont utilisé la règle de convexité afin d'obtenir des parties significatives [78, 135]. Liu *et al.* [140] ont formulé la décomposition convexe de formes comme un problème de programmation linéaire tout en estimant la règle du raccourci.

2.4. TRAVAUX ANTÉRIEURS

Ren *et al.* [189] et Jiang *et al.* [94] ont proposé des méthodes de décomposition qui utilisent le concept de la règle des minima dans leurs travaux. Ren *et al.* [190] ont proposé une décomposition optimale quasi-convexe qui bénéficie des caractéristiques des trois règles de perception pour diviser arbitrairement des formes en un nombre minimum de parties quasi-convexes. Quand la plupart des travaux basés sur la règle de convexité tendent à générer des parties redondantes afin de satisfaire la convexité demandée, Ma *et al.* [146] ont proposé une nouvelle contrainte quadratique exécutée par une programmation non linéaire afin d'obtenir une décomposition cohérente aux parties similaires.

L'étude de Cohen *et al.* [50] sur la perception humaine a montré que les propriétés de la courbure du contour et les propriétés de la zone de région fermée affectent l'identification des parties de formes. Similairement aux travaux de Siddiqi et Kimia [212] qui montrent que les minima de courbures négatives de contours déterminent les fragments significatifs de contours, l'identification des parties peut être déterminée par la longueur du contour ou par la fonction angulaire représentative des fragments de contours. Ainsi, nous pouvons dire que la majorité des méthodes de décomposition de forme existantes peuvent être classées en des méthodes basées-contour et des méthodes basées-région :

- Les méthodes basées-contour utilisent seulement les informations du contour, la frontière limitant l'objet dans l'image, pour extraire des parties représentatives de forme.
- Les méthodes basées-région utilisent les informations codées par l'ensemble de tous les pixels de l'objet pour le décomposer. Seul l'intérieur de la forme est pris en compte pour obtenir des parties de forme significatives.

La méthode la plus connue dans la première catégorie est vraisemblablement celle proposée par Hoffman et Richards [89]. Dans cette méthode, ils affirment que le système visuel humain divise un contour en des parties délimitées par des points minima de courbures négatives car, selon eux, on ne peut jamais voir la forme entière en un seul coup d'oeil (figure 2.2). Après avoir filtré le contour à l'aide des B-Splines, Saber *et al.* [198] ont défini la partie de forme par le segment limité par deux points de courbure maximale locale non nécessairement consécutifs (figure 2.3). Cependant, Sun et Super [223] ont décomposé le contour en se basant sur les points critiques. Les segments du contour sont définis entre toute paire de points non consécutifs de courbure extrême (minimale ou maximale). Une fois le contour est lissé par CSS¹, par convolution avec une gaussienne, les points critiques résultants sont filtrés par une technique de comparaison de la surface du triangle formé par trois segments consécutifs avec un seuil donné (figure 2.4). Dans [27], les points de courbures nulles issues d'un lissage du

¹Curvature Scale Space (Courbure multi-échelle) est une méthode de représentation de formes basée sur la recherche de points d'inflexion sur le contour en plusieurs échelles.

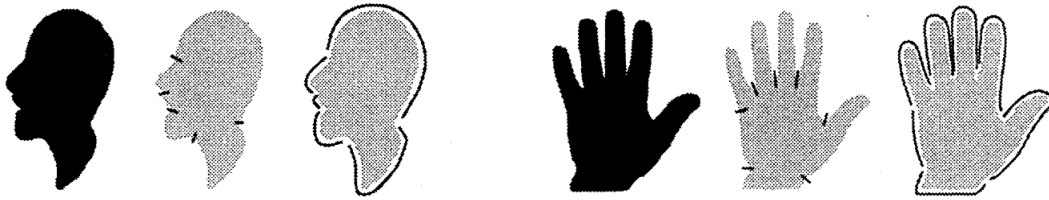


Figure 2.2 – Décomposition des contours de formes en des segments délimités par les points minima de courbures négatives consécutives (théorie de Hoffman et Richards [89]). Dans chaque triplet, la forme originale est en noir (à gauche), les points minima de courbures négatives sont marqués par des barres obliques (milieu) et les segments de contour sont représentés par des courbes en gras (à droite) [89].

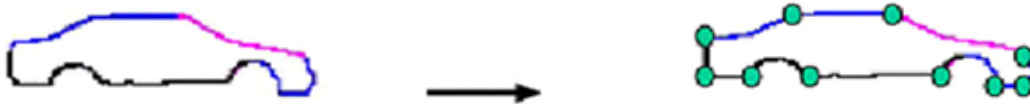


Figure 2.3 – Extraction des points maxima locaux de courbure positive (représentés par des cercles verts au contour à droite) d'un contour filtré à l'aide des B-Splines (à gauche) [198].

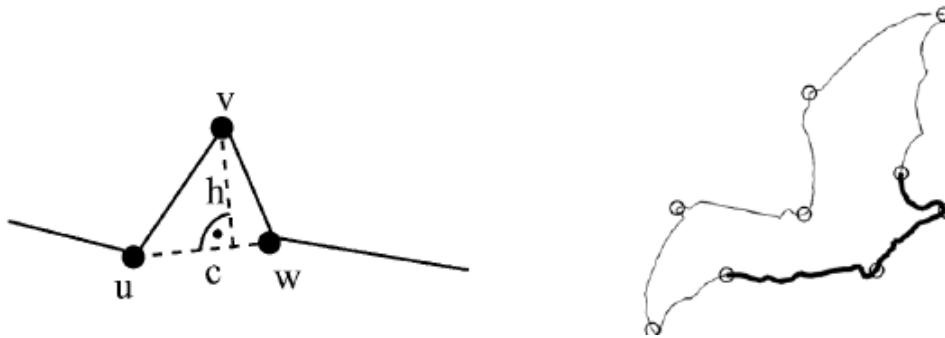


Figure 2.4 – À gauche, le triangle formé par trois points critiques consécutifs d'un contour. À droite, un exemple d'un segment complexe de la forme d'une chauve-souris déterminé par une paire de points critiques de la forme [223].

2.4. TRAVAUX ANTÉRIEURS

contour par une gaussienne sont utilisés pour partitionner ce contour.

Par ailleurs, Latecki et Lakämper [120] ont proposé une méthode de décomposition hiérarchique basée sur une évolution de la courbe discrète (DCE²) représentative du contour. Grâce à l'harmonie de cette méthode de simplification avec la perception humaine et la stabilité des points DCE en présence d'éventuelles déformations importantes du contour, plusieurs travaux [14, 19, 119, 123, 247] ont réutilisé le même processus de simplification dans le but de préserver les parties visuelles essentielles. Par exemple, la simplification DCE d'un contour entier d'éléphant (figure 2.5a) donne un polygone représenté avec 13 sommets (voir figure 2.5b). Pas loin du concept de décomposition DCE, l'approximation polygonale de la courbe a été adoptée dans plusieurs approches selon différentes techniques [65, 101, 111, 200] pour obtenir une décomposition significative d'une courbe donnée.



Figure 2.5 – Illustration de la simplification DCE. (a) : un contour entier d'éléphant; (b): le polygone qui représente le contour en (a) simplifié par DCE; et (c) : les sommets communs de (a) et (b) sont colorés en bleu.

Des études psychophysiques [62] ont montré que les parties dépendent radicalement à la fois du contour et de la géométrie de la région. Siddiqi et Kimia [212] ont proposé de combiner les points du contour de courbures minimales avec l'information de la région, car l'utilisation unique de l'information du contour semble insuffisante. Leur technique de décomposition implique deux types de coupure : des membres et des cous (figure 2.6). Formellement, un membre est limité par une paire de minima de courbures et un cou est localisé à un diamètre minimum d'un cercle inscrit au contour. Les membres et les cous sont également associés aux singularités, chocs, issues du processus d'évolution de la courbe par une équation de réaction et de diffusion, équation introduite par Kimia *et al.* [108]. En outre, Singh *et al.* [217] soutiennent une règle de raccourci qui implique le contour et la région. La coupure qu'ils proposent est formée par un axe de symétrie qui passe par un point minimum local de courbure (figure 2.6). D'autres études

²Discrete Curve Evolution (Évolution de la courbe discrète) est une méthode de simplification de formes basée sur une suppression récursive des points ayant la plus petite contribution à la forme (voir détail en section 2.5 suivante).

suggèrent de combiner avec les informations de contour, les informations issues des symétries locales [156] ou du squelette [259] afin de simuler la perception humaine pour la détermination des parties de formes significatives.

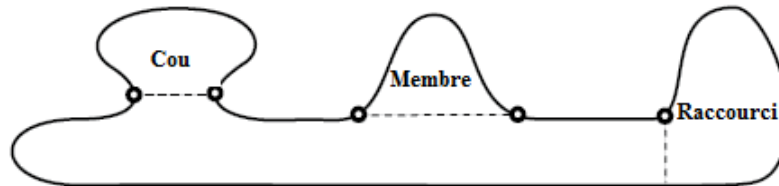


Figure 2.6 – Différents types de coupure : cou, membre, et raccourci [217].

Il existe aussi d'autres approches qui utilisent uniquement les informations de la région pour la décomposition. En 1967, Blum [31] introduit l'axe médian, et une décomposition à ses points de connexions a été suggérée. L'axe médian se caractérise par sa capacité à capturer les indices visuels de la forme, comme la symétrie et la complexité (reflétées par les points de branchements). Sous une forme différente de l'axe médian, plusieurs autres travaux ont utilisé le squelette [5, 35, 108, 185, 193, 197, 230, 232] qui décrit les symétries locales des formes pour le calcul des parties de formes. Par exemple, la méthode de décomposition présentée dans [232] utilise le squelette simplifié de la forme pour fragmenter son contour externe en utilisant les points de ramifications du squelette. Ces points ont une importance particulière car ils peuvent capter les connexions entre les différentes parties de la forme. Ainsi, les noeuds du squelette de la forme sont à la base de la décomposition comme illustré en figure 2.7.

Malgré toutes les méthodes de décomposition citées ci-dessus, d'autres approches plus récentes se sont développées pour introduire des techniques nouvelles pour la décomposition des formes 2D [47, 245, 251].

2.5 Évolution de la courbe discrète (DCE)

Le fait d'identifier les points représentatifs du contour de la forme d'origine revient en quelque sorte à identifier les points de bruit. Le processus de l'évolution de la courbe discrète (DCE) est prouvé théoriquement et expérimentalement comme une méthode efficace pour l'élimination des points de bruit [120–122]. Ce processus élimine ces points par une suppression récursive des sommets du polygone avec la plus petite contribution à la forme, autrement dit, qui sont les plus susceptibles d'entraîner du bruit. À la suite de cette évolution, on obtient un sous-ensemble de sommets qui représentent au mieux le contour d'une forme donnée. Ce sous-ensemble peut être exploité pour la décomposition du contour original en des segments de contour définis par les sommets du polygone simplifié.

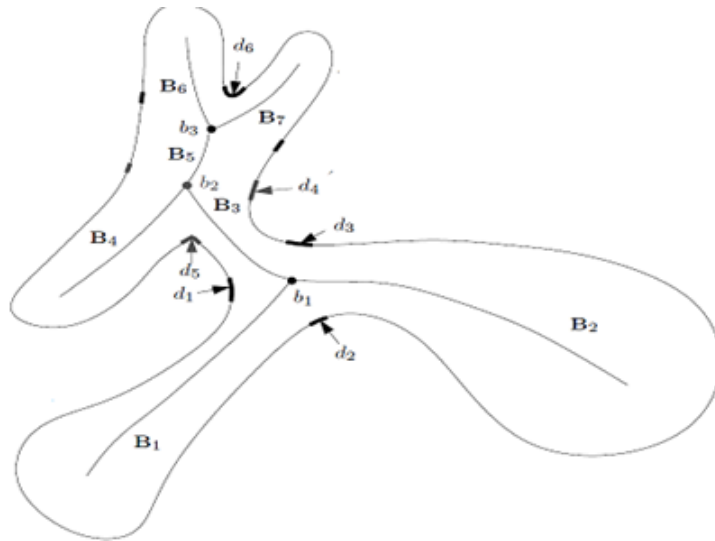


Figure 2.7 – Les deux points d_1 et d_3 du contour correspondant au noeud b_1 induisent une décomposition de la forme illustrée en deux parties ($d_1:d_3$) et ($d_3:d_1$). La décomposition correspondant au squelette se compose de deux arbres : une engendrée par B_1 , b_1 , et B_2 et une autre engendrée par b_1 , b_2 , B_4 , b_3 , B_6 et B_7 [232].

2.5.1 Description de la DCE

Pour obtenir un ensemble de fragments significatifs de contour d'une forme donnée, appliquer l'évolution de la courbe discrète au contour original permet d'acquérir un polygone simplifié, considéré comme une approximation du contour extérieur de l'objet.

2.5.1.1 Objectifs

La motivation principale de l'utilisation du processus de l'évolution de la courbe discrète (DCE) est le fait que le contour d'un objet est souvent déformé par le bruit issu de la numérisation et des erreurs de segmentation. Afin de contrer ces irrégularités, il est souhaitable d'utiliser la DCE pour raisons suivantes :

- réduire l'influence du bruit (figure 2.8), et
- simplifier la représentation de la forme en négligeant les déformations et en éliminant les éléments non significatifs du contour tout en préservant les caractéristiques pertinentes de la forme (figure 2.8).

2.5.1.2 Caractéristiques

Les caractéristiques du processus de l'évolution DCE sont :

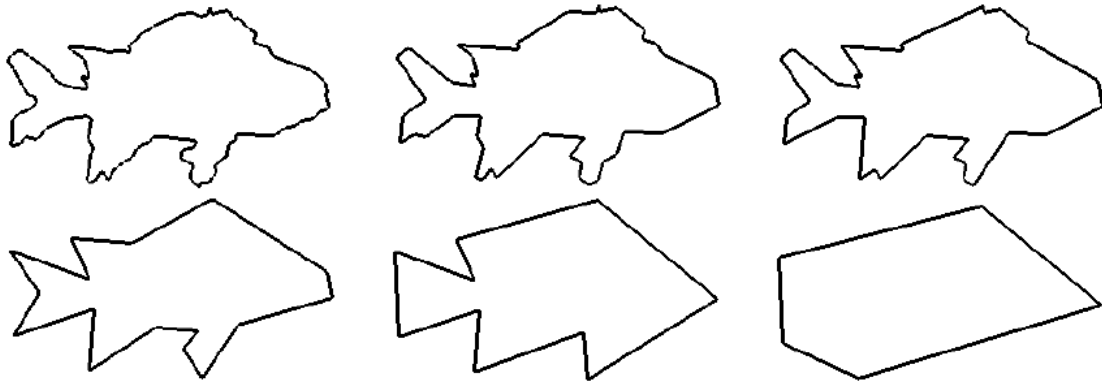


Figure 2.8 – Quelques étapes de l'évolution d'un contour de poisson par DCE. Le premier contour est une version bruitée du contour sur www-site [1].

- La sensibilité au contexte, car les composantes pertinentes (c'est-à-dire les sommets les plus relevants) et celles qui sont sans importance ne peuvent pas être décidées sans contexte. Autrement dit, l'évolution est guidée par une mesure de pertinence qui n'est pas une propriété locale à toute étape d'évolution.
- La stabilité des positions des points restant lors du processus de l'évolution (figure 2.8).
- L'invariance à la translation, rotation, réflexion, et au changement d'échelle.

2.5.1.3 Mesure de relevance

Sans perte d'information toute courbe numérique peut être considérée comme un polygone, avec éventuellement un grand nombre de sommets. On peut donc étudier les évolutions des formes polygonales. L'idée clé de l'évolution proposée est très simple. À chaque étape d'évolution, une paire de segments consécutifs s_1 et s_2 est remplacée par un segment droit unique reliant les extrémités de $s_1 \cup s_2$. Par exemple, les deux segments consécutifs, marqués en noir dans la figure 2.9a, du contour de Cerf simplifié à l'étape d'évolution e est remplacé par un seul segment, marqué en noir dans la figure 2.9b, à l'étape d'évolution $e+1$. En fait, ce phénomène de remplacement est le résultat d'une élimination du point jaune (figure 2.9a) lors du passage de l'étape d'évolution e à l'étape d'évolution $e+1$. Le point éliminé correspond au point qui a la plus faible contribution avec ses segments adjacents à la forme simplifiée à l'étape d'évolution e . Par conséquent, à chaque étape d'évolution la forme converge davantage vers un polygone convexe qui définit le plus haut niveau dans la hiérarchie de la forme. Par exemple, la dernière courbe polygonale, marquée en rouge dans la figure 2.10, représente le polygone convexe à l'étape d'évolution maximale de la courbe originale en noir.

2.5. ÉVOLUTION DE LA COURBE DISCRÈTE (DCE)

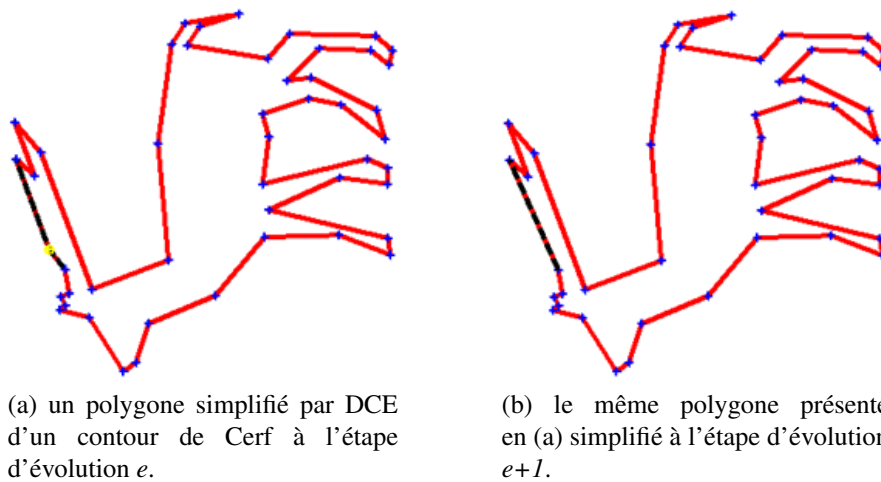


Figure 2.9 – Remplacement de deux segments consécutifs (marqués en noir en (a)) par un seul segment (b) lors du passage de l'étape d'évolution e à l'étape d'évolution $e+1$.

L'évolution par DCE se caractérise par la précision de l'ordre de substitution, car l'obtention d'une évolution intuitive de la forme nécessite la définition d'un ordre correct de substitution. La substitution par DCE se fait en se référant à une mesure de pertinence K donnée par :

$$K(s_1, s_2) = \frac{\beta(s_1, s_2) \cdot l(s_1) \cdot l(s_2)}{l(s_1) + l(s_2)},$$

où $\beta(s_1, s_2)$ est l'angle au sommet commun des segments s_1 et s_2 ; cet angle est considéré entre 0 et 2π . l est la fonction de la longueur normalisée par rapport à la longueur totale d'une courbe polygonale. La propriété principale de cette mesure de pertinence est que plus la valeur de $K(s_1, s_2)$ est élevée plus la contribution du sommet commun des arcs s_1 et s_2 à la forme est grande.



Figure 2.10 – Quelques étapes d'évolution d'un contour de Cerf de la base d'images MPEG-7.

2.5.1.4 Critère d'arrêt de l'évolution

L'arrêt de l'évolution par DCE peut être contrôlé par un des critères d'arrêt suivants :

- Le nombre de sommets imposé, $v > 3$, à garder à la fin de l'évolution (figure 2.11).
- La valeur maximale de relevance, K_{max} , avec laquelle l'évolution continue jusqu'à n'avoir que des points de relevance K supérieure à K_{max} (tableau 2.1).
- L'évolution jusqu'au polygone convexe qui consiste à évoluer jusqu'à avoir la première forme simplifiée convexe (figure 2.10).

Il est à noter qu'on peut imposer en plus à l'évolution de garder les points extrémités de la courbe originale au cours de l'évolution.

Avec l'utilisation de chacun des critères d'arrêt précédemment cités, l'évolution se comporte différemment. À cause des variations inter-individuelles, il est presque impossible de représenter les polygones simplifiés de différentes formes avec un nombre fixe de sommets v tout en gardant les caractéristiques de la forme. Par exemple, seize points sont suffisants pour représenter la forme d'une *Tasse* (1^{ère} ligne de la figure 2.11) et la forme d'un *Papillon* (3^{ème} ligne de la figure 2.11) sans quasiment aucune perte d'information. Cependant, la forme d'un autre *Papillon* (2^{ème} ligne de la figure 2.11) nécessite plus de sommets ($v = 22$) pour la représenter en son étape d'évolution optimale (figure 2.12). De même, la complexité de la forme d'un *Éléphant* (4^{ème} ligne de la figure 2.11) nécessite 25 sommets ($v = 25$) pour atteindre son évolution optimale (figure 2.13 suivante).

Avec une valeur maximale de relevance K_{max} , l'évolution s'achève en éliminant tous les points dont la valeur de pertinence K est inférieure à K_{max} . Plus la valeur de K_{max} est élevée plus l'évolution est avancée. Pour illustrer cette idée, nous considérons trois formes différentes (*Virgule-12*, *Jarre-15*, et *Octopus-18*) de la base d'images MPEG-7 (1^{ère} ligne du tableau 2.1). Une forme simple comme celle de *Virgule-12* nécessite un nombre faible de points (v) pour représenter son polygone simplifié. La courbe originale de cette forme atteint son évolution optimale avec une valeur K_{max} relativement élevée ($K_{max} = 0.5$) et seulement 9 sommets ($v = 9$) représentatifs. Tandis que l'évolution optimale des formes plus complexes de *Jarre* et d'*Octopus* est atteinte avec une valeur K_{max} comprise entre 0,2 et 0,3. Par conséquent, il est évident que plus la forme est complexe (c'est-à-dire avec plus de détails) plus elle nécessite de points pour la représenter en son étape d'évolution optimale et plus la valeur de K_{max} utilisée pour la simplifier doit être faible.

L'évolution jusqu'au polygone convexe exige que la simplification s'arrête avec la première courbe convexe obtenue. Par définition, une courbe discrète est dite convexe si elle conserve la même direction de courbure en tous les points. Le nombre minimal de sommets d'une courbe convexe est égal à 3. Par exemple, la première courbe convexe (à gauche) de la figure 2.14 est formée de trois points DCE, la deuxième (au milieu) est constituée de quatre sommets, tandis que la troisième courbe convexe (à droite) comprend six points DCE. Par conséquent, un polygone convexe résulte d'une évolution maximale qui ne signifie pas forcément une évolution optimale.

2.5. ÉVOLUTION DE LA COURBE DISCRÈTE (DCE)

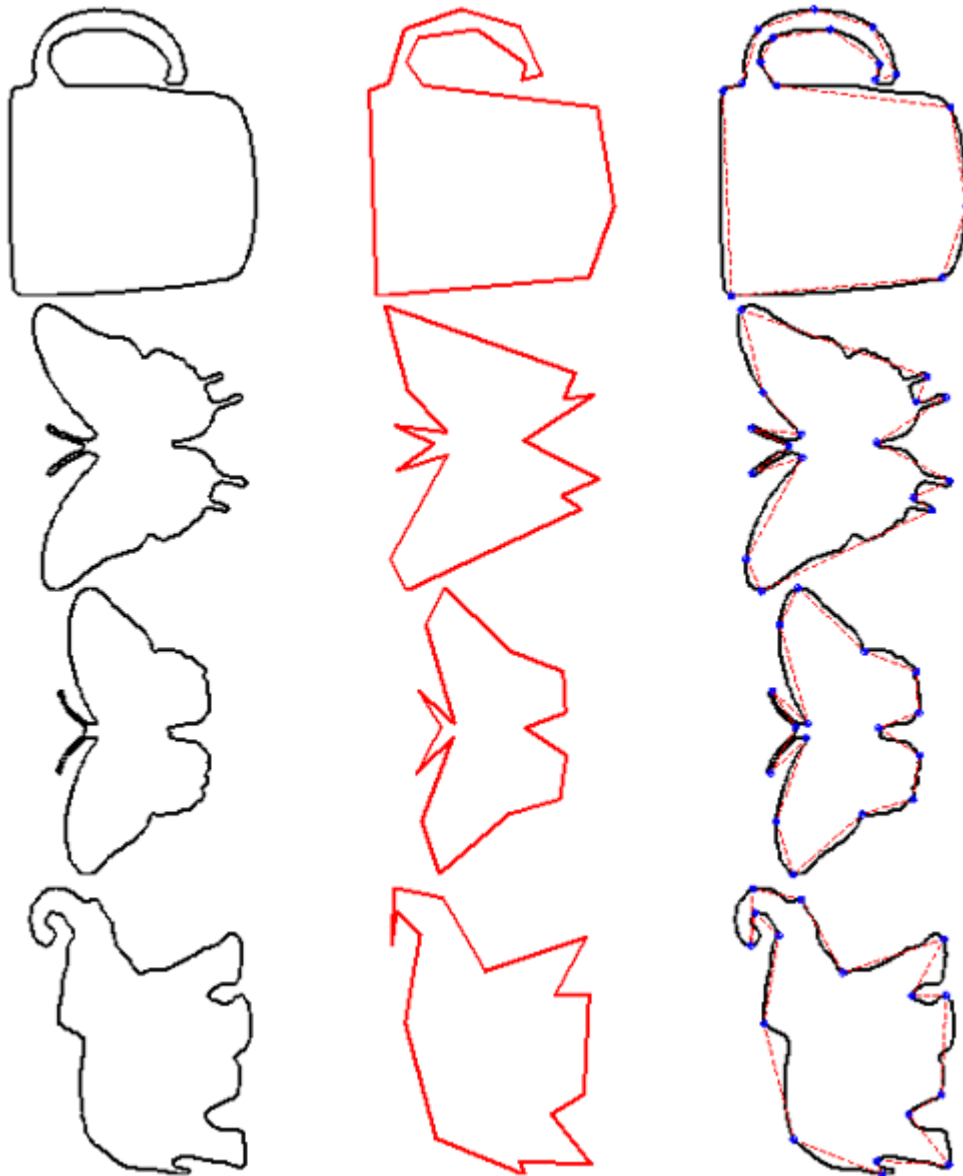


Figure 2.11 – La première colonne représente 4 formes de la base d’images MPEG-7. La deuxième colonne représente leurs formes finales simplifiées avec 16 sommets ($v = 16$). La troisième colonne représente la projection de la courbe simplifiée (2^{ème} colonne) sur la courbe originale (1^{ère} colonne) de chaque forme, où les sommets DCE sont marqués en bleu.

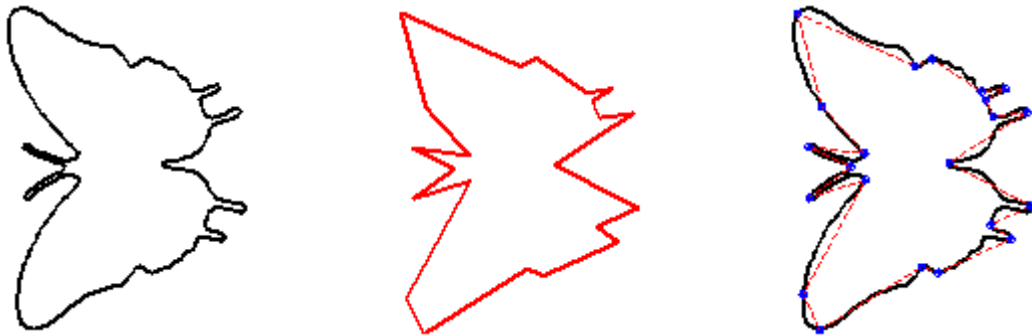


Figure 2.12 – L'évolution optimale de la même forme du *Papillon* de la figure 2.11 atteinte avec 22 sommets (polygone rouge à 22 sommets).

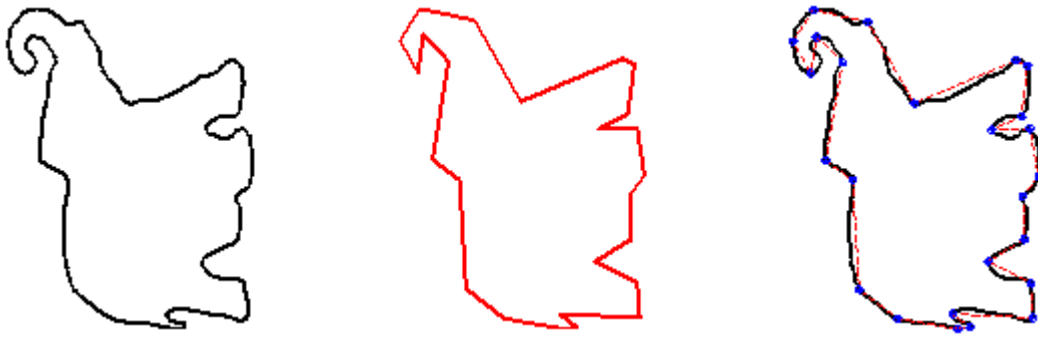


Figure 2.13 – L'évolution optimale de la même forme de l'*Éléphant* de la figure 2.11 atteinte avec 25 sommets (polygone rouge à 25 sommets).

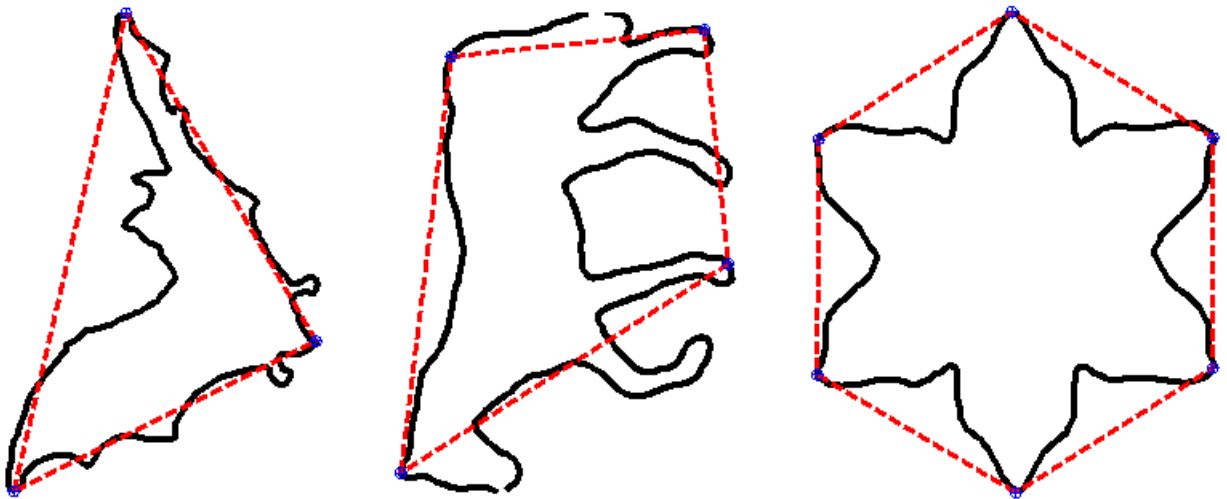
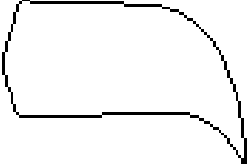
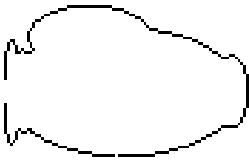
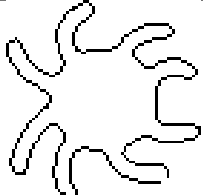
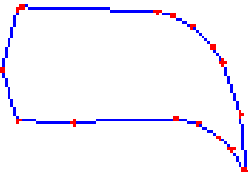
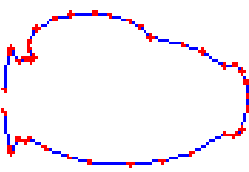
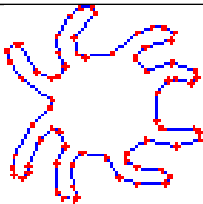
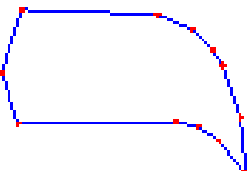
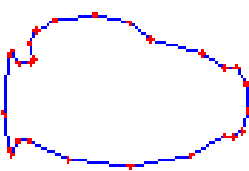
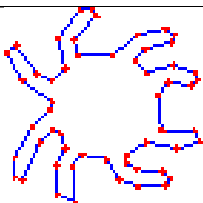
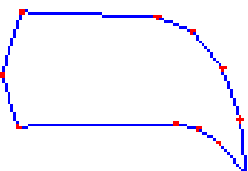
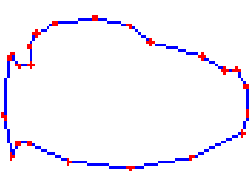
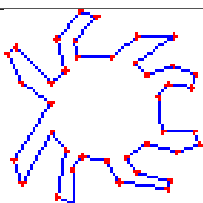
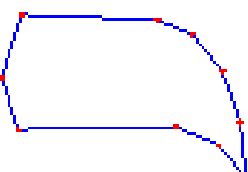
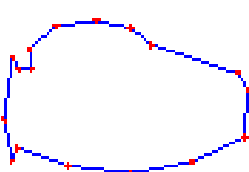
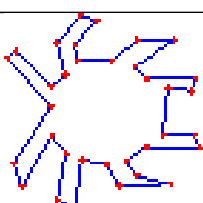
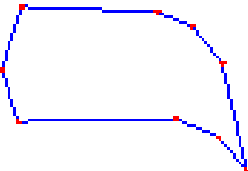
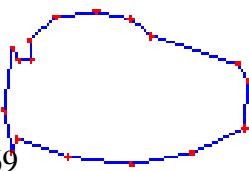
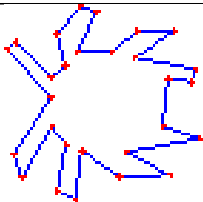


Figure 2.14 – Les trois polygones en rouge représentent les courbes convexes relatives aux trois formes différentes, marquées en noir, de la base d'images MPEG-7.

2.5. ÉVOLUTION DE LA COURBE DISCRÈTE (DCE)

Tableau 2.1 – Évolution de trois formes différentes : *Comma-12*, *Jarre-15*, et *Octopus-18* de la base MPEG-7 avec l'utilisation de cinq différentes valeurs de K_{max} , la valeur maximale de relevance imposée à tous les points restants.

K_{max}	 <i>Virgule-12</i>	 <i>Jarre-15</i>	 <i>Octopus-18</i>
0,1	 17 sommets	 43 sommets	 72 sommets
0,2	 13 sommets	 27 sommets	 55 sommets
0,3	 11 sommets	 23 sommets	 40 sommets
0,4	 10 sommets	 17 sommets	 34 sommets
0,5	 9 sommets	 17 sommets	 28 sommets

Il est aussi intéressant de discuter le critère supplémentaire qui permet de garder les extrémités exclues de toute simplification. En effet, dans le cas d'évolution de formes complètes (courbes fermées), il est intuitivement intéressant de considérer les extrémités parmi les points à éliminer. Par exemple, la deuxième colonne de la figure 2.15 montre que la conservation des points extrémités au cours du processus de simplification DCE n'apporte rien à la forme simplifiée. Au contraire, cette contrainte ajoute un point additionnel qui ne sert pas dans la représentation de la courbe. Cependant, il est nécessaire de garder les points extrémités lors de la simplification des courbes ouvertes représentatives des parties de formes. Les extrémités d'une partie de forme constituent le point initial et le point final qui la caractérisent.

2.5.2 Décomposition par DCE

La décomposition par DCE consiste à décomposer les formes entières en des parties de formes en se basant sur les points DCE issues d'une évolution optimale des formes données. Les deux méthodes les plus utilisées sont la méthode des arcs convexes/concaves maximaux [120, 123, 124] et la méthode des segments délimités par des sommets DCE non consécutifs [14, 16, 19, 247]. Les deux méthodes supposent qu'une forme simplifiée par DCE en son étape d'évolution optimale reste assez représentative de son contour original.

2.5.2.1 Étape d'évolution optimale

Généralement, toute forme simplifiée obtenue à l'étape d'évolution optimale doit être en accord avec la perception visuelle humaine. La recherche d'une telle évolution nécessite la détermination d'un paramètre d'arrêt qui définit le nombre d'itérations de l'évolution à exécuter. Les travaux de l'état de l'art ont étudié ce paramètre d'arrêt de deux manières différentes. Récemment, certains travaux [14, 247] ont directement utilisé le paramètre de la relevance maximale K_{max} (section 2.5.1.4) pour atteindre l'évolution optimale des formes à comparer. Tandis que, d'autres travaux [120, 123] ont cherché l'étape d'évolution optimale par un **retour en arrière** suivant le sens opposé du chemin de l'évolution (figure 2.16) tout en calculant une erreur quadratique partielle.

Recherche par retour en arrière : Pour un contour original P^0 d'une forme donnée, l'ensemble des polygones ordonné P^0, \dots, P^m représente les polygones simplifiés du P^0 à toutes les étapes d'évolution jusqu'à l'évolution maximale atteinte par le premier polygone convexe P^m . Soit P^e le polygone simplifié de P^0 obtenu à l'étape d'évolution e . La figure 2.17 décrit le schéma de la recherche de l'étape d'évolution optimale par retour en arrière.

La recherche de l'évolution optimale de P^0 suit les étapes suivantes :

- Étape 1 : Calculer l'erreur globale entre le polygone convexe P^m et le contour original P^0 à l'aide de la mesure de dissimilarité di_{ss} donnée par l'équation

2.5. ÉVOLUTION DE LA COURBE DISCRÈTE (DCE)

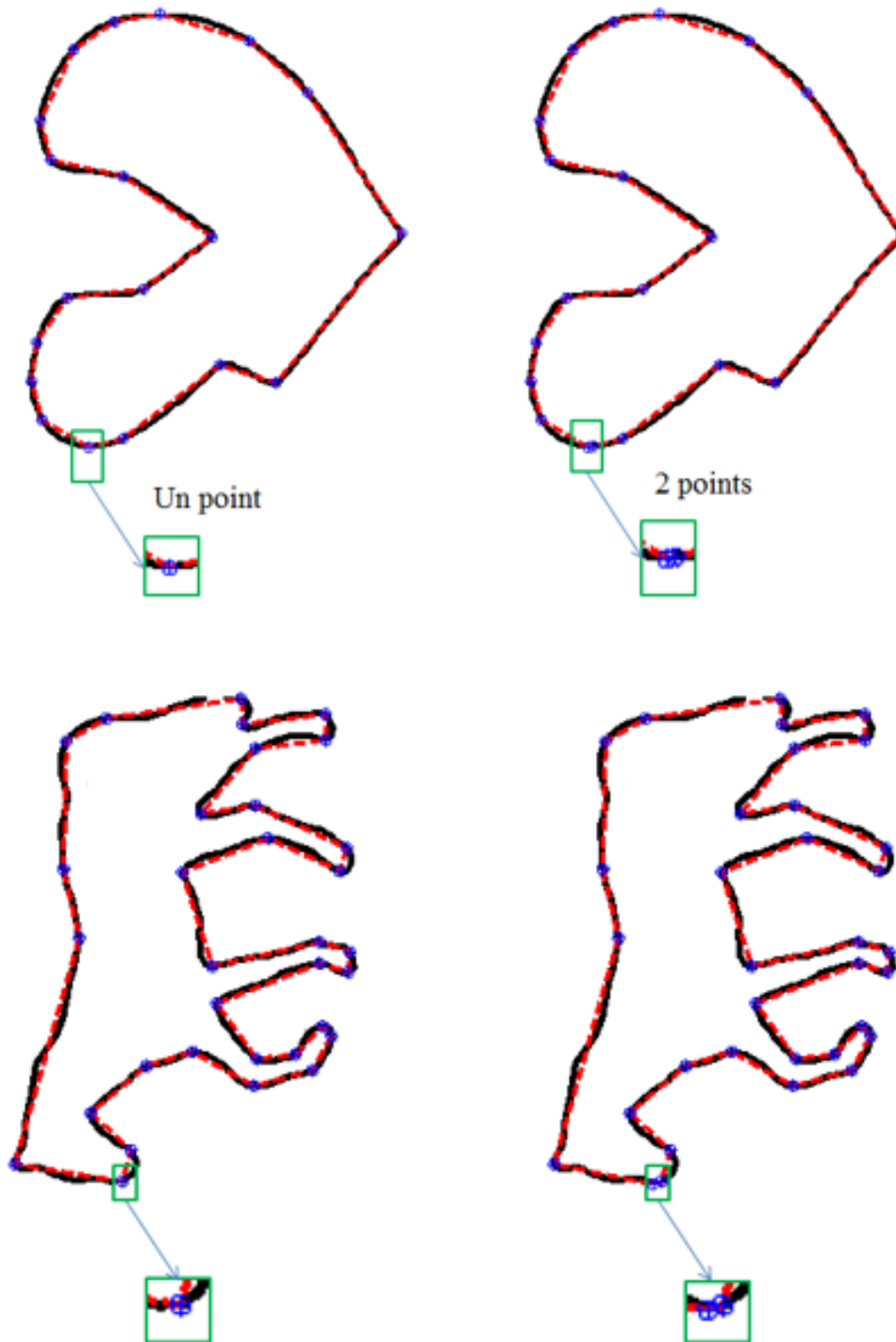


Figure 2.15 – En première colonne, deux formes simplifiées par DCE sans rétention des points extrémités. En deuxième colonne, les mêmes formes de la première colonne sont simplifiées tout en gardant les extrémités.

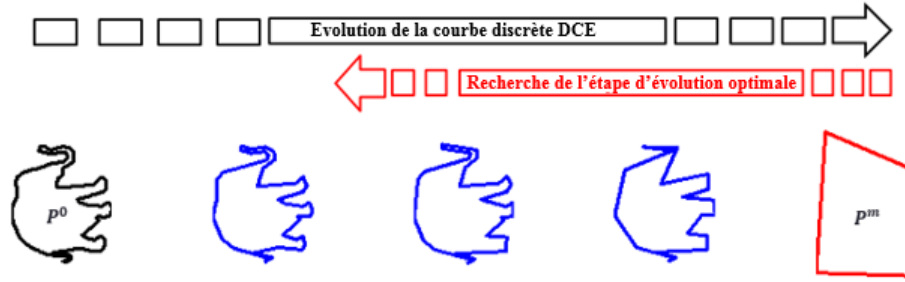


Figure 2.16 – Le sens de la recherche de l'étape d'évolution optimale est opposé au sens de l'évolution.

suivante :

$$diss(P^m, P^0) = \int_0^1 (\theta_m(s) - \theta_0(s))^2 ds, \quad (2.1)$$

où θ_m et θ_0 sont les fonctions d'angle, paramétrées par s l'abscisse curviligne normalisée à l'unité, représentatives de P^m et de P^0 respectivement.

- Étape 2 : Comparer l'erreur globale $diss(P^m, P^0)$ à un seuil fixe S expérimentalement trouvé. Si l'erreur globale $diss(P^m, P^0)$ ne dépasse pas S , l'évolution optimale est atteinte. Si non, on passe alors à examiner l'étape d'évolution précédente $e = m - 1$ en passant à l'étape 3 suivante.
- Étape 3 : Décomposer le polygone simplifié P^e en des parties (polygones).
- Étape 4 : Trouver l'erreur partielle entre toute partie p_e du polygone P^e et sa partie correspondante p_0 de P^0 en calculant la dissimilarité $diss(p_e, p_0)$ donnée par l'équation 2.1. Les parties p_e et leurs parties correspondantes p_0 sont représentées respectivement par leurs fonctions tangentes θ_e et θ_0 paramétrées par s l'abscisse curviligne normalisée à l'unité.
- Étape 5 : Comparer l'erreur partielle $diss(p_e, p_0)$ entre chaque partie p_e et sa partie correspondante p_0 au même seuil S . Si l'une au moins des erreurs partielles dépasse le seuil S , il s'agit de répéter les étapes 3, 4 et 5 pour le polygone simplifié P^e avec $e = e - 1$. L'évolution optimale est atteinte lorsque toutes les erreurs partielles correspondantes sont inférieures à S .

2.5.2.2 Méthode des arcs convexes/concaves

Suite à la projection du polygone simplifié sur le contour original, deux différentes manières ont été utilisées pour extraire les parties visuelles du contour d'origine correspondant. Latecki et Lakämper [120, 122–124] proposent des coupures partielles

2.5. ÉVOLUTION DE LA COURBE DISCRÈTE (DCE)

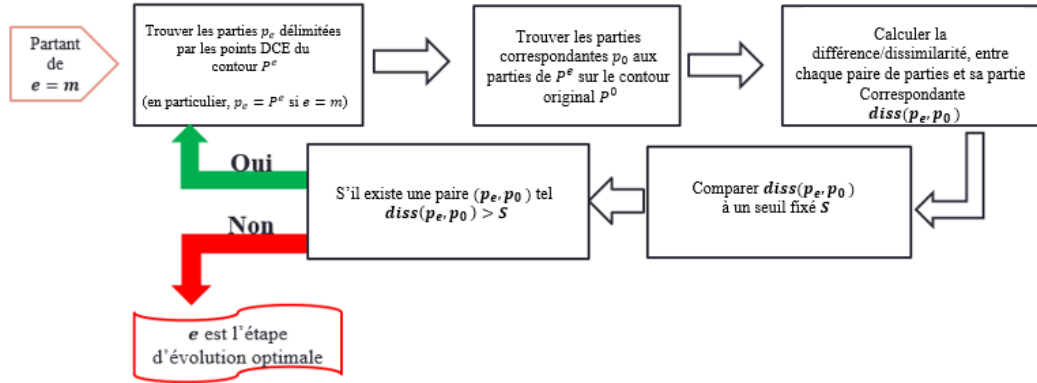


Figure 2.17 – Schéma descriptif de la recherche de l'étape de l'évolution optimale d'une courbe originale P^0 donnée.

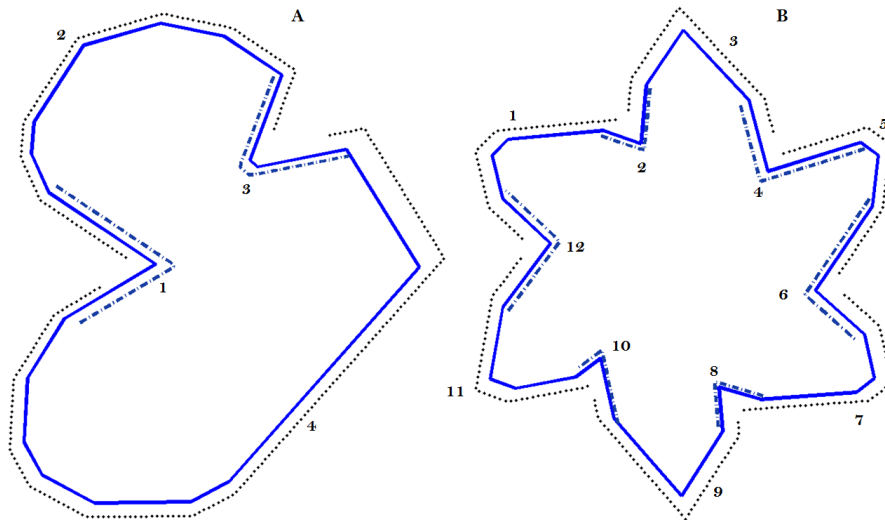


Figure 2.18 – Arcs convexes/concaves maximaux de deux formes de la base d'images MPEG-7. Les arcs internes pointillés indiquent les arcs concaves tandis que les arcs externes pointillés se rapportent aux arcs convexes.

pour décomposer une forme donnée. Les coupures sont définies par les parties de forme limitées par des arcs convexes/concaves maximaux. Par exemple, la figure 2.18 montre quelques formes avec les arcs convexes/concaves maximaux correspondants. Chaque arc joignant deux minima de courbure négative consécutifs est appelé un arc convexe maximal (arcs 2 et 4 de la forme A et arcs 1, 3, 5, 7, 9 et 11 de la forme B). L'arc limité par deux maxima de courbure positive consécutifs est appelé un arc concave maximal (arcs 1 et 3 dans la forme A et arcs 2, 4, 6, 8, 10 et 12 de la forme B). Les parties de forme sont simplement obtenues en joignant les extrémités des arcs convexes/concaves maximaux sur la forme originale. Pour plus de détails sur la règle de la convexité hiérarchique, le lecteur peut se référer à [120].

2.5.2.3 Méthode des segments aux sommets

D'autres travaux [14, 16–19, 247] ont proposé une définition plus large des parties de forme qui n'est pas limitée par la notion des arcs convexes/concaves maximaux. Dans cette approche, une partie de forme est définie comme un fragment du contour initial limité par une paire non-consécutive de sommets DCE obtenus. Par exemple, le contour fermé d'une souris de la figure 2.19 (a) est représenté par le polygone simplifié illustré en figure 2.19 (b). La deuxième ligne de la figure 2.19 montre certaines parties extraites en s'appuyant sur les sommets du polygone simplifié (b) et sur le contour d'origine (a).

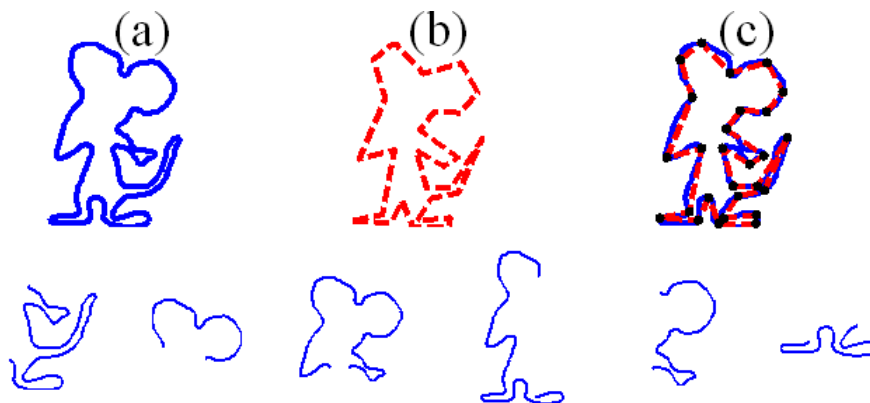


Figure 2.19 – En (a) le contour d'une forme d'une souris, en (b) le polygone simplifié par DCE et en (c) le polygone de sommets placé sur le contour d'origine (marqués par des cercles noirs). La deuxième ligne illustre certaines parties extraites du contour initial (a).

2.6 Conclusion

Dans ce chapitre nous avons présenté le processus de décomposition de formes. La décomposition des formes devient l'un des problèmes fondamentaux en la reconnaissance de formes, car cette décomposition fournit souvent des descriptions locales qui sont très proches de la perception humaine. Un argument en faveur de l'utilisation des descriptions locales provient du fait que le système visuel humain utilise une représentation partielle des formes pour la reconnaissance de formes à partir des parties de forme. Les parties sont également utiles pour représenter les formes des objets qui ne sont pas rigides. Le contour d'un chat assis et le contour d'un chat en marche se correspondent mal lors d'une correspondance globale des contours. Par conséquent, une décomposition des formes en parties établit une correspondance partielle des formes au lieu de correspondre les formes entières.

Plusieurs méthodes de décomposition de forme ont été proposées dans la littérature. Celles-ci peuvent généralement être classées comme "basées-région" ou "basées-contour". Les méthodes basées-région partitionnent la forme en tenant compte de son intérieur. Les méthodes basées-contour fournissent une décomposition du contour de la forme en parties, en tenant compte seulement des caractéristiques du contour de la forme. Comme la forme d'un objet se distingue essentiellement par le contour externe, les méthodes basées-contour sont plus investies que les méthodes basées-région. Cependant, l'évaluation des algorithmes de décomposition de forme est encore subjective, car la décomposition elle-même n'est pas un problème bien posé. Cela peut être la raison principale pour laquelle aucune évaluation quantitative n'a été proposée au niveau des algorithmes de décomposition de formes.

Les contours des objets dans les images numériques sont souvent déformés par le bruit de la numérisation et de la segmentation. L'évolution discrète de la courbe (DCE) a été proposée dans la littérature comme une méthode efficace pour éliminer ces distorsions tout en préservant les aspects perceptifs suffisants pour la reconnaissance de formes. C'est pourquoi, dans ce chapitre, nous avons présenté et détaillé le processus de décomposition par DCE. Selon la géométrie de la forme, la DCE évalue l'information globale du contour afin de générer le contour simplifié. En plus, la simplification par DCE est guidée par une mesure de pertinence basée sur les caractéristiques globales du contour de la forme. Ainsi, ce processus est moins influencé par le bruit que les autres techniques de simplification basées sur des caractéristiques locales (CSS, courbure locale, etc.). Le chapitre suivant sera dédié aux approches géodésiques développées dans le cadre de cette thèse pour la reconnaissance de parties de formes et pour la reconnaissance de formes partielles.

Partie II

Méthodologies et validation

Chapitre 3

Reconnaissance de parties de formes basée sur les géodésiques de formes

Sommaire

3.1	Introduction	80
3.2	Recalage de parties de contours 2D	81
3.2.1	Mesure de similarité	82
3.2.2	Recalage robuste	83
3.2.3	Métrique proposée	85
3.2.4	Minimisation par programmation dynamique	86
3.3	Système de reconnaissance de parties de formes	88
3.3.1	Construction de la base des parties de formes étiquetées	88
3.3.2	Stratégie de sous-sélection	91
3.3.3	Schéma proposé pour la reconnaissance	92
3.3.4	Évaluation des performances	92
3.3.4.1	Depuis la base de formes MPEG-7	92
3.3.4.2	Sur des images réelles	98
3.4	Reconnaissance de formes partielles	103
3.4.1	Stratégie de reconnaissance proposée	103
3.4.2	Évaluation depuis la base de formes MPEG-7	104
3.4.3	Évaluation sur des images réelles	106
3.5	Conclusion	116

3.1 Introduction

La segmentation d'images, en particulier la détection de contours, est un problème très important qui se pose dans le domaine de la vision par ordinateur. Les performances du système de reconnaissance de formes dépendent en bonne partie de la qualité de l'image segmentée. Le traitement des images a connu ces dernières années un essor considérable et un nombre conséquent de travaux de recherche ont été proposés au niveau de la détection de contours. Malgré les progrès effectués [207], une segmentation complète (c'est-à-dire avec des contours entiers) ne peut pas être toujours atteinte et reste un problème difficile à résoudre. Les opérateurs classiques de détection de contours ont beaucoup de limitations en présence de bruit, de texture, de changement d'éclairage, d'occlusion ou d'objets en contact dans les images à segmenter. En raison de ces différents facteurs, il est souvent impossible de segmenter l'objet en entier et seulement certaines parties peuvent être détectées (voir figure 3.1 suivante). Ceci justifie pleinement la nécessité d'un système de reconnaissance de parties de formes et d'un système de reconnaissance de formes partielles.

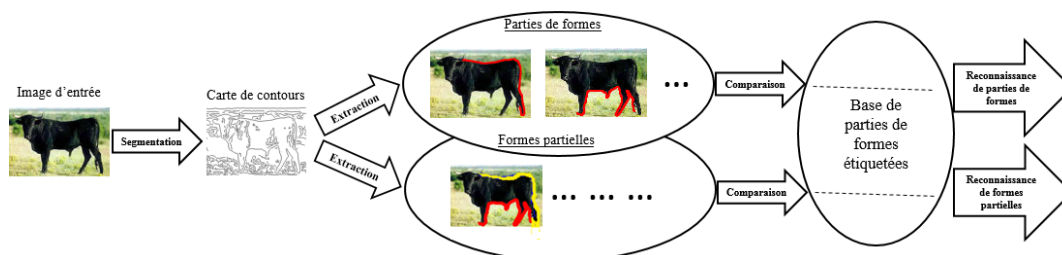


Figure 3.1 – Schéma proposé pour la reconnaissance de parties de formes et de formes partielles.

Dans ce chapitre, nous proposons deux approches de reconnaissance basées-contour qui sont fondées sur un processus de comparaison entre les parties de formes par mise en correspondance (autrement dit recalage) : une approche de reconnaissance de parties de formes et une approche de reconnaissance de formes partielles¹. La méthode de recalage proposée consiste à optimiser une mesure de similarité basée sur les géodésiques dans l'espace de formes. Afin d'atténuer l'influence des variations inter-individuelles, nous proposons d'utiliser un recalage robuste.

Nous commençons par décrire en section 3.2 la distance géodésique utilisée et l'approche proposée pour le recalage robuste de parties de contours 2D. En section 3.3, nous proposons une nouvelle métrique distance géodésique pour la classification et la recherche de parties de formes. Nous décrivons ensuite la stratégie de sous-sélection et le schéma de reconnaissance que nous proposons avec la distance basée sur les géodésiques de formes. Cette section présente une comparaison de performance notre

¹Il s'agit, dans les deux approches proposées, de formes planaires.

approche basée recalage et géodésiques de formes à l'état de l'art des approches de reconnaissance de parties de formes (classification et recherche). Ces mesures de performance seront établies à partir de différents tests effectués à la fois sur les images de la base MPEG-7 et sur des images réelles. Dans la section 3.4, nous exploitons la distance proposée avec une nouvelle stratégie développée pour la reconnaissance de formes partielles. Il est montré ensuite que l'approche géodésique proposée surpasse encore les approches de l'état de l'art pour la reconnaissance de formes partielles.

3.2 Recalage de parties de contours 2D

Nous nous intéressons à une approche robuste de recalage de parties de formes, avec un traitement symétrique, sans contraintes spécifiques et qui soit spatialement local. Pour ce faire, nous proposons de développer une approche basée sur les géodésiques dans l'espace de formes où le problème de recalage revient à un problème de recherche d'un chemin géodésique.

Il est connu que les calculs géodésiques constituent la base des analyses statistiques sur les espaces de formes [9, 241]. Ils généralisent également les outils standards de la géométrie euclidienne comme les moyennes [10], la régression linéaire [179], les splines cubiques [237] et d'autres. Par la suite, les géodésiques dans l'espace de formes ont été récemment étudiées dans de nombreuses applications d'imagerie telles que le recalage des formes [225, 237, 241], le suivi [225] et la déformations des formes [106].

Dans notre approche, nous proposons d'exploiter les géodésiques dans l'espace de formes pour l'analyse des courbes ouvertes, plus précisément pour la reconnaissance de parties de formes et la reconnaissance de formes partielles. L'étude quantitative d'une partie de forme commence par la formulation d'un espace de formes \mathcal{C} , dans lequel chaque partie de forme est représentée par un point. Étant donné un contour planaire ouvert, le groupe des translations, rotations et changements d'échelle crée une famille de contours ouverts. La seule propriété commune de tous les éléments de cette famille est leur forme. L'ensemble de toutes ces familles est appelé l'espace de formes. En général, l'étude de la géométrie de l'espace de forme est basée sur l'exploration des espaces non-linéaires, les variétés, qui sont généralement de grande dimension.

Notre utilisation de l'approche de l'espace et les géodésiques de formes pour la comparaison des parties de formes est montrée par le fait que la métrique donnée par la distance géodésique est invariante par rapport à un ensemble de transformations géométriques (rotation, translation, changement d'échelle, ...). Aussi, la fonction de correspondance trouvée entre les formes peut être contrainte d'être un difféomorphisme (c'est-à-dire lisse et inversible, avec une inverse lisse) pour que les parties connexes restent connectées, les ensembles disjoints restent disjoints, la régularité des caractéristiques telles que la courbure soit préservée, et les coordonnées soient transformées de manière cohérente.

3.2.1 Mesure de similarité

La définition d'une métrique sur l'espace de formes \mathcal{C} permet de mesurer la similarité entre deux parties de formes données. Lorsque l'espace de formes est muni de la structure d'une variété Riemannienne, cette distance est définie par le chemin géodésique le plus court, autrement dit, le chemin géodésique qui a la longueur minimale (voir figure 3.2). Les géodésiques dans l'espace de formes sont définies comme les chemins les plus courts entre les parties de formes par rapport à une métrique riemannienne donnée. Par exemple, les géodésiques sont les segments de lignes droites sur un espace euclidien, et elles sont les arcs de grands cercles sur une sphère.

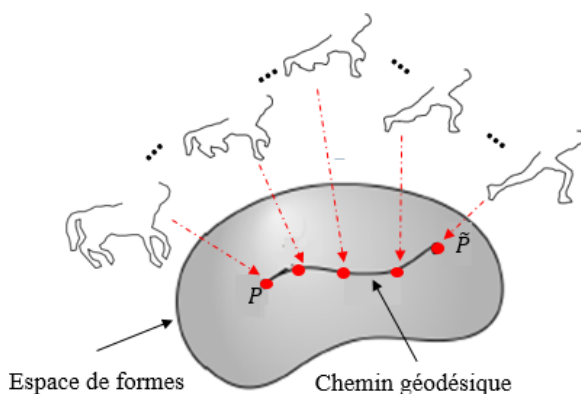


Figure 3.2 – Les déformations continues entre toute paire de parties de formes P et \tilde{P} sont représentées par des chemins géodésiques dans l'espace de formes.

Généralement, la modélisation des courbes sur une variété Riemannienne consiste à définir une métrique hilbertienne sur chaque plan tangent² de l'espace des courbes. Par conséquent, les courbes dans l'espace de formes \mathcal{C} sont représentées comme des déformations des parties de formes de référence, ce qui rend le calcul des géodésiques comme l'optimisation d'un critère de similarité avec des contraintes sur un groupe des déformations G . Un exemple de ces contraintes pourrait être d'empêcher le croisement des courbes. Le groupe G agit sur \mathcal{C} , de sorte que toute partie de forme de référence P peut être déformée en une autre $aP \in \mathcal{C}$, ($a \in G$). Ainsi, les géodésiques dans l'espace de formes correspondent aux déformations optimales qui alignent une courbe à l'autre comme décrit dans [161].

Nombreux sont les travaux antérieurs qui ont proposé des métriques pour l'analyse des courbes continues fermées [68, 69, 75, 103, 161, 204, 256]. Mais, seulement quelques chercheurs ont décrit l'analyse des courbes ouvertes. D'abord Younes [255, 256] a défini une métrique sur un groupe de difféomorphismes pour analyser et étudier les courbes 2D ouvertes qui ne présentent pas de croisements. Une autre représentation des courbes

²l'ensemble des champs de vecteurs qui déforment infinitésimalement une courbe donnée

ouvertes basée sur la racine carrée de la vitesse (*SRV : Square Root Velocity*) a été utilisée dans [152,249]. Dans l'étude qui nous intéresse, les courbes étudiées sont ouvertes et sans croisement, c'est pour cela que nous allons considérer dans notre approche la métrique proposée par Younes [255, 256] qui est stable et symétrique par définition.

3.2.2 Recalage robuste

Considérons deux parties de formes P et \tilde{P} normalisées en utilisant la méthode d'alignement décrite dans la section 4.2 du chapitre 4. L'utilisation de cette méthode sera validée dans le chapitre suivant (chapitre 4). D'abord, P et \tilde{P} sont localement représentées par l'angle formé entre la tangente à la courbe et l'axe horizontal. Ces fonctions, autrement dit signatures, angulaires $\theta(s)$ et $\tilde{\theta}(s)$ sont paramétrées par l'abscisse curviligne s normalisée à l'unité (voir figure 3.3). L'utilisation de la signature angulaire $\theta(s)$ (respectivement $\tilde{\theta}(s)$) réalise naturellement l'invariance à la translation et au changement d'échelle. En effet, toute translation de la courbe n'a aucune incidence sur θ , et toute homothétie n'a aucun effet sur le paramètre normalisé s . En plus, la normalisation de P et \tilde{P} via notre méthode d'alignement (section 4.2) basée sur la redétermination des points extrémités de P et \tilde{P} suivie par la paramétrisation avec l'abscisse curviligne s normalisée à l'unité ont géré le problème des différentes longueurs des parties de formes données.

Ensuite, nous proposons d'établir un recalage robuste entre ces deux courbes ouvertes. Plus précisément, ce recalage consiste à chercher la fonction de correspondance optimale $\phi^*(s)$ entre P et \tilde{P} , qui fait correspondre les points de caractéristiques similaires comme illustré en figure 3.3. Cette correspondance entre P et \tilde{P} est établie en minimisant une mesure de similarité avec une norme robuste MS_R donnée par l'équation suivante :

$$MS_R^{P,\tilde{P}}(\phi) = \arccos \int_0^1 \sqrt{\phi_s(s)} \left| \cos \frac{\|\theta(s) - \tilde{\theta}(\phi(s))\|_\rho}{2} \right| ds, \quad (3.1)$$

$$\phi^* = \underset{\phi}{\operatorname{argmin}} MS_R^{P,\tilde{P}}(\phi),$$

où le terme $\sqrt{\phi_s(s)}$ pénalise la torsion et l'étirement tout au long des courbes et $\|\cdot\|_\rho$ est l'estimateur robuste utilisé. Dans [30], plusieurs normes robustes ρ sont proposées pour l'estimation robuste. La minimisation de MS_R est effectuée par programmation dynamique avec une contrainte qui exige une monotonie croissante pour ϕ , c'est-à-dire $\phi_s = \frac{d\phi(s)}{ds} > 0$. Dans le but d'établir une correspondance qui ne soit pas sensible aux variabilités inter-individuelles, nous utilisons l'estimateur robuste de Leclerc, une des fonctions les plus simples et efficaces permettant de pénaliser des valeurs relativement aberrantes. Il est donné par :

$$\rho(r) = 1 - \exp(-r^2/(2\sigma^2)),$$

CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR
LES GÉODÉSQUES DE FORMES

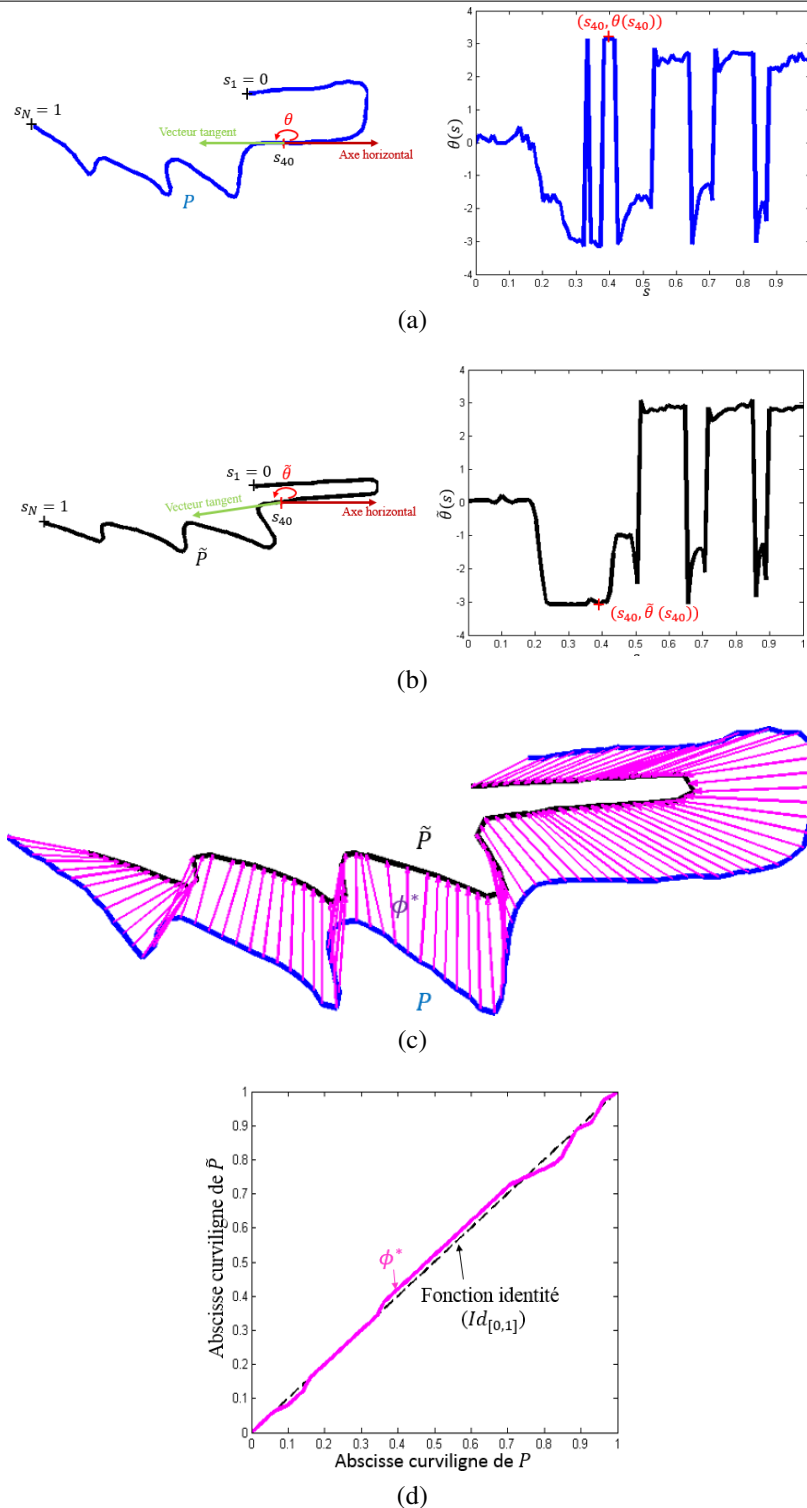


Figure 3.3 – Deux parties d’arbres P et \tilde{P} et leurs signatures angulaires $\theta(s)$ et $\tilde{\theta}(s)$ en (a) et (b) respectivement. En (c) et (d), la fonction de correspondance entre P et \tilde{P} obtenue par l’approche de recalage proposée.

3.2. RECALAGE DE PARTIES DE CONTOURS 2D

avec σ est l'écart-type des erreurs $r = \theta - \tilde{\theta}$. La figure 3.4b montre que la valeur de la norme augmente et tend vers l'infini avec la fonction quadratique. Cependant, l'estimateur robuste de Leclerc tend vers une limite envisageable quand l'erreur augmente. Considérée comme une fonction de coût, la mesure de similarité utilisée $MS_R^{P, \tilde{P}}$ correspond au coût de déformation $\delta(a, P)$ recalant P et \tilde{P} , tel que $\tilde{P} = a.P$. Ce coût est calculé en intégrant les coûts des petites déformations définissant un chemin $g : [0, 1] \rightarrow G$, dite chemin de déformation. Le groupe d'action G est muni d'une métrique Riemannienne appropriée.

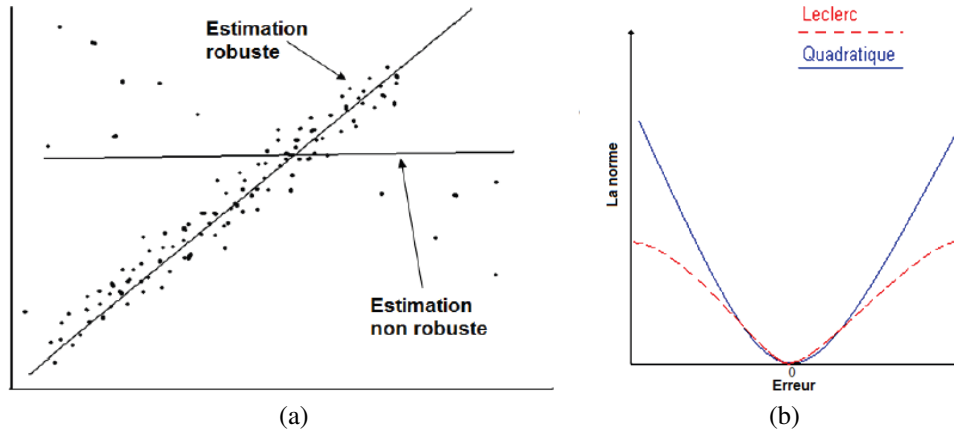


Figure 3.4 – Comparaison entre un estimateur robuste et un estimateur non robuste. En (a) l'estimation robuste est une technique voulue insensible aux petites variations et robuste aux données aberrantes. En (b) le principe de la norme robuste consiste à atténuer la contribution des points de variation élevée.

3.2.3 Métrique proposée

Une fois la fonction de correspondance optimale ϕ^* est obtenue, notre distance basée sur les géodésiques de formes d_{Geod} entre les deux parties de formes P et \tilde{P} est définie par :

$$d_{Geod}(P, \tilde{P}) = MS(\phi^*) = \arccos \int_0^1 \sqrt{\phi_s^*(s)} \left| \cos \left(\frac{\theta(s) - \tilde{\theta}(\phi^*(s))}{2} \right) \right| ds. \quad (3.2)$$

La distance proposée d_{Geod} est définie comme l'erreur résiduelle du recalage robuste établi entre P et \tilde{P} . Elle est également définie comme le coût de déformation minimal parmi tous les chemins de déformation $g : [0, 1] \rightarrow G$, tels que $g(0) = id_G$ et $g(1) = a$. Ceci peut être exprimé comme suit:

$$d_{Geod}(P, \tilde{P}) = SM(\phi^*) = \inf_g \left\{ \delta(a, P), \tilde{P} = a.P \right\} \quad (3.3)$$

CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR LES GÉODÉSIQUES DE FORMES

Vue comme une fonction de coût, cette métrique d_{Geod} respecte toutes les conditions nécessaires pour avoir une vraie distance entre les courbes $2D$. La distance proposée d_{Geod} , basée sur le résiduel du recalage présente plusieurs avantages :

- elle s'appuie sur une comparaison locale ;
- elle est symétrique ;
- elle est robuste aux données aberrantes et aux inter-variabilités ;
- elles autorise la définition de points repères.

Toutefois, cette distance présente un principal inconvénient par rapport aux métriques usuelles. Étant basée sur le résiduel de recalage, elle présente un important degré de complexité de calcul.

3.2.4 Minimisation par programmation dynamique

Nous proposons pour la résolution du problème d'optimisation d'utiliser la technique de programmation dynamique dont le principe est inventé par Richard Bellmann [23]. En général, cette technique est définie comme une méthode d'optimisation pour répondre à un problème de planification posé lorsque des décisions doivent intervenir à des périodes discrètes et qu'à chaque période, un nombre fini (petit) d'options de décisions peuvent être prises. Une décision à une étape est un point de départ pour l'étape suivante. L'algorithme de programmation dynamique est généralement connu pour son application aux problèmes de recherche de plus court chemin dans un graphe.

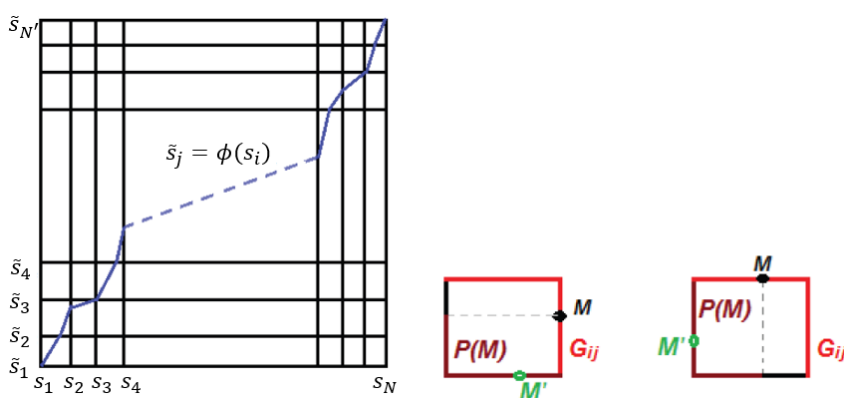


Figure 3.5 – Illustration du schéma de minimisation par programmation dynamique avec un exemple de fonction de correspondance.

En particulier, la programmation dynamique est souvent utilisée pour résoudre des problèmes de DTW à des fins de classification [86, 93, 186]. Étant données les deux

3.2. RECALAGE DE PARTIES DE CONTOURS 2D

signatures angulaires discrétisées $\theta(s_i)_{i=1\dots N}$ et $\tilde{\theta}(\tilde{s}_j)_{j=1\dots N'}$, l'algorithme de programmation dynamique considère dans le plan $[s_1, s_N] \times [\tilde{s}_1, \tilde{s}_{N'}]$ la grille G qui contient tous les points (x, y) tels que $x = s_i \in [s_1, s_N]$ et $y = \tilde{s}_j \in [\tilde{s}_1, \tilde{s}_{N'}]$ (figure 3.5). Tout point (s_i, \tilde{s}_j) fait référence à une correspondance des instants s_i et \tilde{s}_j . La programmation dynamique est utilisée pour trouver par parcours inverse la correspondance optimale qui désigne le trajet de coût minimal. Toute correspondance admissible entre $[s_1, s_N]$ et $[\tilde{s}_1, \tilde{s}_{N'}]$ est une fonction :

- non-linéaire (en global),
- continue (les pas consécutifs doivent être faits entre des points de cellules adjacentes),
- croissante,
- linéaire sur chaque cellule rectangulaire de la grille (c'est-à-dire localement linéaire), et
- contrainte de respecter les conditions aux limites (point de départ et point d'arrivée aux coins opposés de la grille).

La valeur du coût est calculée en chaque point de la grille en fonction des valeurs aux points précédents.

Sur la grille G , on appelle $H[i-1, i][j]$ le segment horizontal $\tilde{s} = \tilde{s}_j$; $s_{i-1} \leq s \leq s_i$. De même, on appelle $V[i][j-1, j]$ le segment vertical $s = s_i$; $\tilde{s}_{j-1} \leq \tilde{s} \leq \tilde{s}_j$. Soit $G_{ij} = H[i-1, i][j] \cup V[i][j-1, j]$. Soit $M(s, \tilde{s})$ un point de la grille tel que $M \in G_{ij}$. Soit $M' = (s', \tilde{s}')$ un autre point de la grille G , on dit que $M' < M$ si $s' < s$ et $\tilde{s}' < \tilde{s}$. Soit $P(M)$ l'ensemble de points $M' < M$, précédents de M , tels que

$$M' \in H[i-1, i][j-1] \cup V[i-1][j-1, j].$$

Pour tout point $M' \in P(M)$, on définit le coût de M' vers M par :

$$W_{M'M} = \rho(\theta(s_i) - \tilde{\theta}(\tilde{s}_j))(s - s'),$$

$$\rho(\theta(s_i) - \tilde{\theta}(\tilde{s}_j)) = \arccos \sqrt{\phi_s(s_i)} \left| \cos \frac{\|\theta(s_i) - \tilde{\theta}(\tilde{s}_j)\|_\rho}{2} \right|.$$

Le problème de minimisation de $MS_R^{P\tilde{P}}$ pourrait être reformulé comme le problème de trouver un entier k et une séquence de points $M_1 = (s_1, \tilde{s}_1), M_2, \dots, M_{k-1}, M_k = (s_N, \tilde{s}_{N'})$, avec $M_{i-1} \in P(M_i)$, qui minimise un coût global $L(M_1, \dots, M_k) = \sum_{i=2}^k W_{M_{i-1}M_i}$. Ce chemin peut être trouvé efficacement en utilisant le programme dynamique qui évalue d'une façon récurrente le coût au point

M , $W(M)$, comme une mesure cumulative du coût au point précédent M' et du coût de M' vers M :

$$W(M) = \inf_{M' \in P(M)} (W(M') + W_{M'M}).$$

3.3 Système de reconnaissance de parties de formes

Comme nous l'avons déjà mentionné, la plupart des systèmes de reconnaissance utilisent des techniques de correspondance de formes entières ou de formes partielles, mais la correspondance de parties de forme n'est que très peu abordée dans la littérature. Récemment, un schéma de reconnaissance de parties de formes basé sur le contexte de forme [19] et un autre fondé sur l'analyse Procuste [71, 154] ont été proposés. Les formes sont alignées par une représentation dans le repère canonique pour le premier et par l'alignement Procuste pour le second. La distance basée sur le contexte de forme (voir section 1.6.5) et la distance Procuste (voir section 1.6.4) sont respectivement utilisées pour mesurer la similarité entre les parties de formes. Pour notre part, nous allons exploiter la métrique basée sur les géodésiques de formes d_{Geod} proposée dans la section 3.2.1, pour la classification et la recherche de parties de formes. Comme nous allons le montrer par la suite, l'utilisation de notre métrique présente l'avantage de prendre en compte la correspondance des points de caractéristiques similaires. Aussi, la métrique proposée est symétrique, dans le sens où le recalage d'une partie de forme P par rapport à une autre partie de forme \tilde{P} est identique au recalage de \tilde{P} par rapport à P . En fait, dans les deux cas on cherche le trajet de coût de déformation minimal recalant les deux parties de formes.

3.3.1 Construction de la base des parties de formes étiquetées

En absence d'une base standard de parties de formes, nous proposons de construire une base de parties de formes étiquetées à partir d'une base de formes entières (voir figure 3.6). Les étapes de la construction sont les suivantes :

- Pour chaque forme entière, simplifier le contour fermé en un polygone représentatif via la méthode de l'évolution par retour en arrière de la courbe discrète (DCE) décrite en section 2.5.2 du chapitre 2.
- Pour chaque forme entière, utiliser les sommets du polygone simplifié, trouvé par évolution optimale, pour obtenir les parties visuelles de son contour d'origine. Une partie de forme est définie comme un fragment du contour initial limité par une paire non-consécutive de sommets DCE obtenus.
- Exclure de la base des parties de formes étiquetées les parties qui ont une faible x -variance ou y -variance dans le repère canonique (x, y) .

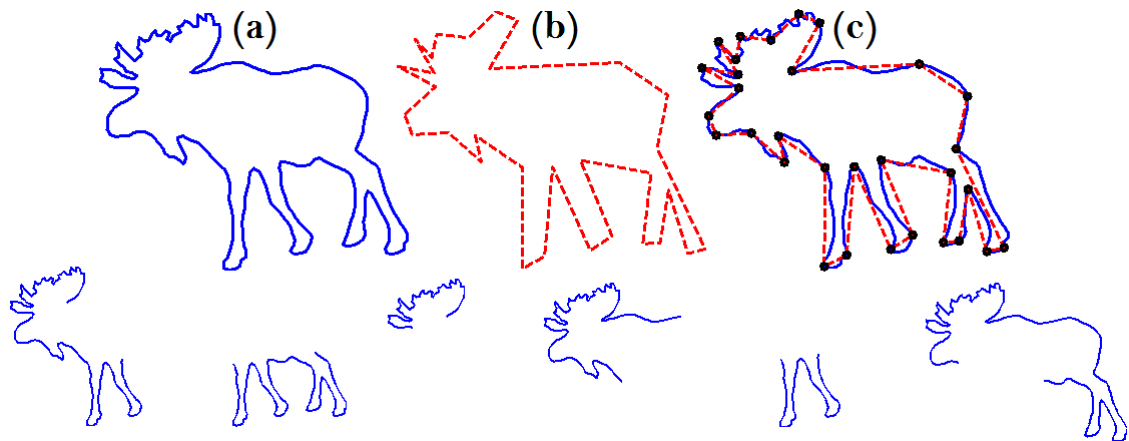


Figure 3.6 – En (a) le contour d’une forme d’un cerf, en (b) le polygone simplifié par DCE et en (c) le polygone de sommets placé sur le contour d’origine (marqués par des cercles noirs). La deuxième ligne illustre certaines parties extraites du contour initial (a).

Cependant, une décomposition (DCE) significative nécessite une recherche de l’étape de l’évolution optimale la plus réaliste. En utilisant la mesure de similarité³ m_s basée sur la fonction d’angle [123], il est impossible de trouver avec un même seuil S un niveau d’évolution intuitivement correct pour différentes classes de formes. Expérimentalement, avec un seuil fixe S , une simplification homogène est loin d’être atteinte pour la totalité de la base MPEG-7 par exemple. La figure 3.7 illustre des polygones simplifiés (en rouge) de trois formes différentes (en noir) de la classe *Cerf* de la base MPEG-7. Le polygone simplifié (en rouge) de la forme du *Cerf-6*, obtenu à son

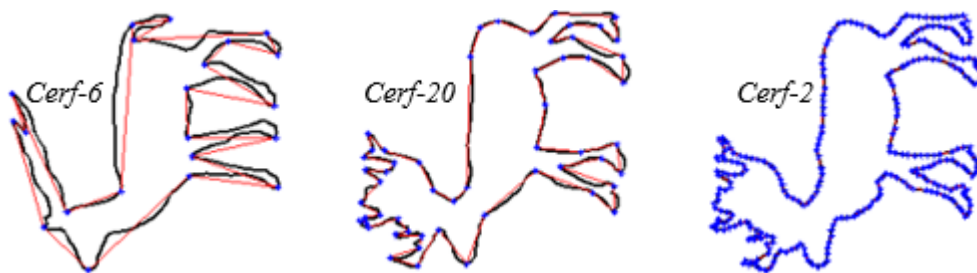


Figure 3.7 – Évolution optimale de trois formes de *Cerf* tirées de la base MPEG-7. Chaque forme est représentée par le polygone en rouge dont les points DCE (sommets) sont représentés en bleu. L’optimalité est obtenue par la méthode de retour en arrière en utilisant la mesure de similarité partielle m_s (section 2.5.2) avec un seuil fixe ($S = 0.4$).

étape d’évolution optimale, indique que cette forme est abusivement simplifiée avec un seuil fixe $S = 0.4$. En revanche avec le même seuil, la simplification de la forme du

³définie par l’équation 2.1, en section 2.5.2 du chapitre 2.

CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR LES GÉODÉSQUES DE FORMES

Cerf-2 est insuffisante. Ainsi, les deux formes de *Cerf-6* et *Cerf-2* sont mal simplifiées avec le seuil $S = 0.4$. Au contraire, la simplification obtenue de la forme du *Cerf-20* est un polygone significatif et très représentatif avec un nombre minimal de points DCE (en bleu). Cette différence extrême qui existe au niveau de la simplification est due à la grande sensibilité de la fonction d'angle au bruit. Pour cette raison, nous avons proposé dans notre méthode de remplacer la mesure de similarité partielle ms (équation 2.1, section 2.5.2) basée sur la fonction d'angle par une mesure de similarité partielle ms' basée sur les coordonnées cartésiennes des fragments de contours :

$$ms'(P_e, P_0) = \int_0^1 [(x_e(s) - x_0(s))^2 + (y_e(s) - y_0(s))^2] / A ds, \quad (3.4)$$

où (x_e, y_e) et (x_0, y_0) désignent les coordonnées cartésiennes des segments de formes P_e et P_0 respectivement. Et, nous désignons par A l'aire du contour fermé P_0 .

La mesure de similarité partielle ms' (équation 3.4) est donnée par la distance \mathbb{L}^2 entre la partie P_e du polygone simplifié à l'étape d'évolution e et la partie correspondante P_0 sur la forme originale ; P_e et P_0 sont définies par leurs coordonnées cartésiennes (x_e, y_e) et (x_0, y_0) respectivement, paramétrées par l'abscisse curviligne normalisée s . Comme l'erreur partielle proposée dans ms' est définie par l'intermédiaire des coordonnées des parties de formes, elle est donc proportionnelle au changement d'échelle. Une fois divisée par l'aire de la forme originale (avant simplification), elle devient invariante au changement d'échelle. L'utilisation de cette erreur quadratique au lieu de celle basée sur la fonction d'angle a montré la possibilité d'avoir un seul seuil au moins pour les formes d'une même classe au sein de la base MPEG-7. Par exemple, en reprenant les mêmes formes de la classe *Cerf*, la figure 3.8 montre que l'utilisation de la mesure de similarité partielle proposée ms' avec un seuil fixe $S = 0.0006$ conduit à l'évolution optimale de chacune des trois formes.

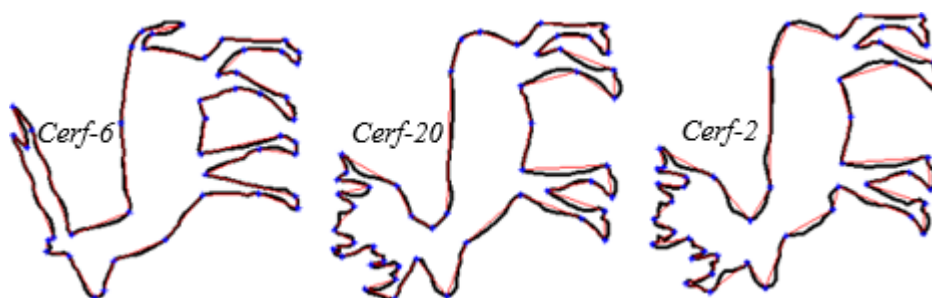


Figure 3.8 – Évolution optimale des trois formes de *Cerf*, présentées en figure 3.7, en utilisant la mesure de similarité partielle proposée ms' (équation 3.4) avec un seuil fixe ($S = 0.0006$).

Expérimentalement nous avons trouvé qu'à travers un seuil fixe S égal à 0.0003, l'évolution optimale de la plupart des formes de la base MPEG-7 peut être atteinte. Afin

3.3. SYSTÈME DE RECONNAISSANCE DE PARTIES DE FORMES

d'atteindre l'évolution optimale de toutes les formes de la base MPEG-7, nous avons proposé de diviser manuellement la base de formes MPEG-7 en deux catégories. Une catégorie de formes simples qui renferme les formes présentant peu de détails telles que les formes des classes *Pomme*, *Cloche*, *Os*, *Fontaine*, *Chaussure*, *Camion*, *Marteau*, *Bouteille*, *Montre*, etc. Une catégorie de formes complexes avec beaucoup de détails telles que les *Dendroctone*, *Éléphant*, *Cerf*, *Souris*, *Octopus*, *Papillon*, *Chameau*, *Hélicoptère*, etc. Par suite, nous avons cherché expérimentalement un seuil fixe pour chaque catégorie. Le seuil S que nous avons trouvé pour les formes complexes vaut 0.0006 tandis que celui qui correspond à la catégorie des formes simples est égal à 0.0002.

Plus généralement, nous supposons que la base de formes entières d'origine contient n_C classes, $(C_k)_{k=1,\dots,n_C}$; chaque classe se compose de n_I instances/formes différentes, $(I_j)_{j=1,\dots,n_I}$. Nous désignons par (I_j, C_k) une instance I_j qui appartient à la classe C_k . Ici, nous décomposons chaque instance en n_P parties de forme significatives ; n_P dépend de l'instance à fragmenter. Selon ce schéma, on obtient alors un ensemble de parties de formes étiquetées $(P_i)_{i=1,\dots,n_P}$ où P_i désigne la $i^{\text{ème}}$ partie de l'instance (I_j, C_k) . Chaque partie est représentée par N points d'échantillonnage. Pour décomposer une instance donnée en parties, le contour fermé correspondant est simplifié en un polygone représentatif via la méthode de l'évolution par retour en arrière de la courbe discrète (DCE) en utilisant la mesure de similarité proposée ms' avec un seuil S convenablement choisi. Par exemple, le contour d'une forme d'un cerf de la figure 3.6a est représenté par le polygone simplifié illustré en figure 3.6b.

Ensuite, nous utilisons les sommets noirs (Fig. 3.6c) du polygone simplifié pour obtenir les parties visuelles de son contour d'origine correspondant. Une partie de forme est définie comme un fragment du contour initial limité par une paire non-consécutive de sommets obtenus. Dans la deuxième ligne de la figure 3.6, nous montrons certaines parties extraites en s'appuyant sur les sommets du polygone simplifié (b) et sur le contour d'origine (a). Enfin, nous excluons de notre base de données toutes les parties de contour non significatives ayant de petites x-variance (v_x) et y-variance (v_y) dans le repère canonique. Nous avons expérimentalement trouvé qu'une partie de forme de variance v_x et v_y inférieure à 0.11 est presque une ligne droite et par suite non significative. À l'issue de ces étapes, nous construisons une base de parties de formes étiquetées pour la reconnaissance de parties de formes.

3.3.2 Stratégie de sous-sélection

À cause de la grande taille de la base construite de parties de formes, nous proposons une méthode de sous-sélection. Cette stratégie consiste à utiliser la distance proposée dans une analyse multi-échelle. La distance proposée est exploitée tout d'abord à un niveau bas d'échantillonnage avec M points représentatifs ($M < N$) pour sélectionner un premier ensemble de parties de formes les plus similaires à la requête donnée. Puis au

niveau d'échantillonnage original avec N points représentatifs, nous comparons la requête donnée avec les parties sélectionnées au lieu de la comparer à tous les éléments de la base. L'efficacité de cette stratégie est expérimentalement validée dans une section suivante.

3.3.3 Schéma proposé pour la reconnaissance

On suppose avoir une base de parties de formes étiquetées, (P_l, C_l) , où P_l est la $l^{\text{ième}}$ partie de forme de la base et C_l est la classe correspondante. La classification d'une nouvelle partie de forme P est issue du classifieur plus proche voisin en fonction de la distance d_{Geod} considérée.

Nous abordons également le problème de la recherche qui consiste à déterminer quelles parties de formes de la base d'apprentissage sont les plus similaires à une partie requête donnée. Nous mesurons la précision de la recherche (PR) par l'équation 3.5 en tant que pourcentage des parties correctes.

$$PR_{(en\%)} = \frac{\text{Nbre des parties correctes}}{\text{Nbre des parties extraites}} \times 100. \quad (3.5)$$

3.3.4 Évaluation des performances

Afin d'évaluer les performances de l'approche proposée, nous avons construit la base de parties de formes étiquetées en utilisant la partie B de la base d'images MPEG-7. Cette base d'images est composée d'un grand nombre de formes réelles et synthétiques : 70 classes de formes avec 20 exemples de chaque classe. En utilisant les formes numérotées de 1 à 10 de chaque classe, la méthode de construction de la base d'apprentissage décrite dans la section 3.3.1 aboutit à une base de près de 60 000 parties de formes étiquetées ($|P_l| \approx 60000$). Pour tester l'efficacité de notre système de reconnaissance de parties de formes, nous avons réalisé une évaluation de performance à partir de la base MPEG-7 et sur des images réelles segmentées.

3.3.4.1 Depuis la base de formes MPEG-7

Nous avons utilisé deux ensembles de requêtes, chacun composé de 69 parties de formes (une requête par classe à l'exception de la classe de "Device-6", la classe 29 de la base de données MPEG-7). Plus précisément, pour chaque classe de la base MPEG-7, une requête est sélectionnée manuellement à partir de différentes instances en utilisant une des formes numérotées de 11 à 20. Chaque requête correspond à une partie significative qui caractérise sa propre classe. En raison de l'existence d'une forte variabilité intra-classe pour la forme *Device-6*, nous avons exclu cette requête par manque d'une partie significative commune aux formes de cette classe. Pour évaluer les performances de notre système de reconnaissance en fonction de la longueur de la requête, les requêtes du second ensemble (figure 3.9b) sont les parties de formes étendues des mêmes requêtes du premier

3.3. SYSTÈME DE RECONNAISSANCE DE PARTIES DE FORMES

ensemble (figure 3.9a). Dans la figure 3.9, nous illustrons les requêtes utilisées ; ces requêtes sont données dans le même ordre que les différentes classes dans la base de référence MPEG-7.

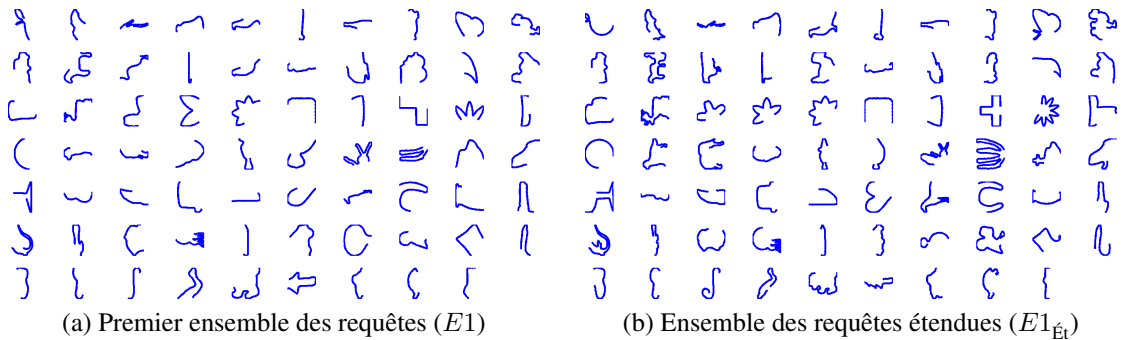


Figure 3.9 – Illustration des bases de test. En (a): Les 69 petites requêtes extraites des exemples 11-20 de la base MPEG-7. En (b): Les 69 requêtes étendues de mêmes requêtes données en (a).

L'approche proposée pour la correspondance et la reconnaissance de parties de formes est ainsi validée par une expérience quantitative sur les requêtes illustrées en figure 3.9 et la base construite des parties de formes étiquetées. À des fins de comparaison, nous avons implémenté l'approche basée sur le contexte de forme et l'approche Procuste présentées dans [19] et [71] respectivement. Chaque partie de contour est représentée par 100 points échantillonnés ($N = 100$). La méthode de sous-sélection proposée dans la section 3.3.2 est appliquée pour les trois approches afin de les placer dans un même cadre comparatif, où pour une requête donnée nous avons sélectionné à un niveau bas d'échantillonnage avec $N/2$ points un petit ensemble de 2% des parties de formes qui sont les plus semblables à cette requête.

La correspondance des points : Avant d'analyser les résultats de reconnaissance (recherche et classification) sur les requêtes données, nous montrons ici la différence majeure entre la méthode géodésique proposée et les deux autres méthodes de l'état de l'art. Cette différence réside dans la correspondance obtenue des points échantillonnés des parties de formes considérées. Dans la méthode utilisant la distance de Procuste, la correspondance est de type point-à-point. À l'encontre, la méthode géodésique et la méthode du contexte de forme traitent la correspondance des points à l'intérieur de chaque paire comme un problème d'optimisation. Dans la figure 3.10, nous illustrons les différentes correspondances sur une partie de la *Mouche-2* (C.38) (courbe continue en rouge) et la requête de la Mouche donnée en figure 3.9b. Avec une métrique, la méthode géodésique établit une correspondance correcte et ordonnée entre les points ayant des caractéristiques similaires (figure 3.10a). Cependant, l'approche basée sur le contexte de forme effectue une correspondance non ordonnée et localement aberrante entre les points

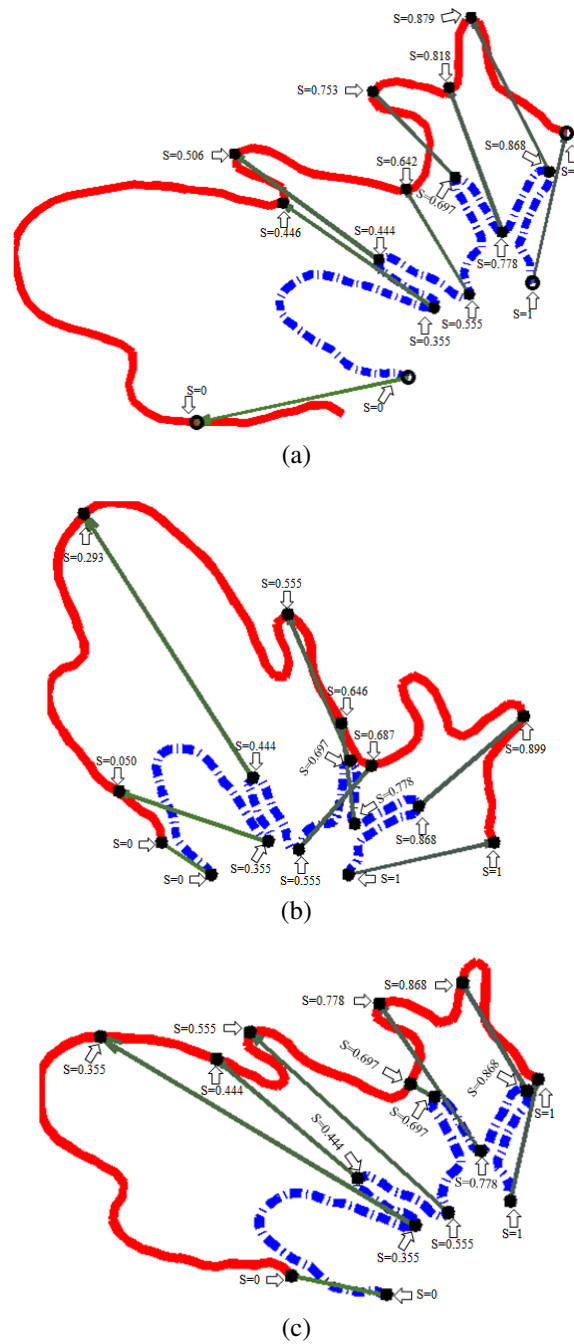


Figure 3.10 – La correspondance entre une partie de forme de la *Mouche-2* (C.38) (courbe continue en rouge) et la requête de *Mouche* (courbe pointillée en bleu). Correspondance avec notre méthode et avec notre métrique en (a), avec la méthode du contexte de forme en (b) et avec la méthode Procruste en (c).

3.3. SYSTÈME DE RECONNAISSANCE DE PARTIES DE FORMES

de la paire des parties de formes (figure 3.10b). On peut également voir en figure 3.10c la fausse correspondance obtenue avec la distance de Procuste aux points de mêmes abscisses curvilignes s .

La recherche : Pour évaluer la recherche des parties de formes, le taux de précision de recherche de chaque requête (PR) est mesuré comme le pourcentage de parties de formes qui appartiennent à la même classe parmi les 10 parties recherchées les plus similaires à la requête en utilisant l'équation 3.5. Tout d'abord, nous cherchons à évaluer l'efficacité de notre approche pour la reconnaissance de parties de formes de petite longueur. Les petites requêtes de l'ensemble $E1$ (figure 3.9a) sont utilisées à cette fin. Sur cet ensemble, grâce à l'utilisation de notre métrique basée sur les géodésiques dans l'espace de formes, le taux moyen de précision de recherche monte à 60.58 %. Ce taux diminue pour atteindre respectivement 52.90 % et 46.09 % avec l'approche basée sur le contexte de forme et l'approche Procuste. Pour plus de détails, le taux de précision de recherche de chaque requête est rapporté dans la figure 3.11a. Ensuite, nous appliquons les mêmes approches sur l'ensemble des requêtes étendues $E1_{\text{ét}}$, où chaque requête est la version prolongée de celle de l'ensemble $E1$. Avec ces requêtes de longueurs plus importante (figure 3.9b), nous notons une amélioration considérable des taux en recherche de parties de formes. L'approche géodésique conduit à une hausse des taux de précision de recherche avec une moyenne qui atteint 92.46 % de recherche correcte, alors que cette moyenne diminue pour atteindre respectivement 76.81 % et 72.03 % lors de l'utilisation de l'approche du contexte de forme et de l'approche Procuste (figure 3.11b). La comparaison quantitative des performances de recherche des trois méthodes est résumée en figure 3.12. Il apparaît clairement remarquable que la méthode géodésique atteint une meilleure performance en recherchant des requêtes significatives. Nous pouvons ainsi déduire que notre approche géodésique surpasse la méthode du contexte de forme et la méthode Procuste en recherche de parties de formes.

Concernant les résultats individuels de recherche, nous allons analyser le comportement des trois méthodes en présence de la variabilité intra-classe, lorsque la requête est une partie de forme significative (c'est-à-dire une requête de l'ensemble de la figure 3.9b). Par exemple, dans la figure 3.13, nous présentons les requêtes de l'*O*s et du *Cheval* avec des réponses extraites de la base d'apprentissage de formes déformées et articulées provenant des mêmes classes. En utilisant notre méthode géodésique, ces parties sont extraites parmi les dix premières parties de formes les plus similaires tandis que la méthode du contexte de forme et la méthode Procuste ne les ont pas trouvées. En effet, la description d'une partie de forme par la méthode du contexte de forme dépend de la répartition de ses points par rapport au point de référence. La méthode Procuste et la méthode du contexte de forme sont aussi sensibles aux déformations élastiques et aux articulations, alors que notre méthode géodésique peut les gérer efficacement.

La classification : La méthode proposée a été également exploitée pour la classification de parties de formes sur les requêtes des deux ensembles (figure 3.9) précédemment

CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR
LES GÉODÉSQUES DE FORMES

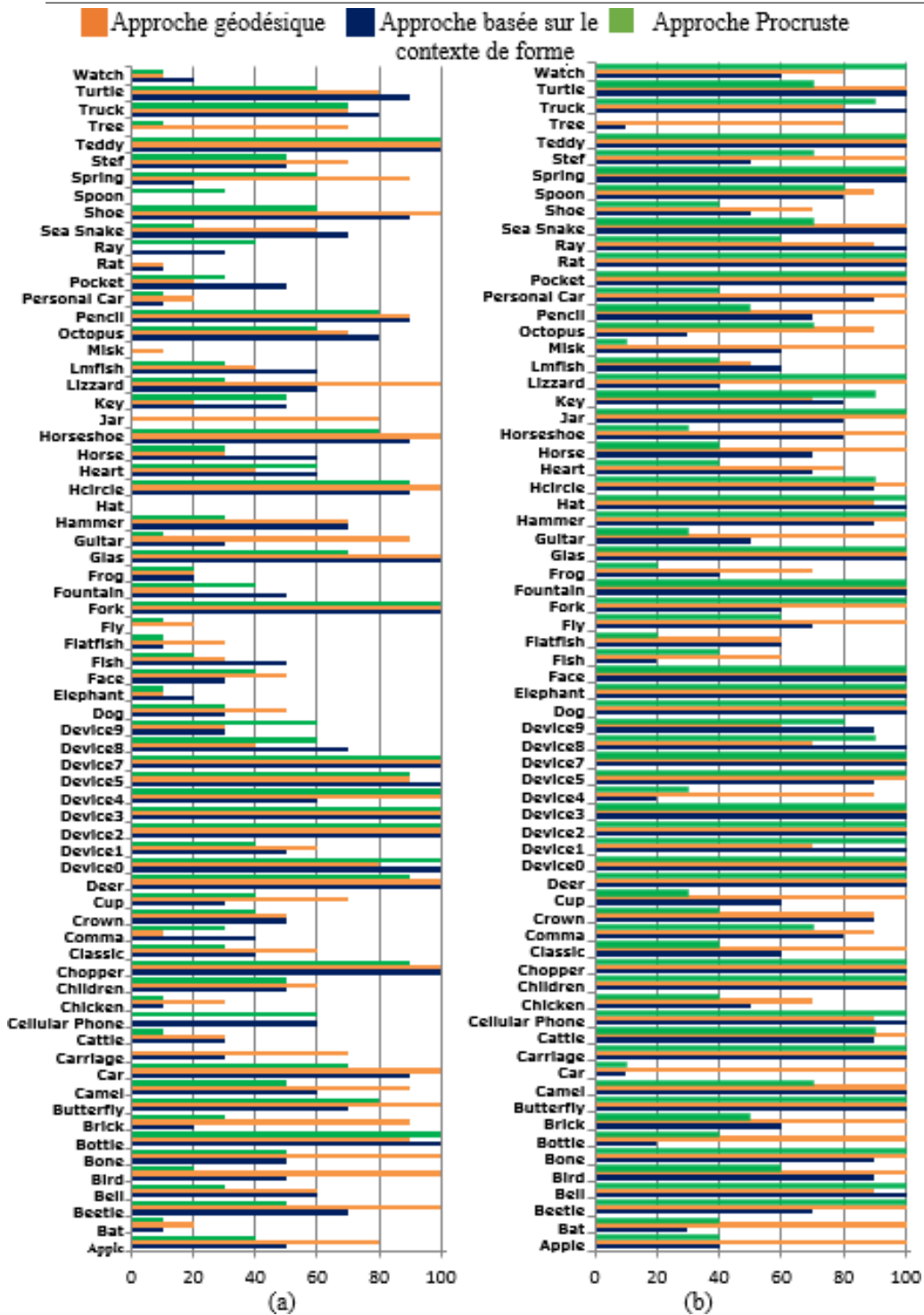


Figure 3.11 – Taux de recherche correcte de chaque requête obtenus par la méthode géodésique, la méthode du contexte de forme et la méthode Procuste. En (a) pour l'ensemble $E1$ des requêtes de petite longueur (figure 3.9a). En (b) pour l'ensemble $E1_{Ét}$ des requêtes de longueur plus importante (figure 3.9b)

3.3. SYSTÈME DE RECONNAISSANCE DE PARTIES DE FORMES

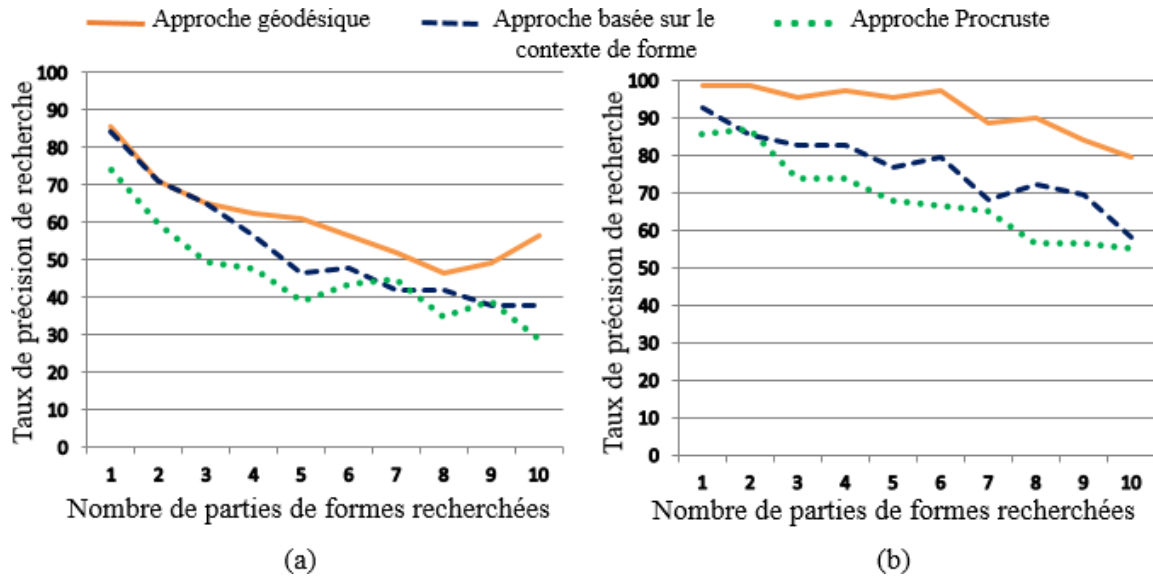


Figure 3.12 – Comparaison des performances de la méthode géodésique, la méthode du contexte de forme et la méthode Procuste en recherche, réalisée avec l'ensemble des petites requêtes (a) et avec l'ensemble des requêtes étendues (b).

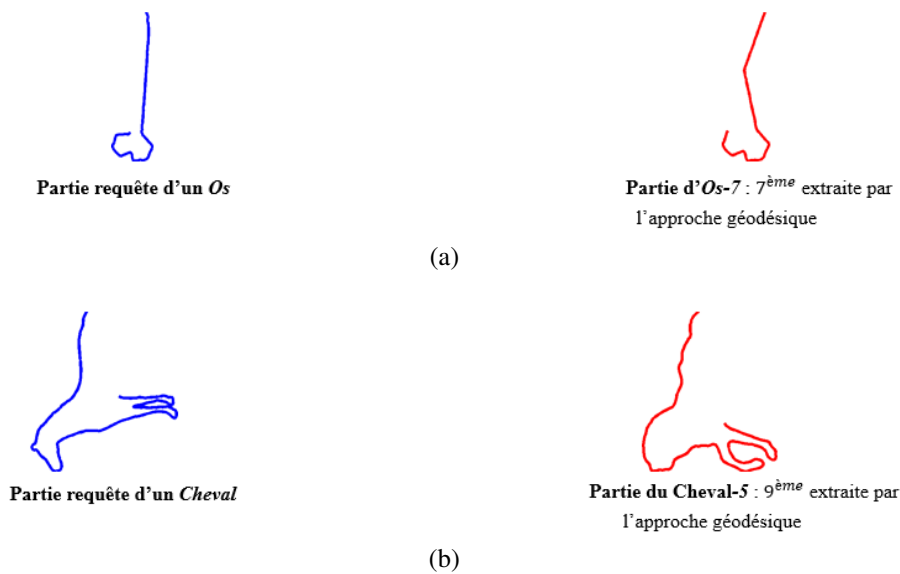


Figure 3.13 – Exemples de parties de formes extraites démontrant que notre approche est robuste aux déformations élastiques et aux articulations. Une partie de *Os-7* et une partie de *Cheval-5* sont parmi les 10 parties de formes trouvées avec notre méthode mais pas avec les deux autres méthodes.

*CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR
LES GÉODÉSIQUES DE FORMES*

utilisés pour la recherche de parties de formes.

Tableau 3.1 – Taux de classification correcte obtenus avec les classifieurs *1-NN*, *5-NN* et *10-NN* en utilisant les approches géodésique, du contexte de forme et Procuste sur l'ensemble des requêtes *E1* (figure 3.9a) et sur l'ensemble des requêtes étendues $E1_{\text{ét}}$ (figure 3.9b).

Classifieur <i>K-NN</i>	Approche géodésique		Approche du contexte de forme		Approche Procuste	
	<i>E1</i>	$E1_{\text{ét}}$	<i>E1</i>	$E1_{\text{ét}}$	<i>E1</i>	$E1_{\text{ét}}$
<i>1-NN</i>	85.51%	98.55%	85.51%	92.75%	73.91%	88.41%
<i>5-NN</i>	76.81%	100%	73.91%	89.86%	63.77%	91.30%
<i>10-NN</i>	68.12%	100%	68.12%	89.86%	60.87%	84.06%

Avec l'utilisation du classifieur des plus proches voisins (*K-NN*), nous avons attribué à chaque requête la classe majoritaire parmi les classes auxquelles appartiennent les *K* parties de formes les plus similaires. Sur le premier ensemble des requêtes de petite longueur *E1* (figure 3.9a), l'approche géodésique et l'approche du contexte de forme atteignent un taux de classification correcte égal à 85.51% avec le classificateur *1-NN* alors que l'approche Procuste atteint un taux de classification correcte bien inférieur (73.91%). En se référant au tableau 3.1, ce taux de classification correcte diminue avec le classifieur *10-NN* pour atteindre 68.12% avec la méthode géodésique et avec la méthode du contexte de forme, alors qu'il décroît à 60.87% avec l'approche Procuste. En appliquant le même schéma de classification sur l'ensemble des requêtes étendues $E1_{\text{ét}}$ (figure 3.9b), les valeurs des taux de classification correcte (tableau 3.1) montrent une amélioration des performances des trois approches. Avec les classifieurs *5-NN* et *10-NN*, l'approche géodésique et l'approche du contexte de forme ont présenté une performance stable de 100% et 89.86% respectivement. Cependant avec l'approche Procuste, ce taux de classification correcte diminue à 91.30% avec le classifieur *5-NN* et 88.41% puis à 84.06% avec les classifieurs *1-NN* et *10-NN*. Par conséquent, avec l'utilisation de tous les classifieurs *K-NN* sur les deux ensembles *E1* et $E1_{\text{ét}}$, notre méthode surpasse la méthode du contexte de forme et la méthode Procuste en classification de parties de formes. À la lumière des performances de chacune des trois méthodes utilisées, il apparaît clairement qu'il est plus facile de reconnaître une longue requête qu'une requête de petite longueur. En effet, lorsque la requête est une partie de forme significative, les trois approches conduisent à des taux de reconnaissance acceptables malgré la présence de la variabilité inter-classe.






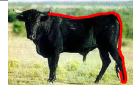









3.3.4.2 Sur des images réelles

Nous avons également testé l'efficacité de notre approche sur des parties de formes issues d'images segmentées. Nous avons utilisé les 5 parties les plus significatives des 5 images segmentées (une partie de chaque image) données dans le tableau 3.2. Ces

3.3. SYSTÈME DE RECONNAISSANCE DE PARTIES DE FORMES

parties sont utilisées comme des requêtes à reconnaître (à classifier et à rechercher) par notre système.

Tableau 3.2 – Exemples d’images réelles (en première colonne) et leurs cartes de contours (en deuxième colonne). Dans la troisième colonne, nous représentons en rouge la partie de contour la plus significative de l’image segmentée. Les colonnes restantes donnent la classe assignée par chacune des trois méthodes utilisées (géodésique, contexte de forme et Procuste) à chaque objet de l’image.










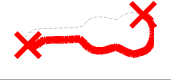



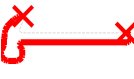




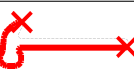
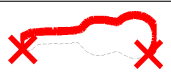

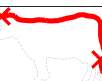

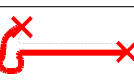
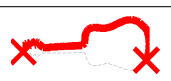

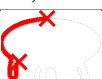

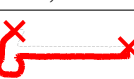


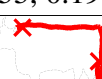

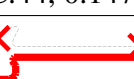
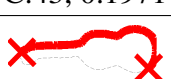
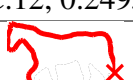
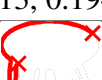
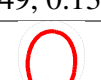
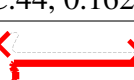
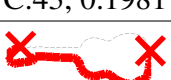

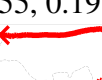
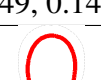
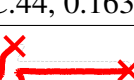
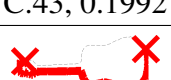
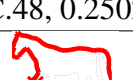
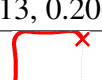
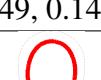
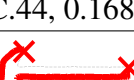
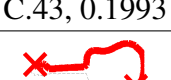
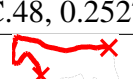
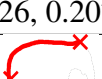

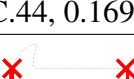
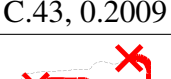
Image d’entrée	Carte de contours	Partie de contour retenue	Classe assignée (10-NN)		
			Approche géodésique	Approche du contexte de forme	Approche Procuste
			→Cheval (C.48)	→Chien (C.33)	→Fontaine (C.40)
			→Vache (C.13)	→Dispositif-3 (C.26)	→Dispositif-3 (C.26)
			→Fer à cheval (C.49)	→Fer à cheval (C.49)	→Fer à cheval (C.49)
			→Marteau (C.44)	→Os (C.6)	→Os (C.6)
			→Guitare (C.43)	→Guitare (C.43)	→Clé (C.51)

La recherche : Les résultats de recherche des 5 parties issues des 5 images réelles segmentées du tableau 3.2 sont donnés par les tableaux 3.3, 3.4 et 3.5. En première ligne du tableau 3.3, nous avons donné les requêtes utilisées. Les autres lignes du tableau présentent les résultats de recherche obtenus par notre approche. Les résultats obtenus par la méthode du contexte de forme et la méthode Procuste sur les mêmes requêtes sont donnés dans le tableau 3.4 et le tableau 3.5 respectivement.

Le taux de précision de recherche de chaque requête (*PR*) est mesuré comme le pourcentage de parties de formes qui appartiennent à la même classe parmi les 10 parties recherchées les plus similaires à la requête. En utilisant notre approche, le taux de précision de recherche atteint est toujours meilleur que celui obtenu par l’approche basée sur le contexte de forme ou par la méthode Procuste. Par exemple pour la requête partie de *Marteau* (C.44), la méthode du contexte de forme et la méthode Procuste donnent respectivement 30% et 0% de recherche correcte tandis que la méthode géodésique donne 100%. Même si la requête de *Marteau* est une partie représentative de sa forme originale, la méthode du contexte de forme a mal classé des parties de formes de différentes classes (*Os* (C.6) et *Tortue* (C.69)) parmi les 10 parties les plus similaires à la





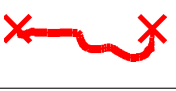



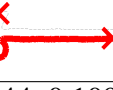
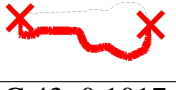




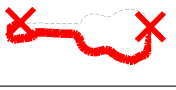




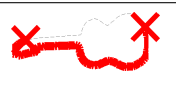



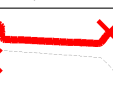




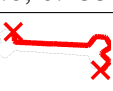
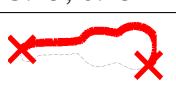



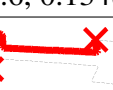
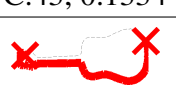
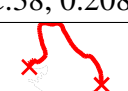


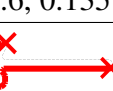
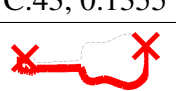
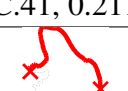


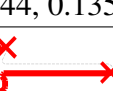
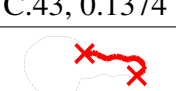



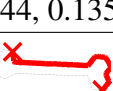
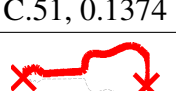
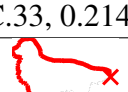


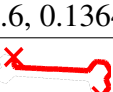
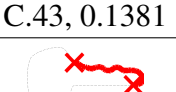
**CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR
LES GÉODÉSIQUES DE FORMES**

Tableau 3.3 – Résultats de la recherche en utilisant notre méthode sur les 5 requêtes retenues du tableau 3.2. En 2^{ème} ligne : le taux de précision de la recherche des requêtes. Dans les lignes restantes, nous montrons les 10 parties de formes les plus similaires à chaque requête. (C.x, D) : x est le numéro de la classe dans la base MPEG-7 et D est la distance attribuée à la requête.

				
PR=70%	PR=40%	PR=90%	PR=100%	PR=100%
				
C.64, 0.2424	C.22, 0.1758	C.49, 0.0885	C.44, 0.1436	C.43, 0.1877
				
C.48, 0.2437	C.22, 0.1758	C.49, 0.1097	C.44, 0.1436	C.43, 0.1881
				
C.48, 0.2441	C.13, 0.1789	C.49, 0.1149	C.44, 0.1465	C.43, 0.1908
				
C.12, 0.2442	C.13, 0.1873	C.58, 0.1217	C.44, 0.1465	C.43, 0.1944
				
C.48, 0.2458	C.55, 0.1915	C.49, 0.1304	C.44, 0.1472	C.43, 0.1971
				
C.12, 0.2495	C.13, 0.1946	C.49, 0.1351	C.44, 0.1629	C.43, 0.1981
				
C.48, 0.2496	C.55, 0.1971	C.49, 0.1412	C.44, 0.1639	C.43, 0.1992
				
C.48, 0.2502	C.13, 0.2067	C.49, 0.1434	C.44, 0.1688	C.43, 0.1993
				
C.48, 0.2527	C.26, 0.2072	C.49, 0.1443	C.44, 0.1699	C.43, 0.2009
				
C.48, 0.2536	C.34, 0.209	C.49, 0.1461	C.44, 0.1748	C.43, 0.2062





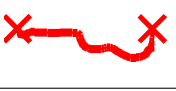





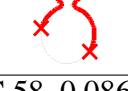



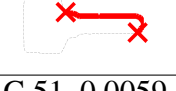
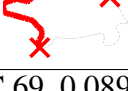


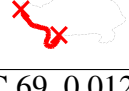
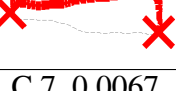
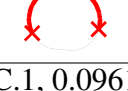



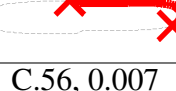
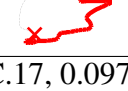


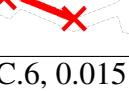
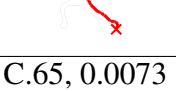
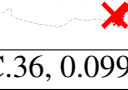


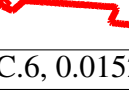

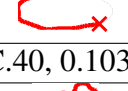


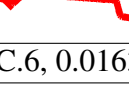
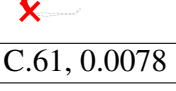
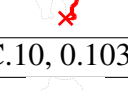


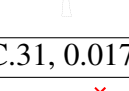
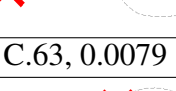
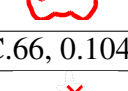


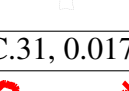
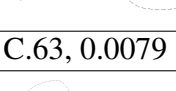
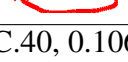


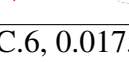
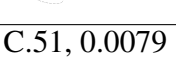
3.3. SYSTÈME DE RECONNAISSANCE DE PARTIES DE FORMES

Tableau 3.4 – Résultats de la recherche en utilisant l'approche du contexte de forme sur les 5 requêtes retenues du tableau 3.2. En 2^{ème} ligne : le taux de précision de la recherche des requêtes. Les lignes restantes sont organisées de la même manière que celles du tableau 3.3.

				
<i>PR=0%</i>	<i>PR=0%</i>	<i>PR=90%</i>	<i>PR=30%</i>	<i>PR=80%</i>
				
C.33, 0.1855	C.26, 0.0841	C.49, 0.0834	C.44, 0.1092	C.43, 0.1017
				
C.5, 0.1871	C.26, 0.0853	C.49, 0.1076	C.69, 0.1186	C.43, 0.1248
				
C.33, 0.198	C.26, 0.0864	C.49, 0.1148	C.69, 0.1239	C.43, 0.129
				
C.33, 0.1991	C.26, 0.087	C.49, 0.1173	C.6, 0.1334	C.43, 0.1314
				
C.5, 0.2042	C.26, 0.087	C.49, 0.1199	C.6, 0.1348	C.43, 0.1334
				
C.38, 0.2089	C.26, 0.0885	C.55, 0.1213	C.6, 0.1357	C.43, 0.1355
				
C.41, 0.211	C.26, 0.0898	C.49, 0.1236	C.44, 0.1359	C.43, 0.1374
				
C.41, 0.2118	C.26, 0.0902	C.49, 0.1238	C.44, 0.1359	C.51, 0.1374
				
C.33, 0.2142	C.26, 0.0904	C.49, 0.1314	C.6, 0.1364	C.43, 0.1381
				
C.33, 0.2145	C.26, 0.0905	C.49, 0.1329	C.6, 0.1367	C.51, 0.1393

**CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR
LES GÉODÉSIQUES DE FORMES**

Tableau 3.5 – Résultats de la recherche en utilisant la distance de Procruste sur les 5 requêtes retenues du tableau 3.2. En 2^{ème} ligne : le taux de précision de la recherche des requêtes. Les lignes restantes sont organisées de la même manière que celles du tableau 3.3.

				
PR=0%	PR=0%	PR=60%	PR=0%	PR=10%
				
C.66, 0.0754	C.60, 0.0074	C.49, 0.0164	C.69, 0.0067	C.31, 0.0056
				
C.58, 0.0865	C.34, 0.0081	C.49, 0.0248	C.69, 0.0069	C.51, 0.0059
				
C.69, 0.0898	C.34, 0.0089	C.49, 0.0261	C.69, 0.0121	C.7, 0.0067
				
C.1, 0.0961	C.26, 0.0089	C.49, 0.0287	C.17, 0.0137	C.56, 0.007
				
C.17, 0.0974	C.26, 0.0092	C.47, 0.0294	C.6, 0.0151	C.65, 0.0073
				
C.36, 0.0991	C.26, 0.0096	C.49, 0.0297	C.6, 0.0152	C.43, 0.0078
				
C.40, 0.1034	C.26, 0.0097	C.55, 0.0309	C.6, 0.0163	C.61, 0.0078
				
C.10, 0.1035	C.26, 0.0097	C.21, 0.0327	C.31, 0.0173	C.63, 0.0079
				
C.66, 0.1045	C.26, 0.0099	C.52, 0.035	C.31, 0.0174	C.63, 0.0079
				
C.40, 0.106	C.26, 0.01	C.49, 0.0356	C.6, 0.0175	C.51, 0.0079

requête. Pour la même requête de *Marteau*, l'approche Procuste n'a réussi à rechercher aucune partie de forme de la même classe. En revanche, notre méthode a correctement classé des parties même déformées et articulées de la requête parmi les plus similaires. D'où la grande sensibilité de la méthode du contexte de forme et de la méthode Procuste aux articulations et aux déformations élastiques. Même si la méthode du contexte de forme est aujourd'hui la plus utilisée dans la littérature pour la reconnaissance de parties de formes, elle est caractérisée par sa correspondance désordonnée des points.

Reste à souligner que dans le cas où la requête n'est pas assez représentative de sa classe, il pourrait y avoir un grand nombre de parties de formes qui la ressemblent. Par exemple, la requête partie de *Vache* (C.13) est peu significative car elle est commune à plusieurs formes d'animaux. Pour cette requête, la méthode du contexte de forme et la méthode Procuste échouent totalement à rechercher les parties de formes de la même classe qui la ressemblent ($PR=0\%$). Notre méthode géodésique donne cependant 40% de recherche correcte. Malgré cette diminution dans le taux de recherche, il apparaît clairement que dans le cas où notre approche n'arrive pas à reconnaître la requête partie de *Vache*, elle fait correspondre les parties de formes qui sont visuellement similaires (parties de formes d'autres animaux). Cependant, la méthode du contexte de forme et la méthode Procuste ne le font pas.

La classification : La méthode proposée a été également testée pour la classification de ces mêmes requêtes issues des images segmentées (tableau 3.2). Avec l'utilisation du classifieur 10 plus proches voisins, notre méthode classe correctement toutes les requêtes tandis que la méthode du contexte de forme a mal classé 3 requêtes : la partie de *Cheval* (C.48), la partie de *Vache* (C.13) et la partie de *Marteau* (C.44). Aussi, la méthode Procuste a réussi à classer seulement la requête de la partie de *Fer à cheval* (C.49). Par conséquent, notre méthode est aussi plus efficace que ces deux autres méthodes pour la classification de parties de formes.

3.4 Reconnaissance de formes partielles

Dans cette section, nous abordons le problème de la reconnaissance de formes partielles (bidimensionnelles). Ce problème se pose lorsque l'on ne dispose que d'un ensemble disjoint de parties de la forme à reconnaître.

3.4.1 Stratégie de reconnaissance proposée

Supposons que nous disposons, à partir d'une image segmentée, d'un ensemble de n_Q parties composant la forme partielle à reconnaître ; chaque partie étant représentée par N points d'échantillonnage. Nous avons donc un ensemble de requêtes $\{Q_r\}_{r=1, \dots, n_Q}$. Nous considérons que les parties disponibles ne recouvrent pas le contour entier de la forme. Dans un premier temps, chaque requête est comparée séparément aux parties de la base

CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR LES GÉODÉSQUES DE FORMES

construite des parties de formes étiquetées. Ensuite, nous combinons les informations issues de la reconnaissance des n_Q parties de forme pour reconnaître la forme entière. La stratégie de reconnaissance de formes partielles comprend les six étapes suivantes :

- Étape 1 : Pour chaque requête Q , trouver l'ensemble S des M parties de formes les plus similaires en utilisant la distance géodésique d_{Geod} à un niveau bas d'échantillonnage où les parties de formes sont représentées par L points ($L < N$).
- Étape 2 : Déterminer ensuite, pour chaque requête Q , l'ensemble \mathcal{C} des classes auxquelles appartiennent les éléments de S :

$$\mathcal{C} = \{C_l / \exists P_l \in S\}.$$

- Étape 3 : Identifier l'ensemble commun \mathcal{E} des classes candidates comme suit :

$$\mathcal{E} = \mathcal{C}_1 \cap \mathcal{C}_2 \cap \dots \cap \mathcal{C}_{n_Q}.$$

- Étape 4 : Comparer chaque requête Q au sous ensemble des parties de formes, $\{P_l / C_l \in \mathcal{E}\}$, en utilisant la distance géodésique d_{Geod} à un niveau d'échantillonnage plus élevé avec N points.
- Étape 5 : Pour chaque instance/forme complète (I_j, C_k) et sous la contrainte $C_k \in \mathcal{E}$, calculer la distance globale D à partir de l'ensemble des requêtes $\{Q_r\}$ de la façon suivante :

$$D(\{Q_r\}, (I_j, C_k)) = \sum_{r=1}^{n_Q} \min_{1 \leq i \leq n_P} d_{Geod}(Q_r, P_i), \quad (3.6)$$

où P_i désigne la $n^{ième}$ partie de l'individu (I_j, C_k) .

- Étape 6: La reconnaissance de la forme partielle donnée est effectuée enfin en utilisant la distance globale D (Eq. 3.6) via le critère des K plus proches voisins.

3.4.2 Évaluation depuis la base de formes MPEG-7

Afin de valider expérimentalement le système de reconnaissance de formes partielles proposé, nous utilisons la même base construite des parties de formes étiquetées décrite dans la section 3.3.1. Pour les tests, nous utilisons 69 paires de requêtes (Q_1, Q_2) ; une paire de chaque classe à l'exception de la classe de "Device-6". Les éléments de chaque paire sont des parties de la même forme sélectionnée parmi les formes numérotées de 11 à 20 de la base MPEG-7. Le premier élément (Q_1) est une requête du premier ensemble

3.4. RECONNAISSANCE DE FORMES PARTIELLES

$E1$ de petites requêtes (figure 3.9a), et le deuxième élément (Q_2) est une requête du deuxième ensemble $E2$ de petites requêtes donné par la figure 3.14 suivante. Les deux ensembles sont disposés dans le même ordre de sorte que les éléments (Q_1) et (Q_2) d'une paire de requêtes ont le même classement dans $E1$ et $E2$ respectivement.

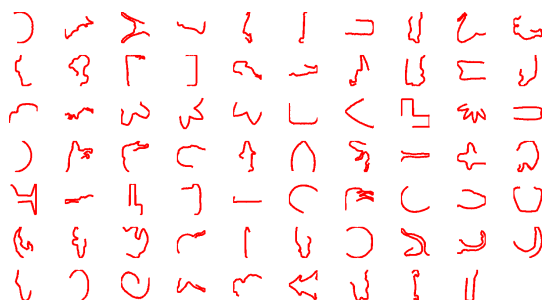


Figure 3.14 – Le deuxième ensemble ($E2$) de 69 petites requêtes extraites des mêmes formes utilisées pour l'obtention des requêtes de l'ensemble $E1$ (figure 3.9a).

Pour la première étape de notre algorithme décrit dans la section 3.4.1, le niveau bas d'échantillonnage adopté est de 50 points ($L = N/2 = 50$) ; la reconnaissance de formes partielles est basée sur les 20 parties de formes les plus similaires à chaque requête ($M = 20$). Pour chaque paire de requêtes (Q_1, Q_2), le taux de recherche de formes partielles basée sur une seule partie est mesuré comme le pourcentage des formes fermées appartenant à la même classe de Q_1 , respectivement de Q_2 , parmi les 10 formes fermées les plus proches. Globalement, plus la requête est une partie importante qui caractérise sa forme originale, plus le taux de recherche de la forme partielle est élevé. Mais lorsque la requête n'est pas si représentative, elle devient peu discriminante et beaucoup de formes fermées (de différentes classes) seront candidates pour elle. Afin d'améliorer les taux de réussite de recherche, nous appliquons notre stratégie de combinaison. Ainsi, chaque paire de requêtes est comparée aux formes de l'ensemble commun des classes candidates \mathcal{E} .

La recherche : Pour toutes les requêtes de $E1$, respectivement de $E2$, les taux de précision de recherche de formes partielles, basée sur une seule partie sont donnés par le tableau 3.6, pour notre méthode géodésiques, la méthode basée sur le contexte de forme et la méthode Procruste. Pour chaque paire de requêtes (Q_1, Q_2), le taux de précision de recherche surpasse les taux obtenus à partir de Q_1 et de Q_2 lorsque chaque requête est étudiée seule. Par exemple en utilisant notre distance géodésique, les taux moyens de précision de recherche des formes partielles à une seule partie atteignent 48.99% et 46.09% pour les requêtes de $E1$ et de $E2$ respectivement. La précision de recherche des formes partielles croit pour atteindre un taux moyen de 77.25% avec la combinaison des informations issues des deux éléments de chaque paire de requêtes. Ce taux moyen de précision de recherche diminue à 67.97% et 48.84% lors de l'utilisation de la distance basée sur le contexte de forme et la distance Procruste respectivement dans notre système

**CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR
LES GÉODÉSQUES DE FORMES**

de recherche des formes partielles.

Tableau 3.6 – Taux de précision de recherche des formes partielles en utilisant notre méthode géodésique, la méthode basée sur le contexte de forme et la méthode Procuste. Nous donnons le taux moyen de précision de recherche basée sur une seule partie (les requêtes de $E1$ ou de $E2$) et celui de recherche basée sur la combinaison de deux parties (de $E1$ et de $E2$).

Taux de recherche	Approche géodésique			Approche du contexte de forme			Approche Procuste		
	Q_1	Q_2	(Q_1, Q_2)	Q_1	Q_2	(Q_1, Q_2)	Q_1	Q_2	(Q_1, Q_2)
	48.99%	46.09%	77.25%	41.45%	41.59%	67.97%	36.38%	36.81%	48.84%

Tableau 3.7 – Taux de classification des formes partielles en utilisant notre méthode géodésique, la méthode basée sur le contexte de forme et la méthode Procuste avec le classifieur K - NN . Nous donnons le taux de classification correcte des formes partielles à une seule partie (les requêtes de $E1$ ou de $E2$) et celui des formes partielles à deux parties par combinaison de parties de $E1$ et de $E2$.

Taux de classification	Approche géodésique			Approche du contexte de forme			Approche Procuste		
	Q_1	Q_2	(Q_1, Q_2)	Q_1	Q_2	(Q_1, Q_2)	Q_1	Q_2	(Q_1, Q_2)
1 - NN	84.06%	78.26%	91.30%	84.06%	76.81%	91.30%	73.91%	72.46%	82.61%
5 - NN	68.12%	76.81%	91.30%	72.46%	65.22%	84.06%	63.77%	68.12%	72.46%
10 - NN	63.77%	65.22%	88.44%	63.77%	57.97%	81.16%	60.87%	62.32%	63.77%

La classification : Notre système de reconnaissance de formes partielles est également évalué pour la classification des formes partielles sur les mêmes 69 paires de requêtes précédemment utilisées pour la recherche. Avec la classification K - NN , nous attribuons à chaque forme partielle la classe majoritaire parmi les K formes fermées les plus proches. En utilisant notre distance géodésique, la distance basée sur le contexte de forme, ou la distance Procuste, les taux moyens de classification des formes partielles à deux parties dépassent toujours les taux moyens de classification des formes partielles à une seule partie de forme (tableau 3.7). Avec le classifieur 1 - NN , le taux moyen de classification des formes partielles monte à 91.30% avec notre approche géodésique et avec l'approche du contexte de forme, et il diminue à 82.61% avec l'approche Procuste. Une comparaison plus approfondie des résultats du tableau 3.7 montre bien l'efficacité de notre stratégie de reconnaissance des formes partielles lors de l'utilisation de la distance géodésique avec les classifieurs 5 - NN et 10 - NN .

3.4.3 Évaluation sur des images réelles

Afin de valider notre approche de reconnaissance, nous évaluons ses performances sur quelques images réelles. Pour ce faire, nous utilisons les cinq images données en première colonne du tableau 3.8. De la carte de contours, nous considérons les deux

3.4. RECONNAISSANCE DE FORMES PARTIELLES

parties les plus importantes du contour comme forme partielle à reconnaître (ces requêtes sont illustrées en troisième colonne du tableau 3.8).

La recherche : Les résultats de recherche des formes partielles issues des images réelles sont illustrés dans les tableaux 3.9, 3.10, et 3.11 respectivement pour l'approche géodésique, l'approche de contexte de forme et l'approche Procuste. Ces résultats montrent également une amélioration des performances de recherche lors de la combinaison des requêtes : dans la plupart des cas, le taux de recherche des requêtes combinées (R) surpasse le taux de recherche de chaque requête considérée seule (R_1 ou R_2). L'approche basée sur le contexte de forme échoue à rechercher correctement les requêtes du *Papillon* et de la *Vache* ($R = 0\%$). L'approche Procuste échoue à rechercher correctement la requête de la *Guitare*, de la *Vache* et du *Papillon* ($R = 0\%$) tandis que notre méthode géodésique réussit à les retrouver avec respectivement des taux de 100%, 70% et 60% de recherche correcte.

Au niveau de la correspondance des parties de formes qui composent les formes partielles, l'approche géodésique établit une bonne correspondance des parties dans la plupart des cas (tableau 3.9). Il faut noter que même une reconnaissance correcte d'une forme partielle peut être accompagnée d'une correspondance non optimale des parties, en particulier lorsque la forme partielle étudiée renferme des parties symétriques comme dans le cas des formes de *Papillon* (en première colonne du tableau 3.9) et de *Guitare* (en cinquième colonne du tableau 3.9). À partir des tableaux 3.9, 3.10 et 3.11, on peut aussi remarquer que dans le cas où notre approche géodésique n'arrive pas à reconnaître une paire de requêtes donnée, elle fait correspondre des parties de formes qui sont visuellement similaires tandis que les méthodes contexte de forme et Procuste ne le font pas.



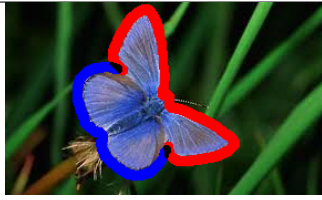

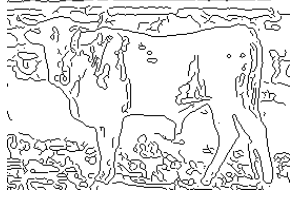
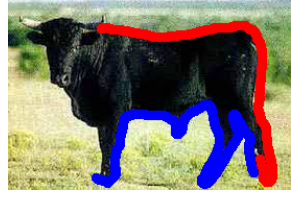


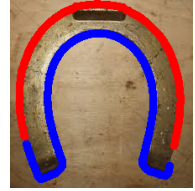

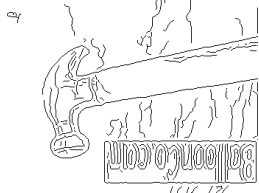
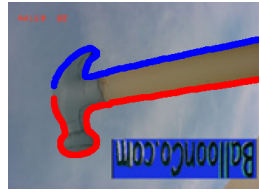

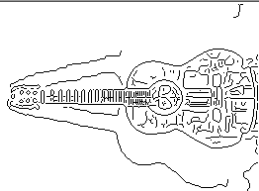

La classification : Le système proposé est également évalué pour la classification de formes partielles sur les mêmes paires précédemment utilisées pour la recherche de formes (tableau 3.8). Dans le tableau 3.12, les résultats montrent que seule notre méthode aboutit toujours à une classification correcte ($I-NN$). Par exemple, la méthode basée sur le contexte de forme attribue de fausses classes pour les requêtes de *Papillon* et de *Vache*, alors que la méthode Procuste ne parvient pas à classer correctement les requêtes de *Papillon*, de *Vache* et de *Guitare*. Par conséquent, notre méthode géodésique surpasse la méthode du contexte de forme et la méthode Procuste en reconnaissance de formes partielles. Nous pouvons en déduire la grande capacité de notre méthode géodésique à reconnaître des formes partielles qui présentent des déformations élastiques.

Comparaison à une autre stratégie de combinaison

Dans la littérature, la reconnaissance de formes partielles s'appuyant sur une stratégie de combinaison a été proposée par Tanase et Veltkamp [231] avec une approche de programmation dynamique appliquée directement sur la fonction d'angle du contour de
























































*CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR
LES GÉODÉSQUES DE FORMES*

Tableau 3.8 – Exemples d’images réelles (en première colonne) et leurs cartes de contours (en deuxième colonne). Dans la troisième colonne, nous donnons les paires de parties de contours les plus significatives (marquées en rouge et en bleu), retenues pour la reconnaissance des formes dans les images segmentées.

Image d’entrée	Carte de contours	Parties de contour retenues
		
		
		
		
		
























































3.4. RECONNAISSANCE DE FORMES PARTIELLES

Tableau 3.9 – Résultats de la recherche sur des images réelles en utilisant notre méthode géodésique. En première ligne : les paires de parties de contour retenues du tableau 3.8. En deuxième ligne : la classe correspondante. En troisième ligne : les taux de recherche R_1 et R_2 de Q_1 et Q_2 respectivement. En quatrième ligne : le taux de recherche R des requêtes combinées (Q_1, Q_2). Dans les lignes restantes, nous donnons les 10 formes recherchées les plus similaires aux requêtes combinées (Q_1, Q_2). (C.x, D) : x est le numéro de la classe de la base MPEG-7 et D est la distance globale attribuée à la requête (équation 3.6).

				
<i>Papillon, C.9</i>	<i>Vache, C.13</i>	<i>Fer à cheval, C.49</i>	<i>Marteau, C.44</i>	<i>Guitare, C.43</i>
$R_1 = 10\%, R_2 = 0\%$	$R_1 = 30\%, R_2 = 30\%$	$R_1 = 30\%, R_2 = 70\%$	$R_1 = 70\%, R_2 = 70\%$	$R_1 = 80\%, R_2 = 0\%$
$R = 60\%$	$R = 70\%$	$R = 100\%$	$R = 70\%$	$R = 100\%$
				
C.9, 0.4593	C.13, 0.4532	C.49, 0.1549	C.44, 0.3063	C.43, 0.4748
				
C.9, 0.4753	C.13, 0.4845	C.49, 0.1836	C.44, 0.3063	C.43, 0.4805
				
C.9, 0.4989	C.13, 0.4850	C.49, 0.1876	C.44, 0.3645	C.43, 0.5059
				
C.26, 0.5067	C.13, 0.5259	C.49, 0.2030	C.44, 0.3652	C.43, 0.5081
				
C.26, 0.5099	C.13, 0.5297	C.49, 0.2455	C.44, 0.33652	C.43, 0.5085
				
C.26, 0.5119	C.22, 0.5327	C.49, 0.2667	C.44, 0.3652	C.43, 0.5160
				
C.9, 0.5189	C.22, 0.5379	C.49, 0.2674	C.44, 0.3652	C.43, 0.5426
				
C.9, 0.5243	C.13, 0.5594	C.49, 0.2708	C.6, 0.4154	C.43, 0.5632
				
C.26, 0.5345	C.13, 0.5623	C.49, 0.3087	C.6, 0.4180	C.43, 0.5757
				
C.9, 0.5477	C.22, 0.5676	C.49, 0.3106	C.6, 0.4258	C.43, 0.5790
























































*CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR
LES GÉODÉSQUES DE FORMES*

Tableau 3.10 – Résultats de la recherche sur les mêmes paires de requêtes retenues du tableau 3.8 en utilisant l’approche du contexte de forme. Les lignes sont organisées de la même manière que celles du tableau 3.9.

				
<i>Papillon, C.9</i>	<i>Vache, C.13</i>	<i>Fer à cheval, C.49</i>	<i>Marteau, C.44</i>	<i>Guitare, C.43</i>
$R_1 = 0\%, R_2 = 0\%$	$R_1 = 20\%, R_2 = 0\%$	$R_1 = 30\%, R_2 = 70\%$	$R_1 = 30\%, R_2 = 30\%$	$R_1 = 50\%, R_2 = 0\%$
$R = 0\%$	$R = 0\%$	$R = 100\%$	$R = 30\%$	$R = 50\%$
				
C.26, 0.4643	C.34, 0.2997	C.49, 0.1485	C.44, 0.2451	C.43, 0.2577
				
C.26, 0.4900	C.34, 0.3228	C.49, 0.1538	C.6, 0.2796	C.43, 0.3030
				
C.26, 0.5098	C.34, 0.3451	C.49, 0.1650	C.44, 0.2948	C.62, 0.3124
				
C.26, 0.5167	C.34, 0.3773	C.49, 0.1655	C.44, 0.2948	C.43, 0.3281
				
C.26, 0.6194	C.34, 0.4006	C.49, 0.1737	C.6, 0.2956	C.36, 0.3300
				
C.26, 0.6827	C.34, 0.4224	C.49, 0.1737	C.6, 0.2986	C.62, 0.3399
				
C.26, 0.6888	C.34, 0.4530	C.49, 0.1801	C.6, 0.3040	C.62, 0.3456
				
C.26, 0.7069	C.34, 0.4589	C.49, 0.1884	C.69, 0.3074	C.43, 0.3481
				
C.26, 0.7114	C.34, 0.4599	C.49, 0.2531	C.6, 0.3096	C.43, 0.3483
				
C.26, 0.7221	C.34, 0.4940	C.49, 0.2607	C.6, 0.3112	C.62, 0.3513


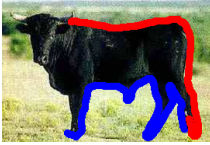



3.4. RECONNAISSANCE DE FORMES PARTIELLES

Tableau 3.11 – Résultats de la recherche sur les mêmes paires de requêtes retenues du tableau 3.8 en utilisant l’approche Procuste. Les lignes sont organisées de la même manière que celles du tableau 3.9.

				
<i>Papillon, C.9</i>	<i>Vache, C.13</i>	<i>Fer à cheval, C.49</i>	<i>Marteau, C.44</i>	<i>Guitare, C.43</i>
$R_1 = 0\%, R_2 = 0\%$	$R_1 = 30\%, R_2 = 0\%$	$R_1 = 20\%, R_2 = 60\%$	$R_1 = 0\%, R_2 = 50\%$	$R_1 = 10\%, R_2 = 0\%$
$R = 0\%$	$R = 0\%$	$R = 80\%$	$R = 50\%$	$R = 0\%$
				
C.26, 0.0674	C.60, 0.1053	C.49, 0.0214	C.44, 0.0300	C.62, 0.0198
				
C.26, 0.0719	C.60, 0.1092	C.49, 0.0275	C.44, 0.0300	C.62, 0.0299
				
C.26, 0.0725	C.60, 0.1323	C.49, 0.0300	C.44, 0.0300	C.62, 0.0341
				
C.26, 0.0776	C.60, 0.1515	C.49, 0.0300	C.44, 0.0300	C.62, 0.0343
				
C.26, 0.1495	C.60, 0.1680	C.49, 0.0316	C.6, 0.0380	C.62, 0.0420
				
C.26, 0.2168	C.60, 0.1700	C.47, 0.0413	C.6, 0.0383	C.62, 0.0444
				
C.26, 0.2284	C.60, 0.1745	C.49, 0.0415	C.6, 0.0383	C.62, 0.0495
				
C.26, 0.3179	C.60, 0.1745	C.49, 0.0480	C.44, 0.0400	C.62, 0.0577
				
C.26, 0.3320	C.60, 0.1796	C.47, 0.0560	C.6, 0.0410	C.62, 0.0612
				
C.26, 0.0333	C.60, 0.2256	C.49, 0.0570	C.6, 0.0426	C.62, 0.0860

*CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR
LES GÉODÉSIQUES DE FORMES*

Tableau 3.12 – Résultats de classification sur les mêmes paires de requêtes retenues du tableau 3.8. Nous donnons la classe assignée par chacune des trois méthodes utilisées : géodésique, contexte de forme et Procuste.






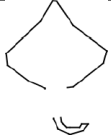

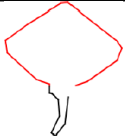






parties de contour	Classe assignée (<i>I-NN</i>)		
	<i>Approche géodésique</i>	<i>Approche du contexte de forme</i>	<i>Approche Procuste</i>
	→ <i>Papillon (C.9)</i>	→ <i>Dispositif-3 (C.26)</i>	→ <i>Dispositif-3 (C.26)</i>
	→ <i>Vache (C.13)</i>	→ <i>Éléphant (C.34)</i>	→ <i>Raie (C.60)</i>
	→ <i>Fer à cheval (C.49)</i>	→ <i>Fer à cheval (C.49)</i>	→ <i>Fer à cheval (C.49)</i>
	→ <i>Guitare (C.43)</i>	→ <i>Guitare (C.43)</i>	→ <i>Marteau (C.44)</i>
	→ <i>Fer à cheval (C.49)</i>	→ <i>Fer à cheval (C.49)</i>	→ <i>Chaussure (C.62)</i>

3.4. RECONNAISSANCE DE FORMES PARTIELLES





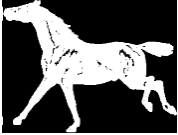


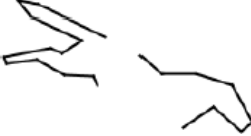






l'objet. Cette approche est connue par l'approche des fonctions d'angle multiple (MTA: *The Multiple Turning Angle*). La distance quadratique définie par l'intégrale des différences est utilisée comme mesure de similarité lors de l'établissement de la correspondance entre les formes partielles. Ainsi, nous testons les performances de notre schéma de reconnaissance de formes partielles sur la même application de recherche de formes partielles donnée dans [231] afin de mettre les deux schémas dans un cadre comparatif. Dans cette application, toutes les formes de la base de données MPEG-7 sont incluses dans la construction de la base des parties de formes étiquetées. Pour appliquer notre système de reconnaissance, nous avons utilisé la méthode de construction décrite dans la section 3.3.1 pour construire une nouvelle base de parties de formes étiquetées à partir des 20 formes de chaque classe pour les 70 classes de la base MPEG-7. Par conséquent, nous avons étudié les mêmes requêtes utilisées dans les travaux de Tanase et Veltkamp [231]. Les images utilisées et les parties requêtes des formes partielles sont illustrées en première et en deuxième colonnes du tableau 3.13. Les taux de précision de recherche calculés sur les 20 premières formes recherchées (les plus similaires à la requête) sont rapportés dans les deux dernières colonnes du même tableau. Notre approche géodésique surpasse la méthode quadratique proposée par Tanase et Veltkamp sur les 13 premières images requêtes, mais elle échoue sur la dernière (forme du *Chien-11*). Ici, les parties requêtes ne sont pas suffisantes pour caractériser la classe de *Chien* ; ces parties sont communes à de nombreuses formes d'autres classes. Par exemple, les parties qui correspondent aux membres de chien peuvent également être trouvés dans les formes des autres classes, en particulier dans les classes de *Cheval*, de *Cerf*, et de *Vache*. Dans la figure 3.15, nous montrons les 20 formes les plus similaires aux parties requêtes de la forme du *Chien-11* recherchées par notre méthode géodésique. Parmi les formes recherchées, nous trouvons quelques exemples de la classe de *Cheval*. Ceci est conforme à la comparaison visuelle car les parties requêtes sont visuellement très similaires aux parties des formes de *Cheval*. Les sous-performances de l'approche MTA peuvent être en partie dues à la non-invariance à l'échelle et sa grande sensibilité aux distorsions. En revanche, l'invariance à l'échelle de notre méthode n'a été défavorable que dans le cas de *Chien/Cheval*.

*CHAPITRE 3. RECONNAISSANCE DE PARTIES DE FORMES BASÉE SUR
LES GÉODÉSIQUES DE FORMES*

Tableau 3.13 – Comparaison des performances de recherche de formes partielles entre notre approche géodésique et l’approche basée sur les fonctions d’angle MTA [231]. Les formes partielles requêtes sont celles utilisées dans [231]. Les taux de recherche obtenus avec les deux méthodes sont donnés en troisième et quatrième colonnes du tableau.

Image requête	Parties requêtes	Taux de recherche des formes partielles	
		<i>Schéma de MTA</i>	<i>Notre schéma de combinaison</i>
 <i>Insecte-20 (C.3)</i>		55%	100%
 <i>Insecte-10 (C.3)</i>		55%	100%
 <i>Raie-3 (C.60)</i>		60%	100%
 <i>Raie-17 (C.60)</i>		40%	100%
 <i>Cerf-15 (C.22)</i>		35%	100%
 <i>Cerf-13 (C.22)</i>		40%	90%
 <i>Oiseau-11 (C.5)</i>		35%	80%

3.4. RECONNAISSANCE DE FORMES PARTIELLES

 <i>Oiseau-17 (C.5)</i>		40%	100%
 <i>Oiseau-9 (C.5)</i>		40%	80%
 <i>Cheval-17 (C.48)</i>		60%	70%
 <i>Cheval-3 (C.48)</i>		45%	55%
 <i>Couronne-13 (C.20)</i>		40%	75%
 <i>Papillon-11 (C.9)</i>		50%	100%
 <i>Chien-11 (C.33)</i>		45%	20%

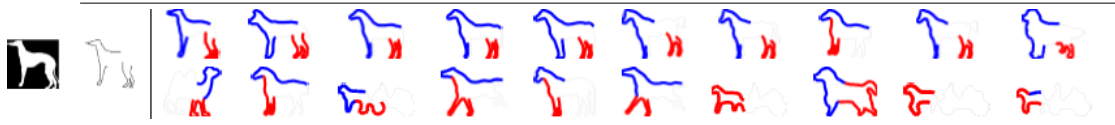


Figure 3.15 – Les 20 formes les plus similaires à la paire des parties requêtes de la forme *Chien-11* (dernière ligne du tableau 3.13) recherchées par notre schéma de combinaison basée sur les géodésiques.

Temps d'exécution : Le temps moyen d'exécution est un critère important d'évaluation de tout système de reconnaissance. Dans ce travail, l'implémentation est réalisée sur une machine Intel Core i5-2400 CPU @ 3,10 GHz en utilisant Matlab R2010b. En ce qui concerne la complexité de calcul des trois systèmes étudiés pour la reconnaissance des parties de formes, le système basé sur l'analyse procustéenne est le plus efficace en termes de calcul. Par exemple, la comparaison de deux parties de forme échantillonnées avec 100 points nécessite approximativement 0.00023, 0.2855 et 1.2039 secondes avec les méthodes Procuste, du contexte de forme et géodésique respectivement. Grâce à la stratégie de sous-sélection proposée, le temps de calcul nécessaire pour comparer une requête à toutes les instances de la base des parties de formes étiquetée est réduit à la moitié. Cependant, dans l'étape actuelle de développement, la vitesse de l'implémentation de l'algorithme géodésique proposé n'est pas encore optimisé et le temps de fonctionnement peut être amélioré.

3.5 Conclusion

Dans ce chapitre, nous avons proposé un schéma de recalage robuste de courbes ouvertes 2D en introduisant un estimateur robuste pour que le processus de recalage soit indépendant aux données aberrantes. Cette approche consiste à minimiser un critère de similarité issu de l'analyse des géodésiques de formes. Par la suite, nous avons exploité le résidu du recalage pour définir une nouvelle métrique pour la reconnaissance de parties de formes. Nous avons décrit notre stratégie de sous-sélection et la méthode de construction de la base des parties de formes étiquetées. Nous avons également proposé une stratégie de combinaison avec cette même métrique pour la reconnaissance des formes partielles. Nous avons montré l'efficacité de cette métrique par différents tests de reconnaissance (classification et recherche). Les tests ont été effectués sur des parties requêtes de la base d'images MPEG-7 et sur des images réelles segmentées. Nous avons comparé notre système de reconnaissance de parties de formes avec la méthode basée sur le contexte de forme et avec la méthode Procuste, bien connues dans la littérature pour leur utilisation dans la reconnaissance de parties de formes. Ces résultats expérimentaux ont montré la supériorité de notre méthode grâce à sa grande capacité à gérer les articulations et les déformations élastiques par rapport aux autres méthodes de

3.5. CONCLUSION

l'état de l'art. Nous avons aussi démontré expérimentalement que notre stratégie de reconnaissance de formes partielles donne de meilleurs résultats avec notre distance géodésique tandis que ses performances diminuent avec la distance basée sur le contexte de forme et la distance Procuste. Enfin, cette stratégie a été favorablement comparée avec la méthode *MTA* de Tanase et Veltkamp pour la recherche de formes partielles. Dans la plupart des cas, notre approche fonctionne mieux. Dans le chapitre suivant, nous avons étudié le problème de l'orientation des parties de formes. Une nouvelle méthode a été proposée et ainsi comparé aux méthodes de l'état de l'art les plus utilisées pour assurer l'invariance à la rotation des parties de formes.

Chapitre 4

Invariance à la rotation des parties de formes

Sommaire

4.1	Introduction	120
4.2	Méthode proposée	121
4.3	Critères proposés pour la comparaison des méthodes d'orientation	122
4.3.1	Critère basé sur la correspondance	122
4.3.2	Critère basé sur la distance géodésique	125
4.4	Comparaison des méthodes d'orientation	126
4.4.1	Évaluation sur des paires de parties de formes identiques	127
4.4.2	Évaluation sur des parties de formes synthétiques	131
4.4.3	Évaluation sur des parties de formes de la base des formes MPEG-7	142
4.5	Conclusion	145

4.1 Introduction

L'orientation est l'un des éléments de synthèse de la base de la géométrie d'une forme ou d'une partie de forme. Pour garantir l'invariance par rotation, il est nécessaire de convertir une forme orientée arbitrairement dans une orientation commune unique.

Les différentes méthodes de la décomposition des formes adoptées dans la plupart des travaux qui s'intéressent à analyser et à étudier les parties de formes sont basées sur l'utilisation des points caractéristiques (minima locaux de courbures négatives, maxima locaux de courbures positives, etc.) pour décomposer toute forme entière donnée. Une petite articulation ou déformation de la forme d'origine à décomposer peut être à l'origine d'un déplacement significatif des points caractéristiques qui affectent à leur tour les parties de formes résultantes de la décomposition. De même, la méthode de la décomposition DCE adoptée dans ce travail est aussi basée sur la recherche des points caractéristiques des formes à décomposer. Par conséquent, une telle décomposition de formes visuellement similaires produit naturellement des parties de formes similaires ayant souvent des perturbations d'extrémités (extrémités de différentes longueurs et de différentes orientations). Par exemple, la figure 4.1 montre deux parties de deux instances différentes de la forme d'une chauve-souris (*Bat-3* et *Bat-6*) trouvées par la technique de la décomposition DCE des formes d'origine (courbes pointillées noires en figure 4.1a et figure 4.1b). Il est visuellement clair que l'extrémité gauche de la partie de la chauve-souris en figure 4.1a (courbe bleue) est plus longue que l'extrémité gauche de la partie de la chauve-souris en figure 4.1b (courbe pointillée rouge). Dans une telle situation, les points extrémités trouvés ne se correspondent pas dans la réalité. D'où la nécessité de développer une méthode d'orientation qui fait correspondre au mieux les points extrémités des courbes avant de comparer les deux parties de formes données.



Figure 4.1 – Les deux parties (courbes en gras) de deux différentes chauves-souris (*Bat-3* en (a) et *Bat-6* en (b)), extraites par l'utilisation de la technique de la décomposition DCE, sont visuellement très similaires. Une petite perturbation de longueur existe au niveau des extrémités gauches des deux parties de formes.

Dans ce chapitre, nous commençons par décrire en section 4.2 la méthode proposée

pour assurer l'invariance à la rotation des parties de formes. En section 4.3, nous présentons les critères proposés pour comparer les méthodes d'orientation les plus utilisées dans la littérature à notre méthode proposée. Dans la section 4.4, nous présentons une comparaison des méthodes d'orientation avec des tests comparatifs établis sur des parties de formes identiques, synthétiques et sur des parties de formes de la base des formes MPEG-7.

4.2 Méthode proposée

Afin de résoudre le problème de l'orientation des parties de formes, disant en présence surtout des perturbations aux extrémités, nous avons proposé un nouvel algorithme basé sur l'alignement Procuste et la redétermination des points extrémités.

Description : Dans cet algorithme, nous nous basons sur le fait que les deux parties de formes à comparer sont globalement similaires. Nous nous intéressons maintenant à la différence sur les extrémités et les transformations rigides seulement. Par exemple, une des deux parties de formes peut avoir au moins une extrémité de différente longueur ou de différente direction que celle de l'autre partie de forme (voir la figure 4.1). Donc, il est évident de supposer que les points extrémités des deux parties de formes à comparer ne se correspondent pas nécessairement. Pour cela, afin de faire face au problème des extrémités, nous proposons de commencer par une étape de recalage préliminaire pour déterminer les nouveaux points extrémités. Ce recalage est relativement compliqué pour la tâche de redétermination des extrémités, d'où l'étape préliminaire par le recalage rigide issu d'une analyse procustéenne (AP) pour l'alignement des parties de formes. Ce choix est motivé par la faible complexité de l'implémentation et l'efficacité de l'algorithme efficace de l'alignement Procuste. Dans cet algorithme, nous proposons de retrouver les points extrémités des parties de formes initialement identifiés en leur permettant de coulisser librement le long du contour de la partie de forme dans une fenêtre à proximité des points extrémités originaux. Les nouveaux points extrémités définissent les parties de formes qui présentent le recalage optimal résultant d'un alignement Procuste des parties de formes à comparer (voir Fig. 4.2).

Étant données deux courbes Q et P représentées par N points échantillonnés, nous cherchons les nouvelles extrémités en leur permettant de glisser librement le long du contour dans $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$ à proximité des points extrémités d'origine afin de minimiser l'erreur quadratique ou simplement dit la distance Procuste D_{Pr} (Eq. 1.21). Soit $W < N$ la taille de la fenêtre coulissante, les sous-ensembles des indices des points $\mathcal{F}_1 = \{1, \dots, W\}$ et $\mathcal{F}_2 = \{N + 1 - W, \dots, N\}$ représentent les fenêtres au voisinage des deux extrémités de chaque partie de forme d'origine.

Nous notons que nous déterminons les nouveaux points extrémités uniquement pour l'une des parties de formes, tandis que l'autre partie reste fixe avec ses extrémités initiaux. Par conséquent, l'alignement des deux parties de formes Q et P par notre

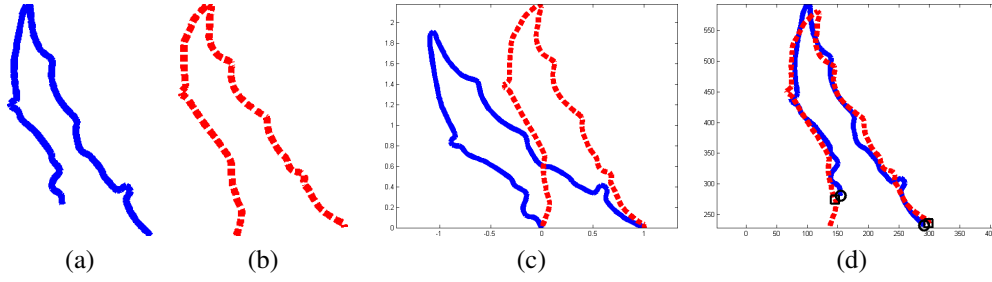


Figure 4.2 – Deux parties de formes de chauves-souris (a) et (b) (mêmes parties de la figure 4.1) et leur représentation invariante par la méthode du repère canonique en (c) et par notre méthode en (d).

méthode proposée est donnée par l'équation suivante:

$$(a, t, \theta, e)^* = \underset{(a, t, \theta, e) \in \mathbb{R}^+ \times \mathbb{R}^2 \times [0, 2\pi[\times \mathcal{F}}{\text{argmin}} \{D_{Pr}(Q_e, P), D_{Pr}(Q, P_e)\}, \quad (4.1)$$

où Q_e (respectivement P_e) est la nouvelle partie de forme Q (respectivement P) définie avec les nouveaux points extrémités $(q_{i_1}; q_{i_2})$ (respectivement $(p_{i_1}; p_{i_2})$) et $e = (i_1; i_2) \in F$ sont les indices des points de référence sur la partie de forme originale Q (respectivement P). Par exemple, dans la figure 4.2d, les nouveaux points extrémités des parties de formes sont marqués par des cercles noirs pour la partie de forme en bleu et par des carrés noirs pour la partie de forme en rouge.

4.3 Critères proposés pour la comparaison des méthodes d'orientation

Dans cette étude comparative, nous avons proposé deux critères afin de comparer quantitativement et qualitativement la performance de chacune des quatre méthodes d'orientation des parties de formes précédemment décrites et leur stabilité aux perturbations d'extrémité. Le premier critère est basé sur la fonction de la correspondance optimale obtenue par un recalage robuste établi entre deux parties de formes données via notre algorithme proposé. Le résidu de ce recalage est exploité pour calculer la distance géodésique qui constitue le deuxième critère évaluatif.

4.3.1 Critère basé sur la correspondance

Afin d'évaluer la stabilité des méthodes d'orientation en présence des perturbations d'extrémité, nous proposons d'utiliser l'erreur quadratique entre la correspondance

4.3. CRITÈRES PROPOSÉS POUR LA COMPARAISON DES MÉTHODES D'ORIENTATION

imposée et la correspondance estimée entre deux parties de formes synthétiques (une partie de contour est obtenue en appliquant une transformation connue à l'autre).

Premièrement, nous désignons par la correspondance imposée la transformation $\phi_{imposée}$ appliquée à la partie de forme d'origine Q , paramétrée comme étant un signal 1D, pour obtenir une partie de forme P . Pour contrôler et étudier l'influence de toute perturbation apportée à l'extrémité de parties de formes sur la performance des méthodes d'orientation, nous proposons d'utiliser une fonction de transformation qui garde les extrémités de la partie de forme d'origine. Par exemple, la courbe pointillée rouge de la figure 4.3b représente la transformation utilisée dans notre étude. L'identité aux extrémités et la fonction sigmoïde ailleurs conduit à transformer la partie de forme Q (figure 4.3a) à une partie de forme P (figure 4.3c) tout en gardant les mêmes extrémités. L'utilisation d'une telle fonction de transformation nous donne la possibilité d'avoir des parties de formes différentes dépourvues des problèmes aux extrémités. Ceci va nous aider à ajouter librement des perturbations aux extrémités afin d'interpréter le comportement de toute méthode d'orientation étudiée en présence de tout changement apporté aux extrémités.

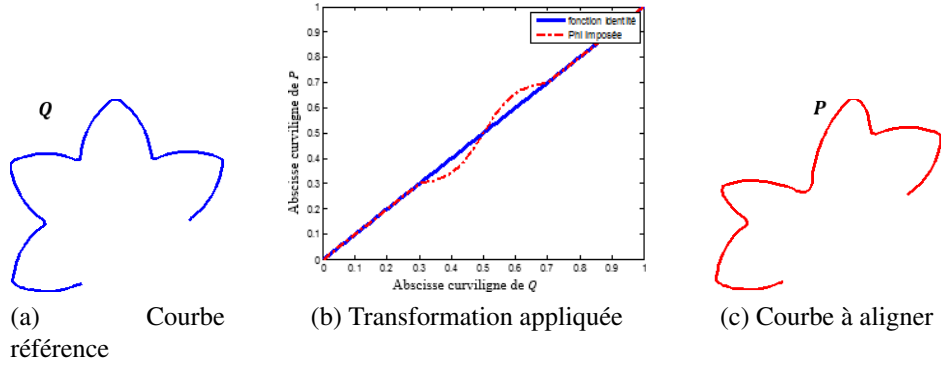


Figure 4.3 – La courbe ouverte P en (c) est obtenue en appliquant la transformation connue, donnée par la courbe rouge pointillée en (b), sur la courbe référence Q en (a). Les abscisses curvilignes de Q et P sont normalisées

Deuxièmement, la différence d'orientation entre toute paire de parties de formes (Q,P) est éliminée en utilisant chacune des quatre méthodes d'orientation. Ensuite, un recalage robuste entre toute paire de parties de formes (\tilde{Q}, \tilde{P}) bien orientées est établi. Plus précisément, nous cherchons la fonction de correspondance optimale $\phi_{estimée}$ entre \tilde{Q} et \tilde{P} , qui fait correspondre les points de caractéristiques similaires. Cette correspondance entre les abscisses curvilignes normalisées de \tilde{Q} et \tilde{P} est donnée par l'équation suivante:

$$\phi^* = \phi_{estimée} = \arg \min_{\phi} \left(\arccos \int_0^1 \sqrt{\phi_s(s)} \left| \cos \frac{\|\theta_Q(s) - \theta_P(\phi(s))\|_p}{2} \right| ds \right),$$

où les courbes \tilde{Q} et \tilde{P} sont représentées par leurs fonctions tangentes θ_Q et θ_P respectivement. Ces fonctions sont paramétrées par l'abscisse curviligne normalisée à l'unité. L'estimateur robuste de Leclerc $\|\cdot\|_\rho$ est utilisé pour pénaliser des valeurs relativement aberrantes. Le terme $\sqrt{\phi_s(s)}$ pénalise la torsion et l'étirement tout au long des courbes et $\phi_s(s) = \frac{d\phi(s)}{ds}$. Une fois $\phi_{estimée}$ est obtenue, le critère d'évaluation proposé vaut l'erreur quadratique moyenne $EQM_\phi = E(|\phi_{imposée} - \phi_{estimée}|^2)$ entre la fonction de transformation appliquée et celle trouvée par notre algorithme robuste de correspondance.

Difficulté et limitation

–*Repère commun* : La fonction de correspondance entre toute paire de parties de formes (\tilde{Q}, \tilde{P}) n'est pas toujours trouvée dans le bon repère. Il faut donc la ramener au repère convenable pour calculer correctement l'erreur quadratique EQM_ϕ . Les méthodes d'orientation de l'alignement Procuste, de l'alignement angulaire, et de la représentation dans un repère invariant exigent la correspondance des points extrémités des parties de formes. Tandis que notre méthode d'orientation proposée consiste à redéterminer des points extrémités des parties de formes à un voisinage choisi de tout point extrémité d'origine. Les points éliminés sont exclus de toute correspondance établie des points. La recherche de la correspondance optimale entre les points de deux parties de formes nécessite une re-normalisation des abscisses curvilignes de la partie de forme redéfinie. Dans le cas où les nouveaux points extrémités trouvés se différencient de ceux de la partie d'origine, la fonction de correspondance trouvée n'appartient pas au même repère de celui de la correspondance imposée. Donc, une étude comparative nécessite de mettre les fonctions de correspondance dans le même repère.

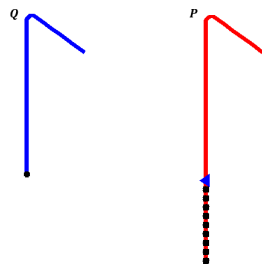


Figure 4.4 – À la base de la fonction angulaire, la correspondance optimale peut correspondre le point extrémité (en noir) de la courbe Q à l'un des points marqués en noir et en bleu de la courbe P . Alors que, la vérité terrain exige de correspondre le point extrémité (en noir) de la courbe Q au point marqué par un triangle bleu de la courbe P .

–*Points caractéristiques bornes* : La correspondance optimale représentée par la fonction $\phi_{estimée}$ établie la correspondance entre les points des parties de formes ayant

4.3. CRITÈRES PROPOSÉS POUR LA COMPARAISON DES MÉTHODES D'ORIENTATION

des caractéristiques angulaires similaires. Et, tout problème de correspondance à proximité des points extrémités peut aboutir à un calcul erroné de l'erreur EQM_ϕ . En fait, un point extrémité d'une partie de forme peut avoir une fonction d'angle approximativement égale à celle de plus qu'un point situant à proximité du point extrémité de l'autre partie de forme. Souvent dans une telle situation, la correspondance optimale des points au voisinage des points extrémités ne reflète pas la vérité terrain. Par exemple dans la figure 4.4, le point extrémité de la courbe Q (marqué en noir) a une fonction d'angle égale à celle de l'ensemble des points de l'extrémité de P marqués en noir et en bleu. Donc, la correspondance optimale peut correspondre le point extrémité de la courbe Q à un point de l'ensemble des points situé à proximité du point extrémité. Cependant, la vérité terrain suppose que la correspondance optimale doit correspondre le point extrémité (marqué par un rond noir) de la courbe Q au point marqué par un triangle bleu de la courbe P situé au voisinage du point extrémité d'origine. En effet, tout calcul de l'erreur quadratique moyenne de correspondance consiste à choisir deux points caractéristiques situés à proximité des points extrémités d'origine pour délimiter la correspondance à prendre en compte et pour éviter toute correspondance des points qui ne ressemble pas à la vérité terrain. Par exemple, la figure montre l'une des propositions des points caractéristiques (points noirs). Ces points bornes ont des abscisses curvilignes égal à 0.126 (s_{inf}) et 0.874 (s_{sup}) respectivement de chacune des courbes Q et P .

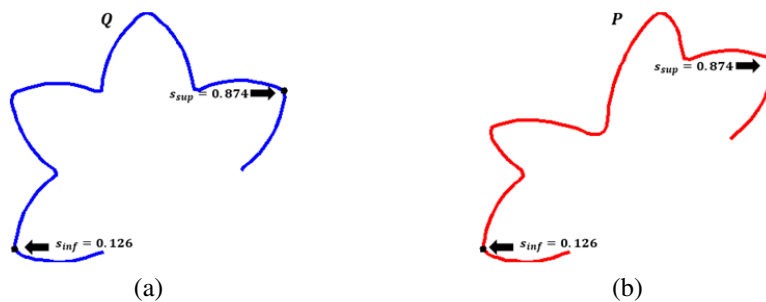


Figure 4.5 – Les courbes en figure 4.5a et figure 4.5b représentent les mêmes parties de formes Q et P de la figure 4.3, où s_{inf} et s_{sup} désignent les abscisses curvilignes de l'une des paires des points caractéristiques bornes possibles.

4.3.2 Critère basé sur la distance géodésique

L'utilisation de l'erreur quadratique moyenne EQM_ϕ en tant qu'un critère comparatif entre les méthodes d'orientation est soumise aux différentes contraintes. Un repère commun des fonctions de correspondance et des points caractéristiques bornes sont nécessaires pour un calcul correct de cette erreur. Pour cela, nous proposons d'utiliser la distance géodésique D entre toute paire des parties de formes après alignement, comme un critère

comparatif plus pratique que l'erreur EQM_ϕ . En plus, le but final est de trouver cette distance géodésique en utilisant cet alignement préliminaire. Pour cela l'alignement sera évalué aussi en fonction de la distance obtenue.

4.4 Comparaison des méthodes d'orientation

Afin d'étudier la stabilité de chaque méthode d'orientation en présence des perturbations d'extrémité, nous avons proposé d'ajouter des perturbations à l'une des extrémités de l'une des parties de formes. L'extrémité à perturber a été choisie aléatoirement de la partie de forme redéfinie. Il est intuitivement raisonnable d'admettre que la partie incluse dans l'autre partie est considérée comme la partie de forme dont l'extrémité à perturber. Tout changement ou toute perturbation de l'extrémité se traduit par un étirement de l'extrémité en question dans cinq différentes directions. Par exemple dans la figure 4.6, nous illustrons telles perturbations apportées à l'une des extrémités de la partie Q de la figure 4.3. La direction de l'extrémité d'origine est l'une des directions choisies pour réaliser les changements. Les perturbations y apportées sont représentées par des lignes pointillées en mauve dans la figure 4.6.

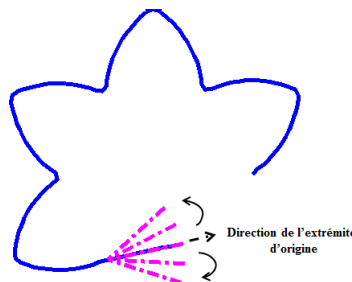


Figure 4.6 – Les lignes pointillées en mauve représentent les perturbations d'extrémité ajoutées à la courbe en bleu, même courbe de la figure 4.3a, dans cinq différentes directions dont l'une est la direction de l'extrémité d'origine.

Dans ce qui suit, nous avons comparé notre méthode d'orientation aux celles des trois autres méthodes utilisées dans des travaux antérieures. En fait, nous avons exploité les deux critères évaluatifs proposés (c'est-à-dire l'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D (décrite dans la section 3.2.3)) pour évaluer la performance et la stabilité des quatre méthodes d'orientation face aux perturbations d'extrémités. Tout d'abord la comparaison est réalisée sur des paires de parties de formes identiques. La comparaison est apportée ensuite sur des paires de parties de formes synthétiques dont l'une est obtenue par une déformation de l'autre par une transformation donnée $\phi_{imposée}$. Enfin, la distance géodésique seule est utilisée pour une comparaison plus générale effectuée sur 70 paires des parties de formes obtenues par

4.4. COMPARAISON DES MÉTHODES D'ORIENTATION

la décomposition DCE des formes de la base MPEG-7. Dans la suite, nous allons désigner par :

- la méthode 1 : l'alignement Procuste,
- la méthode 2 : l'alignement angulaire,
- la méthode 3 : la représentation dans un repère invariant et
- la méthode 4 : la méthode d'orientation proposée.

4.4.1 Évaluation sur des paires de parties de formes identiques

Afin de valider la méthode proposée comme une méthode d'orientation capable de résoudre le problème de la rotation des parties de formes et la comparer aux trois autres, nous avons réalisé une comparaison préliminaire sur des paires de parties de formes identiques. Par exemple, la figure 4.7 montre une paire des parties de formes identiques Q et P . La courbe P est obtenue par une transformation identique (c'est-à-dire $\phi_{imposée} = Id_{[0,1]}$) de la courbe Q comme indique la figure 4.7b.

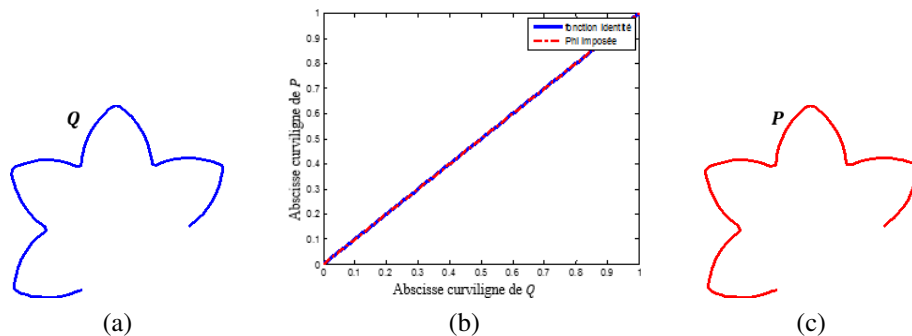


Figure 4.7 – Paires de parties de formes identiques. La courbe P (c) est obtenue par l'application de la transformation identique (b) sur la courbe Q (a).

En absence de toute perturbation d'extrémité, les quatre méthodes réussissent à mettre les deux courbes Q et P avec la même orientation. L'erreur moyenne de correspondance EQM_ϕ et la distance géodésique D sont égales à 0 pour toute méthode d'orientation utilisée.

Ensuite, des perturbations d'extrémité sont apportées sur la partie de forme Q dans les différentes directions (voir la figure 4.6). La première perturbation d'extrémité appliquée est une extension de l'extrémité d'origine dans sa même direction initiale, voir la figure 4.8a. Chaque perturbation d'extrémité ajoutée impose la nécessité de mettre la fonction de correspondance initialement imposée dans son nouveau repère où le point

extrémité d'origine du côté perturbé ne correspond plus au même abscisse curviligne dans l'ancien repère. La figure 4.8 montre la différence entre la fonction de correspondance initialement imposée (figure 4.8b) et son équivalent (figure 4.8c) vérité-terrain. Précisément, le point extrémité d'origine ($s_{0_{orig}} = 0$) correspond au point dont l'abscisse curviligne est différente de 0 ($s = 0.092$) sur la nouvelle courbe. Les deux parties de formes sont alignées ensuite par chacune des quatre méthodes d'orientation et la fonction de correspondance trouvée est comparée à la vérité-terrain.

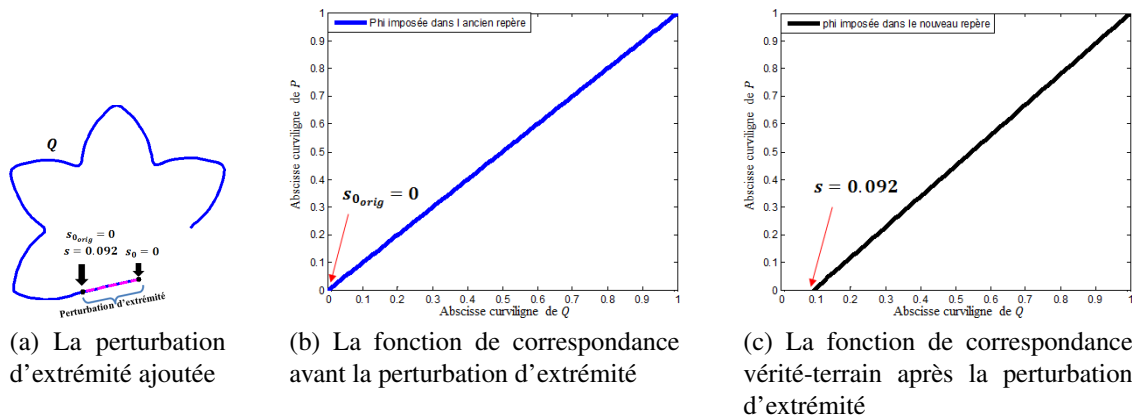
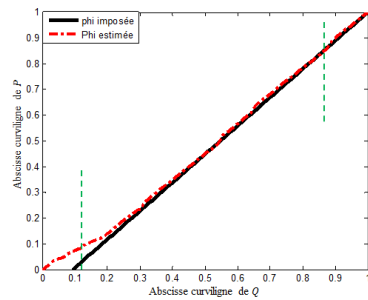


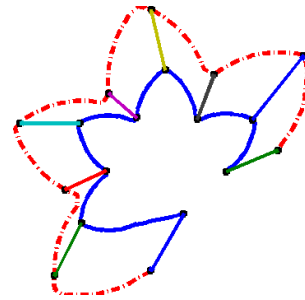
Figure 4.8 – La perturbation d'extrémité ajoutée à la même courbe Q de la figure 4.7a exige de remettre la fonction de correspondance initialement imposée (4.8b) dans le nouveau repère pour avoir la vérité-terrain (c).

Enfin, l'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D sont calculées entre les deux parties de formes alignées Q et P (voir la section 4.3). Les graphes en (a), (c) et (e) de la figure 4.9 illustrent la fonction de correspondance estimée (courbe rouge pointillée) entre la partie de forme Q de la figure 4.8a et la partie de forme P (courbe rouge) de la figure 4.7c après l'alignement par chacune des trois méthodes d'orientation. Les effets de la rotation sont supprimés par l'utilisation de chacune des méthodes suivantes: alignement Procuste (méthode 1), alignement angulaire (méthode 2), la représentation dans un repère invariant (méthode 3). L'erreur quadratique moyenne EQM_ϕ est calculée en prenant en compte tous les points homologues ayant l'abscisse curviligne comprise entre les valeurs 0.126 et 0.874 (marquées par des lignes verticales pointillées en vert en (a), (c) et (e) de la figure 4.9). Une telle correspondance bornée entre deux points caractéristiques des parties de formes est très importante pour calculer l'erreur EQM_ϕ correctement (voir la figure 4.5). Dans la figure 4.9 ((b),(d) et (f)), il est clairement remarquable que les points caractéristiques (points noirs) de la courbe bleue ne correspondent pas vraiment aux points de caractéristiques similaires de la courbe rouge à cause du déclin de la stabilité de la performance des méthodes d'orientation (1, 2 et 3) utilisées en présence de la

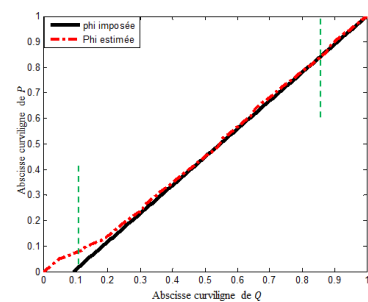
4.4. COMPARAISON DES MÉTHODES D'ORIENTATION



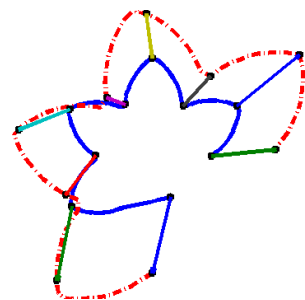
(a) Méthode 1 $\rightarrow EQM_{\phi} = 0.00012$



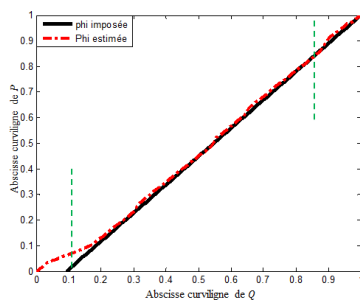
(b) Méthode 1 $\rightarrow D = 0.20584$



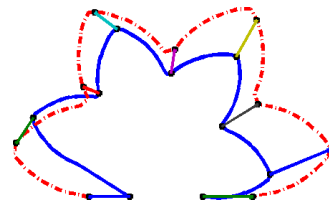
(c) Méthode 2 $\rightarrow EQM_{\phi} = 0.00012$



(d) Méthode 2 $\rightarrow D = 0.20983$



(e) Méthode 3 $\rightarrow EQM_{\phi} = 0.00009$



(f) Méthode 3 $\rightarrow D = 0.20427$

Figure 4.9 – Les courbes pointillées rouges en figures (a), (c), et (e) représentent les fonctions de correspondance trouvées $\phi_{estimée}$ à la suite de l'utilisation respective de chacune de méthodes d'orientation 1-3. De même les flèches colorées dans les figures (b), (d), et (f) illustrent la correspondance entre des points caractéristiques de Q et P , obtenue respectivement par chacune des méthodes d'orientation (1-3).

perturbation d'extrémité. L'erreur moyenne de correspondance trouvée suite à l'utilisation de chacune des trois premières méthodes d'orientation est relativement supérieure (Fig. 4.9 (a), (c) et (e)) à celle trouvée en appliquant notre méthode

d'orientation (Fig. 4.10). La bonne correspondance des points de caractéristiques similaires (points noirs dans la figure 4.10d) des deux parties de formes Q et P obtenue par notre méthode d'orientation a abouti à une distance géodésique ($D = 0.08391$) relativement inférieure aux valeurs obtenues par les trois autres méthodes. En appliquant la méthode proposée d'orientation, la réinitialisation des point extrémités (Fig. 4.10d) nécessite de remettre la fonction de correspondance trouvée (Fig. 4.10b) dans le nouveau repère qui correspond à la vérité-terrain (Fig. 4.10).

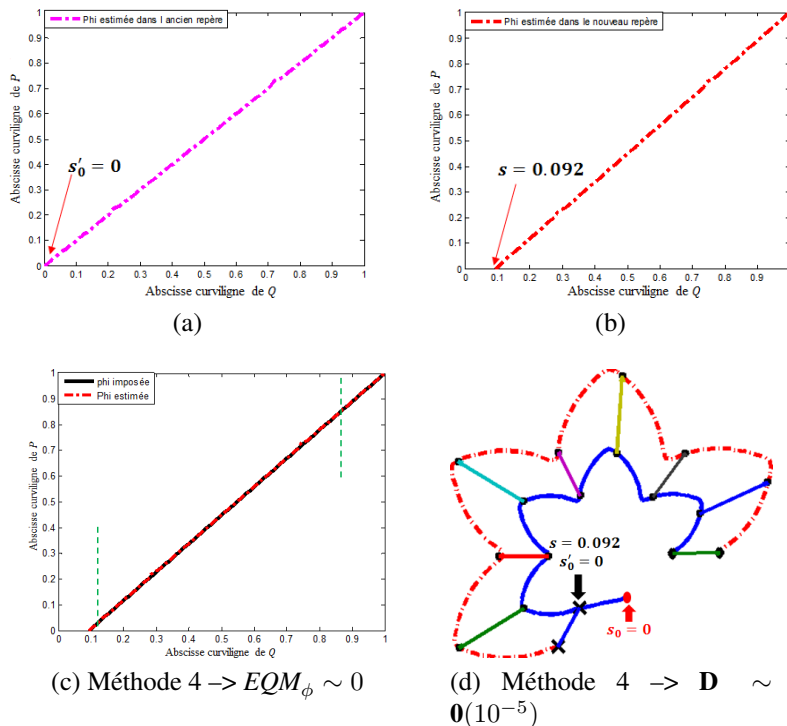


Figure 4.10 – L'étape de la réinitialisation des points extrémités, adoptée dans notre méthode d'orientation (Fig. (d)), exige de remettre la fonction de correspondance estimée (Fig. (a)) dans le nouveau repère de la vérité-terrain (Fig. (b)).

Suite aux perturbations d'extrémité dans les quatre autres directions, l'erreur moyenne de correspondance et la distance géodésique sont recalculées en utilisant chacune des quatre méthodes d'orientation. Donc, nous désignons par les directions 1, 2, 3, 4 et 5, les extensions d'extrémité qui font l'angle $0^\circ, 15^\circ, 30^\circ, -15^\circ, -30^\circ$ respectivement avec la direction de l'extrémité d'origine. En comparant les valeurs obtenues, données dans le tableau 4.1, notre méthode d'orientation paraît la plus stable aux perturbations d'extrémité apportées dans les cinq différentes directions. En présence des perturbations d'extrémité, notre méthode réussit toujours à correspondre correctement les points ayant de caractéristiques similaires de deux parties de formes Q

4.4. COMPARAISON DES MÉTHODES D'ORIENTATION

et P .

Tableau 4.1 – L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues par l'application de chacune des 4 méthodes d'orientation pour l'alignement des courbes Q et P de la figure 4.7, où la partie de forme Q présente une perturbation d'extrémité dans 5 différentes directions (figure 4.6).

	Méthode 1	Méthode 2	Méthode 3	Méthode 4
Direction 1 (angle=0°)	$EQM_\phi = 0.00012$ $D = 0.20584$	$EQM_\phi = 0.00012$ $D = 0.20983$	$EQM_\phi = 0.00009$ $D = 0.20427$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $\mathbf{D} \sim \mathbf{0}(10^{-5})$
Direction 2 (angle=15°)	$EQM_\phi = 0.00012$ $D = 0.20293$	$EQM_\phi = 0.00012$ $D = 0.21063$	$EQM_\phi = 0.00005$ $D = 0.17548$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $\mathbf{D} \sim \mathbf{0}(10^{-5})$
Direction 3 (angle=30°)	$EQM_\phi = 0.00012$ $D = 0.20464$	$EQM_\phi = 0.00012$ $D = 0.21560$	$EQM_\phi = 0.00006$ $D = 0.20205$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $\mathbf{D} \sim \mathbf{0}(10^{-5})$
Direction 4 (angle=-15°)	$EQM_\phi = 0.00012$ $D = 0.21365$	$EQM_\phi = 0.00012$ $D = 0.21397$	$EQM_\phi = 0.000012$ $D = 0.25753$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $\mathbf{D} \sim \mathbf{0}(10^{-5})$
Direction 5 (angle=-30°)	$EQM_\phi = 0.00012$ $D = 0.22473$	$EQM_\phi = 0.00009$ $D = 0.20950$	$EQM_\phi = 0.00014$ $D = 0.28946$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $\mathbf{D} \sim \mathbf{0}(10^{-5})$

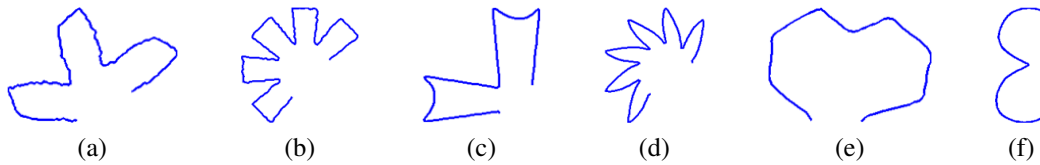


Figure 4.11 – Parties de formes références utilisées pour l'évaluation de la performance des 4 méthodes d'orientation en présence des perturbations d'extrémité.

En plus, nous avons montré sur six autres paires de parties de formes identiques l'efficacité et la stabilité de notre méthode d'orientation en présence des perturbations d'extrémité. Les six parties de formes références utilisées sont données par la figure 4.11. Une telle comparaison, évaluée sur des paires des parties de formes identiques, pour cela nous passons à une évaluation sur des formes synthétiques avec des formes non identiques.

4.4.2 Évaluation sur des parties de formes synthétiques

Après avoir testé les quatre méthodes d'orientation sur des paires des parties de formes identiques, une comparaison plus générale des mêmes méthodes est établie sur des paires de formes synthétiques. Chaque paire ici se compose de deux parties de formes différentes dont l'une est la transformée de l'autre par une déformation attribuée par l'application d'une fonction de correspondance ($\phi_{imposée}$) qui est différente de la fonction identique

(c.à-d. $\phi_{imposée} \neq Id_{[0,1]}$). Par exemple dans la figure 4.3, la courbe P (figure 4.3c) est obtenue en appliquant une déformation $\phi_{imposée}$ (figure 4.3b) à la courbe référence Q (figure 4.3a).

Tableau 4.2 – L’erreur quadratique moyenne de correspondance EQM_{ϕ} et la distance géodésique D obtenues par l’application de chacune des 4 méthodes d’orientation pour l’alignement des courbes Q (figure 4.12a) et P (Fig. 4.3c), où la partie de forme Q présente une perturbation d’extrémité dans 5 différentes directions (Fig. 4.6).

	Méthode 1	Méthode 2	Méthode 3	Méthode 4
Sans perturbation d’extrémité	$EQM_{\phi}=\mathbf{0.00003}$ $D = 0.19572$	$EQM_{\phi}=\mathbf{0.00003}$ D=0.19417	$EQM_{\phi} = 0.00007$ $D = 0.19782$	$EQM_{\phi} = 0.00011$ $D = 0.22507$
Direction 1 (angle=0°)	$EQM_{\phi} = 0.00034$ $D = 0.27490$	$EQM_{\phi} = 0.00026$ $D = 0.25663$	$EQM_{\phi} = 0.00023$ $D = 0.24393$	$EQM_{\phi}=\mathbf{0.00007}$ D=0.22702
Direction 2 (angle=15°)	$EQM_{\phi} = 0.00033$ $D = 0.26287$	$EQM_{\phi} = 0.00026$ $D = 0.25773$	$EQM_{\phi} = 0.00009$ $D = 0.22917$	$EQM_{\phi}=\mathbf{0.00007}$ D=0.22702
Direction 3 (angle=30°)	$EQM_{\phi} = 0.00028$ $D = 0.25427$	$EQM_{\phi} = 0.00031$ $D = 0.26206$	$EQM_{\phi} = 0.00041$ $D = 0.27207$	$EQM_{\phi}=\mathbf{0.00007}$ D=0.22702
Direction 4 (angle=-15°)	$EQM_{\phi} = 0.00036$ $D = 0.25153$	$EQM_{\phi} = 0.00036$ $D = 0.25656$	$EQM_{\phi} = 0.00030$ $D = 0.28220$	$EQM_{\phi}=\mathbf{0.00007}$ D=0.22702
Direction 5 (angle=-30°)	$EQM_{\phi} = 0.00036$ $D = 0.26045$	$EQM_{\phi} = 0.00036$ $D = 0.26256$	$EQM_{\phi} = 0.00030$ $D = 0.30644$	$EQM_{\phi}=\mathbf{0.00007}$ D=0.22702

Similairement à l’étude comparative des méthodes d’orientation sur des paires des parties de formes identiques, nous suivons les mêmes étapes de calcul de l’erreur quadratique moyenne de correspondance et la distance géodésique en absence et en présence de perturbations d’extrémité. Les valeurs obtenues sont données par le tableau 4.2.

En absence des perturbations d’extrémité à Q , les deux premières méthodes d’orientation, l’alignement Procruste et l’alignement angulaire, ont réussi à trouver une correspondance $\phi_{estimée}$ presque identique à celle initialement imposée (référant à la première ligne du tableau 4.2) car les erreurs moyennes de correspondance sont presque nulles à 10^{-5} près. En se référant aux valeurs obtenues au-delà de la deuxième ligne du tableau 4.2, nous concluons que notre méthode d’orientation est la plus stable à toute perturbation d’extrémité de Q parmi les quatre méthodes appliquées. Ainsi, la méthode d’orientation proposée aboutit à une correspondance optimale des points caractéristiques des deux parties de formes données. Par exemple en présence d’une perturbation d’extrémité de la courbe Q dans la direction quatre (Fig. 4.12a), les figures 4.13b, 4.13d, 4.13f et 4.14b montrent que la correspondance la plus optimale des points caractéristiques des deux parties de formes Q et P est atteint par l’utilisation de notre méthode d’orientation. La supériorité de performance de la méthode d’orientation proposée a été ainsi montrée sur six autres paires de parties de formes (voir les tableaux 4.3, 4.4, 4.5, 4.6, 4.7, 4.8). Les paires de parties de formes utilisées pour la

4.4. COMPARAISON DES MÉTHODES D'ORIENTATION

comparaison sont obtenues en déformant les mêmes parties de formes références de la figure 4.11 par la transformation connue donnée dans la figure 4.3b.

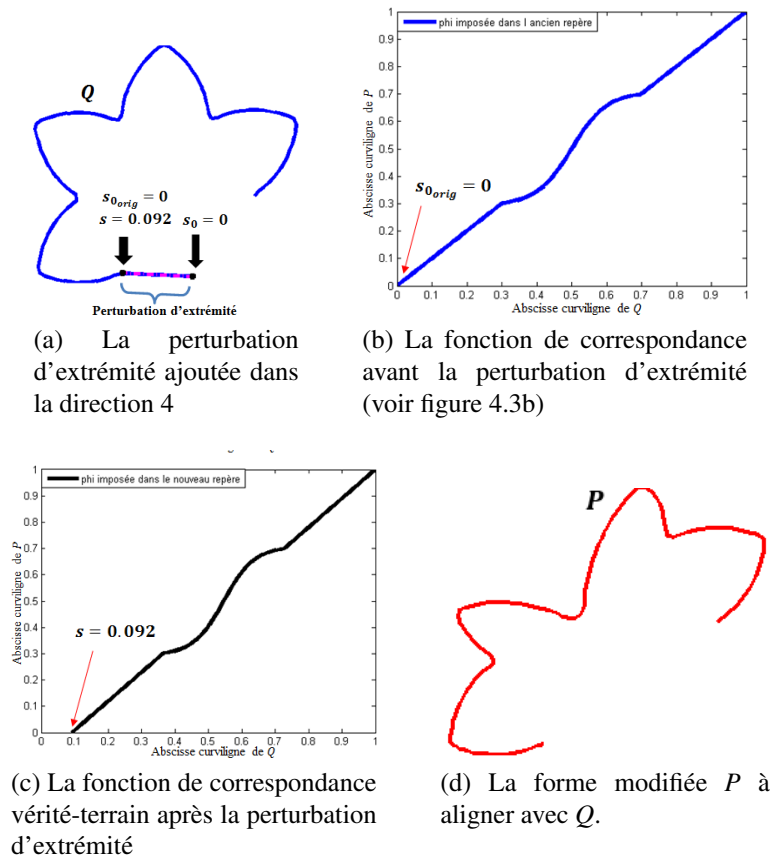
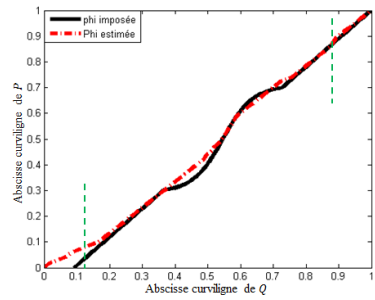
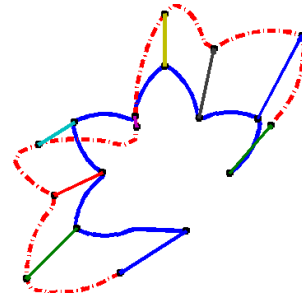


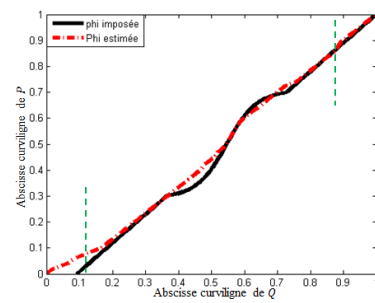
Figure 4.12 – La perturbation d'extrémité ajoutée à la même courbe Q de la figure 4.7a exige de remettre la fonction de correspondance initialement imposée (b) dans le nouveau repère de la vérité-terrain (c).



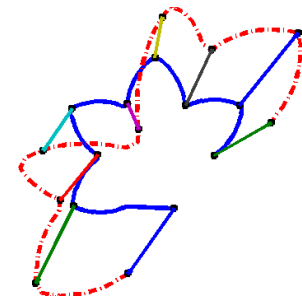
(a) Méthode 1 $\rightarrow EQM_{\phi} = 0.00036$



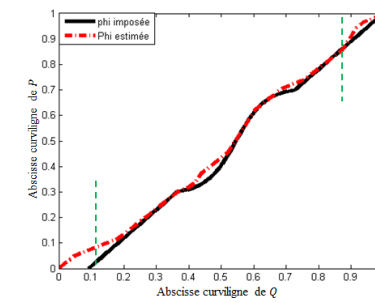
(b) Méthode 1 $\rightarrow D = 0.26045$



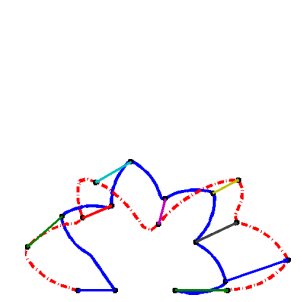
(c) Méthode 2 $\rightarrow EQM_{\phi} = 0.00036$



(d) Méthode 2 $\rightarrow D = 0.26256$



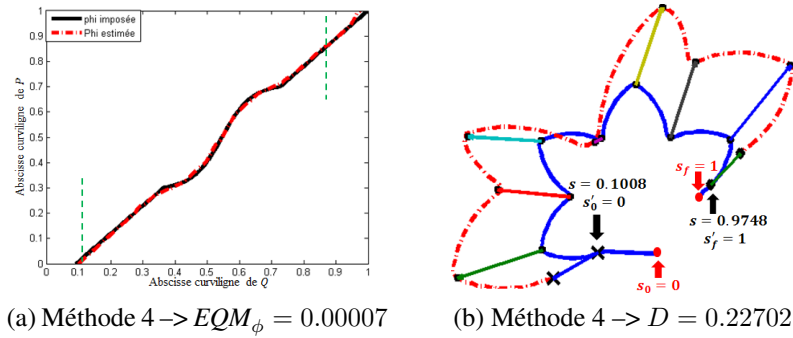
(e) Méthode 3 $\rightarrow EQM_{\phi} = 0.00030$



(f) Méthode 3 $\rightarrow D = 0.30644$

Figure 4.13 – En chaque ligne, les deux figures représentent la fonction de correspondance $\phi_{estimée}$ et la correspondance entre des points caractéristiques de Q (figure 4.12a) et P (figure 4.12d), obtenues respectivement par chacune des méthodes d'orientation (1-3).

4.4. COMPARAISON DES MÉTHODES D'ORIENTATION



(a) Méthode 4 $\rightarrow EQM_\phi = 0.00007$

(b) Méthode 4 $\rightarrow D = 0.22702$

Figure 4.14 – La fonction de correspondance obtenue suite à l'utilisation de notre méthode d'orientation est la courbe rouge pointillée de la figure (a). Elle coïncide bien avec la vérité-terrain. La figure (b) illustre la correspondance entre des points caractéristiques de Q (figure 4.12a) et P (figure 4.12d).

Tableau 4.3 – L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre une paire de courbes (Q, P) alignées avec chacune des 4 méthodes d'orientation. La figure 4.11a montre la partie de forme Q qui va subir une perturbation d'extrémité dans 5 différentes directions. La partie de forme P est la partie déformée de Q par la transformation $\phi_{imposée}$ de la figure 4.3b.

		Méthode 1	Méthode 2	Méthode 3	Méthode 4
$P \equiv Q$ ($\phi_{imposée} \equiv Id_{[0,1]}$)	Direction 1 (angle=0°)	$EQM_\phi = 0.00016$ $D = 0.24077$	$EQM_\phi = 0.00016$ $D = 0.24011$	$EQM_\phi = 0.001$ $D = 0.43769$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 2 (angle=15°)	$EQM_\phi = 0.0001$ $D = 0.21907$	$EQM_\phi = 0.00011$ $D = 0.22328$	$EQM_\phi = 0.00033$ $D = 0.41816$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 3 (angle=30°)	$EQM_\phi = 0.0001$ $D = 0.21677$	$EQM_\phi = 0.00016$ $D = 0.22859$	$EQM_\phi = 0.00029$ $D = 0.34758$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 4 (angle=-15°)	$EQM_\phi = 0.00016$ $D = 0.25059$	$EQM_\phi = 0.00016$ $D = 0.24669$	$EQM_\phi = 0.00096$ $D = 0.44049$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 5 (angle=-30°)	$EQM_\phi = 0.00016$ $D = 0.26334$	$EQM_\phi = 0.00015$ $D = 0.24272$	$EQM_\phi = 0.00033$ $D = 0.4380$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	$P \neq Q$ ($\phi_{imposée} \neq Id_{[0,1]}$) (Fig. 4.3b)	Sans perturbation d'extrémité	$EQM_\phi = \mathbf{0.0003}$ $D = 0.16629$	$EQM_\phi = \mathbf{0.0003}$ $D = 0.16525$	$EQM_\phi = \mathbf{0.0003}$ $D = 0.18891$
Direction 1 (angle=0°)		$EQM_\phi = 0.0005$ $D = 0.24182$	$EQM_\phi = \mathbf{0.0003}$ $D = 0.25976$	$EQM_\phi = 0.0017$ $D = 0.41293$	$EQM_\phi = \mathbf{0.0003}$ $D = \mathbf{0.17056}$
Direction 2 (angle=15°)		$EQM_\phi = 0.0005$ $D = 0.22923$	$EQM_\phi = 0.0005$ $D = 0.23494$	$EQM_\phi = 0.0017$ $D = 0.39275$	$EQM_\phi = \mathbf{0.0003}$ $D = \mathbf{0.17056}$
Direction 3 (angle=30°)		$EQM_\phi = 0.0005$ $D = 0.22161$	$EQM_\phi = 0.0006$ $D = 0.23760$	$EQM_\phi = 0.0009$ $D = 0.34159$	$EQM_\phi = \mathbf{0.0003}$ $D = \mathbf{0.17056}$
Direction 4 (angle=-15°)		$EQM_\phi = 0.0005$ $D = 0.25077$	$EQM_\phi = 0.0005$ $D = 0.24986$	$EQM_\phi = 0.0017$ $D = 0.41644$	$EQM_\phi = \mathbf{0.0003}$ $D = \mathbf{0.17056}$
Direction 5 (angle=-30°)		$EQM_\phi = 0.0005$ $D = 0.2627$	$EQM_\phi = 0.0005$ $D = 0.2585$	$EQM_\phi = 0.0016$ $D = 0.4380$	$EQM_\phi = \mathbf{0.0003}$ $D = \mathbf{0.17056}$

Tableau 4.4 – L’erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre une paire de courbes (Q, P) alignées avec chacune des 4 méthodes d’orientation. La figure 4.11b montre la partie de forme Q qui va subir une perturbation d’extrémité dans 5 différentes directions. La partie de forme P est la partie déformée de Q par la transformation $\phi_{imposée}$ de la figure 4.3b.

		Méthode 1	Méthode 2	Méthode 3	Méthode 4
$P \equiv Q$ ($\phi_{imposée} \equiv Id_{[0,1]}$)	Direction 1 (angle=0°)	$EQM_\phi = 0.00001$ $D = 0.25052$	$EQM_\phi = 0.00142$ $D = 0.61008$	$EQM_\phi = 0.00746$ $D = 0.72908$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 2 (angle=15°)	$EQM_\phi = 0.00001$ $D = 0.25052$	$EQM_\phi = 0.00142$ $D = 0.61809$	$EQM_\phi = 0.00654$ $D = 0.70063$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 3 (angle=30°)	$EQM_\phi = 0.00001$ $D = 0.25548$	$EQM_\phi = 0.00142$ $D = 0.61839$	$EQM_\phi = 0.00667$ $D = 0.69223$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 4 (angle=-15°)	$EQM_\phi = 0.00001$ $D = 0.25527$	$EQM_\phi = 0.00142$ $D = 0.61209$	$EQM_\phi = 0.01227$ $D = 0.75565$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 5 (angle=-30°)	$EQM_\phi = 0.00001$ $D = 0.26401$	$EQM_\phi = 0.00142$ $D = 0.61583$	$EQM_\phi = 0.03711$ $D = 0.64216$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	$P \neq Q$ ($\phi_{imposée} \neq Id_{[0,1]}$)(Fig. 4.3b)	Sans perturbation d’extrémité	$EQM_\phi = \mathbf{0.00002}$ $D = 0.22019$	$EQM_\phi = \mathbf{0.00002}$ $D = \mathbf{0.22005}$	$EQM_\phi = \mathbf{0.00002}$ $D = 0.22472$
Direction 1 (angle=0°)		$EQM_\phi = \mathbf{0.00003}$ $D = 0.30341$	$EQM_\phi = 0.00236$ $D = 0.5759$	$EQM_\phi = 0.0058$ $D = 0.72849$	$EQM_\phi = \mathbf{0.00003}$ $D = \mathbf{0.23518}$
Direction 2 (angle=15°)		$EQM_\phi = \mathbf{0.00003}$ $D = 0.30328$	$EQM_\phi = 0.00255$ $D = 0.57357$	$EQM_\phi = 0.00509$ $D = 0.69434$	$EQM_\phi = \mathbf{0.00003}$ $D = \mathbf{0.23502}$
Direction 3 (angle=30°)		$EQM_\phi = \mathbf{0.00003}$ $D = 0.30727$	$EQM_\phi = 0.0027$ $D = 0.57644$	$EQM_\phi = 0.00518$ $D = 0.68419$	$EQM_\phi = \mathbf{0.00003}$ $D = \mathbf{0.23459}$
Direction 4 (angle=-15°)		$EQM_\phi = \mathbf{0.00003}$ $D = 0.30727$	$EQM_\phi = \mathbf{0.00003}$ $D = 0.33012$	$EQM_\phi = 0.00612$ $D = 0.48027$	$EQM_\phi = \mathbf{0.00003}$ $D = \mathbf{0.23502}$
Direction 5 (angle=-30°)		$EQM_\phi = \mathbf{0.00003}$ $D = 0.31427$	$EQM_\phi = 0.00004$ $D = 0.33588$	$EQM_\phi = 0.0197$ $D = 0.68004$	$EQM_\phi = \mathbf{0.00003}$ $D = \mathbf{0.23459}$

4.4. COMPARAISON DES MÉTHODES D'ORIENTATION

Tableau 4.5 – L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre une paire de courbes (Q, P) alignées avec chacune des 4 méthodes d'orientation. La figure 4.11c montre la partie de forme Q qui va subir une perturbation d'extrémité dans 5 différentes directions. La partie de forme P est la partie déformée de Q par la transformation $\phi_{imposée}$ de la figure 4.3b.

		Méthode 1	Méthode 2	Méthode 3	Méthode 4
$P \equiv Q$ ($\phi_{imposée} \equiv Id_{[0,1]}$)	Direction 1 (angle=0°)	$EQM_\phi = 0.00003$ $D = 0.24385$	$EQM_\phi = 0.00003$ $D = 0.23806$	$EQM_\phi = 0.00002$ $D = 0.53918$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 2 (angle=15°)	$EQM_\phi = 0.00003$ $D = 0.23144$	$EQM_\phi = 0.00003$ $D = 0.22933$	$EQM_\phi = 0.00235$ $D = 0.56326$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 3 (angle=30°)	$EQM_\phi = 0.00003$ $D = 0.22245$	$EQM_\phi = 0.00003$ $D = 0.22495$	$EQM_\phi = 0.00008$ $D = 0.58646$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 4 (angle=-15°)	$EQM_\phi = 0.00003$ $D = 0.25902$	$EQM_\phi = 0.00003$ $D = 0.250143$	$EQM_\phi = 0.00008$ $D = 0.45741$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 5 (angle=-30°)	$EQM_\phi = 0.00003$ $D = 0.27636$	$EQM_\phi = 0.00002$ $D = 0.26502$	$EQM_\phi = 0.00033$ $D = 0.44312$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	$P \neq Q$ ($\phi_{imposée} \neq Id_{[0,1]}$)(Fig. 4.3b)	Sans perturbation d'extrémité	$EQM_\phi=0.00083$ $D = 0.11084$	$EQM_\phi=\mathbf{0.00069}$ $D=\mathbf{0.10246}$	$EQM_\phi=0.00071$ $D = 0.26814$
Direction 1 (angle=0°)		$EQM_\phi=0.00091$ $D = 0.2294$	$EQM_\phi=0.00085$ $D = 0.23322$	$EQM_\phi = 0.00128$ $D = 0.49794$	$EQM_\phi=\mathbf{0.00069}$ $D=\mathbf{0.11316}$
Direction 2 (angle=15°)		$EQM_\phi=0.00091$ $D = 0.21722$	$EQM_\phi = 0.00085$ $D = 0.22447$	$EQM_\phi = 0.00225$ $D = 0.51326$	$EQM_\phi=\mathbf{0.00069}$ $D=\mathbf{0.11316}$
Direction 3 (angle=30°)		$EQM_\phi=0.00084$ $D = 0.20883$	$EQM_\phi = 0.00085$ $D = 0.21995$	$EQM_\phi = 0.00242$ $D = 0.53249$	$EQM_\phi=\mathbf{0.00069}$ $D=\mathbf{0.11316}$
Direction 4 (angle=-15°)		$EQM_\phi=0.00084$ $D = 0.24466$	$EQM_\phi=0.00082$ $D = 0.24508$	$EQM_\phi = 0.00114$ $D = 0.37323$	$EQM_\phi=\mathbf{0.00069}$ $D=\mathbf{0.11316}$
Direction 5 (angle=-30°)		$EQM_\phi=0.00085$ $D = 0.26231$	$EQM_\phi = 0.00085$ $D = 0.26022$	$EQM_\phi = 0.001$ $D = 0.34794$	$EQM_\phi=\mathbf{0.00069}$ $D=\mathbf{0.11316}$

Tableau 4.6 – L’erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre une paire de courbes (Q, P) alignées avec chacune des 4 méthodes d’orientation. La figure 4.11d montre la partie de forme Q qui va subir une perturbation d’extrémité dans 5 différentes directions. La partie de forme P est la partie déformée de Q par la transformation $\phi_{imposée}$ de la figure 4.3b.

		Méthode 1	Méthode 2	Méthode 3	Méthode 4
$P \equiv Q$ ($\phi_{imposée} \equiv Id_{[0,1]}$)	Direction 1 (angle=0°)	$EQM_\phi = 0.00007$ $D = 0.28109$	$EQM_\phi = 0.00311$ $D = 0.81274$	$EQM_\phi = 0.00016$ $D = 0.46004$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 2 (angle=15°)	$EQM_\phi = 0.00007$ $D = 0.27151$	$EQM_\phi = 0.00394$ $D = 0.75756$	$EQM_\phi = 0.00486$ $D = 0.72753$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 3 (angle=30°)	$EQM_\phi = 0.00007$ $D = 0.26574$	$EQM_\phi = 0.00402$ $D = 0.7495$	$EQM_\phi = 0.00469$ $D = 0.73336$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 4 (angle=-15°)	$EQM_\phi = 0.00007$ $D = 0.29352$	$EQM_\phi = 0.00159$ $D = 0.74871$	$EQM_\phi = 0.02854$ $D = 0.54212$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 5 (angle=-30°)	$EQM_\phi = 0.00007$ $D = 0.30791$	$EQM_\phi = 0.00157$ $D = 0.72381$	$EQM_\phi = 0.02794$ $D = 0.54131$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
$P \neq Q$ ($\phi_{imposée} \neq Id_{[0,1]}$)(Fig. 4.3b)	Sans perturbation d’extrémité	$EQM_\phi = \mathbf{0.00003}$ $D = \mathbf{0.20927}$	$EQM_\phi = \mathbf{0.00003}$ $D = 0.20943$	$EQM_\phi = \mathbf{0.00003}$ $D = 0.2102$	$EQM_\phi = 0.00007$ $D = 0.24538$
	Direction 1 (angle=0°)	$EQM_\phi = 0.00016$ $D = 0.3279$	$EQM_\phi = 0.00177$ $D = 0.73777$	$EQM_\phi = 0.00046$ $D = 0.50696$	$EQM_\phi = \mathbf{0.00005}$ $D = \mathbf{0.24538}$
	Direction 2 (angle=15°)	$EQM_\phi = 0.00016$ $D = 0.31335$	$EQM_\phi = 0.00177$ $D = 0.73415$	$EQM_\phi = 0.00615$ $D = 0.7366$	$EQM_\phi = \mathbf{0.00005}$ $D = \mathbf{0.24538}$
	Direction 3 (angle=30°)	$EQM_\phi = 0.00016$ $D = 0.31335$	$EQM_\phi = 0.00177$ $D = 0.72945$	$EQM_\phi = 0.00617$ $D = 0.74039$	$EQM_\phi = \mathbf{0.00005}$ $D = \mathbf{0.24538}$
	Direction 4 (angle=-15°)	$EQM_\phi = 0.00018$ $D = 0.36549$	$EQM_\phi = 0.00177$ $D = 0.74337$	$EQM_\phi = 0.03412$ $D = 0.575$	$EQM_\phi = \mathbf{0.00005}$ $D = \mathbf{0.24538}$
	Direction 5 (angle=-30°)	$EQM_\phi = 0.00018$ $D = 0.37778$	$EQM_\phi = 0.00177$ $D = 0.72421$	$EQM_\phi = 0.00839$ $D = 0.61969$	$EQM_\phi = \mathbf{0.00005}$ $D = \mathbf{0.24538}$

4.4. COMPARAISON DES MÉTHODES D'ORIENTATION

Tableau 4.7 – L'erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre une paire de courbes (Q, P) alignées avec chacune des 4 méthodes d'orientation. La figure 4.11e montre la partie de forme Q qui va subir une perturbation d'extrémité dans 5 différentes directions. La partie de forme P est la partie déformée de Q par la transformation $\phi_{imposée}$ de la figure 4.3b.

		Méthode 1	Méthode 2	Méthode 3	Méthode 4
$P \equiv Q$ ($\phi_{imposée} \equiv Id_{[0,1]}$)	Direction 1 (angle=0°)	$EQM_\phi = 0.00029$ $D = 0.30885$	$EQM_\phi = 0.00022$ $D = 0.26198$	$EQM_\phi = 0.00659$ $D = 0.42078$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 2 (angle=15°)	$EQM_\phi = 0.00049$ $D = 0.31$	$EQM_\phi = 0.00029$ $D = 0.29748$	$EQM_\phi = 0.00659$ $D = 0.4238$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 3 (angle=30°)	$EQM_\phi = 0.00049$ $D = 0.30677$	$EQM_\phi = 0.00049$ $D = 0.3051$	$EQM_\phi = 0.00658$ $D = 0.40643$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 4 (angle=-15°)	$EQM_\phi = 0.00026$ $D = 0.2746$	$EQM_\phi = 0.00007$ $D = 0.23487$	$EQM_\phi = 0.00563$ $D = 0.4084$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 5 (angle=-30°)	$EQM_\phi = 0.00011$ $D = 0.27262$	$EQM_\phi = 0.00007$ $D = 0.22781$	$EQM_\phi = 0.00563$ $D = 0.40113$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
$P \neq Q$ ($\phi_{imposée} \neq Id_{[0,1]}$)(Fig. 4.3b)	Sans perturbation d'extrémité	$EQM_\phi = \mathbf{0.00008}$ $D = 0.19022$	$EQM_\phi = \mathbf{0.00008}$ $D = \mathbf{0.18906}$	$EQM_\phi = 0.00012$ $D = 0.20152$	$EQM_\phi = \mathbf{0.00008}$ $D = 0.19742$
	Direction 1 (angle=0°)	$EQM_\phi = 0.00012$ $D = 0.28794$	$EQM_\phi = \mathbf{0.00007}$ $D = 0.26888$	$EQM_\phi = 0.00644$ $D = 0.39943$	$EQM_\phi = \mathbf{0.00007}$ $D = \mathbf{0.20732}$
	Direction 2 (angle=15°)	$EQM_\phi = 0.00012$ $D = 0.28196$	$EQM_\phi = 0.00012$ $D = 0.28151$	$EQM_\phi = 0.00644$ $D = 0.3994$	$EQM_\phi = \mathbf{0.00007}$ $D = \mathbf{0.20732}$
	Direction 3 (angle=30°)	$EQM_\phi = 0.00012$ $D = 0.27925$	$EQM_\phi = 0.00012$ $D = 0.28249$	$EQM_\phi = 0.00644$ $D = 0.38493$	$EQM_\phi = \mathbf{0.00007}$ $D = \mathbf{0.20732}$
	Direction 4 (angle=-15°)	$EQM_\phi = \mathbf{0.00007}$ $D = 0.28238$	$EQM_\phi = \mathbf{0.00007}$ $D = 0.27490$	$EQM_\phi = 0.00566$ $D = 0.39068$	$EQM_\phi = \mathbf{0.00007}$ $D = \mathbf{0.20732}$
	Direction 5 (angle=-30°)	$EQM_\phi = \mathbf{0.00007}$ $D = 0.2942$	$EQM_\phi = 0.00008$ $D = 0.27679$	$EQM_\phi = 0.0051$ $D = 0.38982$	$EQM_\phi = \mathbf{0.00007}$ $D = \mathbf{0.20732}$

Tableau 4.8 – L’erreur quadratique moyenne de correspondance EQM_ϕ et la distance géodésique D obtenues entre une paire de courbes (Q, P) alignées avec chacune des 4 méthodes d’orientation. La figure 4.11f montre la partie de forme Q qui va subir une perturbation d’extrémité dans 5 différentes directions. La partie de forme P est la partie déformée de Q par la transformation $\phi_{imposée}$ de la figure 4.3b.

		Méthode 1	Méthode 2	Méthode 3	Méthode 4
$P \equiv Q$ ($\phi_{imposée} \equiv Id_{[0,1]}$)	Direction 1 (angle=0°)	$EQM_\phi = 0.00037$ $D = 0.14746$	$EQM_\phi = 0.00037$ $D = 0.14906$	$EQM_\phi = 0.00069$ $D = 0.24096$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 2 (angle=15°)	$EQM_\phi = 0.00037$ $D = 0.15047$	$EQM_\phi = 0.0008$ $D = 0.22192$	$EQM_\phi = 0.00043$ $D = 0.22906$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 3 (angle=30°)	$EQM_\phi = 0.00033$ $D = 0.16004$	$EQM_\phi = 0.00104$ $D = 0.22364$	$EQM_\phi = 0.00024$ $D = 0.17514$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 4 (angle=-15°)	$EQM_\phi = 0.00037$ $D = 0.151$	$EQM_\phi = 0.00031$ $D = 0.15038$	$EQM_\phi = 0.00076$ $D = 0.23156$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
	Direction 5 (angle=-30°)	$EQM_\phi = 0.00031$ $D = 0.161$	$EQM_\phi = 0.0002$ $D = 0.1858$	$EQM_\phi = 0.00071$ $D = 0.21814$	$EQM_\phi \sim \mathbf{0}(10^{-5})$ $D \sim \mathbf{0}(10^{-5})$
$P \neq Q$ ($\phi_{imposée} \neq Id_{[0,1]}$)(Fig. 4.3b)	Sans perturbation d’extrémité	$EQM_\phi=0.00079$ $D=0.11669$	$EQM_\phi=0.00141$ $D=0.12345$	$EQM_\phi=0.00016$ $D = 0.23110$	$EQM_\phi=0.00141$ $D=0.11546$
	Direction 1 (angle=0°)	$EQM_\phi=0.00142$ $D = 0.15567$	$EQM_\phi=0.00141$ $D = 0.15071$	$EQM_\phi = 0.00154$ $D = 0.23785$	$EQM_\phi=0.00141$ $D=0.09793$
	Direction 2 (angle=15°)	$EQM_\phi=0.00142$ $D = 0.1554$	$EQM_\phi = 0.00142$ $D = 0.15366$	$EQM_\phi = 0.00111$ $D = 0.23603$	$EQM_\phi=0.00141$ $D=0.09793$
	Direction 3 (angle=30°)	$EQM_\phi=0.000142$ $D = 0.16144$	$EQM_\phi = 0.0016$ $D = 0.17017$	$EQM_\phi = 0.00071$ $D = 0.23122$	$EQM_\phi=0.00141$ $D=0.09793$
	Direction 4 (angle=-15°)	$EQM_\phi=0.00141$ $D = 0.16146$	$EQM_\phi=0.00141$ $D = 0.1534$	$EQM_\phi = 0.00156$ $D = 0.25109$	$EQM_\phi=0.00141$ $D=0.09793$
	Direction 5 (angle=-30°)	$EQM_\phi=0.00141$ $D = 0.17386$	$EQM_\phi = 0.00094$ $D = 0.1683$	$EQM_\phi = 0.00153$ $D = 0.22042$	$EQM_\phi=0.00141$ $D=0.09793$

4.4. COMPARAISON DES MÉTHODES D'ORIENTATION

En résumé, l'erreur quadratique moyenne de correspondance EQM_{ϕ} et la distance géodésique D ont été validées comme deux critères significatifs pour l'évaluation et la comparaison des quatre méthodes d'orientation.

La seule limitation qui se présente lors du calcul de l'erreur quadratique moyenne de correspondance réside dans les points caractéristiques bornes à choisir. Chaque paire des parties de formes a ses propres points caractéristiques bornes conduisant à calculer correctement l'erreur de correspondance EQM_{ϕ} . La pertinence du choix de ses points est demandée afin de s'éloigner de toutes les correspondances erronées à proximité de l'extrémité. D'où, afin de généraliser cette étude comparative des méthodes d'orientation, il nous faut d'utiliser la distance géodésique comme critère d'évaluation de performance car l'utilisation de ce critère ne pose aucune contrainte. En plus, nous rappelons qu'il s'agit ici d'une étape préliminaire pour calculer cette distance qui devrait être la plus petite pour les parties supposées similaires ou de même classe. Ensuite, la comparaison est réalisée sur 70 paires de parties visuellement similaires de formes dont chacune appartient à l'une de 70 classes de la base des formes *MPEG-7*. Les parties de formes analysées sont obtenues par une décomposition DCE des formes originales de cette base.

4.4.3 Évaluation sur des parties de formes de la base des formes MPEG-7



Figure 4.15 – Les 70 paires de parties de formes utilisées dans l'étude comparative des 4 méthodes d'orientation.

Une évaluation de la robustesse et de la consistance de la performance de chacune des quatre méthodes d'orientation, en présence des perturbations d'extrémité, est réalisée sur un ensemble de 70 paires de parties de formes de la base des formes *MPEG-7* (figure 4.15). Chaque paire est composée de deux parties de formes visuellement similaires qui appartiennent à la même classe. Donc, une paire de parties de formes de chaque classe des 70 classes de la base des formes *MPEG-7* est utilisée dans l'étude comparative. La distance géodésique est le seul critère évaluatif dans cette comparaison.

Afin d'évaluer la performance de toute méthode d'orientation en présence des perturbations d'extrémité, chaque paire des parties de formes de la figure 4.15 est manipulée comme il est décrit précédemment (rallongement à différentes orientations). À la suite de toute perturbation d'extrémité apportée à l'une des deux parties de formes similaires, la distance géodésique est calculée entre les deux courbes après alignement. Les perturbations d'extrémité sont créées suivant les mêmes cinq directions déjà mentionnées. En total, nous obtenons 5 valeurs de la distance géodésique pour chaque paire de parties de formes. On cherche dans ce cas à caractériser la distribution des 5 valeurs obtenues de la distance géodésique, en évaluant la valeur moyenne et l'écart-type de cette distribution.

4.4. COMPARAISON DES MÉTHODES D'ORIENTATION

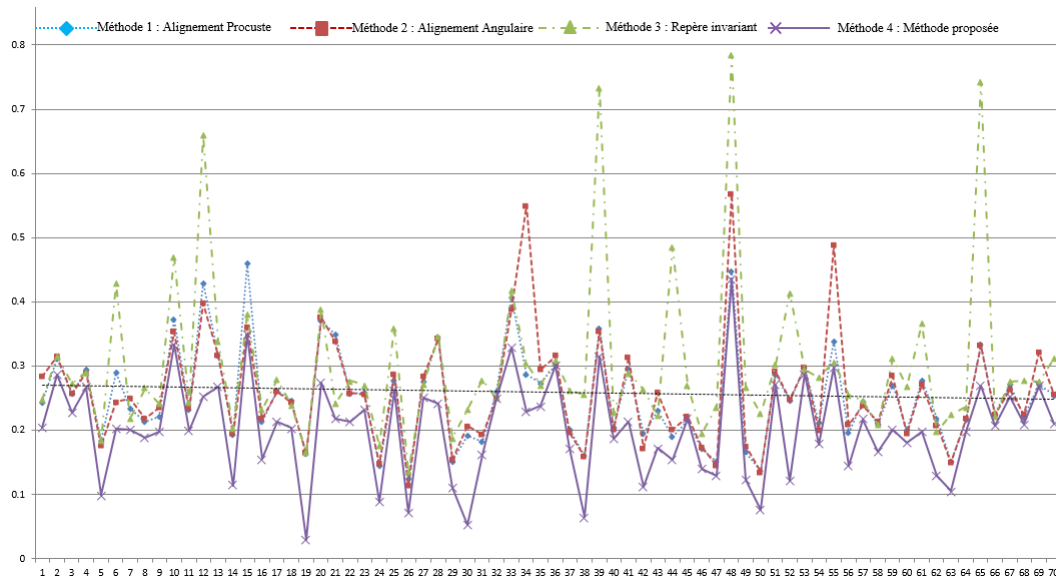


Figure 4.16 – Les valeurs moyennes de la distance géodésique calculées entre les 70 paires des parties de formes (figure 4.15) obtenues en présence des perturbations d'extrémité à la suite d'un alignement par les 4 méthodes d'orientation.

Les valeurs moyennes de la distance géodésique entre les 70 paires de parties de formes (figure 4.15) en présence des perturbations d'extrémité après alignement par chacune des quatre méthodes d'orientation sont affichées sur les quatre courbes de la figure 4.16. En comparant les valeurs moyennes de la distance géodésique obtenues par les quatre méthodes d'orientation, nous pouvons voir que la distance minimale entre les paires est toujours obtenue après l'alignement réalisé par notre méthode. En plus, il est remarquable que l'alignement par la méthode proposée d'orientation a abouti à une distance moyenne minimale qui vaut 0.20193 entre toutes les paires de parties de formes en présence des perturbations d'extrémité. Alors que des distances moyennes supérieures égales à 0.25231, 0.25950 et 0.30173 sont atteintes par la méthode 1, la méthode 2 et la méthode 3 respectivement.

Outre la différence significative entre les valeurs moyennes de la distance géodésique trouvées entre les 70 paires de parties de formes en présence des perturbations d'extrémité, il serait intéressant d'analyser les écarts entre les distances obtenues pour chaque paire. Cet écart reflète la robustesse des méthodes analysées. Pour chaque paire de parties de formes (figure 4.15), nous proposons d'analyser la variation de l'écart-type σ_D de l'ensemble des distances géodésiques, obtenue par les différentes perturbations d'extrémité, d'une méthode d'orientation à l'autre. Les figures 4.17a, 4.17b, 4.17c et 4.17d visualisent chacune la bande de la variation $[\bar{D} - \sigma_D, \bar{D} + \sigma_D]$ des distances D calculées entre toute paire de parties de formes soumises aux différentes perturbations suite à l'utilisation de l'alignement Procuste (méthode 1), l'alignement angulaire

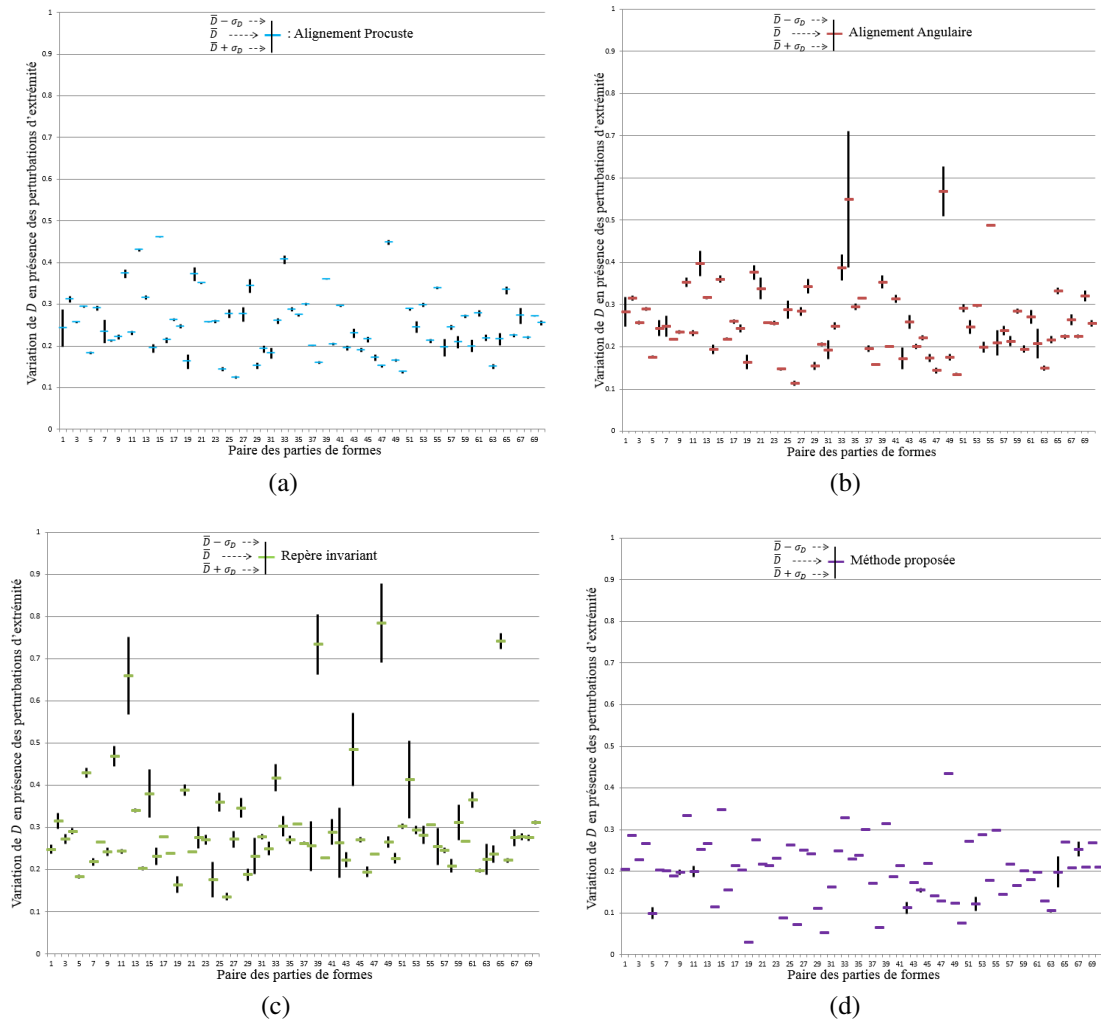


Figure 4.17 – À chaque paire de parties de formes des figures (a), (b), (c) et (d), une bande de la variation $[\bar{D} - \sigma_D \bar{D} + \sigma_D]$ des distances D calculées en présence de différentes perturbations a été attribuée pour chacune des 4 méthodes d'orientation.

(méthode 2), le repère invariant (méthode 3) et la méthode proposée (méthode 4) respectivement. Chaque bande est centrée par la valeur moyenne \bar{D} de l'ensemble des distances obtenues. Plus la bande est large, plus important est l'écart-type σ_D et inversement. Dans notre étude, si l'écart-type est faible, cela reflète la grande stabilité de la méthode d'orientation utilisée aux perturbations d'extrémité. Une lecture comparative des 4 diagrammes à bandes de la figure 4.17 montre la supériorité de la méthode d'orientation proposée (figure 4.17d) bien indiquée par l'absence presque totale des bandes de variation large qui apparaissent fréquemment par les autres méthodes. D'où, la grande stabilité de l'alignement des parties de formes réalisé par notre méthode

d'orientation aux perturbations d'extrémité.

4.5 Conclusion

Le fait d'orienter similairement la direction des parties de formes, courbes ouvertes, est une étape fondamentale de tout problème de comparaison des parties de formes. L'alignement Procuste, l'alignement angulaire et la représentation dans le repère invariant sont les seules méthodes de la littérature qui ont répondu au sujet de l'orientation des parties de formes.

Dans ce chapitre, nous avons proposé une nouvelle méthode d'orientation basée sur l'alignement Procuste et sur la redétermination des points extrémités, en se basant sur le fait qu'une partie de forme est extraite soit par une fragmentation d'une forme entière, soit par une segmentation appliquée. Donc, le trajet jusqu'à l'obtention d'une partie de forme peut être affecté par différentes complexités telles que le bruit et l'occlusion conduisant à des perturbations aux extrémités des parties extraites. En effet, il est inconvenable de faire correspondre les points extrémités des deux parties de formes à comparer. Au travers ce chapitre, nous avons aussi présenté une étude comparative établie entre la méthode d'orientation proposée et les trois méthodes de la littérature. Notre méthode a montré une grande stabilité aux perturbations d'extrémité et une supériorité par rapport aux trois autres méthodes, ce qui justifie son utilisation comme une étape préliminaire efficace au calcul de la distance géodésique présentée dans le chapitre précédent.

Comme une extension de ce travail sur la reconnaissance de parties de formes et de formes partielles, nous proposons dans le chapitre suivant de traiter le problème de la reconnaissance de formes fermées à partir des parties de formes.

Chapitre 5

Reconnaissance des formes entières à partir des parties de formes

Sommaire

5.1	Introduction	148
5.2	Travaux antérieurs	149
5.3	Système de correspondance basé sur la métrique quadratique	150
5.3.1	Décomposition des formes $2D$ fermées	150
5.3.2	Mesure de similarité quadratique	151
5.4	Approche proposée	153
5.4.1	Méthode de décomposition proposée	153
5.4.2	Distance utilisée	154
5.5	Applications sur la base de formes MPEG-7	155
5.5.1	Description des bases de test	155
5.5.2	Résultats et discussion	156
5.6	Conclusion	164

5.1 Introduction

Ce chapitre traite le sujet de la reconnaissance des formes 2D fermées basée sur la correspondance de leurs parties significatives. L'objectif de ce chapitre est de prouver la grande capacité cognitive des parties de formes et des mesures de similarité pour les applications les plus pratiques de reconnaissance de formes. L'utilisation d'une configuration significative des parties de formes est d'une grande importance pour l'analyse des formes qui contiennent de petites variations ou de grandes variations dues aux différentes poses, aux déformations ou aux occlusions. En effet, même si la différence ou la similarité entre deux formes données réside dans une petite partie des formes, l'utilisation des informations locales est capable de différencier ou de correspondre entre ces formes. D'une part, dans la réalité, nous pouvons rencontrer des formes qui appartiennent aux différentes classes qui se ressemblent globalement (Fig. 5.1a). D'autre part, il existe ainsi des formes de la même classe qui sont jugées comme différentes du point de vue global (Fig. 5.1b).

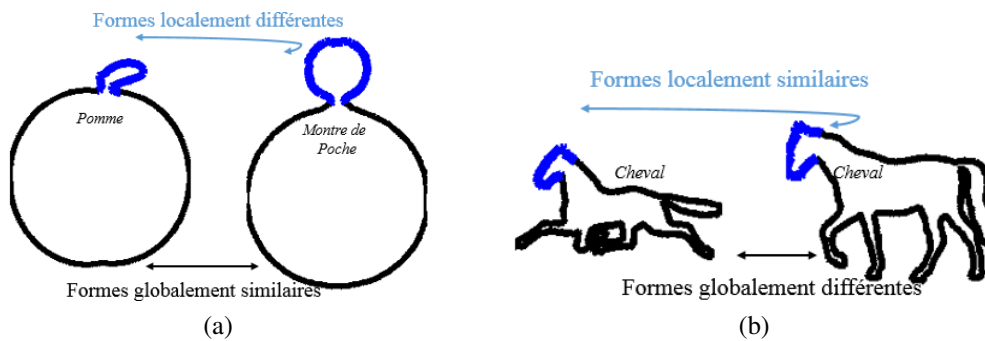


Figure 5.1 – (a) : Formes de différentes classes peuvent être globalement similaires. (b) : Formes de la même classe qui sont différentes du point de vue global.

Dans ce chapitre, une méthode robuste pour faire correspondre les parties de formes 2D fermées est proposée. Une méthode de décomposition des formes basée sur le principe de décomposition en des arcs convexes/concaves maximaux [123, 124] est tout d'abord représentée en section 5.4.1 et la distance robuste est ensuite développée en section 5.4.2. La mesure de similarité utilisée est basée sur l'approche géodésique proposée (au chapitre 3) pour la comparaison des segments de contours de formes. Enfin, la section 5.5 est consacrée à démontrer par des résultats expérimentaux la capacité de notre approche géodésique pour trouver la correspondance optimale entre les parties de formes 2D fermées.

5.2 Travaux antérieurs

Au cours des dernières années, de nombreux algorithmes de correspondance ont été proposés pour la classification et/ou la recherche des formes $2D$ fermées. La plupart des méthodes existantes dans la littérature pour la reconnaissance de formes $2D$ fermées sont essentiellement basées soit sur le contour soit sur le squelette. Par exemple, Bai *et al.* [15] ont proposé d'utiliser les chemins géodésiques pour faire correspondre les graphes de squelettes de formes. Malgré les résultats prometteurs démontrés par des nombreuses approches de correspondance de formes basées-squelette [12, 15, 208, 213], comme nous avons déjà expliqué, nous nous intéressons dans cette étude aux méthodes basées-contour. Toutefois, la complémentarité d'information que peut apporter la squelette pourrait être abordée comme perspective de ce travail en combinant notre méthode basée-contour à d'autres méthodes basées-squelette.

Entre temps, certaines méthodes basées-contour sont également développées. Belongie *et al.* [25] ont utilisé le contexte de forme comme descripteur de forme; Ling et Jacobs [139] ont remplacé ensuite la distance euclidienne utilisée dans l'approche du contexte de forme par la distance géodésique dans la forme pour traiter les articulations. Une fois le contour est représenté par une chaîne de symboles [58], une distance d'édition modifiée est utilisée pour calculer la similarité entre les chaînes de symboles dans [59, 60]. Dans [176], une nouvelle méthode variationnelle basée sur les géodésiques de formes dans l'espace de formes a été développée pour la classification et la recherche des formes $2D$ fermées. En plus, Wang *et al.* [246] ont défini des fonctions de chaleur pour décrire le contour de la forme afin de l'utiliser après pour la reconnaissance de formes $2D$. Dans [17], la combinaison de contour et de squelette pour la classification de forme a été bien étudiée en utilisant un modèle de mélange de gaussiennes. Toutes ces méthodes utilisent des algorithmes de correspondance globale qui peuvent mener à une reconnaissance inefficace dans le cas de formes qui sont similaires du point de vue global mais qui se différent localement en une partie de forme (Fig. 5.1a). En effet, la dissimilarité locale entre formes de différentes classes ne serait pas mise en évidence devant la similarité globale des formes entières. En plus, les méthodes de reconnaissance globale deviennent inutiles quand une partie significative de la forme donnée est articulée.

En général, la perception humaine de la forme est basée sur des parties visuelles des objets. Dans ce concept, les parties visuelles significatives sont essentielles pour la reconnaissance de l'objet entier. Par exemple, une main sur la porte signifie qu'un être humain se trouve derrière la porte. Cette hypothèse cognitive a motivé récemment les chercheurs à proposer des systèmes de reconnaissance de formes à partir des parties de formes [14, 17, 71, 119, 123, 124, 223, 247]. Latecki *et al.* [119, 123, 124] ont développé une méthode de correspondance qui cherche à trouver la correspondance optimale entre les parties significatives de formes, obtenues par la méthode DCE, en utilisant un algorithme de programmation dynamique. En se basant sur le fait que tous les fragments

de contour contiennent à la fois des informations locales et globales, les parties de forme sont utilisées en tant que des caractéristiques clés de formes pour l'apprentissage d'un livre-code¹ de formes [14, 247]. Une nouvelle représentation de formes, appelée sac de fragments de contour (*BCF: Bag of contour fragments*), inspirée du sac classique de mots (*BoW: Classical bag of words*) est récemment développée dans [247]. De même, un nouveau descripteur de forme, le vecteur de fragments de contour agrégées (*VACF: Vector of aggregated contour fragments*), est appliqué pour la classification de formes [14]. Étant représenté toute partie de forme par une courbe 2D ouverte, différentes difficultés apparaissent lors de l'analyse des formes via leurs parties. Parmi ces difficultés, nous notons :

- la longueur de toute partie de forme,
- les transformations géométriques (translation, rotation et facteur d'échelle) et
- le problème de distorsion.

Dans ce chapitre, nous visons à développer un système de reconnaissance de formes à partir des parties de formes qui consiste à calculer la similarité globale à partir des similarités locales. Nous nous intéressons ainsi à trouver et à illustrer la correspondance optimale entre les parties significatives des formes données. Le système de correspondance proposé par Latecki et Lakamper [123] semble d'être le plus applicable à notre cas grâce à sa capacité : (a) à définir une configuration significative des parties de formes, (b) à gérer les différents problèmes indiqués ci-dessus et (c) à fournir la meilleure correspondance entre les différentes parties significatives des formes à comparer. Pour cela, nous résumons dans la section suivante le système de correspondance et la métrique quadratique proposée dans [123].

5.3 Système de correspondance basé sur la métrique quadratique

La recherche d'une correspondance optimale entre les parties de formes 2D fermées exige une compréhension approfondie de la structure locale de formes. Grâce à sa capacité à donner des parties significatives, nous nous sommes intéressés à la méthode de décomposition développée par Latecki et Lakamper [123].

5.3.1 Décomposition des formes 2D fermées

Étant donné un contour 2D fermé, le procédé de la décomposition en des parties significatives de formes selon l'algorithme de Latecki et Lakamper [123] comprend les étapes suivantes :

¹Codebook dans la littérature anglo-saxonne.

5.3. SYSTÈME DE CORRESPONDANCE BASÉ SUR LA MÉTRIQUE QUADRATIQUE

1. Trouver le polygone simplifié par la méthode DCE d'une forme fermée. Il est obtenu à son étape d'évolution optimale (Voir la section 2.5.2.1).
2. Extraire les arcs convexes/concaves maximaux constituant le polygone simplifié (Fig 5.2). Les arcs convexes/concaves maximaux ont été définies dans la section 2.5.2.2 du chapitre 2 de ce rapport.

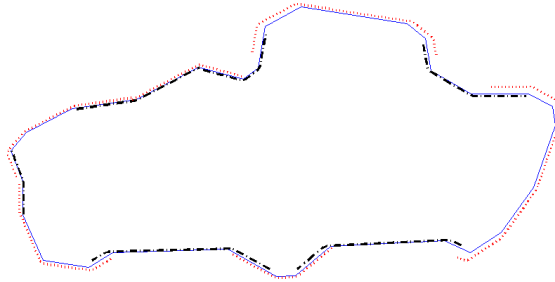


Figure 5.2 – Arcs convexes/concaves maximaux d'une forme d'une voiture de la base de données MPEG-7. Les arcs internes pointillés indiquent les arcs concaves tandis que les arcs externes pointillés se rapportent aux arcs convexes.

Une fois les formes sont décomposées en parties, la comparaison se fait par une mesure de similarité adaptée. Pour des fins de comparaison, nous donnons dans ce qui suit un rappel de la définition de la mesure de similarité quadratique adoptée dans [123].

5.3.2 Mesure de similarité quadratique

La plupart des approches proposées utilisent la norme \mathbb{L}^2 pour la correspondance et la comparaison des parties de formes [119, 123, 198, 231]. Étant donné un arc/partie d'un objet, une valeur significative doit être associée à chaque point échantillonné de la partie afin de la caractériser. La fonction d'angle [19, 123, 126, 231] et la fonction de courbure [57] sont les signatures de forme les plus utilisées en travaillant avec une distance quadratique. Grâce à sa capacité à gérer les transformations géométriques, nombreux sont les travaux qui ont adopté la fonction d'angle comme une signature descriptive. Dans la partie expérimentale, nous allons comparer notre système de reconnaissance de formes basée sur la correspondance des parties de formes avec celui proposé par Latecki et Lakamper [123] qui est basé sur une distance quadratique. Par conséquent, il semble important de donner un bref aperçu de leur mesure de similarité.

Étant donné deux courbes polygonales fermées \mathcal{P}_1 et \mathcal{P}_2 , la similarité entre deux arcs a_1 et a_2 de \mathcal{P}_1 et de \mathcal{P}_2 respectivement, est définie comme suit:

$$m_{partielle}^q(a_1, a_2) = d_q(a_1, a_2) \times \max(l(a_1), l(a_2)) \times \max\left(\frac{l(a_1)}{l(a_2)}, \frac{l(a_2)}{l(a_1)}\right), \quad (5.1)$$

$$\text{avec } d_q(a_1, a_2) = \int_0^1 (\theta_1(s) - \theta_2(s) + \theta_0)^2 ds \quad (5.2)$$

où θ_1 et θ_2 représentent les fonctions d'angle des arcs a_1 et a_2 respectivement. A noter que ces fonctions sont paramétrées par la longueur d'arc s normalisée. Cependant, l est la longueur d'arc relative à la courbe fermée. L'intégrale, autrement dite la distance quadratique angulaire $d_q(a_1, a_2)$, est pondérée par la longueur d'arc pénalisée par la différence des longueurs des arcs correspondants. L'invariance par rotation est apportée en minimisant l'intégrale avec la constante $\theta_0 \in [0, 2\pi)$.

Une comparaison une à une de toutes les paires des arcs convexes/concaves maximaux est inutile et coûteuse en temps d'exécution. Pour établir la correspondance la plus significative entre les formes, une partie/arc est considérée comme groupe, autrement dit une union, d'une séquence non-vide consécutive des arcs convexes/concaves maximaux. Pour une courbe \mathcal{P}_i , un groupement G_i est composé d'un ensemble ordonné de groupes consécutifs couvrant la courbe toute entière. Étant désigné par $\mathcal{C}(\mathcal{P}_i)$ l'ensemble de tous les groupements possibles d'une courbe \mathcal{P}_i , le fait de trouver la correspondance optimale entre les parties des formes fermées revient à optimiser la mesure de similarité globale définie par l'équation suivante :

$$m_{globale}^q(\mathcal{P}_1, \mathcal{P}_2) = \min_{(G_1, G_2) \in \mathcal{C}(\mathcal{P}_1, \mathcal{P}_2)} \left\{ \sum_{a \in G_1} m_{partielle}^q(a, f_{(G_1, G_2)}(a)) \right\}, \quad (5.3)$$

où $\mathcal{C}(\mathcal{P}_1, \mathcal{P}_2)$ est l'ensemble de toutes les paires correspondantes (G_1, G_2) dans $\mathcal{C}(\mathcal{P}_1) \times \mathcal{C}(\mathcal{P}_2)$ sous la contrainte qu'au moins un de chaque deux arcs correspondants doit être un arc convexe/concave maximal. La bijection $f_{(G_1, G_2)}$ est la fonction de correspondance entre G_1 et G_2 . La minimisation de la mesure de similarité globale (Eq. 5.3) est résolue par l'utilisation d'une méthode de programmation dynamique.

Limitations : Plusieurs limitations se posent à l'application de ce système de correspondance; les principales sont :

1. la dépendance du nombre des arcs convexes/concaves vis à vis l'abondance de détails dans les formes à décomposer. Plus la forme contient de détails, plus le nombre des arcs convexes/concaves maximaux est grand, ce qui aboutit à un grand nombre de groupes et par suite un grand nombre de groupements à tester en minimisant la mesure de similarité globale $m_{globale}^q$.
2. la non optimalité de la méthode (méthode angulaire) utilisée pour assurer l'invariance aux transformations géométriques (voir le chapitre 4).

3. la grande similitude entre les courbes polygonales dans les cas où les formes à comparer sont des formes de différentes classes mais qui sont très similaires globalement.

5.4 Approche proposée

D'une part, des formes d'une même classe peuvent présenter de grandes variations à cause des problèmes de pose, de déformation et d'occlusion. D'autre part, des formes de différentes classes peuvent avoir plusieurs parties similaires en commun, et la variation peut être représentée par des petits détails qui constituent les caractéristiques de discrimination locaux de formes. Pour cela, nous exigeons que toute partie de forme, cependant petite ou grande, intervienne en correspondance des formes. Pour cette raison, nous proposons ici une méthode de décomposition de formes basée sur le DCE qui cherche à représenter toute forme $2D$ fermée en ses parties significatives. A noter que l'ensemble des parties de formes extraites forme un recouvrement de la forme entière. Dans ce qui suit, nous développons une nouvelle méthode de décomposition des formes $2D$ fermées et une nouvelle métrique basée sur les géodésiques de formes pour les utiliser en reconnaissance de formes à partir des parties de formes.

5.4.1 Méthode de décomposition proposée

Afin de franchir les limitations (citées ci-dessus) de la méthode de décomposition utilisée dans [123], nous proposons une nouvelle configuration élémentaire des parties significatives de formes $2D$ fermées. Nous avons déjà expliqué que l'utilisation des arcs convexes/concaves maximaux comme des configurations élémentaires des parties de formes aboutit à un temps de calcul énorme dans les cas où les formes à décomposer comprennent beaucoup de détails. Nous allons définir nos arcs élémentaires de façon à limiter leur nombre quelle que soit la forme. Étant données deux formes $2D$ fermées F et \tilde{F} , nous développons les nouvelles étapes suivantes :

1. Normaliser les formes par rapport aux transformations géométriques afin de trouver leurs origines optimales. L'utilisation de la signature angulaire $\theta(s)$ (voir la section 1.5 du chapitre 1) réalise naturellement l'invariance à la translation et au changement d'échelle. En effet, toute translation de la courbe n'a aucune incidence sur θ , et toute homothétie n'a aucun effet sur le paramètre normalisé s . Une rotation d'angle c transforme la fonction $\theta(s)$ en $\theta(s) + c$ modulo 2π . Pour assurer l'invariance par rotation, la minimisation de $MS^{F\tilde{F}}(\phi)$ (Eq. 3.1) pour tous les choix d'origines des courbes fermées F et \tilde{F} est considérée.
2. Décomposer le polygone simplifié d'une forme fermée donnée en des arcs convexes/concaves maximaux (voir la section 5.3.1).

3. Définir les arcs élémentaires a_{el} comme des arcs du polygone simplifié composés de l'union d'une séquence consécutive des arcs convexes/concaves maximaux à condition que sa longueur $l(a_{el})$ vaille 10% fois la longueur L du polygone simplifié ($0,1.L - \epsilon \leq l(a_{el}) \leq 0,1.L + \epsilon$). Noter que cette combinaison des arcs convexes/concaves commence par l'arc initial auquel appartient l'origine de la courbe (voir l'étape 1).
4. Extraire les parties élémentaires P_{el} du contour de la forme d'origine délimitées par les arcs élémentaires a_{el} du polygone simplifié.

Nous avons défini ainsi les arcs élémentaires a_{el} de manière que quel que soit la quantité de détails d'une forme donnée, elle se décompose en moins de onze arcs élémentaires a_{el} . En plus, comme l'objectif de notre travail est de chercher la correspondance optimale des parties significatives des formes fermées, nous avons proposé de travailler avec les parties P_{el} correspondantes aux arcs a_{el} sur le contour de la forme d'origine et non pas sur les arcs du polygone simplifié comme ont fait Latecki et Lakamper [123].

5.4.2 Distance utilisée

Dans cette section, nous introduisons une nouvelle méthode pour la correspondance des parties de formes 2D fermées qui exploite une métrique robuste basée sur les géodésiques de formes. Étant donné deux parties P_1 et P_2 de deux courbes fermées \mathcal{P}_1 et \mathcal{P}_2 respectivement, nous utilisons tout d'abord notre algorithme, proposé dans la section 4.3.4 du chapitre 4, basé sur l'alignement Procuste et la redétermination des points extrémités pour atteindre l'invariance des parties de formes aux transformations géométriques. Dans le système de correspondance proposé dans [123], une comparaison une à une de toutes les paires des parties élémentaires P_{el} est inutile et coûteuse en temps de calcul. Pour cela, nous considérons qu'une partie significative P d'une forme donnée est une union d'une séquence non-vide consécutive de parties élémentaires de la courbe d'origine. Par suite, la mesure de similarité basée sur les géodésiques de formes entre les deux parties significatives P_1 et P_2 est définie comme suit :

$$m_{partielle}^g(P_1, P_2) = d_{Geod}(P_1, P_2) \times \max(l(P_1), l(P_2)) \times \max\left(\frac{l(P_1)}{l(P_2)}, \frac{l(P_2)}{l(P_1)}\right), \quad (5.4)$$

où $d_{Geod}(P_1, P_2)$ (équation 3.2) est la métrique *élastique* basée sur les géodésiques de formes proposée pour comparer les parties de formes.

Finalement, la mesure de similarité *élastique* globale entre les courbes non polygonales \mathcal{P}_1 et \mathcal{P}_2 est donnée par l'équation suivante :

$$m_{globale}^g(\mathcal{P}_1, \mathcal{P}_2) = \min_{(G_1, G_2) \in \mathcal{C}(\mathcal{P}_1, \mathcal{P}_2)} \left\{ \sum_{P \in G_1} m_{partielle}^g(P, f_{(G_1, G_2)}(P)) \right\}, \quad (5.5)$$

Pour une courbe \mathcal{P}_i , nous désignons par $\mathcal{C}(\mathcal{P}_i)$ l'ensemble de tous les groupements possibles de la courbe dont la première partie commence par la partie élémentaire comprenant l'origine de la courbe. Ceci restreint la minimisation de la mesure de similarité globale (équation 5.5) à l'ensemble $\mathcal{C}(\mathcal{P}_1, \mathcal{P}_2)$ des paires de groupements qui débutent chacun par la partie qui contient l'origine. Ainsi, toutes les permutations avec d'autres initialisations sont exclues. Par conséquent, notre système de correspondance est beaucoup moins coûteux en temps d'exécution que celui de Latecki et Lakamper [123].

5.5 Applications sur la base de formes MPEG-7

L'approche proposée pour la correspondance des parties significatives de formes 2D fermées (section 5.4) est validée par une expérience quantitative sur des formes de la base d'images MPEG-7. Aux fins d'une évaluation comparative, nous avons utilisé notre méthode de décomposition et notre système de correspondance, proposés dans la section 5.4, avec trois distances différentes qui sont :

- la distance angulaire quadratique d_q (equation 5.2),
- la distance de contexte de forme d_{SC} (equation 1.22) et
- la distance de Procuste d_{Pr} (equation 1.21).

Même si la première étape de notre méthode de décomposition exige une normalisation des formes fermées, les articulations et les déformations locales de formes nécessitent l'alignement des parties de formes à comparer. Dans la suite, l'alignement angulaire [231], la représentation dans un repère canonique [19] et l'alignement Procuste [71, 154] sont utilisés avec la distance quadratique angulaire, la distance de contexte de forme et la distance de Procuste respectivement.

5.5.1 Description des bases de test

Dans ce travail, nous ne nous intéressons pas à la classification et la recherche sur la base MPEG-7 toute entière, car un tel travail nécessite une optimisation du choix des parties significatives, qui ne fait pas partie de ce travail. Par contre, nous nous intéressons ici aux sous-ensembles de la base où il y a eu de confusions avec la classification sur la forme entière directement [178]. Pour cela, nous allons étudier précisément des paires de classes où il existe des formes globalement similaires. En effet, nous utilisons trois sous-bases des formes de la base MPEG-7. Chaque sous-base se compose des formes de deux classes différentes. Plus précisément :

- Sous-base 1 : La première contient la classe des *pommes* et la classe des *montres de poche* (Fig. 5.3a).

CHAPITRE 5. RECONNAISSANCE DES FORMES ENTIÈRES À PARTIR DES PARTIES DE FORMES

- Sous-base 2 : La deuxième contient la classe des *éléphants* et la classe des *octopus* (Fig. 5.3b).
- Sous-base 3 : La troisième contient la classe des *chevaux* et la classe des *chiens* (Fig. 5.3c).

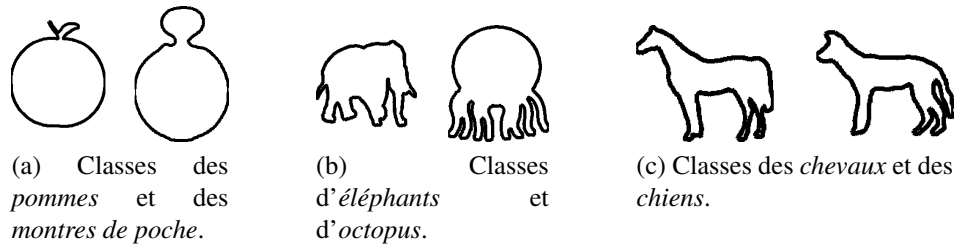


Figure 5.3 – Trois sous-bases de la base d’images MPEG-7 dont chacune contient deux classes différentes de formes qui sont similaires du point de vue globale.

L’approche proposée pour la correspondance et la reconnaissance de formes 2D fermées à partir des parties de formes est implémentée avec les quatre distances citées ci-dessus. Ensuite, nous comparons la performance de chacune des distances pour la recherche et la classification des formes des sous-bases considérées où chaque forme est considérée comme une requête.

5.5.2 Résultats et discussion

Lors de la recherche des formes fermées, le taux de précision de recherche de chaque requête est mesuré comme le pourcentage des formes qui appartiennent à la même classe parmi les 20 formes recherchées les plus similaires à la requête en utilisant l’équation 3.5.

Tableau 5.1 – Taux de recherche obtenus en utilisant les approches : géodésique, angulaire, contexte de forme et Procuste sur les formes de chaque classe des trois sous-bases (Fig. 5.3).

Taux de recherche	Sous-base 1		Sous-base 2		Sous-base 3	
	Classes		Classes		Classes	
	<i>Pommes</i>	<i>Montres de poche</i>	<i>Éléphants</i>	<i>Octopus</i>	<i>Chiens</i>	<i>Chevaux</i>
Approche géodésique	89.50%	94.25%	80.75%	57.50%	76.50%	57.75%
Approche angulaire	69.75%	93.50%	72.50%	57.25%	87.75%	60.75%
Approche du contexte de forme	87%	91.25%	67.25%	51.50%	73.75%	61.75%
Approche Procuste	72.75%	91.25%	68.25%	45.25%	78.25%	49.75%

Grâce à l’utilisation de notre métrique basée sur les géodésiques dans l’espace de formes (équation 5.5), sur les formes des trois sous-bases (Fig. 5.3), les taux moyens de précision

5.5. APPLICATIONS SUR LA BASE DE FORMES MPEG-7

de recherche montent à 89.50%, 94.25%, 80.75% et 57.50% pour les classes des *pommes*, *montres de Poche*, *éléphants* et des *octopus* respectivement (voir le tableau 5.1). Ces moyennes atteignent des taux inférieurs avec l'approche angulaire, l'approche basée sur le contexte de forme et l'approche Procuste. En appliquant les mêmes approches sur l'ensemble des requêtes des deux classes (*chiens* et *chevaux*) de la sous-base 3, l'approche géodésique conduit à un taux de précision de recherche plus bas que celui obtenu avec l'approche angulaire et l'approche Procuste sur la classe des *chiens*. De même, la méthode angulaire et la méthode du contexte de forme atteignent des taux de précision de recherche qui dépassent celui obtenu par notre méthode géodésique. Ce manque de performance de notre méthode pour la recherche des formes des classes de *chiens* et de *chevaux* est expliqué par une interprétation approfondie sur le degré d'optimalité de la correspondance entre les parties de formes trouvée par les métriques étudiées.

Dans les figures 5.4 et 5.5, nous illustrons quelques exemples des correspondances optimales obtenues par les trois méthodes. Les parties similaires dans les paires des formes ont été correctement associées par la méthode géodésique. Les correspondances obtenues par les autres méthodes s'éloignent de la vraie similitude entre les parties.

Comme les formes de ces deux animaux sont très similaires à un facteur d'échelle près, l'invariance à l'échelle de notre méthode rend la différence entre ces formes minimale, ce qui explique la forte confusion entre ces deux classes. Cette confusion a été plus petite avec les autres méthodes pas à cause d'une comparaison meilleure, mais à cause de correspondances des parties non optimales. Pour plus de détails sur les correspondances optimales obtenues, nous donnons dans les tableaux 5.2 à 5.5 des évaluations des correspondances optimales obtenues entre différentes requêtes et individus de la base. Le tableau 5.2 et le tableau 5.3 contiennent les paires (requête, individu) de formes de la classe de *chiens* pour lesquelles la méthode angulaire et la méthode Procuste attribuent respectivement une mauvaise correspondance entre leurs parties. À noter que chaque individu associé à une requête est recherché parmi les 20 formes les plus similaires à cette même requête. De même, le tableau 5.4 et le tableau 5.5 présentent les paires de formes de la classe de *chevaux* qui se trouvent avec une mauvaise correspondance de parties de formes lors de l'utilisation de la méthode angulaire et la méthode du contexte de forme. Cependant, la correspondance obtenue par notre métrique est toujours correcte et optimale.

Suite à une comparaison quantitative des performances de recherche suivie d'une comparaison qualitative des performances de correspondance des méthodes étudiées, nous pouvons en déduire que notre approche géodésique surpasse la méthode angulaire, la méthode du contexte de forme et la méthode Procuste en recherche des formes fermées et en correspondance de leurs parties significatives.

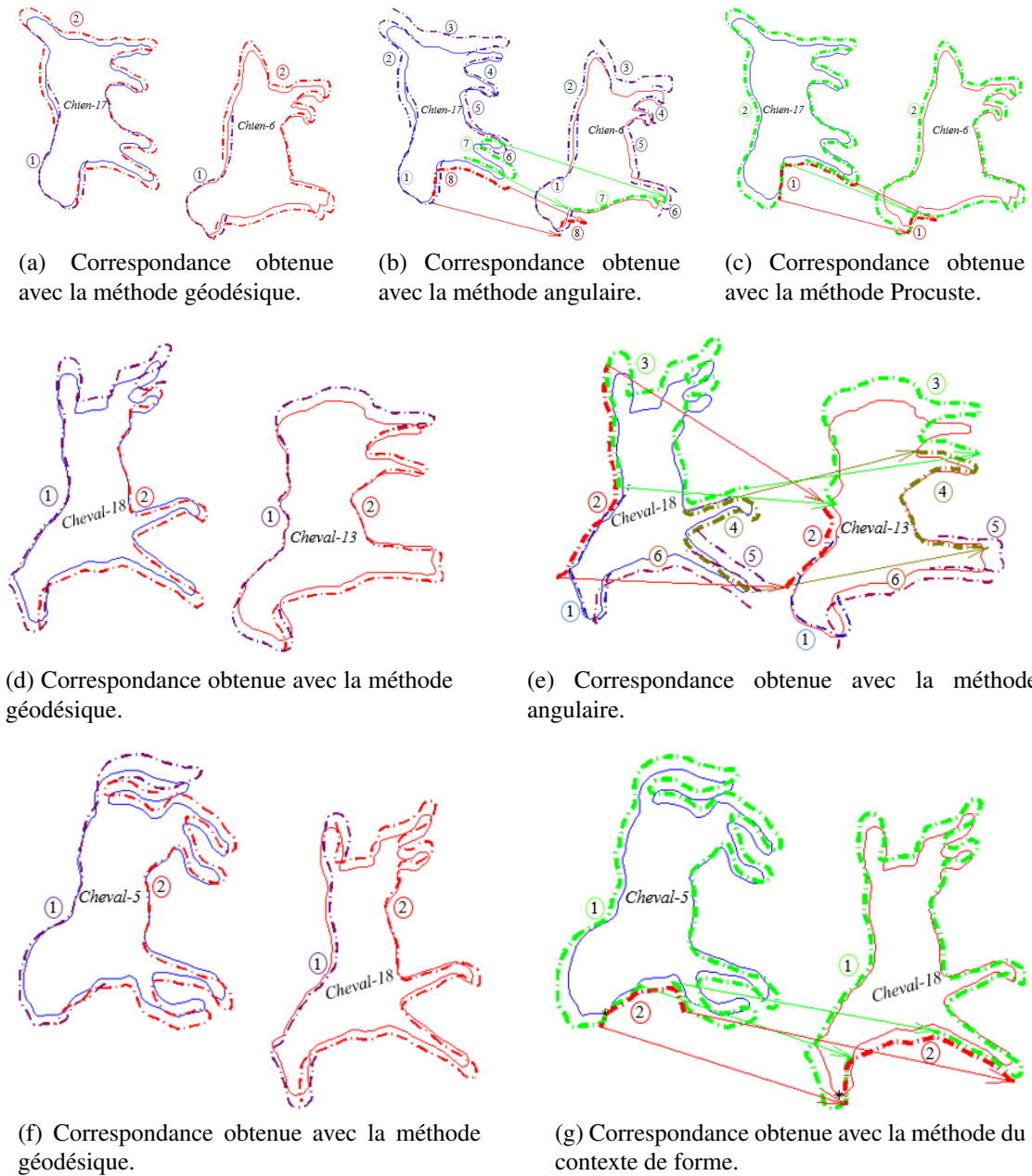


Figure 5.4 – Visualisation de la correspondance optimale entre les parties de différentes formes obtenue par la méthode géodésique (a-d-f), la méthode angulaire (b-e), la méthode du contexte de forme (g) et la méthode Procuste (c).

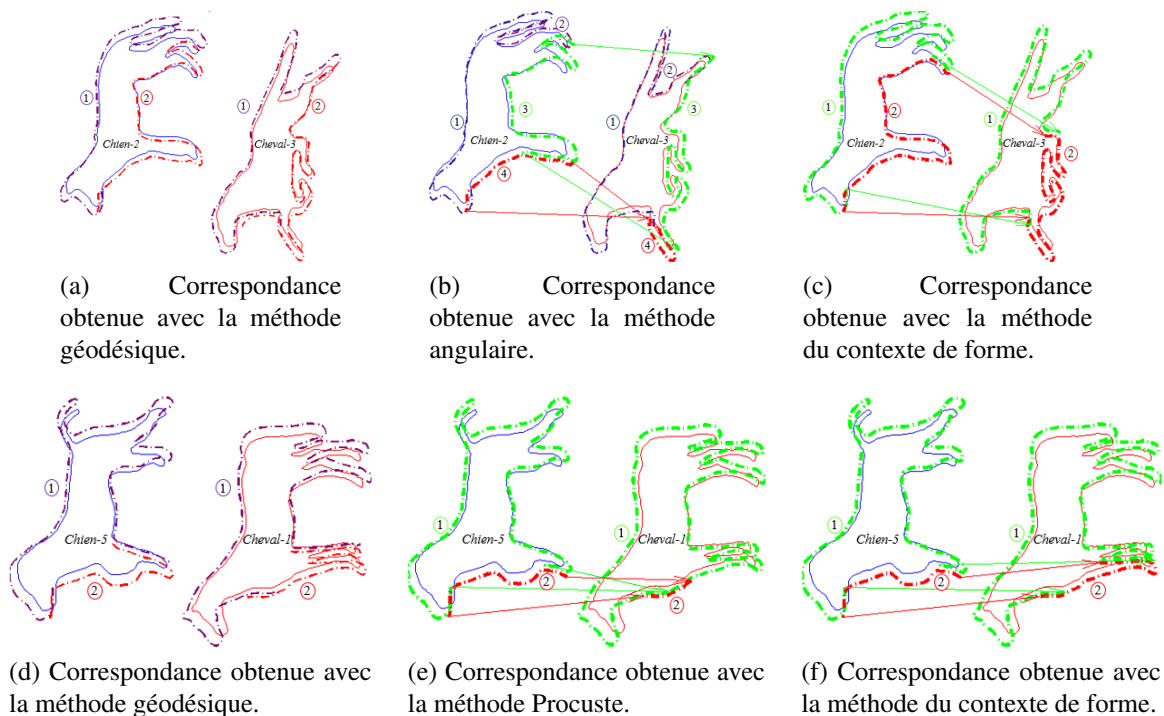


Figure 5.5 – Visualisation de la correspondance optimale entre les parties de différentes formes obtenue par la méthode géodésique (a-d), la méthode angulaire (b), la méthode du contexte de forme (c-f) et la méthode Procuste (e).

Notre système proposé pour la reconnaissance des formes fermées à partir des parties de formes est également exploité pour la classification des mêmes formes des trois sous-bases de la figure 5.3. Le tableau 5.6 montre qu’avec les classifieurs *1-NN*, *5-NN*, *10-NN*, *15-NN* et *20-NN*, les taux moyens de classification des formes montent respectivement à 100%, 100%, 100%, 93.33%, 92.5% avec notre approche géodésique et ils diminuent pour atteindre des taux inférieurs avec l’approche angulaire, l’approche du contexte de forme et l’approche Procuste.

Notons enfin que la performance de la reconnaissance a été améliorée pour ces classes en utilisant les parties par rapport à la reconnaissance en utilisant la forme entière [178]. Le tableau 5.7 compare les taux de recherche sur ces sous-bases entre les deux stratégies. À partir de ces résultats, nous pouvons constater que la comparaison basée sur les parties de formes peut réduire les erreurs de reconnaissance lorsque les formes de différentes classes ne diffèrent qu’en des parties de la forme.

Tableau 5.2 – Les paires (requête, individu) des formes de la classe des *Chiens* qui réalisent une mauvaise correspondance entre leurs parties en utilisant la méthode angulaire et une bonne correspondance des parties de formes avec l’approche géodésique. A noter que chaque individu est recherché parmi les 20 formes les plus similaires à sa propre requête.

Correspondance des parties			
Requête	Individu	Méthode angulaire	Méthode géodésique
<i>Chien-5</i>	<i>Chien-17</i>	Mauvaise	Bonne
<i>Chien-6</i>			
<i>Chien-8</i>			
<i>Chien-9</i>			
<i>Chien-11</i>			
<i>Chien-18</i>			
<i>Chien-10</i>	<i>Chien-8</i>	Mauvaise	Bonne
	<i>Chien-11</i>	Mauvaise	Bonne
<i>Chien-17</i>	<i>Chien-5</i>	Mauvaise	Bonne
	<i>Chien-6</i>	Mauvaise	Bonne
	<i>Chien-8</i>	Mauvaise	Bonne
	<i>Chien-9</i>	Mauvaise	Bonne
	<i>Chien-11</i>	Mauvaise	Bonne
	<i>Chien-20</i>	Mauvaise	Bonne

Tableau 5.3 – Les paires (requête, individu) des formes de la classe des *Chiens* qui réalisent une mauvaise correspondance entre leurs parties en utilisant la méthode Procuste et une bonne correspondance des parties de formes avec l’approche géodésique. A noter que chaque individu est recherché parmi les 20 formes les plus similaires à sa propre requête.

Correspondance des parties			
Requête	Individu	Méthode Procuste	Méthode géodésique
<i>Chien-2</i>	<i>Chien-18</i>	Mauvaise	Bonne
<i>Chien-3</i>			
<i>Chien-5</i>			
<i>Chien-10</i>			
<i>Chien-11</i>			
<i>Chien-12</i>			
<i>Chien-13</i>			
<i>Chien-14</i>			
<i>Chien-16</i>			
<i>Chien-17</i>			
<i>Chien-20</i>			
<i>Chien-11</i>	<i>Chien-6</i>	Mauvaise	Bonne
<i>Chien-12</i>			
<i>Chien-13</i>			
<i>Chien-17</i>			

CHAPITRE 5. RECONNAISSANCE DES FORMES ENTIÈRES À PARTIR DES PARTIES DE FORMES

Tableau 5.4 – Les paires (requête, individu) des formes de la classe des *Chevaux* qui réalisent une mauvaise correspondance entre leurs parties en utilisant la méthode angulaire et une bonne correspondance des parties de formes avec l’approche géodésique. A noter que chaque individu est recherché parmi les 20 formes les plus similaires à sa propre requête.

Correspondance des parties			
Requête	Individu	Méthode angulaire	Méthode géodésique
<i>Cheval-2</i>	<i>Cheval-3</i>	Mauvaise	Bonne Bonne Bonne
	<i>Cheval-17</i>	Mauvaise	
	<i>Cheval-18</i>	Mauvaise	
<i>Cheval-3</i> <i>Cheval-17</i>	<i>Cheval-2</i>	Mauvaise	Bonne
<i>Cheval-13</i> <i>Cheval-20</i>	<i>Cheval-17</i>	Mauvaise	Bonne
<i>Cheval-13</i> <i>Cheval-20</i>	<i>Cheval-18</i>	Mauvaise	Bonne
<i>Cheval-18</i>	<i>Cheval-2</i>	Mauvaise	Bonne Bonne
	<i>Cheval-13</i>	Mauvaise	

Tableau 5.5 – Les paires (requête, individu) des formes de la classe des *Chevaux* qui réalisent une mauvaise correspondance entre leurs parties en utilisant la méthode du contexte de forme et une bonne correspondance des parties de formes avec l’approche géodésique. A noter que chaque individu est recherché parmi les 20 formes les plus similaires à sa propre requête.

Correspondance des parties			
Requête	Individu	Méthode du contexte de forme	Méthode géodésique
<i>Cheval-5</i>	<i>Cheval-17</i>	Mauvaise	Bonne Bonne
	<i>Cheval-18</i>	Mauvaise	
<i>Cheval-8</i> <i>Cheval-11</i> <i>Cheval-12</i>	<i>Cheval-9</i>	Mauvaise	Bonne
<i>Cheval-9</i>	<i>Cheval-11</i>	Mauvaise	Bonne Bonne
	<i>Cheval-12</i>	Mauvaise	
<i>Cheval-18</i>	<i>Cheval-9</i>	Mauvaise	Bonne Bonne
	<i>Cheval-14</i>	Mauvaise	

5.5. APPLICATIONS SUR LA BASE DE FORMES MPEG-7

Tableau 5.6 – Taux de classification correcte obtenus avec les classifieurs *1-NN*, *5-NN*, *10-NN*, *15-NN* et *20-NN* en utilisant les approches : géodésique, angulaire, contexte de forme et Procuste sur chaque classe des trois sous-bases (Fig. 5.3).

Classification	Approche géodésique					Approche angulaire					Approche du contexte de forme					Approche Procuste						
	<i>1-NN</i>	<i>5-NN</i>	<i>10-NN</i>	<i>15-NN</i>	<i>20-NN</i>	<i>1-NN</i>	<i>5-NN</i>	<i>10-NN</i>	<i>15-NN</i>	<i>20-NN</i>	<i>1-NN</i>	<i>5-NN</i>	<i>10-NN</i>	<i>15-NN</i>	<i>20-NN</i>	<i>1-NN</i>	<i>5-NN</i>	<i>10-NN</i>	<i>15-NN</i>	<i>20-NN</i>		
Sous-base 1	Classes	Pommes	100%	100%	100%	100%	100%	95%	100%	90%	90%	100%	100%	100%	100%	100%	100%	100%	100%	100%	95%	95%
		Montres de poche	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Sous-base 2	Classes	Éléphants	100%	100%	100%	100%	100%	100%	100%	95%	100%	100%	100%	100%	95%	95%	100%	100%	100%	95%	95%	
		Octopus	100%	100%	100%	90%	80%	100%	95%	95%	80%	80%	100%	100%	90%	85%	60%	100%	90%	85%	60%	45%
Sous-base 3	Classes	Chiens	100%	100%	100%	95%	100%	100%	100%	100%	100%	100%	90%	90%	90%	100%	100%	100%	100%	100%	100%	100%
		Chevaux	100%	100%	100%	75%	75%	100%	100%	85%	85%	85%	100%	95%	90%	80%	85%	100%	95%	95%	80%	60%
Taux total			100%	100%	100%	93.33%	92.5%	100%	98.33%	96.67%	91.67%	92.5%	100%	97.5%	95%	91.67%	90%	100%	97.5%	96.67%	88.33%	82.5%

Tableau 5.7 – Comparaison des taux de recherche sur les mêmes sous-bases étudiées dans le tableau 5.6 obtenues par les stratégies de la reconnaissance basée-parties de formes et la reconnaissance en utilisant la forme entière.

		Comparaison des formes entières		Comparaison basée-parties de formes	
Sous-base 1	Classes	Pommes	79.5%	89.5%	
		Montres de poche	94.25%	94.25%	
Sous-base 2	Classes	Éléphants	78.5%	80.75%	
		Octopus	50.5%	57.5%	
Sous-base 3	Classes	Chiens	79.25%	87.75%	
		Chevaux	60%	61.75%	

5.6 Conclusion

Nous avons proposé dans ce chapitre d'utiliser les géodésiques de formes dans la définition d'une nouvelle métrique robuste globale pour la reconnaissance de formes $2D$ fermées. Cette métrique est basée sur la métrique proposée dans le chapitre 3 pour la reconnaissance des parties de formes et sur la correspondance des parties de formes à comparer. Nous avons montré l'efficacité de cette métrique par des tests de reconnaissance sur des sous-bases de la base d'images MPEG-7 qui comprennent des classes différentes de formes qui sont globalement très similaires.

Au cours du chapitre, nous avons également analysé les bénéfices et les limites de la méthode de correspondance développée et la distance proposée. Pour cette raison, nous avons comparé notre système géodésique de reconnaissance des formes avec trois autres systèmes qui sont basés sur l'utilisation de la distance angulaire, la distance du contexte de forme, et la distance Procuste respectivement. Nous avons démontré expérimentalement que notre stratégie proposée de reconnaissance des formes $2D$ fermées donne des meilleures correspondance entre les parties de formes lorsqu'on l'applique avec notre distance géodésique tandis que cette performance diminue avec l'utilisation des trois autres distances.

Conclusion générale et perspectives

Dans cette thèse, nous avons étudié le problème de correspondance des parties de formes pour des fins de reconnaissance (classification et recherche). Pour cela, nous avons développé de nouveaux systèmes de reconnaissance :

- un système de reconnaissance de parties de formes,
- un système de reconnaissance de formes partielles et
- un système de reconnaissance de formes entières à partir des parties de formes.

Une partie de forme peut être obtenue soit par une décomposition d'une forme entière, soit par une segmentation appliquée à une image et donnant des fragments de contours. La détermination d'une partie de forme peut être affectée par différentes sources de distorsions telles que le bruit et l'occlusion conduisant surtout à des perturbations au niveau des extrémités des parties lors de la comparaison. Afin d'assurer l'invariance des parties de formes aux transformations géométriques (translation, rotation et facteur d'échelle) nous avons proposé tout d'abord une nouvelle méthode de représentation invariante des parties de formes. Dans la littérature, toutes les méthodes utilisées pour la normalisation des parties de forme exigent une correspondance fixée des points extrémités des parties de formes à comparer. Une telle fixation n'est cependant pas toujours adéquate et conduit à un résultat sous optimal. Pour cela, nous avons développé une méthode basée sur l'alignement Procuste et sur la redétermination optimale des points extrémités.

Pour une comparaison locale des parties de formes et des formes, nous avons proposé des distances robustes partielles et globales respectivement, basées sur le résiduel du recalage de contours de parties de formes $2D$. L'approche de recalage utilisée est basée sur le calcul des géodésiques de formes dans l'espace de formes. La robustesse aux inter-variabilités est assurée par l'utilisation d'un critère robuste dans la définition de la mesure de similarité voulue. Les distances proposées, basées sur les géodésiques, permettent aussi de comparer localement les parties de formes dans un cadre multi-échelle.

Au niveau expérimental, nous avons réalisé une étude comparative entre la méthode d'orientation proposée et les méthodes de la littérature sur des données synthétiques puis

sur des parties de formes de la base MPEG-7. Nous avons montré ainsi une grande stabilité aux perturbations des extrémités et une supériorité de notre méthode par rapport aux autres méthodes de l'état de l'art.

Pour les systèmes de reconnaissance proposés, nous avons d'abord comparé l'approche géodésique proposée pour la reconnaissance des parties de formes à la méthode du contexte de forme et la méthode Procuste les plus utilisées dans la littérature pour la reconnaissance des parties de formes. Les résultats expérimentaux obtenus ont montré la supériorité en performance de notre méthode proposée et ainsi sa grande capacité à surmonter les déformations élastiques et les articulations par rapport aux deux autres méthodes de l'état de l'art. Les taux de précision de recherche et de classification correcte atteignent respectivement 92.46% et 98.55% en utilisant notre système de reconnaissance de parties de formes sur l'ensemble des parties requêtes de longueurs importantes $E1_{\text{Ét}}$ (figure 3.9b). Deuxièmement, nous avons expérimentalement montré que l'utilisation de la stratégie de reconnaissance des formes partielles proposée avec la distance géodésique proposée est significativement plus efficace que celle utilisée avec la distance du contexte de forme et la distance Procuste. Notre système de reconnaissance de formes partielles atteint respectivement 77.25% et 91.30% de taux de précision de recherche et de classification correcte sur un ensemble de formes partielles. Dernièrement, nous avons montré l'efficacité de notre système proposé pour la reconnaissance des formes fermées 2D à partir des parties de formes. Les tests de reconnaissance sont réalisés sur des sous-bases de la base d'images MPEG-7 comprenant des classes différentes de formes qui sont globalement très similaires. Le tableau 5.8 reporte les différents résultats obtenus selon la distance et les algorithmes de reconnaissance de parties de formes et de formes partielles utilisées.

Tableau 5.8 – Taux de recherche et de classification correcte obtenus en utilisant l'approche géodésique proposée sur l'ensemble des requêtes $E1$ (figure 3.9a), $E1_{\text{Ét}}$ (figure 3.9b) et sur les formes partielles formées d'une seule partie (Q_1 et Q_2 de l'ensemble des requêtes $E1$ (figure 3.9a) et $E2$ (figure 3.14) respectivement) et d'une combinaison de deux parties ((Q_1, Q_2)).

Taux		Parties de formes		Formes partielles		
		$E1$	$E1_{\text{Ét}}$	Q_1	Q_2	(Q_1, Q_2)
Recherche		60.58%	92.46%	48.99%	46.09%	77.25%
Classification	<i>1-NN</i>	85.51%	98.55%	84.06%	78.26%	91.30%
	<i>5-NN</i>	76.81%	100%	68.12%	76.81%	91.30%
	<i>10-NN</i>	68.12%	100%	63.77%	65.22%	88.44%

A noter que les approches de reconnaissance proposées ont été testées sur des parties de formes et sur des formes de la base test d'images MPEG-7 puis sur des images réelles afin d'évaluer leurs performances. Suite à des comparaisons avec les méthodes déjà

utilisées et proposées dans la littérature, nous avons montré que nos approches sont assez performantes.

Comme perspectives, nous prévoyons les poursuites d'études suivantes :

- Élargir les applications des systèmes de reconnaissance proposés sur toutes les formes de la base de formes MPEG-7 et sur d'autres bases. À ces fins, nous proposons de développer une technique automatisée basée sur une analyse statistique des formes. L'analyse statistique dans une telle technique nous permettra d'extraire les parties de formes les plus significatives d'une forme entière donnée. Ces parties devront être les plus utiles pour la caractérisation d'une classe de formes parmi d'autres et par suite pour la reconnaissance de formes.
- Pour la reconnaissance de la forme entière à partir des parties, il sera intéressant d'évaluer l'avantage d'utiliser le concept de livre-code développé dans [14] avec notre système de reconnaissance. Dans [14], Bai a proposé une nouvelle méthode de classification de formes basée sur le contexte de forme et la création de livre-code pour comparer les formes à partir leurs parties. Pour cela, nous proposons d'utiliser le même concept de livre-code avec notre distance géodésique.
- Étudier l'approche de classification proposée avec l'apprentissage statistique (SVM, forêt) en utilisant les distances géodésiques calculées entre les parties de formes deux à deux. Puis, approfondir l'étude de l'approche de classification afin de trouver d'autres descripteurs de formes, toujours issus des géodésiques, qui soient plus efficaces pour l'apprentissage statistique.
- Généraliser et étendre les systèmes de reconnaissance proposés pour aborder la reconnaissance des formes 3D, où chaque vue d'une forme 3D peut être considérée comme une partie de forme. Pour cela, nous avons besoin d'élaborer une mesure de similarité basée sur les géodésiques pour la comparaison des formes 3D.

En termes d'application, nous pourrions aborder l'extension des méthodes de reconnaissance développées au traitement d'autres données. Par exemple, nous proposons de travailler sur le domaine de l'observation du milieu marin. À noter que la complexité des images issues des observatoires marins impose aux techniques de reconnaissance un défi considérable en termes de capacité d'adaptation au caractère non rigide des objets à reconnaître et aux conditions du milieu perturbé, avec des situations d'occultations et de variations d'illumination. Une autre perspective importante de ce travail est d'étudier l'apport des méthodes proposées dans la reconnaissance de formes sur des images biologiques, où l'existence d'une forte variabilité inter-individuelle est à l'origine des difficultés rencontrées.

Bibliographie

- [1] www site. <http://www.ee.surrey.ac.uk/research/vssp/imagedb/demo.html>.
- [2] Hossam Abdelmunim and Aly A Farag. Elastic shape registration using an incremental free form deformation approach with the icp algorithm. In *Computer and Robot Vision (CRV), 2011 Canadian Conference on*, pages 212–218. IEEE, 2011.
- [3] Yaser S Abu-Mostafa and Demetri Psaltis. Recognitive aspects of moment invariants. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):698–706, 1984.
- [4] Tsunehiro Aibara, Kenji Ohue, and Yasushi Matsuoka. Human face recognition by p-type fourier descriptor. In *Visual Communications,'91, Boston, MA*, pages 198–203. International Society for Optics and Photonics, 1991.
- [5] Oswin Aichholzer and Franz Aurenhammer. *Straight skeletons for general polygonal figures in the plane*. Springer, 1996.
- [6] Muayed S Al-Huseiny, Sasan Mahmoodi, and Mark S Nixon. Robust rigid shape registration method using a level set formulation. In *Advances in Visual Computing*, pages 252–261. Springer, 2010.
- [7] Habib Ammari, Thomas Boulier, Josselin Garnier, and Han Wang. Shape recognition and classification in electro-sensing. *Proceedings of the National Academy of Sciences*, 111(32):11652–11657, 2014.
- [8] Jeffrey Armstrong, Maher Ahmed, and Siu-Cheung Chau. A rotation-invariant approach to 2d shape representation using the hilbert curve. In *Image Analysis and Recognition*, pages 594–603. Springer, 2009.
- [9] Vincent Arsigny, Olivier Commowick, Xavier Pennec, and Nicholas Ayache. A log-euclidean framework for statistics on diffeomorphisms. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*, pages 924–931. Springer, 2006.

-
- [10] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM journal on matrix analysis and applications*, 29(1):328–347, 2007.
- [11] Thierry Artieres, Sanparith Marukatat, and Patrick Gallinari. Online handwritten shape recognition using segmental hidden markov models. *IEEE transactions on pattern analysis and machine intelligence*, 29(2), 2007.
- [12] Cagri Aslan, Aykut Erdem, Erkut Erdem, and Sibel Tari. Disconnected skeleton: Shape at its absolute scale. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(12):2188–2203, 2008.
- [13] Franz Aurenhammer. Weighted skeletons and fixed-share decomposition. *Computational Geometry*, 40(2):93–101, 2008.
- [14] Song Bai, Xinggang Wang, and Xiang Bai. Aggregating contour fragments for shape classification. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pages 5252–5256. IEEE, 2014.
- [15] Xiang Bai and Longin Jan Latecki. Path similarity skeleton graph matching. *IEEE transactions on pattern analysis and machine intelligence*, 30(7):1282–1292, 2008.
- [16] Xiang Bai, Longin Jan Latecki, and Wen-Yu Liu. Skeleton pruning by contour partitioning with discrete curve evolution. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(3):449–462, 2007.
- [17] Xiang Bai, Wenyu Liu, and Zhuowen Tu. Integrating contour and skeleton for shape classification. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 360–367. IEEE, 2009.
- [18] Xiang Bai, Cong Rao, and Xinggang Wang. Shape vocabulary: A robust and efficient shape representation for shape matching. *Image Processing, IEEE Transactions on*, 23(9):3935–3949, 2014.
- [19] Xiang Bai, Xingwei Yang, and Longin Jan Latecki. Detection and recognition of contour parts based on shape similarity. *Pattern Recognition*, 41(7):2189–2199, 2008.
- [20] Erdem Bala and A Enis Cetin. Computationally efficient wavelet affine invariant functions for shape recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (8):1095–1099, 2004.
- [21] Martin Bauer, Philipp Harms, and Peter W Michor. Sobolev metrics on shape space, ii: Weighted sobolev metrics and almost local metrics. *arXiv preprint arXiv:1109.0404*, 2011.

- [22] Martin Bauer, Philipp Harms, Peter W Michor, et al. Sobolev metrics on the manifold of all riemannian metrics. *Journal of Differential Geometry*, 94(2):187–208, 2013.
- [23] Richard Bellman. Dynamic programming princeton university press. *Princeton, NJ*, 1957.
- [24] Serge Belongie, Jitendra Malik, and Jan Puzicha. Matching shapes. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 1, pages 454–461. IEEE, 2001.
- [25] Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape matching and object recognition using shape contexts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(4):509–522, 2002.
- [26] Donald J Berndt and James Clifford. Using dynamic time warping to find patterns in time series. In *KDD workshop*, volume 10, pages 359–370. Seattle, WA, 1994.
- [27] Stefano Berretti, Alberto Del Bimbo, and Pietro Pala. Retrieval by shape similarity with perceptual distance and effective indexing. *Multimedia, IEEE Transactions on*, 2(4):225–239, 2000.
- [28] Irving Biederman. Human image understanding: Recent research and a theory. *Computer vision, graphics, and image processing*, 32(1):29–73, 1985.
- [29] Irving Biederman. Recognition-by-components: a theory of human image understanding. *Psychological review*, 94(2):115, 1987.
- [30] Michael J Black and Anand Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International Journal of Computer Vision*, 19(1):57–91, 1996.
- [31] Harry Blum. A transformation for extracting new descriptors of shape. *MIT Press*, pages 362–380, 1967.
- [32] Mirosław Bober. Mpeg-7 visual shape descriptors. *IEEE Transactions on circuits and systems for video technology*, 11(6):716–719, 2001.
- [33] Fred L Bookstein. Size and shape spaces for landmark data in two dimensions. *Statistical Science*, pages 181–222, 1986.
- [34] Luc Boutsen and Christian Marendaz. Detection of shape orientation depends on salient axes of symmetry and elongation: Evidence from visual search. *Perception & Psychophysics*, 63(3):404–422, 2001.

-
- [35] Michael Brady and Haruo Asada. Smoothed local symmetries and their implementation. *The International Journal of Robotics Research*, 3(3):36–61, 1984.
- [36] Lisa Gottesfeld Brown. A survey of image registration techniques. *ACM computing surveys (CSUR)*, 24(4):325–376, 1992.
- [37] Martin Burger, Andrea CG Mennucci, Stanley Osher, and Martin Rumpf. Level set and pde based reconstruction methods in imaging. *Springer*, 2013.
- [38] David J Burr. A dynamic model for image registration. *Computer Graphics and Image Processing*, 15(2):102–112, 1981.
- [39] Eugenio Calabi, Peter J Olver, Chehrzad Shakiban, Allen Tannenbaum, and Steven Haker. Differential and numerically invariant signature curves applied to object recognition. *International Journal of Computer Vision*, 26(2):107–135, 1998.
- [40] Arthur Cayley. An introductory memoir upon quantics. *Philosophical Transactions of the Royal Society of London*, pages 245–258, 1854.
- [41] M Emre Celebi and Y Alp Aslandogan. A comparative study of three moment-based shape descriptors. In *International Conference on Information Technology: Coding and Computing (ITCC'05)-Volume II*, volume 1, pages 788–793. IEEE, 2005.
- [42] Guillaume Charpiat. Learning shape metrics based on deformations and transport. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 328–335. IEEE, 2009.
- [43] Guillaume Charpiat, Olivier Faugeras, and Renaud Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Foundations of Computational Mathematics*, 5(1):1–58, 2005.
- [44] Guillaume Charpiat, Renaud Keriven, Jean-Philippe Pons, and Olivier Faugeras. Designing spatially coherent minimizing flows for variational problems based on active contours. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1403–1408. IEEE, 2005.
- [45] GY Chen, Tien D Bui, and A Krzyżak. Invariant pattern recognition using radon, dual-tree complex wavelet and fourier transforms. *Pattern Recognition*, 42(9):2013–2019, 2009.
- [46] Longbin Chen, Rogerio Feris, and Matthew Turk. Efficient partial shape matching using smith-waterman algorithm. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*, pages 1–6. IEEE, 2008.

- [47] KL Chiew and YC Wang. Shape feature representation in partial object matching. *Faculty of Information Technology, University of Malaysia Sarawak, Malaysia*, 2006.
- [48] Gene CH Chuang and CC Jay Kuo. Wavelet descriptor of planar curves: Theory and applications. *Image Processing, IEEE Transactions on*, 5(1):56–70, 1996.
- [49] Elias H Cohen and Manish Singh. Perceived orientation of complex shape reflects graded part decomposition. *Journal of Vision*, 6(8):4, 2006.
- [50] Elias H Cohen and Manish Singh. Geometric determinants of shape segmentation: Tests using segment identification. *Vision Research*, 47(22):2825–2840, 2007.
- [51] Isaac Cohen and Isabelle Herlin. Tracking meteorological structures through curves matching using geodesic paths. In *Computer Vision, 1998. Sixth International Conference on*, pages 396–401. IEEE, 1998.
- [52] David Cohen-Steiner, Pierre Alliez, and Mathieu Desbrun. Variational shape approximation. *ACM Transactions on Graphics (TOG)*, 23(3):905–914, 2004.
- [53] Luigi P Cordella and Mario Vento. Symbol and shape recognition. In *International Workshop on Graphics Recognition*, pages 167–182. Springer, 1999.
- [54] Javier Cortadellas, Josep Amat, and Manel Frigola. Robust normalization of shapes. In *Discrete Geometry for Computer Imagery*, pages 255–266. Springer, 2002.
- [55] Fabrice Crivello, Thorsten Schormann, Nathalie Tzourio-Mazoyer, Per E Roland, Karl Zilles, and Bernard M Mazoyer. Comparison of spatial normalization procedures and their impact on functional maps. *Human brain mapping*, 16(4):228–250, 2002.
- [56] William R Crum, Thomas Hartkens, and DLG Hill. Non-rigid image registration: theory and practice. *The British Journal of Radiology*, 2014.
- [57] Ming Cui, John Femiani, Jiuxiang Hu, Peter Wonka, and Anshuman Razdan. Curve matching for open 2d curves. *Pattern Recognition Letters*, 30(1):1–10, 2009.
- [58] Mohammad Reza Daliri and Vincent Torre. Shape recognition and retrieval using string of symbols. In *2006 5th International Conference on Machine Learning and Applications (ICMLA'06)*, pages 101–108. IEEE, 2006.
- [59] Mohammad Reza Daliri and Vincent Torre. Robust symbolic representation for shape recognition and retrieval. *Pattern Recognition*, 41(5):1782–1798, 2008.

-
- [60] Mohammad Reza Daliri and Vincent Torre. Shape recognition based on kernel-edit distance. *Computer Vision and Image Understanding*, 114(10):1097–1103, 2010.
- [61] Christoph Dalitz, Christian Brandt, Steffen Goebbels, and David Kolanus. Fourier descriptors for broken shapes. *EURASIP Journal on Advances in Signal Processing*, 2013(1):1–11, 2013.
- [62] Joeri De Winter and Johan Wagemans. Segmentation of object outlines into parts: A large-scale integrative study. *Cognition*, 99(3):275–325, 2006.
- [63] Leila DeFloriani and Michela Spagnuolo. *Shape analysis and structuring*. Springer, 2008.
- [64] Tamal K Dey, Joachim Giesen, and Samrat Goswami. Shape segmentation and matching with flow discretization. In *Algorithms and Data Structures*, pages 25–36. Springer, 2003.
- [65] R Dinesh, Santhosh S Damle, and DS Guru. A split-based method for polygonal approximation of shape curves. In *Pattern Recognition and Machine Intelligence*, pages 382–387. Springer, 2005.
- [66] Jian-Jiun Ding, Wei-Lun Chao, Jiun-De Huang, and Cheng-Jin Kuo. Anti-symmetric fourier descriptor for non-closed segments.
- [67] Cem Direkoğlu and Mark S Nixon. Shape classification using multiscale fourier-based description in 2-d space. In *Signal Processing, 2008. ICSP 2008. 9th International Conference on*, pages 820–823. IEEE, 2008.
- [68] Ian L Dryden and Kanti V Mardia. *Statistical shape analysis*, volume 4. J. Wiley Chichester, 1998.
- [69] Nicolae Duta, Anil K Jain, and Marie-Pierre Dubuisson-Jolly. Automatic construction of 2d shape models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(5):433–446, 2001.
- [70] Akrem El-ghazal, O Basir, and Saeid Belkasim. A new shape signature for fourier descriptors. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, volume 1, pages I–161. IEEE, 2007.
- [71] Pedro F Felzenszwalb and Joshua D Schwartz. Hierarchical matching of deformable shapes. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.
- [72] Hou-Yuan F Feng and Theodosios Pavlidis. Decomposition of polygons into simpler components: Feature generation for syntactic pattern recognition. *Computers, IEEE Transactions on*, 100(6):636–650, 1975.

- [73] J Flusser. Invariant shape description and measure of object similarity. In *Image Processing and its Applications, 1992., International Conference on*, pages 139–142. IET, 1992.
- [74] Herbert Freeman and A Saghri. Generalized chain codes for planar curves. In *Proceedings of the 4th International Joint Conference on Pattern Recognition*, pages 701–703, 1978.
- [75] Matthias Fuchs, Bert Jüttler, Otmar Scherzer, and Huaiping Yang. Shape metrics based on elastic deformations. *Journal of Mathematical Imaging and Vision*, 35(1):86–102, 2009.
- [76] Thomas Funkhouser, Michael Kazhdan, Philip Shilane, Patrick Min, William Kiefer, Ayellet Tal, Szymon Rusinkiewicz, and David Dobkin. Modeling by example. *ACM Transactions on Graphics (TOG)*, 23(3):652–663, 2004.
- [77] Jared Marshall Glover. *Probabilistic Procrustean models for shape recognition with an application to robotic grasping*. PhD thesis, Massachusetts Institute of Technology, 2008.
- [78] Raghuraman Gopalan, Pavan Turaga, and Rama Chellappa. Articulation-invariant representation of non-planar shapes. In *Computer Vision–ECCV 2010*, pages 286–299. Springer, 2010.
- [79] Ardeshir Goshtasby. Description and discrimination of planar shapes using shape matrices. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):738–743, 1985.
- [80] Ardeshir Goshtasby, George C Stockman, and Carl V Page. A region-based approach to digital image registration with subpixel accuracy. *Geoscience and Remote Sensing, IEEE Transactions on*, (3):390–399, 1986.
- [81] John C Gower and Garnt B Dijkstra. *Procrustes problems*, volume 3. Oxford University Press Oxford, 2004.
- [82] Gösta H Granlund. Fourier preprocessing for hand print character recognition. *Computers, IEEE Transactions on*, 100(2):195–201, 1972.
- [83] Ulf Grenander. *General pattern theory-A mathematical study of regular structures*. Clarendon Press, 1993.
- [84] Ulf Grenander and Michael I Miller. Computational anatomy: An emerging discipline. *Quarterly of applied mathematics*, 56(4):617–694, 1998.

- [85] Ximo Gual Arnau, Silena Herold-García, and Amelia Simó Vidal. Geometric analysis of planar shapes with applications to cell deformations. *International Society for Stereology*, 34(3), 2015.
- [86] Aliaksei Halachkin. Dynamic time warping for vehicle classification. In *Proceedings of the 21st Conference STUDENT EEICT 2015*. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2015.
- [87] Sunao Hashimoto, Joo Kooi, Tan Hyoungseop, and Kim Seiji Ishikawa. A media-art employing virtual shadows with shape recognition. In *Innovative Computing, Information and Control, 2007. ICICIC'07. Second International Conference on*, pages 58–58. IEEE, 2007.
- [88] Anna Hilsmann and Peter Eisert. Tracking deformable surfaces with optical flow in the presence of self occlusion in monocular image sequences. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*, pages 1–6. IEEE, 2008.
- [89] Donald D Hoffman and Whitman A Richards. Parts of recognition. *Cognition*, 18(1):65–96, 1984.
- [90] Frank Höppner. Discovery of temporal patterns. pages 192–203, 2001.
- [91] Ming-Kuei Hu. Visual pattern recognition by moment invariants. *Information Theory, IRE Transactions on*, 8(2):179–187, 1962.
- [92] Jukka Iivarinen and Ari JE Visa. Shape recognition of irregular objects. In *Photonics East'96*, pages 25–32. International Society for Optics and Photonics, 1996.
- [93] Akitoshi Itai and Hiroshi Yasukawa. Footstep classification using simple speech recognition technique. In *Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on*, pages 3234–3237. IEEE, 2008.
- [94] Tingting Jiang, Zhongqian Dong, Chang Ma, and Yizhou Wang. Toward perception-based shape decomposition. In *Computer Vision–ACCV 2012*, pages 188–201. Springer, 2013.
- [95] Shantanu H Joshi, Eric Klassen, Anuj Srivastava, and Ian Jermyn. A novel representation for riemannian analysis of elastic curves in rn. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–7. IEEE, 2007.

- [96] Shantanu H Joshi, Eric Klassen, Anuj Srivastava, and IH Jermyn. An efficient representation for computing geodesics between n-dimensional elastic shapes. In *IEEE conference on computer vision and pattern recognition*, 2007.
- [97] Sagi Katz and Ayellet Tal. *Hierarchical mesh decomposition using fuzzy clustering and cuts*, volume 22. ACM, 2003.
- [98] Hannu Kauppinen, Tapio Seppänen, and Matti Pietikäinen. An experimental comparison of autoregressive and fourier-based descriptors in 2d shape classification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(2):201–207, 1995.
- [99] Lydia E Kavraki, Petr Švestka, Jean-Claude Latombe, and Mark H Overmars. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *Robotics and Automation, IEEE Transactions on*, 12(4):566–580, 1996.
- [100] Y. Kawamura and Y. Yokota. I-type fourier descriptor: a new fourier descriptor applicable to open curves. In *IEICE Transactions on Information and Systems*, volume 88, pages 2021–2028, 2005.
- [101] J Mark Keil. Polygon decomposition. *Handbook of Computational Geometry*, 2:491–518, 2000.
- [102] David G Kendall. Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2):81–121, 1984.
- [103] JT Kent and KV Mardia. Shape, procrustes tangent projections and bilateral symmetry. *Biometrika*, 92(1):249–249, 2005.
- [104] Miss Snehal M Khakse and VK Shandilya. An efficient and elastic approach for partial shape matching using DTW. *International Journal of Engineering Trends and Technology*, 4(4):1117–1121, 2013.
- [105] Alireza Khotanzad and Yaw Hua Hong. Invariant image recognition by zernike moments. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(5):489–497, 1990.
- [106] Martin Kilian, Niloy J Mitra, and Helmut Pottmann. Geometric modeling in shape space. *ACM Transactions on Graphics (TOG)*, 26(3):64, 2007.
- [107] Hae-Kwang Kim and Jong-Deuk Kim. Region-based shape descriptor invariant to rotation, scale and translation. *Signal Processing: Image Communication*, 16(1):87–93, 2000.

- [108] Benjamin B Kimia, Allen R Tannenbaum, and Steven W Zucker. Shapes, shocks, and deformations i: the components of two-dimensional shape and the reaction-diffusion space. *International journal of computer vision*, 15(3):189–224, 1995.
- [109] E Klassen and A Srivastava. Geometric analysis of planar shapes using geodesic paths. In *Signals, Systems and Computers, 2002. Conference Record of the Thirty-Sixth Asilomar Conference on*, volume 1, pages 468–472. IEEE, 2002.
- [110] Eric Klassen, Anuj Srivastava, Washington Mio, and Shantanu H Joshi. Analysis of planar shapes using geodesic paths on shape spaces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(3):372–383, 2004.
- [111] Alexander Kolesnikov and Pasi Fränti. Reduced-search dynamic programming for approximation of polygonal curves. *Pattern Recognition Letters*, 24(14):2243–2254, 2003.
- [112] Frank P Kuhl and Charles R Giardina. Elliptic fourier features of a closed contour. *Computer graphics and image processing*, 18(3):236–258, 1982.
- [113] VP Dinesh Kumar and Thomas Tessamma. Performance study of an improved legendre moment descriptor as region-based shape descriptor. *Pattern Recognition and Image Analysis*, 18(1):23–29, 2008.
- [114] Iivari Kunttu, Leena Lepistö, Juhani Rauhamaa, and Ari Visa. Multiscale fourier descriptors for defect image retrieval. *Pattern Recognition Letters*, 27(2):123–132, 2006.
- [115] Nacéra Laiche, Slimane Larabi, Farouk Ladraa, and Abdelnour Khadraoui. Curve normalization for shape retrieval. *Signal Processing: Image Communication*, 29(4):556–571, 2014.
- [116] Rolf Lakaemper and Marc Sobel. Using the particle filter approach to building partial correspondences between shapes. *International journal of computer vision*, 88(1):1–23, 2010.
- [117] P Lansky, N Yakimoff, T Radil, and L Mitrani. Errors in estimating the orientation of dot patterns. *Perception*, 18(2):237–242, 1989.
- [118] Fredrik Larsson and Michael Felsberg. Using fourier descriptors and spatial models for traffic sign recognition. In *Image Analysis*, pages 238–249. Springer, 2011.
- [119] Longin Jan Latecki, Rolf Lakaemper, and Diedrich Wolter. Optimal partial shape similarity. *Image and vision computing*, 23(2):227–236, 2005.

BIBLIOGRAPHIE

- [120] Longin Jan Latecki and Rolf Lakämper. Convexity rule for shape decomposition based on discrete contour evolution. *Computer Vision and Image Understanding*, 73(3):441–454, 1999.
- [121] Longin Jan Latecki and Rolf Lakämper. Polygon evolution by vertex deletion. In *Scale-Space Theories in Computer Vision*, pages 398–409. Springer, 1999.
- [122] Longin Jan Latecki and Rolf Lakämper. Shape similarity measure based on correspondence of visual parts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(10):1185–1190, 2000.
- [123] Longin Jan Latecki and Rolf Lakämper. Application of planar shape comparison to object retrieval in image databases. *Pattern Recognition*, 35(1):15–29, 2002.
- [124] Longin Jan Latecki, Rolf Lakämper, and Diedrich Wolter. Shape similarity and visual parts. In *Discrete Geometry for Computer Imagery*, pages 34–51. Springer, 2003.
- [125] Longin Jan Latecki, Vasileios Megalooikonomou, Qiang Wang, Rolf Lakaemper, CA Ratanamahatana, and E Keogh. Partial elastic matching of time series. In *Data Mining, Fifth IEEE International Conference on*, pages 4–pp. IEEE, 2005.
- [126] Longin Jan Latecki, Vasileios Megalooikonomou, Qiang Wang, and Deguang Yu. An elastic partial shape matching technique. *Pattern Recognition*, 40(11):3069–3080, 2007.
- [127] Longin Jan Latecki, Qiang Wang, Suzan Koknar-Tezel, and Vasileios Megalooikonomou. Optimal subsequence bijection. In *Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on*, pages 565–570. IEEE, 2007.
- [128] Costa LdF and JRM Cesar. Shape analysis and classification: theory and practice. *CRC Press, Boca Raton*, 2001.
- [129] Sang Hwa Lee, Siddharth Sharma, Linlin Sang, Jong-Il Park, and Yong Gyu Park. An intelligent video security system using object tracking and shape recognition. In *International Conference on Advanced Concepts for Intelligent Vision Systems*, pages 471–482. Springer, 2011.
- [130] Pete E Lestrel, Osamu Takahashi, and Eisaku Kanazawa. A quantitative approach for measuring crowding in the dental arch: Fourier descriptors. *American journal of orthodontics and dentofacial orthopedics*, 125(6):716–725, 2004.
- [131] Hui Li, BS Manjunath, and Smjit K Mitra. A contour-based approach to multisensor image registration. *Image Processing, IEEE Transactions on*, 4(3):320–334, 1995.

- [132] Xuetao Li, Tong Wing Woon, Tiow Seng Tan, and Zhiyong Huang. Decomposing polygon meshes for interactive applications. In *Proceedings of the 2001 symposium on Interactive 3D graphics*, pages 35–42. ACM, 2001.
- [133] Yajun Li. Reforming the theory of invariant moments for pattern recognition. *Pattern Recognition*, 25(7):723–730, 1992.
- [134] Jyh-Ming Lien. Hybrid motion planning using minkowski sums. *Proceedings of robotics: science and systems IV*, 2008.
- [135] Jyh-Ming Lien and Nancy M Amato. Approximate convex decomposition of polygons. In *Proceedings of the twentieth annual symposium on Computational geometry*, pages 17–26. ACM, 2004.
- [136] Jyh-Ming Lien and Nancy M Amato. Approximate convex decomposition of polyhedra. In *Proceedings of the 2007 ACM symposium on Solid and physical modeling*, pages 121–131. ACM, 2007.
- [137] Jyh-Ming Lien and Yanyan Lu. Planning motion in similar environments. In *Proceedings of the Robotics: Science and Systems Conference (RSS), Seattle, Washington*, volume 2, 2009.
- [138] I-Jong Lin, Sun-Yuan Kung, et al. Coding and comparison of dag’s as a novel neural structure with applications to on-line handwriting recognition. *Signal Processing, IEEE Transactions on*, 45(11):2701–2708, 1997.
- [139] Haibin Ling and David W Jacobs. Shape classification using the inner-distance. *IEEE transactions on pattern analysis and machine intelligence*, 29(2):286–299, 2007.
- [140] Hairong Liu, Wenyu Liu, and Longin Jan Latecki. Convex shape decomposition. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 97–104. IEEE, 2010.
- [141] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.
- [142] Guojun Lu and Atul Sajjanhar. Region-based shape representation and similarity measure suitable for content-based image retrieval. *Multimedia Systems*, 7(2):165–174, 1999.
- [143] Simon Lucey, Yang Wang, Jason Saragih, and Jeffery F Cohn. Non-rigid face tracking with enforced convexity and local appearance consistency constraint. *Image and vision computing*, 28(5):781–789, 2010.

- [144] Bin Luo and Edwin R Hancock. Structural graph matching using the em algorithm and singular value decomposition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(10):1120–1136, 2001.
- [145] Bin Luo and Edwin R Hancock. Iterative procrustes alignment with the em algorithm. *Image and Vision Computing*, 20(5):377–396, 2002.
- [146] Chang Ma, Zhongqian Dong, Tingting Jiang, Yizhou Wang, and Wen Gao. A method of perceptual-based shape decomposition. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 873–880. IEEE, 2013.
- [147] Tianyang Ma and Longin Jan Latecki. From partial shape matching through local deformation to robust global shape similarity for object detection. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 1441–1448. IEEE, 2011.
- [148] Diego Macrini, Sven Dickinson, David Fleet, and Kaleem Siddiqi. Bone graphs: Medial shape parsing and abstraction. *Computer Vision and Image Understanding*, 115(7):1044–1061, 2011.
- [149] Diego Macrini, Chris Whiten, Robert Laganieri, and Marshall Greenspan. Probabilistic shape parsing for view-based object recognition. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 2303–2305. IEEE, 2012.
- [150] Siddharth Manay, Daniel Cremers, Byung-Woo Hong, Anthony J Yezzi, and Stefano Soatto. Integral invariants for shape matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(10):1602–1618, 2006.
- [151] Alan P Mangan and Ross T Whitaker. Partitioning 3d surface meshes using watershed segmentation. *Visualization and Computer Graphics, IEEE Transactions on*, 5(4):308–321, 1999.
- [152] Meenakshi Mani. *Quantitative Analysis of Open Curves in Brain Imaging: Applications to White Matter Fibers and Sulci*. PhD thesis, Université Rennes 1, 2011.
- [153] Jorge S Marques and Arnaldo J Abrantes. Shape alignment : optimal initial point and pose estimation. *Pattern Recognition Letters*, 18(1):49–53, 1997.
- [154] Graham McNeill and Sethu Vijayakumar. Hierarchical procrustes matching for shape retrieval. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 885–894. IEEE, 2006.

-
- [155] Vasileios Megalooikonomou, Qiang Wang, Guo Li, and Christos Faloutsos. A multiresolution symbolic representation of time series. In *Data Engineering, 2005. ICDE 2005. Proceedings. 21st International Conference on*, pages 668–679. IEEE, 2005.
- [156] Xiaofeng Mi and Doug DeCarlo. Separating parts from 2d shapes using relatability. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007.
- [157] Damien Michel, Iasonas Oikonomidis, and Antonis Argyros. Scale invariant and deformation tolerant partial shape matching. *Image and Vision Computing*, 29(7):459–469, 2011.
- [158] Peter W Michor and David Mumford. Riemannian geometries on spaces of plane curves. *arXiv preprint math/0312384*, 2003.
- [159] Peter W Michor and David Mumford. An overview of the riemannian metrics on spaces of curves using the hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23(1):74–113, 2007.
- [160] Peter W Michor and David Mumford. A zoo of diffeomorphism groups on \mathbb{R}^n . *Annals of Global Analysis and Geometry*, 44(4):529–540, 2013.
- [161] Peter W Michor, David Mumford, Jayant Shah, and Laurent Younes. A metric on shape space with explicit geodesics. *arXiv preprint arXiv:0706.4299*, 2007.
- [162] Michael I. Miller and Laurent Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, 41(1-2):61–84, 2001.
- [163] Washington Mio, John C Bowers, and Xiuwen Liu. Shape of elastic strings in euclidean space. *International Journal of Computer Vision*, 82(1):96–112, 2009.
- [164] Washington Mio and Anuj Srivastava. Elastic-string models for representation and analysis of planar shapes. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II–10. IEEE, 2004.
- [165] Washington Mio, Anuj Srivastava, and Shantanu Joshi. On shape of plane elastic curves. *International Journal of Computer Vision*, 73(3):307–324, 2007.
- [166] Sushil Mittal and Peter Meer. Conjugate gradient on grassmann manifolds for robust subspace estimation. *Image and Vision Computing*, 30(6):417–427, 2012.

- [167] Mr Akash G Mohod and Dinesh S Datar. An efficient and elastic approach for partial shape matching using dtw. *International Journal*, 1(4), 2013.
- [168] Farzin Mokhtarian, Sadegh Abbasi, and Josef Kittler. Robust and efficient shape indexing through curvature scale space. In *Proceedings of the 1996 British Machine and Vision Conference BMVC*, volume 96, 1996.
- [169] Farzin Mokhtarian, Sadegh Abbasi, Josef Kittler, et al. Efficient and robust retrieval by shape content through curvature scale space. *Series on Software Engineering and Knowledge Engineering*, 8:51–58, 1997.
- [170] Farzin Mokhtarian and Mirosław Bober. *Curvature scale space representation: theory, applications, and MPEG-7 standardization*, volume 25. Springer Science & Business Media, 2013.
- [171] Farzin Mokhtarian and Alan Mackworth. Scale-based description and recognition of planar curves and two-dimensional shapes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (1):34–43, 1986.
- [172] Tanmoy Mondal, Nicolas Ragot, Jean-Yves Ramel, and Umapada Pal. Performance evaluation of dtw and its variants for word spotting in degraded documents. In *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on*, pages 1141–1145. IEEE, 2015.
- [173] KS Moon and JW Sutherland. Improved image information via procrustes analysis. *ASME-PUBLICATIONS-PED*, 56:139–139, 1993.
- [174] Greg Mori and Jitendra Malik. Estimating human body configurations using shape context matching. In *Computer Vision—ECCV 2002*, pages 666–680. Springer, 2002.
- [175] Karsten Müller and Jens-Rainer Ohm. Contour description using wavelets. In *Workshop on Image Analysis for Multimedia Interactive Services*. Citeseer, 1999.
- [176] Kamal Nasreddine. *Recalage de signaux et reconnaissance de formes. Application à l'analyse des otolithes de poissons*. PhD thesis, Université de Bretagne occidentale-Brest, 2010.
- [177] Kamal Nasreddine, Abdesslam Benzinou, and Ronan Fablet. Shape geodesics for the classification of calcified structures: beyond fourier shape descriptors. *Fisheries Research*, 98(1):8–15, 2009.
- [178] Kamal Nasreddine, Abdesslam Benzinou, and Ronan Fablet. Variational shape matching for shape classification and retrieval. *Pattern Recognition Letters*, 31(12):1650–1657, 2010.

- [179] Marc Niethammer, Yang Huang, and François-Xavier Vialard. Geodesic regression for image time-series. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*, pages 655–662. Springer, 2011.
- [180] Mark Nixon. *Feature extraction & image processing*. Academic Press, 2008.
- [181] Sébastien Ourselin, Alexis Roche, Sylvain Prima, and Nicholas Ayache. Block matching: A general framework to improve robustness of rigid registration of medical images. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2000*, pages 557–566. Springer, 2000.
- [182] Ender Ozcan and Chilukuri K Mohan. Partial shape matching using genetic algorithms. *Pattern Recognition Letters*, 18(10):987–992, 1997.
- [183] Nikos Paragios, Mikael Rousson, and Visvanathan Ramesh. Non-rigid registration using distance functions. *Computer Vision and Image Understanding*, 89(2):142–165, 2003.
- [184] Eric Persoon and King-Sun Fu. Shape discrimination using fourier descriptors. *Systems, Man and Cybernetics, IEEE Transactions on*, 7(3):170–179, 1977.
- [185] Lakshman Prasad. Rectification of the chordal axis transform and a new criterion for shape decomposition. In *Discrete Geometry for Computer Imagery*, pages 263–275. Springer, 2005.
- [186] Lawrence Rabiner and Biing-Hwang Juang. *Fundamentals of speech recognition*. 1993.
- [187] Davood Rafiei and Alberto Mendelzon. Similarity-based queries for time series data. In *ACM SIGMOD Record*, volume 26, pages 13–25. ACM, 1997.
- [188] Toni M Rath and Raghavan Manmatha. Word image matching using dynamic time warping. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages II–521. IEEE, 2003.
- [189] Zhou Ren, Junsong Yuan, Chunyuan Li, and Wenyu Liu. Minimum near-convex decomposition for robust shape representation. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 303–310. IEEE, 2011.
- [190] Zhou Ren, Junsong Yuan, and Wenyu Liu. Minimum near-convex shape decomposition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(10):2546–2552, 2013.
- [191] Zhou Ren, Junsong Yuan, Jingjing Meng, and Zhengyou Zhang. Robust part-based hand gesture recognition using kinect sensor. *Multimedia, IEEE Transactions on*, 15(5):1110–1120, 2013.

- [192] Zhou Ren, Junsong Yuan, and Zhengyou Zhang. Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 1093–1096. ACM, 2011.
- [193] Dennie Reniers and Alexandru Telea. Skeleton-based hierarchical shape segmentation. In *Shape Modeling and Applications, 2007. SMI'07. IEEE International Conference on*, pages 179–188. IEEE, 2007.
- [194] Julien Ricard, David Coeurjolly, and Atilla Baskurt. Generalizations of angular radial transform for 2d and 3d shape retrieval. *Pattern Recognition Letters*, 26(14):2174–2186, 2005.
- [195] Stefan Rolfes and M-J Rendas. Fuzzy shape recognition for robot repositioning. In *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on*, volume 2, pages 1715–1717. IEEE, 1998.
- [196] Hillel Rom and Gerard Medioni. Hierarchical decomposition and axial shape description. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 15(10):973–981, 1993.
- [197] Azriel Rosenfeld. Axial representations of shape. *Computer Vision, Graphics, and Image Processing*, 33(2):156–173, 1986.
- [198] Eli Saber, Yaowu Xu, and A Murat Tekalp. Partial shape recognition by sub-matrix matching for partial matching guided image labeling. *Pattern Recognition*, 38(10):1560–1573, 2005.
- [199] Hiroaki Sakoe and Seibi Chiba. Dynamic programming algorithm optimization for spoken word recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 26(1):43–49, 1978.
- [200] Marc Salotti. Optimal polygonal approximation of digitized curves using the sum of square deviations criterion. *Pattern Recognition*, 35(2):435–443, 2002.
- [201] Chafik Samir, Anuj Srivastava, and Mohamed Daoudi. Three-dimensional face recognition using shapes of facial curves. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(11):1858–1863, 2006.
- [202] Pedro V Sander, Zoe J Wood, Steven J Gortler, John M Snyder, and Hugues H Hoppe. Multi-chart geometry images, September 4 2007. US Patent 7,265,752.
- [203] KC Santosh and Laurent Wendling. Graphical symbol recognition. *Wiley Encyclopedia of Electrical and Electronics Engineering*, 2015.

- [204] Frank R Schmidt, Michael Clausen, and Daniel Cremers. Shape matching by variational computation of geodesics on a manifold. In *Pattern Recognition*, pages 142–151. Springer, 2006.
- [205] Philip Schneider and David H Eberly. *Geometric tools for computer graphics*. Morgan Kaufmann, 2002.
- [206] Thomas B. Sebastian, Philip N. Klein, and Benjamin B. Kimia. On aligning curves. *IEEE transactions on pattern analysis and machine intelligence*, 25(1):116–125, 2003.
- [207] N Senthilkumaran and R Rajesh. Edge detection techniques for image segmentation—a survey of soft computing approaches. *International journal of recent trends in engineering*, 1(2), 2009.
- [208] Doron Shaked and Alfred M Bruckstein. Pruning medial axes. *Computer vision and image understanding*, 69(2):156–169, 1998.
- [209] Dinggang Shen and Christos Davatzikos. Hammer: hierarchical attribute matching mechanism for elastic registration. *Medical Imaging, IEEE Transactions on*, 21(11):1421–1439, 2002.
- [210] Ali Shokoufandeh, Lars Bretzner, Diego Macrini, M Fatih Demirci, Clas Jönsson, and Sven Dickinson. The representation and matching of categorical shape. *Computer Vision and Image Understanding*, 103(2):139–154, 2006.
- [211] Jamie Shotton, Andrew Blake, and Roberto Cipolla. Multiscale categorical object recognition using contour fragments. *IEEE transactions on pattern analysis and machine intelligence*, 30(7):1270–1281, 2008.
- [212] Kaleem Siddiqi and Benjamin B Kimia. Parts of visual form: Computational aspects. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(3):239–251, 1995.
- [213] Kaleem Siddiqi, Ali Shokoufandeh, Sven J Dickinson, and Steven W Zucker. Shock graphs and shape matching. *International Journal of Computer Vision*, 35(1):13–32, 1999.
- [214] Kaleem Siddiqi, Kathryn J Tresness, and Benjamin B Kimia. Parts of visual form: Psychophysical aspects. *PERCEPTION-LONDON-*, 25:399–424, 1996.
- [215] Manish Singh and Donald D Hoffman. Completing visual contours: The relationship between relatability and minimizing inflections. *Perception & Psychophysics*, 61(5):943–951, 1999.

BIBLIOGRAPHIE

- [216] Manish Singh and Donald D Hoffman. 13-part-based representations of visual shape and implications for visual cognition. *Advances in psychology*, 130:401–459, 2001.
- [217] Manish Singh, Gregory D Seyranian, and Donald D Hoffman. Parsing silhouettes: The short-cut rule. *Perception & Psychophysics*, 61(4):636–660, 1999.
- [218] Milan Sonka, Vaclav Hlavac, and Roger Boyle. *Image processing, analysis, and machine vision*. Cengage Learning, 2014.
- [219] Anuj Srivastava, Shantanu Joshi, David Kaziska, and David Wilson. Applications of planar shape analysis to image-based inferences. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*, volume 5, pages v–1037. IEEE, 2005.
- [220] Anuj Srivastava, Shantanu H Joshi, Washington Mio, and Xiuwen Liu. Statistical shape analysis: Clustering, learning, and testing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(4):590–602, 2005.
- [221] Anuj Srivastava, Eric Klassen, Shantanu H Joshi, and Ian H Jermyn. Shape analysis of elastic curves in euclidean spaces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(7):1415–1428, 2011.
- [222] Tomáš Suk and Jan Flusser. Vertex-based features for recognition of projectively deformed polygons. *Pattern Recognition*, 29(3):361–367, 1996.
- [223] Kang B Sun and Boaz J Super. Classification of contour shapes using class segment sets. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 727–733. IEEE, 2005.
- [224] Ganesh Sundaramoorthi, Jeremy D Jackson, Anthony Yezzi, and Andrea C Mennucci. Tracking with sobolev active contours. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 674–680. IEEE, 2006.
- [225] Ganesh Sundaramoorthi, Andrea Mennucci, Stefano Soatto, and Anthony Yezzi. A new geometric metric in the space of curves, and applications to tracking deforming objects by prediction and filtering. *SIAM Journal on Imaging Sciences*, 4(1):109–145, 2011.
- [226] Ganesh Sundaramoorthi, Anthony Yezzi, and Andrea C Mennucci. Sobolev active contours. *International Journal of Computer Vision*, 73(3):345–366, 2007.

- [227] Ganesh Sundaramoorthi, Anthony Yezzi, Andrea C Mennucci, and Guillermo Sapiro. New possibilities with sobolev active contours. *International journal of computer vision*, 84(2):113–129, 2009.
- [228] Boaz J Super. Learning chance probability functions for shape retrieval or classification. In *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04. Conference on*, pages 93–93. IEEE, 2004.
- [229] Boaz J Super. Retrieval from shape databases using chance probability functions and fixed correspondence. *International Journal of Pattern Recognition and Artificial Intelligence*, 20(08):1117–1137, 2006.
- [230] Mirela Tanase and Remco C Veltkamp. Polygon decomposition based on the straight line skeleton. In *Proceedings of the nineteenth annual symposium on Computational geometry*, pages 58–67. ACM, 2003.
- [231] Mirela Tanase and Remco C Veltkamp. Part-based shape retrieval. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 543–546. ACM, 2005.
- [232] Mirela Tănase-Avătavului. Shape decomposition and retrieval. *ASCI dissertation series; 112*, 2005.
- [233] Antoine Taza and Ching Y Suen. Discrimination of planar shapes using shape matrices. *Systems, Man and Cybernetics, IEEE Transactions on*, 19(5):1281–1289, 1989.
- [234] Michael Reed Teague. Image analysis via the general theory of moments. *JOSA*, 70(8):920–930, 1980.
- [235] Arasanathan Thayananthan, Bjoern Stenger, Philip HS Torr, and Roberto Cipolla. Shape context and chamfer matching in cluttered scenes. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 1, pages I–127. IEEE, 2003.
- [236] Anthony Tort. Elliptical fourier functions as a morphological descriptor of the genus *stenosarina* (brachiopoda, terebratulida, new caledonia). *Mathematical Geology*, 35(7):873–885, 2003.
- [237] Alain Trouvé and François-Xavier Vialard. Shape splines and stochastic shape evolutions: A second order point of view. *arXiv preprint arXiv:1003.3895*, 2010.
- [238] Jerry L Turney, Trevor N Mudge, and Richard A Volz. Recognizing partially occluded parts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (4):410–421, 1985.

- [239] MT Turvey, Gregory Burton, Christopher C Pagano, H Yosef Solomon, and Sverker Runeson. Role of the inertia tensor in perceiving object orientation by dynamic touch. *Journal of Experimental Psychology: Human Perception and Performance*, 18(3):714, 1992.
- [240] Yoshinori Uesaka. A new fourier descriptor applicable to open curves. *Electronics and Communications in Japan (Part I: Communications)*, 67(8):1–10, 1984.
- [241] Marc Vaillant, Michael I Miller, Laurent Younes, and Alain Trouvé. Statistics on diffeomorphisms via tangent space representations. *NeuroImage*, 23:S161–S169, 2004.
- [242] Remco C Veltkamp. Shape matching: similarity measures and algorithms. In *Shape Modeling and Applications, SMI 2001 International Conference on.*, pages 188–197. IEEE, 2001.
- [243] Michail Vlachos, Marios Hadjieleftheriou, Dimitrios Gunopulos, and Eamonn Keogh. Indexing multi-dimensional time-series with support for multiple distance measures. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 216–225. ACM, 2003.
- [244] Laura L Walker and Jitendra Malik. Can convexity explain how humans segment objects into parts? *Journal of Vision*, 3(9):503–503, 2003.
- [245] Chun Wang and Zhongyuan Lai. Shape decomposition and classification by searching optimal part pruning sequence. *Pattern Recognition*, 2016.
- [246] Junwei Wang, Xiang Bai, Xinge You, Wenyu Liu, and Longin Jan Latecki. Shape matching and classification using height functions. *Pattern Recognition Letters*, 33(2):134–143, 2012.
- [247] Xinggang Wang, Bin Feng, Xiang Bai, Wenyu Liu, and Longin Jan Latecki. Bag of contour fragments for robust shape classification. *Pattern Recognition*, 47(6):2116–2125, 2014.
- [248] Yu-Ping Wang, Seng Luan Lee, and Kazuo Toraichi. Multiscale curvature-based shape representation using b-spline wavelets. *Image Processing, IEEE Transactions on*, 8(11):1586–1592, 1999.
- [249] Linda Crystal Wells. *Shape analysis of curves in higher dimensions*. PhD thesis, THE FLORIDA STATE UNIVERSITY, 2013.
- [250] Roger P Woods, John C Mazziotta, Simon R Cherry, et al. Mri-pet registration with automated algorithm. *Journal of computer assisted tomography*, 17(4):536–546, 1993.

- [251] Cong Yang, Oliver Tiebe, Pit Pietsch, Christian Feinen, Udo Kelter, and Marcin Grzegorzek. Shape-based object retrieval by contour segment matching. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pages 2202–2206. IEEE, 2014.
- [252] Hyun S Yang and Sanjay Sengupta. Intelligent shape recognition for complex industrial tasks. *IEEE Control Systems Magazine*, 8(3):23–30, 1988.
- [253] Anthony Yezzi and Andrea Mennucci. Metrics in the space of curves. *arXiv preprint math/0412454*, 2004.
- [254] Anthony Yezzi and Andrea Mennucci. Conformal metrics and true" gradient flows" for curves. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 913–919. IEEE, 2005.
- [255] Laurent Younes. Computable elastic distances between shapes. *SIAM Journal on Applied Mathematics*, 58(2):565–586, 1998.
- [256] Laurent Younes. Optimal matching between shapes via elastic deformations. *Image and Vision Computing*, 17(5):381–389, 1999.
- [257] Cong Yue, Hai-Jun Su, and Qiaode Jeffrey Ge. Path generator via the type-p fourier descriptor for open curves. In *13th World Congress in Mechanism and Machine Science, Guanajuato, Mexico, Paper AIL506, 6 pages, 19-25 June 2011*.
- [258] Charles T Zahn and Ralph Z Roskies. Fourier descriptors for plane closed curves. *Computers, IEEE Transactions on*, 100(3):269–281, 1972.
- [259] JingTing Zeng, Rolf Lakaemper, XingWei Yang, and Xin Li. 2d shape decomposition based on combined skeleton-boundary features. In *Advances in Visual Computing*, pages 682–691. Springer, 2008.
- [260] Dengsheng Zhang and Guojun Lu. Generic fourier descriptor for shape-based image retrieval. In *Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on*, volume 1, pages 425–428. IEEE, 2002.
- [261] Dengsheng Zhang and Guojun Lu. A comparative study of curvature scale space and fourier descriptors for shape-based image retrieval. *Journal of Visual Communication and Image Representation*, 14(1):39–57, 2003.
- [262] Dengsheng Zhang and Guojun Lu. Review of shape representation and description techniques. *Pattern recognition*, 37(1):1–19, 2004.
- [263] Hao Zhang and Jitendra Malik. Learning a discriminative classifier using shape context distances. In *null*, page 242. IEEE, 2003.

BIBLIOGRAPHIE

- [264] Dongming Zhao and Jie Chen. Affine curve moment invariants for shape recognition. *Pattern recognition*, 30(6):895–901, 1997.
- [265] Z Zheng and Y Tamura. Cultivar identification of lotus (*nelumbo nucifera gaertn*) by p-type fourier descriptors with petal tip contour. *Horticultural Research (Japan)*, 2005.
- [266] Zeyu Zheng, Hiroyoshi Iwata, Yutaka Hirata, and Yoshiyasu Tamura. Quantitative evaluation of the degree of sprout leaf bending of rice cultivars using p-type fourier descriptors and principal component analysis. *Euphytica*, 163(2):259–266, 2008.
- [267] Song Chun Zhu and Alan L Yuille. Forms: a flexible object recognition and modelling system. *International Journal of Computer Vision*, 20(3):187–212, 1996.
- [268] Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003.
- [269] Emanoil Zuckerberger, Ayellet Tal, and Shymon Shlafman. Polyhedral surface decomposition with applications. *Computers & Graphics*, 26(5):733–743, 2002.

Résumé

Les performances d'un système de reconnaissance de formes dépendent en bonne partie de la qualité de l'image segmentée. Malgré les progrès effectués, une segmentation complète (c'est-à-dire avec des contours entiers) ne peut pas être toujours atteinte. Dans un premier temps, nous nous plaçons dans le cas où seulement certaines parties de la forme entière sont disponibles. D'abord, afin d'assurer l'invariance des parties de formes aux transformations géométriques, nous développons une méthode d'optimisation de l'analyse procustéenne qui consiste à retrouver les points extrémités optimaux qui minimisent la distance de Procutse. Ensuite, nous proposons une approche de reconnaissance de parties de formes et une approche de reconnaissance de formes partielles. Ces deux approches, basées-contour, sont fondées sur un recalage robuste entre les parties de formes. La méthode de recalage proposée consiste à optimiser une mesure de similarité basée sur les géodésiques dans l'espace de formes. Ainsi, nous exploitons le résidu du recalage pour définir une nouvelle métrique pour la reconnaissance de parties de formes. Puis, nous décrivons une stratégie de combinaison avec cette même métrique pour la reconnaissance de formes partielles. Par la suite, nous proposons d'utiliser la distance géodésique proposée pour la reconnaissance des parties de formes dans la définition d'une métrique globale pour la reconnaissance de formes entières. Les tests de reconnaissance (classification et recherche) sont effectués sur des parties requêtes et des formes entières de la base d'images MPEG-7, puis sur des images réelles segmentées. Les résultats expérimentaux montrent la supériorité de nos méthodes par rapport aux autres méthodes de l'état de l'art.

Mots clés : Reconnaissance, parties de formes, formes partielles, recalage robuste, distance géodésique, transformations géométriques, invariances.

Abstract

The quality of the segmentation process directly affects the performance of the shape recognition. Despite the progress that has been made, it is often unreachable to segment the entire object (i.e. closed contour). In fact, only some parts/fragments of objects can be detected. We first develop a new alignment method based on Procrustes analysis in order to ensure invariance of shape parts to geometric transformations (translation, rotation and scale factor). The proposed method consists in finding optimal extremities which minimize the Procrustes distance. Then, we propose a shape part recognition approach and a partial shape recognition approach. These two contour-based approaches are based on matching between shape parts to compare. This matching process consists in establishing a robust registration between shape parts based on geodesics in the shape space. Thus, we exploit the registration residual to define a novel distance for shape part recognition. Later, for partial shape recognition, we describe a geodesics-based combining strategy with the same distance. As well, we propose to use the geodesics distance proposed for shape part recognition to define a global distance for entire shape recognition. Experiments are carried out on parts of shapes and entire shapes of the MPEG-7 database, then on parts issued from segmented real images. The obtained results demonstrate the effectiveness of our proposed recognition schemes. The proposed approaches are shown to significantly outperform previous works for classification and retrieval applications.

Keywords: Recognition, shape parts, partial shapes, robust registration, geodesics distance, geometric transformations, invariance.