N° ENSC-2008/113

THÈSE DE DOCTORAT
DE L'ÉCOLE NORMALE SUPÉRIEURE DE CACHAN

Présentée par

Monsieur Arnaud GARCIA

pour obtenir le grade de

DOCTEUR DE L'ÉCOLE NORMALE SUPÉRIEURE DE CACHAN

Domaine :
MATHÉMATIQUES APPLIQUÉES

Sujet de la thèse :

Analyse statistique et morphologique des images multivaluées. Développements logiciels pour les applications cliniques.

Thèse présentée et soutenue à Cachan le 26 juin 2008 devant le jury composé de :

Bernard CHALMOND	Professeur des Universités-Cergy-Pontoise	Président
Laurent NAJMAN	Professeur, ESIEE, Noisy-le-Grand	Rapporteur
Philippe SALEMBIER	Professeur, Universitat Politècnica de Cataluny	Rapporteur
Michel KOCHER	Docteur ingénieur- Université de Genève	Examinateur
Jean-Paul VALLÉE	Professeur - Hôpitaux Universitaires de Genève	Examinateur
Yves MEYER	Professeur Emérite-ENS Cachan	Directeur de thèse
Corinne VACHIER	Maître de conférences- Paris XII	Co-directeur de thèse

Laboratoire Spécification Vérification
(ENS CACHAN/CNRS/UMR 8095)
61, avenue du Président Wilson, 94235 CACHAN CEDEX (France)

À Lucie,

Remerciements

À la lecture d'autres thèses, la page des remerciements me semblait un peu futile. Pourtant, après ce travail de recherche qui s'est répartie sur 4 années, parsemées de haut et de bas, je constate que seul, je n'aurais pas concrétisé ce travail. Aussi, les personnes que je vais remercier ont eu un rôle significatif dans l'accomplissement du présent ouvrage.

Je commencerai par citer Jean-Paul Vallée, professeur à l'Hôpital Universitaire de Genève. En effet, cette aventure scientifique s'est amorcée grâce à la rencontre de cet homme passionné par la recherche en IRM. Il m'a donné l'élan nécessaire et soutenu sans relâche, de la genèse de ce projet à sa soutenance. Véritable source d'énergie, sa passion et son acharnement professionnel m'ont portés lors de mes difficultés à concilier mon travail et ma thèse. Sincèrement merci pour ce soutien inconditionnel.

Toujours sous l'impulsion de Jean-Paul, je suis allé prendre des cours de mathématiques à l'école d'ingénieurs de Genève où j'ai découvert un enseignant qui a modifié le cours de ma vie : Michel Kocher. Grâce à ces enseignements, j'ai développé un enthousiasme intense pour le traitement des images et la morphologie mathématique. Michel a su éveiller et me transmettre sa ferveur pour cette théorie. Je reste également quelque peu nostalgique de nos nombreux échanges informels. J'espère que nos chemins se croiseront à nouveau.

Sans Michel, je n'aurais pas connu Corinne qui m'a fait entièrement confiance afin que cette thèse puisse avoir lieu. Notre rencontre est à l'image de notre collaboration, la bonne humeur et l'échange. Durant toutes ces années, Corinne m'a laissé l'indépendance dont j'avais besoin et donné le soutien scientifique lors des phases clefs de ce travail. Toujours en marge des sentiers, notre collaboration m'a permis une liberté scientifique particulièrement appréciable. De plus, je te suis infiniment reconnaissant pour avoir trouvé les mots dont j'avais besoin lors des phases critiques.

Messieurs Najman et Salembier, je vous remercie d'avoir accepté d'être rapporteurs pour cette thèse. Vos conseils et remarques ont été un plus indéniable dans ce travail. Enfin, Monsieur Chalmond, j'ai apprécié vos remarques sur les statistiques et bien évidemment d'avoir assumé la présidence de cette soutenance.

Enfin, je tiens à remercier Monsieur Meyer pour ses remarques pertinentes lors de l'oral. Je connaissais l'homme de science, mais l'échange que nous avons eu après délibération m'a particulièrement touché et permis d'apercevoir la personne. J'espère que nous nous recroiserons ...

Les derniers mots de cette thèse sont pour mon épouse Lucie qui a été à mes côtés tout au long de ce travail. Je te remercie pour ta patience et ton soutien pendant toutes ces années consacrées à ce travail. C'est avec toi que je mets le point final à ce travail.

Présentation générale

Circonstances de la thèse

Cette thèse a été réalisée au sein du département d'informatique et de radiologie des Hôpitaux Universitaires de Genève (HUG) en collaboration avec le Centre de Mathématiques et Leurs Applications (CMLA) de l'École Normale Supérieure de Cachan. Pendant les deux premières années de ma thèse, j'étais responsable des développements informatiques à l'Unité d'Imagerie Numérique (UIN) des HUG. Je suis à présent responsable informatique au sein d'un groupement de radiologues à Limoges.

Mon travail d'ingénieur en conception logicielle aux HUG consistait à réaliser l'architecture informatique du dossier patient ainsi qu'à mettre en oeuvre des outils informatiques d'aide à l'analyse des images médicales à l'intention des médecins du département de radiologie. Les objectifs en terme de traitement d'image se sont très vite révélés ambitieux. Il s'agissait notamment de mettre en oeuvre des outils de segmentation d'images médicales fort variées (IRM du coeur pour l'analyse du mouvement cardiaque, IRM du foie et images de la peau pour la détection de tumeurs...). L'ensemble des problèmes rencontré en routine clinique m'a engagé à effectuer un travail plus approfondi qu'un simple développement de routines informatiques. Les réflexions menées dans ce contexte ont été au coeur de mon travail de thèse et sont exposées dans ce document.

Dans leurs pratiques quotidiennes, les radiologues sont souvent amenés à évaluer des structures organiques, tumorales ou cavitaires et les outils d'extraction qu'ils utilisent (c'est-à-dire ce qui leur permet de passer de l'observation visuelle de l'image à la quantification précise) sont assez rudimentaires et se résument souvent à un travail manuel et laborieux. La plupart du temps, les médecins ne sont pas désireux de processus automatiques d'analyse des images, mais recherchent plutôt des logiciels d'aide au diagnostic. Notamment, face à la nature très variable des images rencontrées, le développeur constate d'emblée l'impossibilité d'un traitement informatique non supervisé qui aurait de bonnes performances.

Images multivalués, analyse statistique et analyse morphologique

Je me suis intéressé aux images multivaluées, c'est-à-dire aux images associant à chaque point du plan des vecteurs. La première raison est liée à la nature des images désormais utilisées en milieu médical : les images couleurs et les images multimodalités sont de plus en plus fréquentes. Or, le domaine de l'analyse des images vectorielles est un domaine qui reste encore à explorer d'un point de vue théorique. Une grande partie des opérateurs valides dans le cas des fonctions numériques ne sont pas directement transposables au cas des images multivaluées. D'un autre côté, mon travail a pris une orientation plus particulière pour la prise en compte d'a priori sur les images. En effet, de la même façon que les médecins apprennent par apprentissage, il semblait logique de transmettre d'abord de la connaissance à l'analyseur afin d'orienter les traitements *dans le sens* des images et des applications médicales. La détection des lésions de la peau par exemple est grandement facilitée si l'on transmet à la machine un échantillon de peau vierge. De la même façon, l'identification et la segmentation du coeur en IRM pourront s'effectuer à partir d'un échantillon de coeur prélevé numériquement sur une autre IRM correspondante au même ou à un autre patient (de plus en plus de suivis thérapeutiques sont mis en place permettant de disposer d'une grande connaissance sur les images analysées). Je me suis donc intéressé à la possibilité de renforcer les statistiques de la région d'intérêt par le calcul, à partir de l'image et d'un échantillon, d'une carte de similarité permettant de révéler en chaque point du plan image la cohésion avec l'information a priori. Ajouter de l'information à une image, c'est augmenter sa dimension ce qui explique également pourquoi je me suis intéressé aux images vectorielles.

Mon travail de thèse a d'abord consisté à me poser deux grandes questions. Premièrement, comment aborder l'extraction d'une forme dans une image ? Deuxièmement, comment combiner l'analyse statistique et l'analyse morphologique des images vectorielles ?

La question de l'extraction d'une forme dans une image a été largement débattue dans la communauté scientifique du fait de la complexité du problème. Si le mécanisme est très simple pour l'être humain, il est problématique pour les systèmes informatiques. En effet, l'extraction d'une forme est un processus bien plus complexe que l'addition des parties de l'image qui la composent : c'est un processus dit holiste. Chercher à copier l'analyse sémantique humaine c'est faire face à un problème d'une complexité extrême. L'analyse statistique offre sur ce point un cadre théorique bien formalisé et ouvre des perspectives intéressantes.

Le recourt aux statistiques dans le traitement des images n'est certes pas une idée nouvelle, cependant le calcul de cartes de similarité dans le cas d'images multivaluées reste encore marginal. Je me suis particulièrement intéressé au classifieur bayésien qui s'intègre naturellement dans le cas des espaces multidimensionnels. Dans le contexte médical, disposer d'échantillons des structures présentes dans les images est un cas de figure très courant.

L'emploi d'un classifieur est toujours délicat. En effet,la décision finale de classification se fait en affectant le vecteur à la classe dont il est le plus proche, sans pondérer la réponse en fonction de la plus ou moins grande proximité, sans aucune prise en compte de la topologie de la forme recherchée. Mieux qu'un classifieur, nous proposons une analyse morphologique de la carte des distances qui permet de combiner l'avantage de l'analyse statistique à la prise en compte de l'information spatiale dans l'image. Ainsi, la segmentation est déléguée à un outil approprié.

Parmi l'ensemble des opérateurs disponibles pour l'analyse (déterministe) des images, nous nous sommes concentrés sur les opérateurs non linéaires développés par la morphologie mathématique puisqu'ils permettent justement la prise en compte de la forme dans le traitement. Ainsi, en combinant la classification statistique à l'analyse morphologique, des a priori de forme et de contenu peuvent être simultanément pris en compte.

Liste des contributions

Pour analyser les images multivaluées lorsqu'on dispose d'a priori statistiques sur les formes d'intérêt dans l'image, il est proposé de recourir à une stratégie en deux temps. Premièrement, la construction d'une carte des similarités qui permet la prise en compte de l'information statistique et qui transforme les données vectorielles en données scalaires. Deuxièmement, la transformation de la carte des similarités en une carte de rang permettant une association univoque entre les données multivaluées (les vecteurs de l'image originale) et leurs rangs (les scalaires de l'image de rang). Le caractère réversible de la transformation de rang permet de substituer l'analyse de l'image de rang à celle de l'image multivaluée et de tirer ainsi profit de tous les outils valides pour les images scalaires.

– Si l'exploitation de la distance de Mahalanobis n'est pas une originalité en traitement des images (multivaluées ou non), son utilisation dans d'autres contextes que la segmentation n'a pas encore fait l'objet d'une étude approfondie, notamment dans le cas où la référence statistique n'est

pas l'origine. Ici les applications de filtrage et de détection de formes sont abordées de manière systématique.

– Les soustractions posent de sérieuses difficultés dans le domaine des images multivaluées. Une solution originale est proposée pour définir des soustractions sur les vecteurs en accord avec la classification de Mahalanobis. La solution est basée sur la construction d'une image de rang constituée de nombres entiers consécutifs. Ce choix simple offre des avantages algorithmiques importants et propose une extension immédiate des résidus morphologiques (gradients, chapeaux haut de forme) au cas des images multivaluées.
– Enfin, l'analyse des images multivaluées via une carte des similarités induites la définition d'opérateurs spécifiques où chaque point de l'image va subir un traitement spécifique selon sa similarité au modèle statistique. Pour définir de tels opérateurs adaptatifs, nous proposons d'étendre aux images multivaluées la notion d'opérateur morphologique visqueux originellement introduit pour les images à teintes de gris.

Au niveau des interactions avec les applications cliniques, ma collaboration avec les HUG a résulté en la mise en oeuvre d'outils logiciels et dans la proposition de stratégies originales pour la détection des parois du myocarde en IRM conventionnelle et pour la détection des grains de beauté sur des images couleur de la peau. Ces stratégies sont expérimentales. Elles n'ont pas fait l'objet de validations cliniques, mais sont utiles aux médecins dans le cadre de leurs recherches cliniques.

Structure du document

Ce document est structuré en 3 chapitres et 1 annexe.

La première partie est une réflexion autour de l'image et de son analyse. Comment aborder la question de l'analyse statistique et géométrique des images ? Les fondements de la décision bayésienne et de l'analyse morphologique sont rappelés. Les images multivaluées font l'objet d'une réflexion approfondie. Notamment, un état de l'art est dressé sur les solutions envisageables pour étendre l'analyse morphologique usuelle sur les fonctions à valeurs scalaires au cas des fonctions à valeurs vectorielles.

Le chapitre 2 concentre l'essentiel des travaux fondamentaux réalisés pendant la thèse. Les fondements d'une analyse à la fois statistique et morphologique sont établis. Les cartes de distance construites sont interprétées dans le cas des images couleur. Les opérateurs morphologiques usuels sont expérimentés. Enfin, des opérateurs adaptatifs sont définis et expérimentés sur des images couleur.

Le chapitre 3 présente l'ensemble des développements réalisés pour les applications médicales. Il s'agit d'outils logiciels et d'algorithmes d'analyse des images de peau et d'IRM cardiaques.

Enfin, des travaux sur les développements logiciels cliniques non mentionnés dans le corps du document sont présentés sous la forme des communications dont ils ont fait l'objet en annexe.

Table des matières

1 Analyser les images en intégrant de la connaissance a priori **17**

1.1 Réflexions sur l'analyse des images 17

1.1.1 Formes, segmentation et sémantique 17

1.1.2 Formes et parties de l'image 19

1.1.3 Limite conceptuelle 23

1.2 L'analyse statistique et l'analyse géométrique : quels fondements ? 24

1.2.1 Treillis complets et morphologie mathématique 24

1.2.2 Probabilités et classification bayésienne 29

1.2.3 Les stratégies mixtes 34

1.3 Spécificité des images multivaluées 34

1.3.1 De l'image scalaire à l'image multivaluée 34

1.3.2 Les images couleur et les espaces colorimétriques 35

1.4 Analyse morphologique des images multivaluées 38

1.4.1 Etat de l'art sur la morphologie multivaluée 38

1.4.2 Quatre approches pour ordonner des vecteurs 38

1.4.3 Le cas particulier de l'ordre total 44

1.5 Conclusion 46

2 Analyse morphologique des images avec a priori statistique **47**

2.1 Décision bayésienne et ordre total sur les vecteurs 47

2.1.1 De la règle de Bayes à la distance de Mahalanobis 48

2.1.2 Construction d'images de rang 54

2.2 Expoitation directe de la carte des distances : le cas particulier de la segmentation 61

2.2.1 Deux paradigmes morphologiques pour la segmentation 63
2.2.2 Ligne de Partage des Eaux et distance de Mahalanobis 69
2.2.3 Conclusion 76
2.3 Filtrage morphologique des images multivaluées via la transformation de rang 77
2.3.1 Action des opérateurs morphologiques fondamentaux 77
2.3.2 Les résidus morphologiques 82
2.3.3 Simplification des images multivaluées par filtrage morphologique connexe . . . 84
2.3.4 Conclusion 87
2.4 Définition d'opérateurs morphologiques auto-adaptatifs 94
2.4.1 Formulation classique dans le cas des images à valeurs scalaires 95
2.4.2 Extension au cas des images multivaluées avec prise en compte d'a priori statistiques 99
2.4.3 Le calcul effectif des transformations visqueuses 100
2.4.4 Expérimentations 101
2.5 Conclusion 103

3 Développements logiciels pour les applications cliniques 109
3.1 Dépistage du mélanome de la peau : le projet Naeview 110
3.2 Segmentation d'IRM cardiaques 117
3.2.1 L'IRM cardiaque 117
3.2.2 L'activité cardiaque 119
3.3 La segmentation du myocarde : focus sur deux stratégies classiques 121
3.3.1 Les modèles déformables 122
3.3.2 L'utilisation d'atlas 123
3.4 Algorithme proposé 125
3.4.1 Augmenter les dimensions 125
3.4.2 Initialisation 125
3.4.3 Segmentation de l'endocarde 126
3.4.4 Construction du critère de forme 126
3.4.5 Segmentation dans l'espace des similarités 128
3.5 Conclusion : validation et segmentation 3D 129

3.6 Osirix, un outil d'aide à l'analyse des images médicales 130
3.7 vmorph 133
3.7.1 Architecture 133
3.7.2 Principes 133
3.7.3 Implémentation 134
3.7.4 Intégration dans Matlab 135

4 Conclusion et perspectives **139**

A Quelques communications **141**
A.1 Benefits from content-based visual data access in radiology 142
A.2 Building an Enterprise-wide PACS for all diagnostic images 165
A.3 Grid Computing at the University Hospitals of Geneva 172

Chapitre 1

Analyser les images en intégrant de la connaissance a priori

« De la multiplicité des choses provient l'Un, et de l'Un provient la multiplicité »

De l'Univers, Héraclite (VIe siècle)

1.1 Réflexions sur l'analyse des images

Chercher à extraire les éléments qui composent une scène, en faire une analyse qualitative, préciser les zones d'intérêt, délimiter des régions ; voici certaines des tâches auxquelles s'attelle le traitement des images. Les problèmes de restauration mis à part, l'analyse d'une image est rarement un but en soi. L'image est toujours au coeur d'un système plus complexe d'analyse ou de décision et ne peut par conséquent pas se dissocier de son contexte. Dans le domaine médical, les traitements effectués sur les images visent par exemple à améliorer la perception du médecin (reconstruire en 3D la surface d'un vaisseau : confer figure 1.3) ou encore à l'aider dans son diagnostic (évaluer le volume d'une tumeur : confer figure 1.2)...

1.1.1 Formes, segmentation et sémantique

Parmi toutes les tâches réalisées par le traitement des images, la segmentation est certainement la plus délicate (exemple figure 1.1), car elle arrive en bout de chaîne et concentre de nombreuses questions qui gravitent autour de l'image. Examiner la question de la segmentation est fort instructif de la manière avec laquelle l'image et sa sémantique sont abordées.

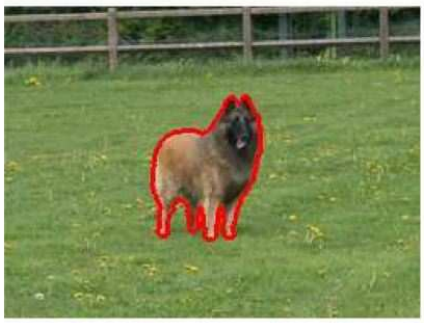


FIG. 1.1: Exemple d'extraction d'une forme. *Peplum* est extrait de l'image. L'image est segmentée en deux zones : l'intérieur contenant *Peplum* et l'extérieur, la frontière étant indiquée en rouge. L'image est étiquetée ou segmentée.

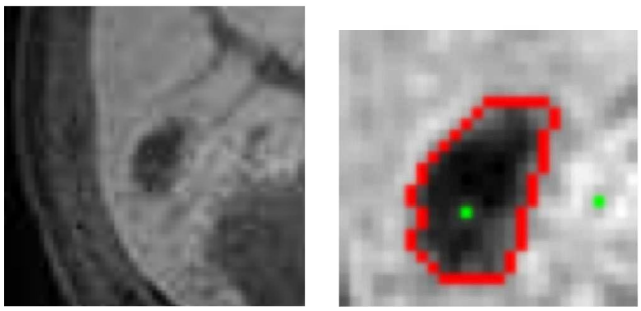


FIG. 1.2: IRM du foie contenant une tumeur. L'image étiquetée en deux zones permet par un simple balayage de calculer le nombre de points à l'intérieur de la tumeur. Connaissant la valeur d'un point en mm^2, il est alors possible d'établir la surface de la tumeur sur cette coupe. Le procédé de segmentation généralisé en 3D sur toute la tumeur permet, en additionnant les coupes, de reconstituer le volume.

Segmenter une image signifie délimiter précisément une ou plusieurs régions d'intérêt dans l'image ; on parle également de partition de l'image [35] :

Définition 1 *La segmentation est un traitement de bas niveau qui consiste à créer une partition de l'image A en sous-ensembles R_i, appelés régions, tels que :*

- $\forall i \; R_i \neq \emptyset$
- $\forall i, j \; ; \; i \neq j \; R_i \cap R_j = \emptyset$
- $A = \bigcup_i R_i$

Cette définition laisse une grande question ouverte : comment choisir parmi toutes les segmentations possibles ?

Si segmenter une image c'est partitionner l'espace en régions connexes et disjointes, le problème sous-jacent et qui justifie très largement la recherche dans ce domaine est bien dans *le sens* des régions obtenues. En effet, lorsque nous effectuons une segmentation sur une image c'est pour définir une entité que l'oeil perçoit, mais qui n'est pas visible dans la matrice de points codant la scène. Ceci est vrai aussi bien en imagerie médicale, lorsque l'on cherche à délimiter le volume d'une tumeur par exemple, que lorsque nous utilisons un logiciel de retouche d'images pour extraire l'arbre d'un paysage ou gommer un défaut.

Une segmentation réussie est par conséquent une segmentation où chaque région à un sens pour l'utilisateur (figure 1.4). Cette précision a le mérite de ne pas occulter le vrai problème de la segmentation qui est d'abord une question de modélisation (sans modèle de la tumeur, comment la trouver dans l'image ?).

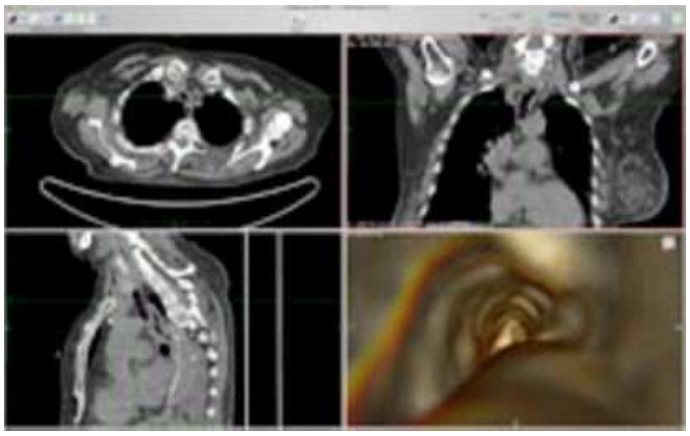
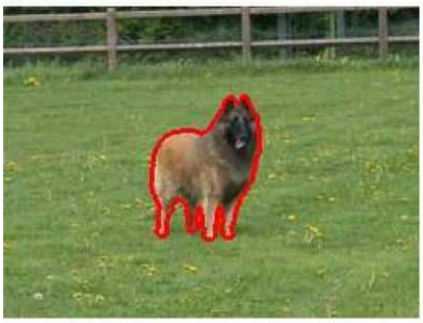


FIG. 1.3: Cette image d'endoscopie virtuelle obtenue à partir du logiciel OsiriX (confer section 3.6) permet d'illustrer le processus de reconstruction de surface. Afin de reconstruire la surface, il est nécessaire de segmenter le vaisseau puis de reconstruire un maillage 3D à partir des points de la frontière.

1.1.2 Formes et parties de l'image

Dans les années 30, un courant psychologique allemand nommé Gestalttheorie[1] ou théorie de la forme met en avant la prédominance du tout sur les parties : « Le tout est différent de la somme des parties ». Par exemple, la molécule d'eau représente autre chose que deux atomes d'hydrogène et un atome d'oxygène. Les propriétés physico-chimiques sont d'ailleurs radicalement différentes prouvant la différence entre les parties et le tout.

[1] Selon wikipedia « On trouve son origine dans quelques idées de Goethe. Aux 19e et 20e siècles ce sont Ernst Mach et surtout Christian von Ehrenfels qui la développent, aussi bien que Max Wertheimer, Wolfgang Köhler, Kurt Koffka et Kurt Lewin »



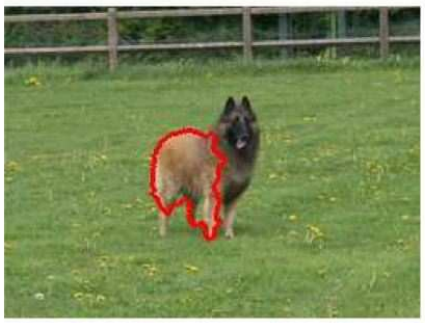


FIG. 1.4: À gauche nous avons une bonne segmentation et à droite une mauvaise segmentation. Pourtant, les deux segmentations sont conformes à la définition 16. Est-ce que segmenter n'implique pas également que la région ait un sens ?

Cette relation qui lie le tout et les parties est une relation holiste [24] :

Définition 2 *Une relation est holiste si elle vérifie les deux propositions suivantes :*

(P1) Le tout est plus que la somme des parties

(P2) Les parties tirent leur possibilité de leur inscription dans un tout

Ces deux propositions sont vérifiées dans le cas de l'image et des formes qui la constituent. En effet, prenons l'exemple d'une image numérique représentant un cercle. Le cercle possède bien une caractéristique supplémentaire par rapport aux n points qui le composent. Le cercle composé de points est donc bien supérieur à ses parties : il ne résulte pas seulement de la somme de ses n points ; le réduire à cette somme serait omettre sa structure circulaire.

La propriété (P2) selon [24], énonce que « La partie trouve son fondement dans le tout ». C'est-à-dire qu'en dehors de son rapport à un tout, la partie n'est qu'un objet indépendant, un objet parmi les objets. Dans une perspective ensembliste, les éléments existent d'une certaine manière avant l'ensemble, alors que dans une perspective holiste, le tout existe en quelque sorte avant ses parties. Dans une image, cet énoncé trouve tout son sens : pour reprendre notre exemple, un point sorti de son contexte ne représente qu'un simple élément alors que ce même point inclus dans la forme circulaire trouve un fondement dans sa relation avec le cercle.

Lorsque nous analysons une image, nous ne pouvons pas omettre l'importance de cette relation holiste. Pourtant sans les points, l'image est vide et la forme disparaît. Les parties sont donc fondamen-

FIG. 1.5: Un chien illustrant le processus holiste de la raison pour extraire une forme. La proposition (P1) est vérifiée, nous avons bien une forme apportant plus que la somme de ses points noirs et blancs. De même (P2) est également vérifiée, les points inclus dans la forme sont différents des points extérieurs alors qu'intrinsèquement identiques.

tales puisqu'elles participent à l'émergence de la forme. Ce que la relation holiste met en avant, c'est le passage d'un ensemble de points à une structure sémantique.

Il y a une forte analogie avec les structures algébriques. En algèbre, un ensemble est égal à la somme de ses éléments. Nous n'avons pas cette supériorité du tout sur les parties. Par contre, si l'ensemble est muni d'une loi de composition interne comme dans le cas d'un groupe [2], nous obtenons cette supériorité. La structure interne prime par conséquent sur les éléments qui la composent (figure 1.6).

Lorsqu'un observateur cherche à mettre en évidence ou à reconnaître une forme dans une image, et ceci malgré le bruit numérique ou les changements de représentations, il est sensible aux structures (alignements, agencements entre formes élémentaires...) plus qu'aux parties qui composent ces structures [3]. C'est le mécanisme d'agencement qui l'emporte. Il y a un fort parallèle avec la notion d'isomorphisme en mathématiques (figure 1.7). Deux groupes peuvent avoir des éléments différents et pourtant avoir la même structure : en mathématiques, nous disons qu'ils sont isomorphes.

La nécessité de se munir d'une structure algébrique adaptée à l'image a été à l'origine de l'invention de la morphologie mathématique par Matheron et Serra qui, à l'instar des techniques issues du traitement

[2] On appelle groupe un ensemble G muni d'une loi de composition interne associative, admettant un élément neutre, tel que tout élément de G possède un inverse.

[3] On retrouve le concept d'agencement chez le philosophe Deleuze [100]

du signal et de l'automatique qui exploitaient au mieux la nature additive des signaux, sortent du cadre de l'espace vectoriel pour se placer dans le cadre des espaces munis d'une relation d'ordre (les treillis) présentée comme la structure mathématique capable de révéler au mieux les propriétés d'assemblage des formes élémentaires dans les images [81].

FIG. 1.6: Représentation de deux formes correspondant à la lettre « a ». Bien que ces deux formes soient composées de parties complètement différentes, elles représentent la même entité. Le tout est bien différent de ses parties. La structure prime sur les éléments.

.	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	(0,0)	(0,1)	(1,0)	(1,1)
(0,1)	(0,1)	(0,0)	(1,1)	(1,0)
(1,0)	(1,0)	(1,1)	(0,0)	(0,1)
(1,1)	(1,1)	(1,0)	(0,1)	(0,0)

o	i	s_1	s_2	s
i	i	s_1	s_2	s
s_1	s_1	i	s	s_2
s_2	s_2	s	i	s_1
s	s	s_2	s_1	i

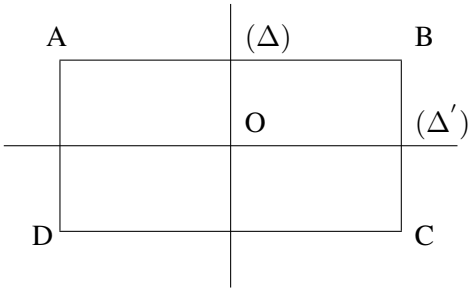


FIG. 1.7: Table d'addition du groupe de Klein, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ en haut à gauche. Si nous définissons par i l'identité, s_1 la réflexion d'axe Δ, s_2 la réflexion de Δ' et enfin s la symétrie de centre O, nous pouvons dresser une table équivalente pour le groupe des isométries laissant A, B, C, D invariants. Les deux groupes sont isomorphes mettant en avant la similitude de leurs structures internes bien que leurs éléments soient différents. En effet, le groupe de Klein est composé des éléments $\{(0,0),(0,1),(1,0),(1,1)\}$, alors que le groupe des isométries contient $\{i, s_1, s_2, s\}$.

1.1.3 Limite conceptuelle

La relation holiste est présente dans de nombreux domaines. En musique par exemple, la mélodie est bien plus qu'une simple somme de notes. Plus généralement, pour l'ensemble de nos sens, nous pouvons retrouver ce type de relation. Elle existe également, au niveau de notre raison dans notre faculté de jugement. Lorsque nous mettons en relation le soleil et la pierre pour déduire que la pierre est chaude, nous effectuons un jugement synthétique. Une étude approfondie de la raison et de notre faculté de jugement a été faite par Kant dans la Critique de la raison pure [66]. Le côté holiste chez Kant a quant à lui été mis en évidence dans la thèse de Yves Bouchard [24].

Kant montre que notre faculté de jugement dérive des Catégories de l'entendement inaccessible par notre raison. Nous ne pouvons appliquer notre raison que dans le champ de l'expérience. Cette conclusion montre la barrière conceptuelle de la raison : la relation holiste présente dans les jugements synthétiques échappe à la raison. On se heurte donc à une limite conceptuelle pour appréhender la forme comme un tout ; seules les parties de l'image sont accessibles.

Dans le domaine de l'image, ces préoccupations ont été à l'origine du développement de nombreuses théories parmi lesquelles la morphologie mathématique (G. Matheron, J. Serra 1967) fortement emprunte de la théorie de la Gestalt et des travaux sur la psycho-vision de Piaget. On peut qualifier la morphologie mathématique de théorie ensembliste dans la mesure où l'image est décomposée en entités élémentaires ou primitives à la manière d'un ensemble complexe qui résulte de l'agencement d'entités géométriques élémentaires. Les fondements de la morphologie mathématique résident dans le choix de la structure qui va servir de modèle d'agencement.

Dans un tout autre domaine et quelque dix années avant l'invention de la morphologie mathématique, en 1956, était organisée aux Etats-Unis la Conférence de Dartmouth, l'événement fondateur qui donnera naissance à l'intelligence artificielle. « Notre but est de procéder comme si tout aspect de l'apprentissage ou de toute autre caractéristique de l'intelligence pouvait être décrit d'une manière suffisamment précise pour être simulé par une machine », écrivaient alors les responsables de l'événement. Dans le domaine de l'image, le fondement théorique de la reconnaissance de forme est la décision bayésienne. L'idée fondamentale de la classification bayésienne est que le problème de décision peut être spécifié en termes de probabilités.

Faire interférer les processus d'analyse et de décision est une question récurrente dans les applications d'imagerie. Notamment, dans le cas où la nature des images et la nature des éléments à étudier

dans l'image sont connues, le dilemme peut se formuler ainsi : doit-on analyser en aveugle pour ensuite conclure sur la base de cette analyse et des connaissances a priori ? ou bien doit-on prendre en compte les a priori avant et/ou dans l'analyse ? La question est bien celle de l'organisation du processus du traitement de l'image. Avant de trancher, nous allons d'abord présenter les fondements de l'analyse morphologique et de la décision bayésienne.

1.2 L'analyse statistique et l'analyse géométrique : quels fondements ?

Pour une description complète des opérateurs morphologiques, le lecteur pourra se référer à l'un ou l'autre des livres de la liste non exhaustive suivante [80, 81, 123, 118, 119, 116, 124]. La présentation qui suit est largement inspirée du cours de Serra (disponible en ligne àl'adresse `http://cmm.ensmp.fr/~serra/cours/index.htm`).

Il existe un grand nombre d'approches dans le domaine du traitement des images. La liste qui suit donne un aperçu de leur diversité : la morphologie mathématique , l'approche variationnelle, l'approche statistique (modèles Bayesien, Markovien ...), les approches géométriques (géométrie projective, calibration, analyse/synthèse d'image ...), l'analyse mathématique (décomposition atomique, ondelette, multiéchelle, fractal ...), la théorie des graphes (appariement de graphes, graphes minimaux ...), les problèmes inverses (régularisations quadratique & semi-quadratique, Maximum d'Entropie ...), la reconnaissance des formes et des déformations (théorie des groupes, courbures, ...) ou encore la théorie psycho-visuelle (Gestalt ...).

Il est important de souligner qu'il existe de nombreuses passerelles entre les différentes méthodes. Il est par conséquent illusoire de penser qu'une méthode permet de résoudre le problème et l'autre non. Par contre, il est fondamental de bien analyser le problème que l'on souhaite résoudre afin d'utiliser la méthode la plus adaptée.

Nous avons souligné précédemment l'importance de la structure interne qui gouverne les assemblages des éléments qui composent le tout. Comment cela se traduit-il pour l'analyse des images ?

1.2.1 Treillis complets et morphologie mathématique

Dans le cas des signaux 1D, l'hypothèse de linéarité (validée par les modèles physiques dans nombre de cas) conduit à se situer dans une structure d'espace vectoriel. Un espace vectoriel sur un corps K est

un ensemble de *vecteurs* V muni d'une addition et muni d'une loi de composition externe entre les scalaires de K et les vecteurs de V tel que

- $(V, +)$ est un groupe commutatif
- V est stable par combinaison linéaire

Dans nombre de signaux physiques, on peut supposer qu'un signal complexe résulte de l'assemblage d'entités élémentaires et que l'assemblage s'effectue par combinaison linéaire (c'est le célèbre théorème de superposition).

En analyse des images, selon le contexte et la nature des images (i.e. leur formation) plusieurs types de structures peuvent avoir du sens. Pour l'analyse des textures, le caractère périodique ou pseudo-périodique de l'information engage plutôt à travailler dans une structure d'espace vectoriel. Pour l'analyse des formes dans les images, la structure linéaire se révèle mal adaptée pour modéliser tout un ensemble de phénomènes, par exemple les occlusions. Dans le cas des images binaires, les occlusions sont parfaitement modélisées par des agencements non linéaires, en l'occurrence le supremum ou l'infimum. Dans ce cas, la structure mathématique convenable n'est plus une structure d'espace vectoriel, mais une structure de treillis.

Définition 3 *Un treillis est un ensemble muni d'une relation d'ordre [80].*

Définition 4 " $\leq$ " *est une relation d'ordre (partiel) sur un ensemble* $\mathcal{E}$ *si* " $\leq$ " *vérifie* $\forall a, b, c$ *dans* $\mathcal{E}$ *:*

1. $a \leq a$ *(*$\leq$ *est réflexive)*
2. $a \leq b \ , \ b \leq a \Rightarrow a = b$ *(*$\leq$ *est antisymétrique)*
3. $a \leq b \leq c \Rightarrow a \leq c$ *(*$\leq$ *est transitive)*

L'ensemble des réels, des entiers ou l'ensemble des fonctions définies sur un espace E fixé munis de la relation d'ordre usuelle entre scalaires sont des treillis. Dans le cas des fonctions, on dira que

$$f \leq g \Leftrightarrow \forall p \in E, \ \ f(p) \leq g(p)$$

L'analyse fondée sur la structure d'espace vectoriel est l'analyse linéaire. L'analyse fondée sur la structure de treillis est l'analyse morphologique. De la même façon que l'analyse linéaire requiert la stabilité par combinaison linéaire, l'analyse morphologique nécessite un treillis complet.

Définition 5 *Un ensemble $\mathcal{E}$ muni d'un ordre partiel $\leq$ est un treillis complet si pour toute famille $\{E_i\}$ d'éléments de $\mathcal{E}$, $\mathcal{E}$ contient au moins un plus grand élément $\vee_i E_i$ et un plus petit élément $\wedge_i E_i$.*

Dans le treillis, les deux lois de composition fondamentales sont le sup ($\vee$) et l'inf ($\wedge$).

En l'analyse linéaire, l'opérateur fondamental (qui préserve la structure de travail et qui commute avec la combinaison linéaire) est la convolution. En morphologie mathématique, les opérateurs fondamentaux qui préservent la structure de travail (le treillis muni de sa relation d'ordre) et qui commutent avec les lois de composition sont la dilatation et l'érosion [80, 118]. Une dilatation est tout opérateur croissant qui commute avec le sup. Une érosion est tout opérateur croissant qui commute avec l'inf. Un opérateur T est croissant s'il préserve l'ordre dans le treillis :

$$\forall a \in \mathcal{E}, b \in \mathcal{E}, \;\; a \leq b \Rightarrow T(a) \leq T(b) \tag{1.1}$$

Définition 6 (dilatation et érosion morphologiques) *Soit $\mathcal{E}$ un treillis complet et T un opérateur sur $\mathcal{E}$ à valeurs dans $\mathcal{E}$:*

$$T : \mathcal{E} \rightarrow \mathcal{E}$$

– *T est une dilatation si T est croissant et commute avec le sup*

$$T(\vee_i E_i) = \vee_i T(E_i)$$

– *T une érosion si T est croissant et commute avec l'inf*

$$T(\wedge_i E_i) = \wedge_i T(E_i)$$

Dans le treillis, les images ne sont plus combinées par combinaison linéaire comme dans l'espace vectoriel. Elles sont combinées par infimum et supremum. La comparaison ne se fait plus via un produit scalaire, mais via la relation d'ordre dans le treillis.

Historiquement, la morphologie mathématique s'applique prioritairement aux images binaires et le treillis des parties d'un espace E est muni de la relation d'ordre usuelle sur les ensembles $\subseteq$. L'extension aux images numériques se fait naturellement en appliquant les opérateurs développés dans le cadre ensembliste aux sous-graphes, surgraphes ou ensembles de niveaux des fonctions numériques (sous

réserve quelles aient les propriétés requises, c'est-à-dire qu'elles soient semi-continues supérieurement ou inférieurement). Le treillis des fonctions définies dans un espace E est alors muni de la relation d'ordre usuelle sur les réels. Dans le cas binaire comme dans le cas numérique, l'ordre est total.

Définition 7 *Une relation d'ordre (au sens large) est totale si elle vérifie :*

$$\forall (x, y) \in V^2, \quad x \leq y \ \vee \ y \leq x \tag{1.2}$$

Dans une relation d'ordre partielle, deux éléments distincts peuvent ne pas pouvoir être mis en relation, ni dans un sens ni dans l'autre :

$$\exists (x, y) \in E^2, x \not\leq y \ \wedge \ y \not\leq x$$

Par contre, un ordre total est un ordre dont tous les éléments sont deux à deux comparables.

L'ensemble $\mathbb{R}$ muni de la relation d'ordre $\leq$ est totalement ordonné, mais l'ensemble des entiers naturels muni de la relation de divisibilité n'est que partiellement ordonné.

D'un point de vue théorique, la définition des opérateurs morphologiques ne nécessite pas de disposer d'un ordre total sur les éléments de l'ensemble. Un ordre partiel convient si en complément le treillis est complet. Dans la pratique la plus courante cependant, le treillis hérite de l'ordre total sur les ensembles ou sur $\mathbb{R}$ (pour les fonctions).

La nature non linéaire des lois de composition dans le treillis est en harmonie avec les agencements ensemblistes : ajouter un élément à une scène c'est occulter l'arrière-plan, la combinaison est bien de nature non linéaire (en l'occurrence une union ou supremum ou, de manière duale, une intersection ou infimum). Dans le cas des images numériques, les phénomènes d'occlusion dans les images sont ainsi (au moins partiellement) modélisés si l'on met à part la question de la profondeur de champ.

Ce qui motive également l'analyse morphologique des images c'est la volonté de tenir compte de la géométrie de l'image et des connaissances a priori sur la forme des objets recherchés dans l'image. Ceci est vrai dans un grand nombre d'applications d'imagerie, et particulièrement en imagerie médicale : en médecine, la géométrie des organes ou des lésions est rarement une inconnue ; la forme du myocarde par exemple est parfaitement connue ; celle d'un grain de beauté l'est également dans une certaine mesure. C'est d'ailleurs par apprentissage que le médecin analyse les images médicales et dépiste les

pathologies. Parce qu'elle place au coeur de son analyse le concept de forme et son rapport à l'image, la morphologie mathématique apparaît comme la démarche la plus à même de révéler la nature des formes qui composent une scène et donc de permettre la prise en compte d'informations morphologiques sur les structures dans l'analyse.

Comme nous l'avons dit, les deux opérateurs morphologiques fondamentaux sont la dilatation (opérateur croissant qui commute avec le sup) et l'érosion (opérateur croissant qui commute avec l'inf) et toute la pyramide des opérateurs morphologiques se construit à partir de ces deux opérateurs de base à partir du moment où le couple forme une adjonction (au sens de Galois) [80, 118].

Définition 8 (Adjonction morphologique) *Le couple (δ, ϵ) où δ est une dilatation et ϵ une érosion forme une adjonction si*

$$f \leq \epsilon(g) \Rightarrow \delta(f) \leq g \tag{1.3}$$

Cette condition satisfaite, les cascades $\delta \circ \epsilon$ et $\epsilon \circ \delta$ forment des ouvertures et fermetures morphologiques qui constituent les opérateurs fondamentaux du *filtrage morphologique*.

Définition 9 (Filtre morphologique) *Un filtre morphologique est tout opérateur de $\mathcal{E}$ dans $\mathcal{E}$ croissant et idempotent.*

La croissance assure la préservation de l'ordre dans le treillis. L'idempotence $T \circ T = T$ assure la stabilité des filtres par l'existence d'invariants ($T(.)$ est invariant par T).

Une ouverture est tout opérateur de $\mathcal{E}$ dans $\mathcal{E}$ croissant, idempotent et anti-extensif. Une fermeture est tout opérateur de $\mathcal{E}$ dans $\mathcal{E}$ croissant, idempotent et extensif.

Le caractère extensif ou anti-extensif du filtre indique comment le résultat du filtre s'ordonne par rapport à l'original dans le treillis. T est extensif si $T \geq Id$ et T est anti-extensif si $T \leq Id$ (Id est l'opérateur identité).

Les opérateurs morphologiques peuvent être regroupés en trois grandes familles : les transformations en tout ou rien (qui généralise les notions d'érosions et de dilatations), les filtres morphologiques (les ouvertures et fermetures, les filtres alternés et filtres alternés séquentiels qui s'obtiennent par cascade d'ouvertures ou fermetures), les résidus morphologiques qui se définissent par différence entre l'identité et un opérateur de type dilatation, érosion, ouverture ou fermeture (comme les gradients morphologiques

et les chapeaux haut de forme par exemple) ou par somme d'opérateurs élémentaires (comme le centre morphologique par exemple) [4].

1.2.2 Probabilités et classification bayésienne

Les a priori sur les images et les objets recherchés dans les images ne sont pas toujours de nature géométrique. Lorsqu'un médecin cherche une tumeur dans une radiographie, il met à contribution de nombreux paramètres pour analyser les lésions dans l'image et conclure sur la nature de la lésion. Tout aussi importants que les paramètres géométriques, les attributs statistiques sont en médical, comme dans d'autres applications souvent cruciaux. Et dans le domaine médical comme dans d'autres applications (en télédétection par exemple), on peut supposer qu'un échantillon (pour utiliser des termes propres à la statistique) de la ou des structure(s) recherchée(s) est disponible. En plus de permettre la prise en compte d'a priori statistique sur les images, l'analyse statistique vise également à combler le vide sémantique de la représentation digitale de l'image. Il s'agit de donner plus d'attributs au point que sa simple luminance ou couleur en lui associant d'autres mesures issues de l'analyse de l'image.

Du point de vue de l'analyse statistique, identifier une forme dans une image c'est décider en chaque point de l'image de son appartenance à une catégorie définie a priori. Plutôt qu'un problème de segmentation, rechercher des formes dans une image s'apparente bien à un problème de classification [5] où l'on affecte une classe à un objet en fonction de ses caractéristiques. Pour donner un exemple concret et simple de classification, prenons le cas d'une plante caractérisée par sa hauteur, sa couleur, sa largeur des pétales que nous souhaitons classer dans une catégorie : marguerites, roses. Toute la question de la classification par apprentissage consiste à décider comment classer de nouveaux objets (dans notre exemple d'une nouvelle plante) à partir des classes construites via les échantillons et des caractéristiques du nouveau candidat.

Le problème de la classification se résout très naturellement via les probabilités. La probabilité a priori d'observer une rose est calculée en divisant le nombre de roses par le nombre de plantes de l'échantillon :

$$P(plante = rose) = \frac{\text{nombre de roses}}{\text{nombre de plantes}}$$

[4] Cette classification écarte certaines transformations singulières comme la ligne de partage des eaux par exemple dont l'issue est binaire.

[5] Nous parlons ici de classification au sens de classification supervisée.

Les probabilités a priori permettent déjà d'obtenir une classification en retenant la probabilité la plus forte. Dans notre exemple, une plante « est » une marguerite si

$$P(plante = marguerite) = \frac{\text{nombre de marguerites}}{\text{nombre de plantes}} > P(plante = rose)$$

Cette solution est évidemment simpliste puisqu'elle ne tient pas compte des caractéristiques mesurées et donne toujours le même résultat. Pour tenir compte des caractéristiques, il faut faire appel aux probabilités conditionnelles : toujours dans notre exemple, la probabilité que la taille soit moyenne sachant que c'est une rose :

$$P(taille = moyenne|plante = rose)$$

et ainsi de suite pour chaque caractéristique. L'intérêt de cette approche en classification repose sur le théorème de Bayes qui donne la probabilité d'appartenance à la classe grâce aux caractéristiques mesurées. On parle alors de probabilité a posteriori.

Proposition 1 (formule de Bayes) *Si $(B_i)_{i\in I}$ est une suite finie ou dénombrable d'événements formant une partition de Ω telle que $P(B_i)$ soit strictement positif pour tous les i et si A est un événement quelconque tel que $P(A)$ soit strictement positif, on a :*

$$\boxed{P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)}} \tag{1.4}$$

Les modèles statistiques en traitement d'image considèrent que la valeur prise en un point du plan image (que cette image soit une fonction scalaire ou vectorielle) est la réalisation d'un processus aléatoire. Dans le cas des images numériques, la variable aléatoire est une application à valeurs dans l'espace des luminances (typiquement [0 ;255]).

Définition 10 *On appelle loi de probabilité d'une variable aléatoire réelle X la probabilité P_X définie pour tout intervalle B de $\mathbb{R}$ par :*

$$P_X(B) = P(\{\omega|X(\omega) \in B\}) \tag{1.5}$$

$$= P(\{X^{-1}(B)\}) \tag{1.6}$$

On utilise souvent la notation $P(X \in B)$ pour $P(\{\omega : X(\omega) \in B\})$. Ainsi que $P(a < X < b)$ pour $P(\{\omega : a < X(\omega) < b\})$.

Dans les cas des images multivaluées, on parle de vecteur aléatoire.

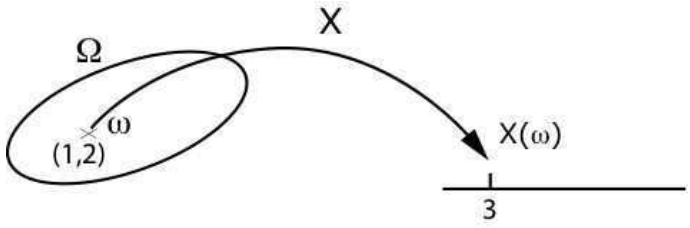


FIG. 1.8: Une variable aléatoire est une application X de Ω dans $\mathbb{R}$. Dans l'exemple ci-dessus d'un lancer de deux dés, nous nous intéressons à la somme. La variable aléatoire associe à chaque événement ω une valeur dans $[2; 12]$.

Définition 11 *On appelle vecteur aléatoire, ou variable aléatoire à valeurs dans $\mathbb{R}^d$, l'application mesurable de $\Omega \rightarrow \mathbb{R}^d$*

La loi de probabilité du vecteur aléatoire X est de la même façon définie par :

$$\forall B \in \mathbb{R}^d,\ P_X(B) = P(X^{-1}(B)) = P(X \in B) \tag{1.7}$$

Avant de classer un élément dans une classe, comme nous l'avons dit dans notre exemple introductif, on peut déjà chercher à savoir « quelle est sa probabilité d'appartenir à la classe C_1 sachant que ses caractéristiques sont X^l ? » ce qui revient à calculer la probabilité a posteriori $P(C_1|X^1)$. Le terme « a posteriori », est à prendre dans le sens de « après », par opposition aux probabilités $P(X^1)$ et $P(X^1|C_1)P(C_1)$ obtenues à partir de l'échantillon. La probabilité renseigne par conséquent, de la similarité avec la classe C_k à partir de ses caractéristiques X^l. La formule de Bayes permet de donner une réponse au problème de classification à partir des mesures effectuées. Le choix de la classe est fait en prenant la probabilité maximale :

$$p \in C_k \mid \max P(C_k|X^l)\ \forall k \tag{1.8}$$

Mais dans quelle mesure cette décision est-elle optimale ?

Comme le souligne [36], l'analyse bayésienne répond à un problème fondamental de l'induction [6] à savoir qu'il existe une manière optimale d'utiliser l'information disponible dans l'échantillon. Cette optimisation est à prendre dans le sens où elle minimise l'espérance du risque [51].

La probabilité d'appartenir à la classe C_k connaissant les caractéristiques X^l est donnée par $P(C_k|X^l)$. Notons L_{ij} le coût d'une mauvaise classification : cas où le classifieur affecte la classe C_j alors que la bonne classe est C_i. L'espérance du risque associée à l'affectation de la classe j pour les N classes est alors :

$$r_j(X^l) = \sum_{k=1}^{N} L_{kj} P(C_k|X^l) \tag{1.9}$$

C'est le risque moyen de se tromper pour la caractéristique X^l sur la classe C_j.

En utilisant la formule de Bayes (proposition 1) dans l'équation (1.9),

$$r_j(X^l) = \frac{1}{P(X^l)} \sum_{k=1}^{N} L_{kj} P(X^l|C_k) P(C_k) \tag{1.10}$$

$\frac{1}{P(X^l)}$ est toujours positif et indépendant de j. Minimiser $r_j(X^l)$ revient donc à minimiser

$$\sum_{k=1}^{N} L_{kj} P(X^l|C_k) P(C_k) \tag{1.11}$$

et la solution h^* qui minimise le risque en moyenne est obtenue en prenant le risque le plus faible parmi toutes les classes C_j possibles :

$$\boxed{h^* = \underset{j \in N}{\operatorname{argmin}} \sum_{k=1}^{N} L_{kj} P(X^l|C_k) P(C_k)} \tag{1.12}$$

Nous pouvons supposer les coûts de mauvaise classification tous égaux. En effet, dans le cas de la reconnaissance d'une forme dans une image, la décision est en tout ou rien : soit le point est bien affecté à la forme, soit il ne l'est pas. En notant par $\delta_{ij} = 1$ si $i = j$ et par $\delta_{ij} = 0$ lorsque $i \neq j$ la fonction de coût devient :

$$L_{ij} = 1 - \delta_{ij}$$

Par conséquent, lorsque la réponse est correcte : $L_{ij} = 0$ (en effet, $i = j$ et $\delta_{ij} = 1$).

[6] D'après [36], l'induction en intelligence artificielle est « le processus par lequel on tire des lois de portée générale en partant de l'observation de cas particuliers. »

L'équation 1.12 peut se réécrire :

$$h^* = \underset{j \in N}{\operatorname{argmin}} \sum_{k=1}^{N} (1 - \delta_{kj}) P(X^l|C_k) P(C_k) \tag{1.13}$$

$$= \underset{j \in N}{\operatorname{argmin}} \sum_{k=1}^{N} P(X^l|C_k) P(C_k) - \sum_{k=1}^{N} \delta_{kj} P(X^l|C_k) P(C_k) \tag{1.14}$$

Cette dernière équation se simplifie, puisque les termes de la somme $\sum_{k=1}^{N} \delta_{kj} P(X^l|C_k) P(C_k)$ sont tous nuls excepté si $\delta_{kj} = 1$, c'est-à-dire si $k = j$. La somme vaut donc $P(X^l|C_j)P(C_j)$. De plus, d'après la formule des probabilités totales, $\sum_{k=1}^{N} P(X^l|C_k)P(C_k) = P(X^l)$. Nous pouvons par conséquent, réécrire l'équation 1.13 :

$$h^* = \underset{j \in N}{\operatorname{argmin}} P(X^l) - P(X^l|C_j)P(C_j) \tag{1.15}$$

$$= \underset{j \in N}{\operatorname{argmax}} P(X^l|C_j)P(C_j) \tag{1.16}$$

Cette dernière équation simplifiée[7], nous obtenons la règle de maximum à posteriori [36].

Définition 12 *La règle de décision bayésienne lorsque les coûts sont égaux devient la règle du Maximum A Posteriori :* (MAP)*.*

$$\boxed{h^* = \underset{j \in N}{\operatorname{argmax}} P(X^l|C_j)P(C_j)} \tag{1.17}$$

La règle du MAP donne bien la meilleure affectation puisqu'elle minimise le risque de mauvaise classification. Disposant d'un échantillon de référence indiquant la nature statistique des objets recherchés dans une image (ou des objets d'intérêt dans une image), le MAP fournit une solution pour classer chaque point de l'image de façon rigoureuse. Cependant, il reste que la stratégie à mettre en oeuvre pour minimiser l'erreur de classification est absolument dépendante de l'échantillon et de ses propriétés statistiques. Comme nous le verrons plus loin, dans le cas où l'échantillon suit une loi gaussienne, le MAP revient à minimiser une distance statistique : la distance de Mahalanobis.

[7] $P(X^l)$ n'influence pas la minimisation, puisque positif et identique quelque soit j. Quant au signe moins, il renverse le problème de minimisation en un problème de maximisation.

1.2.3 Les stratégies mixtes

Les stratégies mixtes en traitement des images sont courantes. De nombreuses coopérations entre l'analyse morphologique et la classification bayésienne ont déjà été imaginées dans différents contextes. Forbes et Raftery proposent de combiner les deux approches afin d'utiliser la vitesse d'exécution de la morphologie mathématique, bien moins coûteuse en temps de calcul que les systèmes de décision bayésiens. Les auteurs parlent de « Bayesian Morphology » [45]. Bien souvent la combinaison des deux techniques est utilisée pour lier l'approche locale apportée par la morphologie mathématique avec les contraintes globales données par l'approche bayésienne. Une utilisation pour la détection de routes est présentée par [68]. L'idée est d'extraire les routes à partir de filtres et d'outils de segmentation morphologiques (en l'occurrence la *LPE*) puis d'affiner la reconstruction des routes à partir d'un modèle Markovien global du réseau routier. La combinaison de la morphologie mathématique avec d'autres méthodes fait l'objet de nombreuses extensions prometteuses, voir notamment les travaux de Bloch [21, 22] sur la morphologie floue.

1.3 Spécificité des images multivaluées

1.3.1 De l'image scalaire à l'image multivaluée

« Transformer c'est comprendre ». Cette phrase énoncée par Jean Serra dans [123] pourrait être mise en exergue de tout livre sur l'analyse des données. Les signaux et les images dans leur représentation d'origine, étant occulte, il faut changer de représentation, les transformer pour percevoir l'information. Ainsi, le filtrage de Fourier puis plus tard les ondelettes, l'analyse granulométrique, l'analyse multiéchelle, l'analyse statistique s'attachent à construire à partir d'un signal d'origine une famille de mesures ou de signaux dérivés susceptibles de faire émerger le sens. En contrepartie, ces transformations modifient souvent le signal en une famille de signaux. On parle alors d'images vectorielles ou multivaluées.

Autre configuration possible : des images multivaluées sont construites afin de renforcer une information déjà présente dans les données. Nous présenterons en section 2.2.2 un exemple illustrant l'exploitation de telles images pour la segmentation d'un « patchwork » : une image binaire 2D est transformée en une pile de 15 bandes à partir d'une analyse statistique de sa granulométrie (voir figure 2.25). Nous avons également dans ce travail modifié l'IRM cardiaque en une image à 2 bandes en intégrant des statistiques sur la géométrie du coeur. Toutes ces modifications surdimensionnent l'image d'origine afin de

renforcer l'information utile et afin d'aider à converger vers la solution. Bien évidemment, cela implique une plus grande complexité des traitements, aussi bien théoriquement que numériquement.

Enfin, si les images multivaluées peuvent émerger de l'analyse, elles peuvent également être générées directement par le système d'acquisition : imagerie médicale PET-CT, image de radar (imagerie hyperspectrale) et bien évidemment images couleur.

Parmi toutes les techniques développées pour l'analyse des images, certaines se transposent directement du cas scalaire au cas vectoriel. C'est le cas notamment des méthodes variationnelles (elles sont basées sur des équations de diffusion et correspondent, dans le cas des images couleur, à un transport local de la couleur), des champs de Markov [35] et plus généralement de toutes les méthodes intégrant des probabilités (celles-ci s'étendant relativement facilement aux opérations matricielles), de l'analyse harmonique...

Il n'en est pas de même des méthodes fondées sur une relation d'ordre comme c'est le cas de la morphologie mathématique. Comme nous l'avons vu, l'analyse morphologique est construite sur des opérations fondamentales de supremum et d'infimum qui n'ont plus de sens dans le cas où les images ont des valeurs vectorielles.

L'énorme développement de l'imagerie multivaluée et notamment de l'imagerie couleur s'est accompagné d'une importante intensification des recherches dans le domaine de l'analyse des images multivaluées et notamment dans le domaine de l'analyse morphologique multivaluée. Si les travaux réalisés sur ce thème depuis une dizaine d'années portent majoritairement sur la couleur, certains travaux de fond ont tenté de clarifier le fondement du passage de la morphologie sur scalaires à la morphologie sur vecteurs. Avant d'en faire état, nous allons d'abord préciser quelques notions importantes concernant les images couleur.

1.3.2 Les images couleur et les espaces colorimétriques

Les images couleur sont devenues très courantes dans le monde domestique comme dans le monde industriel. Nous distinguons environ 24 niveaux de gris différents alors que l'ordre de grandeur pour la couleur est du millier. Nous pouvons par conséquent penser que la majorité du travail en reconnaissance de formes est effectué sur ce type d'imagerie. Cependant, la plupart des algorithmes se font sur des images en niveaux de gris dont l'extension aux images couleur passe souvent par une simple transposition sur chaque composante de la couleur (revue récente sur la segmentation et les images couleur [28]).

Une première raison provient de la représentation mathématique de la couleur dans un espace à trois dimensions, alors qu'il ne suffit que d'une dimension pour représenter le niveau de gris. Mais la principale difficulté réside dans la comparaison des couleurs : comment comparer un rouge d'un bleu ? Dans quelle mesure le vert est-il ressemblant ou dissemblable du jaune ? Comment définir un gradient ? etc.

Pour simplifier ce problème sans le résoudre, divers espaces de couleur sont possibles : le système R.G.B (figures 1.9 et 1.10) est la représentation la plus commune. Les composantes (Red, Green, Blue) permettent une représentation complète du spectre des couleurs grâce aux lois de la colorimétrie [27]. Les raisons du succès de cette représentation tiennent au fait que technologiquement la plupart des écrans produisent la couleur à partir de ces trois composantes. De plus, l'analyse physiologique des cellules de l'oeil montre une sensibilité à ces trois rayonnements, donnant ainsi au système une crédibilité. Cependant, il existe bien d'autres systèmes de représentation. Une présentation très complète sur ces différents espaces colorimétriques est faite à l'adresse suivante :

`http://www.profil-couleur.com/ec/101-espace-couleur-generalite.php`.

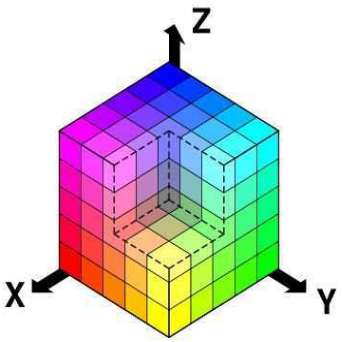


FIG. 1.9: Représentation du cube RGB. Chaque couleur du spectre lumineux peut être obtenue par composition des composantes RGB. (origine de l'image : wikipedia)

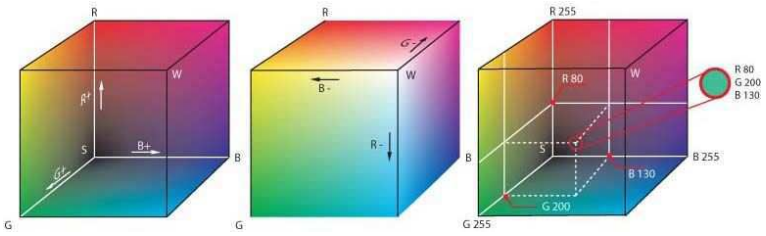


FIG. 1.10: Illustration de la représentation géométrique dans l'espace 3D d'une couleur à partir des composantes RGB. (origine de l'image : wikipedia)

L'étude des systèmes colorimétriques est un sujet d'étude très important dès les années 20 avec la création en 1931 de la commission internationale de l'éclairage (CIE). L'objectif est de créer un espace couleur tel que le cerveau humain le perçoit. Cette modélisation de la vision humaine aboutit à différents modèles dont le CIE L*a*b* adopté en 1976 très couramment utilisé dans le domaine de la segmentation des images couleur (figure 1.11). Un des intérêts de ce modèle tient au fait que la luminosité telle que nous la percevons et non la luminance exacte est dissociée de l'aspect chromatique contenu dans chaque cercle du plan. Ce fonctionnement est très proche de notre représentation visuelle de la couleur permettant ainsi au modèle d'être intuitif. Un autre point intéressant comme le souligne Enficiaud dans sa thèse [43] est que ce système a été défini afin d'appliquer la distance euclidienne. Ainsi, dans l'espace CIE L*a*b* la distance euclidienne entre deux couleurs voisines correspond à un écart tel que nous le percevons. Il faut souligner qu'il s'agit d'écarts de ressemblance dans le voisinage d'une couleur est qu'il ne faudrait surtout pas surinterpréter le sens des distances euclidiennes.

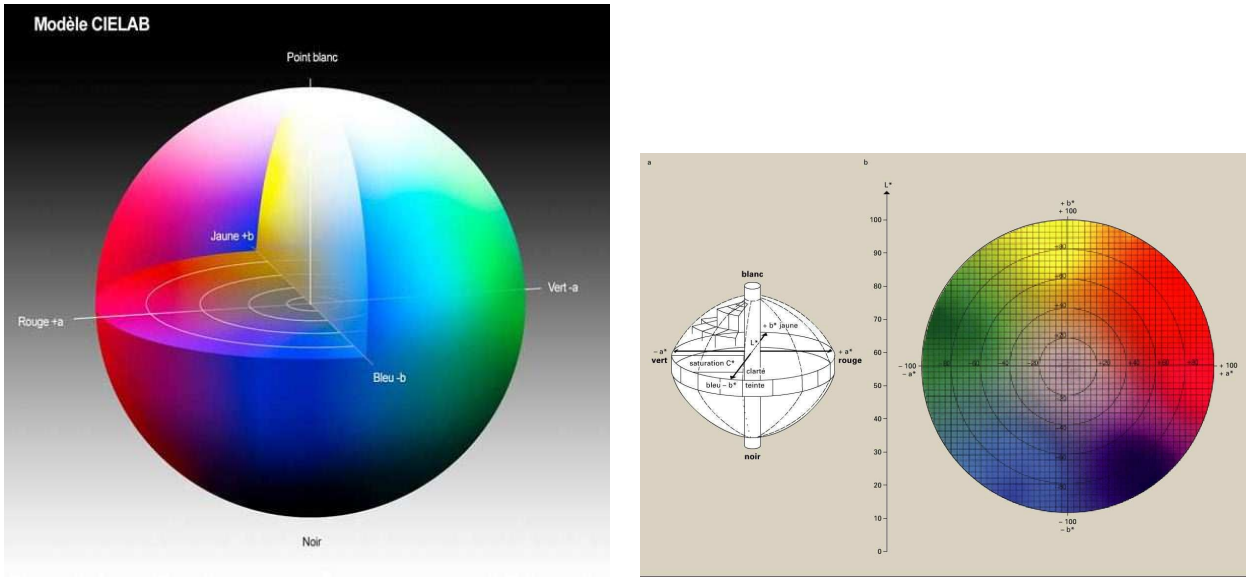


FIG. 1.11: L'espace colorimétrique CIE L*a*b* est représenté par une sphère aplatie. L'axe vertical représente la luminosité obtenue à partir d'une échelle psychométrique variant du noir au blanc. L'aspect chromatique se visualise à travers la teinte et la saturation suivant deux axes horizontaux : vert-rouge (a*/+a*) et bleu-jaune (b*/b+*). La couleur est saturée progressivement en partant du centre neutre (gris) vers les bords, sur chaque plan horizontal. (Origines des images : à gauche, *profil-couleur.com*, à droite Encyclopédie Universalis 2007)

Nous avons présenté les deux principaux systèmes colorimétriques, mais ils sont nombreux et le choix est difficile puisqu'aucun système ne l'emporte sur les autres [50]. Sélectionner le meilleur système pour un problème donné est un des enjeux pour le succès de l'algorithme.

1.4 Analyse morphologique des images multivaluées

1.4.1 Etat de l'art sur la morphologie multivaluée

Comme nous l'avons dit plus haut, l'extension des opérateurs morphologiques valides pour les ensembles et les fonctions scalaires au cas des fonctions vectorielles n'est pas immédiate. Une expérience simple consiste à tester l'approche vectorielle naïve qui construit l'opérateur sur les vecteurs en appliquant l'opérateur scalaire sur chaque composante du vecteur : voir figure 1.12. Cette solution attrayante par sa simplicité n'est pas correcte dès lors que les composantes vectorielles sont corrélées. Nous reviendrons sur ce point lorsque nous présenterons cette approche en détail dans la section 1.4.2.

Pour étendre l'analyse morphologique aux images multivaluées [122, 53, 54], il faut disposer d'un treillis complet des images multivaluées et d'une relation d'ordre sur les vecteurs. En théorie, seul un ordre partiel est nécessaire. Dans [53], Goutsias et Heijmans établissent les fondements de l'analyse morphologique sur images multivaluées et introduisent le concept de h-adjonction pour définir des filtres morphologiques sans ordre total sur les vecteurs. Cette approche est à la base de nombreuses stratégies pour mettre en oeuvre la morphologie mathématique multivaluée. Le principe consiste à définir une relation d'ordre dans l'espace des images multivaluées (notées F) à partir d'une relation d'ordre existante dans un autre treillis, par le biais d'une application h :

$$\forall x, x' \in F,\ x \leq_h x' \Longleftrightarrow h(x) \leq h(x')$$

La solution théorique est attractive, mais sa mise en oeuvre pratique est délicate. En effet, l'absence d'un ordre total pose des difficultés pratiques importantes lors de la mise en oeuvre des filtres morphologiques [127]. Dans [127], Talbot montre qu'il est nécessaire de se munir d'un ordre total si l'on veut éviter de créer de nouveaux vecteurs dans les images par l'action des opérateurs morphologiques.

Nous le voyons, le fait d'ordonner les vecteurs est un point primordial de l'analyse morphologique multivaluée. Voyons maintenant les différentes stratégies qui ont été proposées dans la littérature.

1.4.2 Quatre approches pour ordonner des vecteurs

Selon Barnett, il y a quatre approches différentes pour ordonner des vecteurs : l'approche marginale, l'approche partielle, l'approche conditionnelle et la réduction des dimensions [15]. Elles sont présentées

ci-après. Il existe d'autres stratégies plus exotiques qui ne seront pas exposées ici pour des raisons de concision. Nous renvoyons le lecteur aux travaux d'Aptoula et Lefèvre sur le sujet [12, 10].

Ordre marginal L'approche marginale (M-ordre) est la plus simple puisqu'elle consiste à traiter chaque bande (chaque composante vectorielle) séparément. Cette simplicité est accompagnée d'une rapidité de traitement puisque les opérations sont effectuées composante par composante. L'ordre partiel [8] de la structure se définit par l'équation 1.18.

$$\forall \mathbf{v}, \mathbf{v}^{'} \in \mathbb{R}^n, \quad \mathbf{v} \leq \mathbf{v}^{'} \quad \Leftrightarrow \quad \forall i \in \{1, \ldots, n\}, \quad v_i \leq v_i^{'} \tag{1.18}$$

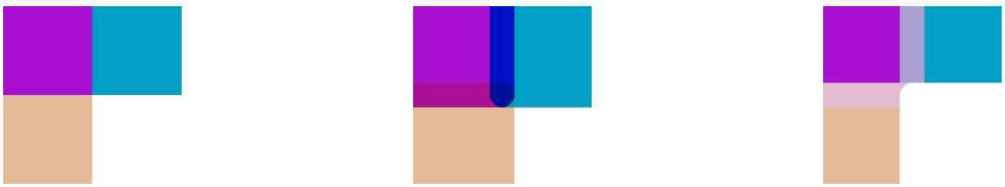


FIG. 1.12: De gauche à droite : l'image originale, son érodée, puis la dilatée avec un disque de taille 10. L'approche marginale génère de nouvelles couleurs.

L'inconvénient majeur de l'approche marginale est la création de nouveaux éléments comme nous le montrons sur la figure 1.12. L'espace des valeurs est préservé par les opérateurs non linéaires (et notamment morphologiques) : dans le cas des images scalaires, une image codée sur 8 bits à valeurs entre 0 et 255 sera transformée en une image ayant encore ses valeurs dans le même domaine. Dans le cas de l'approche marginale, ce n'est pas le cas : de nouvelles valeurs sont créées. Cela dit, nombre de travaux ont été réalisés sur la base de cette approche qui donnent de bons résultats : citons les travaux de Gu pour le suivi d'objet [58] ou encore ceux d'Aptoula pour la segmentation des galaxies [11]). D'autre part, il est encore possible de pallier au problème de corrélation des variables en décorrélant les composantes par différentes stratégies. Des solutions basées sur l'analyse en composantes principales (ACP) sont présentées dans [17, 73].

[8] En effet, tous les éléments ne peuvent pas être comparés, par exemple $\mathbf{v} = [7, 3], \mathbf{v}^{'} = [5, 6]$ ne sont pas comparables car les ordres induits sur chaque composante diffèrent

Ces approches marginales peuvent se généraliser par une approche matricielle (voir les travaux de Wilson [136]) à partir de la formulation proposée par Heijmans et Ronse de [63] ; ces travaux s'insèrent d'ailleurs dans le cadre théorique plus large des h-adjonctions proposé par Goutsias et Heijmans [53].

Ordre partiel L'approche partielle (P-ordre) est basée sur le regroupement des vecteurs en classes d'équivalence en fonction de critères comme l'ordre, le rang ou l'« extremeness » [72]. Cette approche géométrique permet d'ordonner les vecteurs par des enveloppes convexes incluses les unes dans les autres (« convex hull peeling » [15], voir figure 1.13).

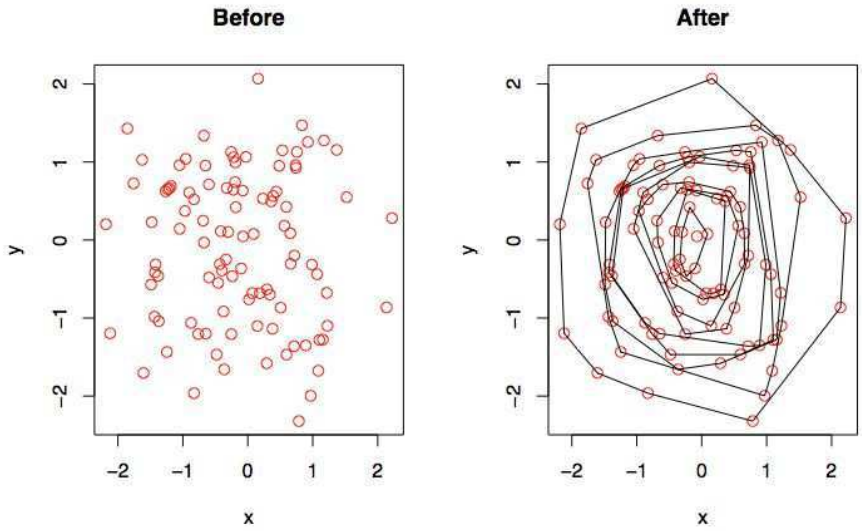


FIG. 1.13: Illustration du principe du « convex hull peeling », les vecteurs sont regroupés en enveloppes convexes incluses les unes dans les autres (image extraite de la présentation sur le sujet par [72]).

Définition 13 (Ensemble convexe) *Un ensemble $C \subseteq \mathbb{R}^n$ est convexe si pour chaque point $x, y \in C$, le segment $[xy]$ est contenu dans C.*

Définition 14 (Enveloppe convexe) *L'enveloppe convexe « convex hull » d'un ensemble de points X de $\mathbb{R}^n$ est notée $CH(X)$, et correspond à l'intersection de tous les ensembles convexes de $\mathbb{R}^n$ contenant X.*

Les vecteurs sont tous comparables, cependant il ne s'agit que d'un pré-ordre. En effet, le regroupement en classe d'équivalence ne permet plus de distinguer deux vecteurs de la même classe, la création d'un ordre total n'est par conséquent pas possible dans ce cadre strict.

Ordre conditionnel L'approche conditionnelle (C-ordre) est basée sur la sélection de composantes afin d'ordonner les vecteurs. L'équation 1.19, traduit un choix sur la première composante.

$$\forall \mathbf{v}, \mathbf{v}^{'} \in \mathbb{R}^n, \quad \mathbf{v} \leq \mathbf{v}^{'} \quad \Leftrightarrow \quad v_1 \leq v_1^{'} \tag{1.19}$$

La difficulté est de choisir les composantes qui permettront de classer les vecteurs. Cette approche est utilisée lorsque le poids de certaines composantes est plus important.

Un exemple particulièrement utilisé de « C-ordre », est l'approche lexicographique née de l'idée de traiter les composantes dans l'ordre alphabétique. Le terme de lexicographique vient d'ailleurs de ce processus de décision qui s'effectue dans l'ordre du dictionnaire ou dans l'ordre alphabétique. Par exemple, lorsque les éléments du produit cartésien de deux ensembles A et B , $A \times B$, doivent être comparés, l'approche lexicographique propose de traiter la lettre a puis b :

$$(a, b) \leq (a^{'}, b^{'}) \quad \text{si et seulement si} \quad a < a^{'} \text{ ou } (a = a^{'} \text{ et } b \leq b^{'})$$

Si maintenant toutes les composantes sont utilisées, on parle de cascade lexicographique et l'ordre est total.

L'ordre lexicographique est couramment utilisé dans le domaine de la couleur [122, 60, 8]. Bien souvent l'espace colorimétrique associé est un espace de type Teinte-Luminance-Saturation (TLS). En effet, le fait d'ordonner les vecteurs par rapport à leur composante est d'autant plus pertinent que les composantes sont significatives. Or, dans l'espace RGB la corrélation est forte entre les composantes. Par contre, dans les espaces de type TLS, la cascade lexicographique est pertinente. Cependant, même dans ce type d'espace, il peut être parfois judicieux de pondérer le poids de la première composante (la teinte) afin de donner davantage d'importance à la seconde ; on parle alors d'ordre lexicographique modulo α [6].

Ordonner les vecteurs reste toujours un problème complexe, qui peut être simplifié en choisissant un espace adapté. Cela dit, même si les espaces de type TLS semblent plus pertinents pour la couleur il ne faut pas omettre le traitement angulaire de la teinte. Les travaux de Hanbury et Serra fournissent un cadre théorique et des solutions à ce problème [59, 60]. L'utilisation d'une référence h_0 permet de mesurer l'accroissement angulaire de la teinte. L'ordre total est obtenu en ajoutant une seconde condition sur les autres composantes.

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \prec \begin{pmatrix} v_1^{'} \\ v_2 \\ \vdots \\ v_n^{'} \end{pmatrix} \Leftrightarrow \begin{array}{l} v_1 < v_1^{'} \quad \text{ou} \\ v_1 = v_1^{'} \quad \text{et} \quad v_2 < v_2^{'} \\ \dots \\ v_1 = v_1^{'}, \dots, v_{n-1} = v_{n-1}^{'} \text{et} v_n < v_n^{'} \end{array} \tag{1.20}$$

Réduction des dimensions Cette approche (encore appelée R-ordre) consiste à réduire les vecteurs de dimension n à un scalaire afin de les ordonner avec la relation d'ordre usuelle. On se donne donc une application h à valeurs dans $\mathbb{R}$.

$$\begin{array}{rccc} h: & \mathbb{R}^n & \rightarrow & \mathbb{R} \\ & \mathbf{v} & \rightarrow & h(\mathbf{v}) \end{array} \tag{1.21}$$

Les approches de R-ordering se divisent généralement en deux groupes : celles basées sur les distances et celles utilisant une projection. L'injectivité de l'application n'étant pas obligatoire, seul un pré-ordre est garanti : deux vecteurs distincts peuvent très bien-être à équidistance de la référence.

$$\forall \mathbf{v}, \mathbf{v}^{'} \in \mathbb{R}^n, \quad \mathbf{v} \leq \mathbf{v}^{'} \quad \Leftrightarrow \quad h(\mathbf{v}) \leq h(\mathbf{v}^{'}) \tag{1.22}$$

L'utilisation d'une distance impose de choisir une référence $\mathbf{v}_0$ et une métrique pour comparer les vecteurs. Alors, l'équation générale 1.22 devient :

$$\forall \mathbf{v}, \mathbf{v}^{'} \in \mathbb{R}^n, \quad \mathbf{v} \leq \mathbf{v}^{'} \quad \Leftrightarrow \quad d(\mathbf{v}, \mathbf{v}_0) \leq d(\mathbf{v}^{'}, \mathbf{v}_0) \tag{1.23}$$

Le R-ordre est très intéressant par son principe de réduction des dimensions. Même si la technologie permet de manipuler plus facilement les vecteurs, le travail sur des scalaires à de nombreux avantages : vitesse de traitement, réutilisation des bibliothèques de traitement d'image existantes, visualisation, nombreux algorithmes... Des applications sur des images spectrales notamment peuvent être citées en exemple [102].

Il est toutefois possible de ne pas s'arrêter au R-ordre partiel et d'ajouter des conditions supplémentaires en cas d'égalité afin d'obtenir un ordre total. Ainsi, les bénéfices sont doubles : réduction des dimensions et création d'un ordre total. Différentes stratégies sont possibles pour départager les vecteurs ayant la même distance : il est possible d'utiliser des critères géométriques [34, 33], ou de combiner le R-ordre avec une approche lexicographique (approche mixte R-ordre et C-ordre) [9, 114]. Différentes

combinaisons en fonction des espaces colorimétriques sont possibles, ainsi Angulo propose les priorités suivantes en fonction de l'espace colorimétrique [9] :

- RGB : G > R > B (importance relative de chaque composante perçue)
- LSH : L>S>$H\%h_0$ (origine de la teinte h_0)
- L*a*b* : L >a* >b*, a* et b* sont pris arbitrairement

Il est toutefois possible dans le cas particulier des images à teintes de gris d'obtenir un ordre total par le simple biais d'une distance. Il suffit de prende comme référence le maximum (ou le mininimum) de l'image et la distance usuelle permet de départager sans ambiguïté chaque point [9]. Bien évidemment, si la référence n'est plus positionnée en un point d'extrémité, l'antisymétrie est brisée.

Dans ces approches mixtes, et plus généralement pour le R-ordre, le choix de la distance est un point crucial. Toutes les distances sont utilisables (les normes L^p, les distances statistiques...). La distance euclidienne est très souvent utilisée dans les espaces L*a*b* puisque justement cet espace est construit de telle sorte que la différence psycho-visuelle et la norme euclidienne soient corrélées : un écart dans l'espace de couleur correspond à un même écart dans l'espace euclidien. L'ensemble des auteurs précédemment cités combinent donc cette distance avec un espace de ce type. L'utilisation de la distance euclidienne dans l'espace RGB est possible ([33]) mais cela reste très rare. La distance de Mahalanobis quant à elle a fait l'objet de plusieurs travaux [3, 53]. Dans [3], la référence est une couleur fixe (le noir). Dans [53], une utilisation plus générale est proposée. Lorsqu'on travaille avec une distance statistique comme la distance de Mahalanobis, l'espace couleur est considéré comme un espace $3D$ standard sans qu'il y ait d'interprétation particulière de la représentation colorimétrique. De fait, la méthode s'exporte bien au cas des images multivaluées non couleur. En toute généralité, les distances statistiques permettent de classifier les vecteurs par rapport à une famille d'échantillons et ceci quelque soit la nature des données vectorielles.

Un autre point fort des distances statistiques est de fournir des cartes de similarité. Si dans l'espace couleur, l'ordre est toujours délicat (dire que le bleu est plus petit que le jaune est toujours discutable), dans l'espace des similarités, les comparaisons ont une interprétation simple et toujours cohérente quelques soient la nature des données (couleur ou autre) : $v_1 < v_2$ signifie que les caractéristiques du vecteur v_1 sont plus proches de l'échantillon de référence que celles du vecteur v_2.

1.4.3 Le cas particulier de l'ordre total

L'ordre total obtenu par les approches présentées ou par d'autres astucieuses techniques (par exemple : le « bit mixing paradigm »de [70, 25]) permet de déduire sur les images vectorielles des opérateurs morphologiques connus pour les images à valeurs scalaires.

Nous avons vu qu'un ordre T sur un ensemble de vecteurs est une transformation injective qui associe à tout vecteur d'un sous-ensemble $\mathcal{V}$ de $\mathbb{R}^p$ une valeur scalaire (par exemple réelle ou entière) :

$$\begin{array}{ccc} \mathcal{V} \subset \mathbb{R}^p & \overset{\mathrm{T}}{\rightarrow} & \mathbb{R} \\ V & \mapsto & \mathrm{T}(V) \end{array}$$

L'ordre est total si en plus, la transformation est surjective, c'est-à-dire que deux vecteurs distincts de $\mathcal{V}$ seront associés à deux ordres différents :

$$V' \neq V \Rightarrow \mathrm{T}(V) \neq \mathrm{T}(V')$$

En d'autres termes, l'ordre est total si T est inversible.

Notons $\mathrm{T}(F)$ l'image de F par T : $\mathrm{T}(F)(p) = \mathrm{T}[F(p)]$. Si T est inversible, l'image multivaluée orginale F peut être reconstruite à partir de la carte $\mathrm{T}(F)$:

$$F \overset{\mathrm{T}}{\longrightarrow} f = \mathrm{T}(F) \overset{\mathrm{T}^{-1}}{\longrightarrow} F \tag{1.24}$$

Dans le cas où T est une bijection, l'image de F par T, notée $f = \mathrm{T}(F)$, sera appelée *image de rang* associée à F. f est alors une représentation sans perte de F.

Toutes les images F ayant leurs valeurs dans $\mathcal{V} \subset \mathbb{R}^p$ peuvent être comparés au sens de l'ordre T [52] :

$$F \leq_{\mathrm{T}} G \Leftrightarrow \forall p \in E,\ \ \mathrm{T}[F(p)] \leq \mathrm{T}[G(p)]$$

(ici, E est le domaine de définition commun à F et G).

Ainsi, la définition d'un opérateur morphologique (dilation, érosion, ouverture, fermeture...) agissant sur l'image vectorielle F résulte du schéma de décomposition/reconstruction 1.24. Par exemple, si φ désigne un des quatre opérateurs morphologiques fondamentaux agissant sur les fonctions à valeurs

scalaires, alors l'opérateur vectoriel associé peut être construit par la cascade :

$$\phi = \mathrm{T}^{-1} \circ \varphi \circ \mathrm{T} \tag{1.25}$$

Le calcul des opérateurs vectoriels se ramène donc au calcul d'un opérateur usuel sur fonction scalaire puis au retour dans l'espace des valeurs vectorielles :

$$\begin{array}{ccc} F & \overset{\phi}{\longrightarrow} & G \\ \mathrm{T}\ \downarrow & & \uparrow\ \mathrm{T}^{-1} \\ f & \underset{\varphi}{\longrightarrow} & g \end{array}$$

Si la transformation T qui associe un ordre aux vecteurs est une bijection, alors tous les opérateurs morphologiques fondamentaux basés sur le supremum et l'infimum se définissent très simplement. Mais ceci n'est plus vrai des résidus (gradients morphologiques et chapeaux haut de forme) puisqu'ils nécessitent d'effectuer des soustractions entre vecteurs.

Deux vecteurs V_1 et V_2 étant ordonnés via leur image par T, $r_1 = \mathrm{T}(V_1)$ et $r_2 = \mathrm{T}(V_2)$ dans l'espace de représentation scalaire, la différence $V_1 - V_2$ est définie via la différence $r_1 - r_2$. Mais que représente cette nouvelle distance dans l'espace original de représentation des vecteurs (dans l'espace couleur par exemple) ? Très classiquement en morphologie couleur, on contourne la difficulté : les soustractions sont effectuées sur une unique composante (correspondant à la luminance dans l'espace L*a*b* par exemple) indépendamment de l'ordre sur les vecteurs [4] ou le gradient est défini en évaluant les sauts entre les zones de couleur uniforme [125]. On peut également calculer la distance entre la dilatation et l'érosion, afin de ne pas avoir recours à un opérateur dans le treillis [5]. Il est à noter que ces gradients sont valables en LSH et L*a*b*, leur extension en RGB est nettement plus problématique en raison de la forte corrélation entre les composantes. La distance peut être calculée sur une seule composante, sur 2 ou 3 ; gradient achromatique, chromatique ou couleur (distance par rapport aux 3 composantes). Cela n'a encore une fois de sens que dans la représentation L*a*b*, là où la distance euclidienne à un sens. Dans le cas des espaces LSH il est toujours possible de mesure l'accroissement angulaire de la teinte et de prendre la différence [60, 61].

Dans l'ensemble des travaux mentionnés, l'opérateur de soustraction n'est jamais défini. Nous verrons au chapitre 2 que le problème peut se résoudre en recourant à une image de rang constituée d'entiers consécutifs.

1.5 Conclusion

Résumons le but et les moyens. Nous cherchons à analyser des images multivaluées, c'est-à-dire notamment à les filtrer, à les segmenter, en tenant compte d'un a priori statistique sur les éléments significatifs de l'image. Nous proposons de fonder la reconnaissance des éléments significatifs sur le principe de la décision bayésienne, l'analyse proprement dite étant réalisée via les opérateurs morphologiques. L'analyse des images multivaluées nécessitant d'ordonner les vecteurs composant l'image, nous proposons de construire une carte de similarité par classification des points de l'image afin d'utiliser cette carte comme moteur d'analyse de l'image. La construction de cette structure et son exploitation sont l'objet du chapitre qui suit.

Chapitre 2

Analyse morphologique des images avec a priori statistique

Pour prendre en compte les a priori statistiques et géométriques dans l'analyse des images (supposées de plus multivaluées), il est proposé de recourir à une analyse en deux temps : la première étape consiste à calculer la similarité de chaque vecteur de l'image multivaluée à l'échantillon (selon la théorie de la décision bayésienne) ; dans la seconde étape, les traitements morphologiques sur les images multivaluées sont définis comme une extension de leur équivalent scalaire. Ce schéma est résumé en figure 2.1. Cette analyse n'est possible que s'il est possible de remplacer sans perte l'image multivaluée par sa carte de similarité.

Comme une suite logique à cette démarche, nous examinons ensuite la possibilité de définir des filtres susceptibles de s'adapter localement en chaque point de l'image à la nature statistique du point, c'est-à-dire à sa similarité. La question à laquelle on se propose de réponde est la suivante : est-il possible de simplifier l'image en préservant intégralement les éléments de la scène proche de la référence statistique tout en simplifiant à l'extrême le reste de l'image (c'est-à-dire l'information considérée comme non significative) ? Ce point fera l'objet de la section 2.4 de ce chapitre.

2.1 Décision bayésienne et ordre total sur les vecteurs

Comme nous l'avons dit en introduction, nous voulons ordonner les valeurs d'une image multivaluée selon leur similarité à un échantillon de référence prédéfini. Ce problème est un problème de classification. Il s'agit de calculer la similarité de chaque point de l'image à la référence statistique. Dans le

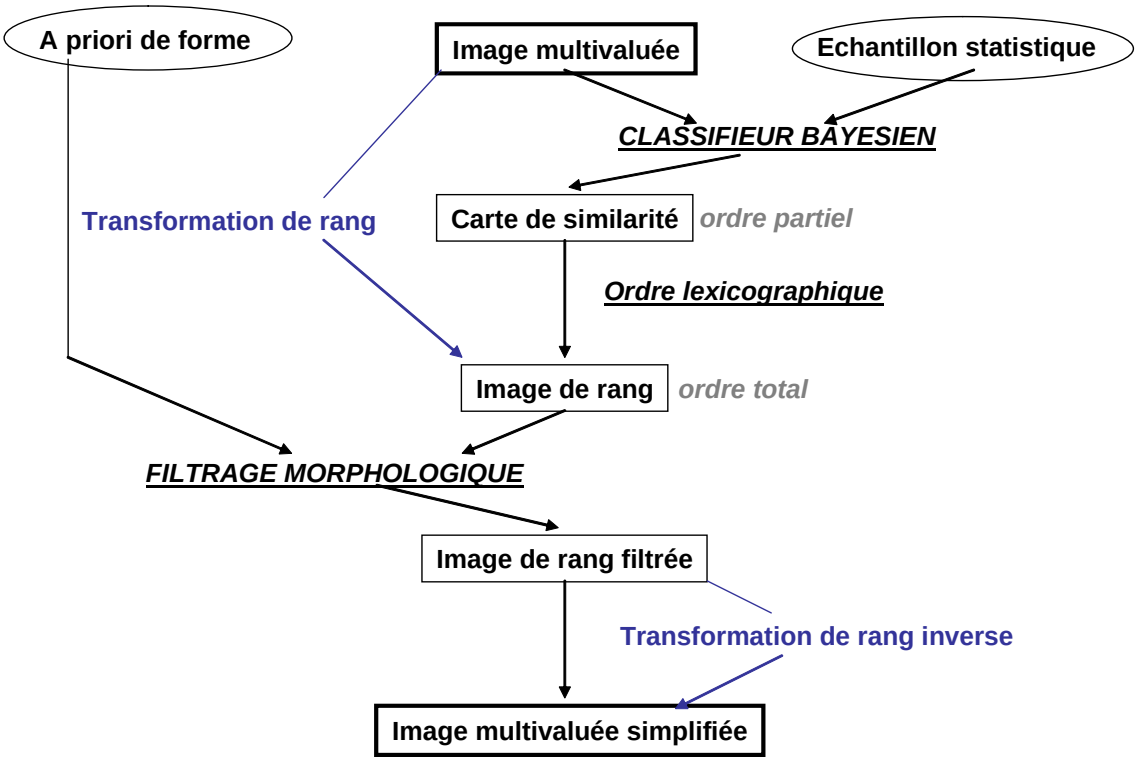


FIG. 2.1: Prise en compte d'informations statistiques et géométriques pour l'analyse des images multivaluées.

chapitre 1, nous avons rappelé que la minimisation du risque de mauvaise classification relevait des règles de la décision bayésienne. Le calcul effectif du maximum a posteriori (MAP) nécessite de faire des hypothèses sur la nature statistique des données. Sous l'hypothèse que les données suivent une loi gaussienne, la règle du MAP revient à minimiser une certaine distance : la distance de Mahalanobis. Ce point est bien connu, mais il mérite d'être rappelé ici de même qu'il importe de bien comprendre les mécanismes intervenant dans le calcul de la distance de Mahalanobis afin de pouvoir ensuite interpréter les cartes de distances qui seront obtenues puis exploitées dans la chaîne de traitement des images.

Dans cette section, nous montrons également comment un ordre total sur les vecteurs est construit à partir des cartes de distance et nous précisons comment l'image de rang résultante peut être mise à profit sur le plan algorithmique et pour l'extension des opérateurs nécessitant des opérations de soustraction (ce qui est le cas des résidus morphologiques).

2.1.1 De la règle de Bayes à la distance de Mahalanobis

La présentation qui suit est extraire des ouvrages [36, 106, 86].

La règle du *MAP* (définition 12) est optimale dans le sens où elle minimise le risque de mauvaise classification. Le calcul nécessite la connaissance de $P(C_j)$ ainsi que de la probabilité conditionnelle $P(X^l|C_j)$. Pour calculer $P(C_j)$, il y a deux solutions classiques. La première consiste à faire l'hypothèse de classes équiprobables (alors $P(C_j) = \frac{1}{N}$ dans le cas de N classes). La deuxième consiste à estimer $P(C_j)$ d'après l'échantillon ou depuis des bases de connaissances sur le problème.

Le calcul de la probabilité conditionnelle $P(X^l|C_j)$ est nettement plus délicat. Souvent, une hypothèse sur la distribution suivie et les paramètres sont estimés. La densité de probabilité gaussienne est largement utilisée. C'est le cas notamment en traitement statistique des images (par exemple pour modéliser le niveau de bruit dans une image...). Les raisons de cette forte utilisation peuvent s'expliquer à l'aide du théorème de la limite centrale. Ce théorème montre que l'addition de phénomènes aléatoires indépendants qui suivent une même loi, convergent vers la loi normale (démonstration complète dans [13]). Le recours à ce théorème est souvent abusif et il ne faut pas croire que tous les phénomènes physiques suivent cette loi. Cependant, elle constitue une bonne hypothèse que nous retiendrons dans ce travail. En effet, disposer d'échantillons pour l'analyse des images, ne signifie pas pour autant disposer de modèle statistique sur ces régions. En l'absence d'information sur la nature des données, on peut soit choisir un modèle a priori (ce sera le cas ici), soit construire un modèle a posteriori (ce qui serait la meilleure solution).

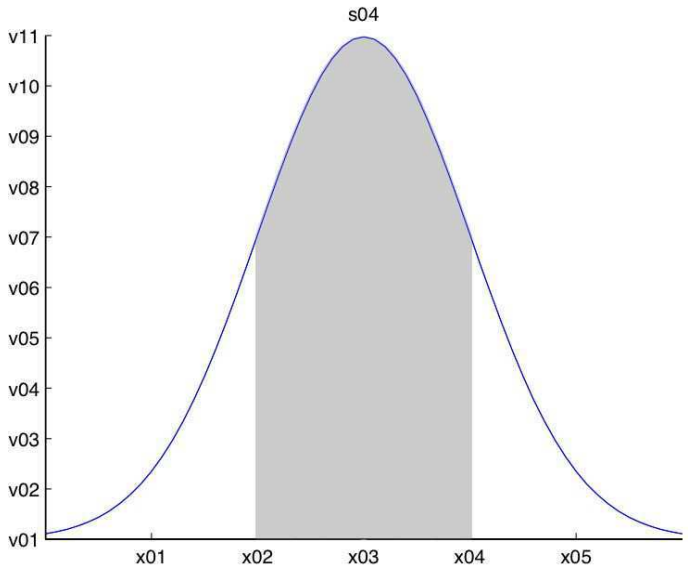


FIG. 2.2: Représentation de la loi normale ($\mu = 0, \sigma^2 = 4$). La probabilité d'appartenance est représentée par la surface sous la courbe. Ainsi, la partie hachurée $\int_{-\sigma}^{+\sigma} f(x) \approx 0.68$, représente la grande majorité des possibilités, c'est-à-dire 68%. Par exemple, si les caractéristiques de la forme suivent une loi normale de moyenne μ et de variance σ^2, alors environ 68% des points auront une valeur entre $\mu \pm \sigma$.

Dans le cas multidimensionnel, la densité de probabilité normale s'écrit :

$$\boxed{p(X^l|C_j) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_j|^{\frac{1}{2}}} e^{-\frac{1}{2}(X^l-m_j)^T\Sigma_j^{-1}(X^l-m_j)}} \tag{2.1}$$

Dans cette définition, C_j l'échantillon, c'est-à-dire la matrice dont les colonnes sont les vecteurs qui composent l'échantillon : $C_j = [\ X_1 \quad X_2 \quad ... \quad X_q\ ]$. X^l est un point de l'échantillon.

m_j est obtenu en calculant la moyenne sur chaque composante de l'échantillon : $m_j = \begin{bmatrix} E(X_{1,1}, X_{2,1}, ..., X_{q,1}) \\ E(X_{1,2}, X_{2,2}, ..., X_{q,2}) \\ ... \\ E(X_{1,p}, X_{2,p}, ..., X_{q,p}) \end{bmatrix}$.

Enfin, Σ_j est la matrice de covariance de l'échantillon :

$$\Sigma_j = \begin{bmatrix} \mathrm{E}[(X_1-\mu_1)(X_1-\mu_1)] & \mathrm{E}[(X_1-\mu_1)(X_2-\mu_2)] & \cdots & \mathrm{E}[(X_1-\mu_1)(X_q-\mu_q)] \\ \mathrm{E}[(X_2-\mu_2)(X_1-\mu_1)] & \mathrm{E}[(X_2-\mu_2)(X_2-\mu_2)] & \cdots & \mathrm{E}[(X_2-\mu_2)(X_q-\mu_q)] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}[(X_q-\mu_q)(X_1-\mu_1)] & \mathrm{E}[(X_q-\mu_q)(X_2-\mu_2)] & \cdots & \mathrm{E}[(X_q-\mu_q)(X_q-\mu_q)] \end{bmatrix}$$

avec $\mu_j = E(X_j)$

Remarquons tout de suite que le terme $(X^l-m_j)^T\Sigma_j^{-1}(X^l-m_j)$ définit bien une distance qui peut d'ailleurs être interpétée comme une distance euclidienne ($(X^l-m_j)^T(X^l-m_j)$, basée sur le théorème de Pythagore) pondérée (par Σ_j^{-1}). Cette distance est la *distance de Mahalanobis.* Nous allons voir que la règle du MAP revient, dans l'hypothèse où les données suivent une loi gaussienne), à minimiser la distance de Mahalanobis.

Appliquons la règle du *MAP* à cette densité de probabilité. Pour des raisons de commodité des calculs, nous prendrons le logarithme de la densité de probabilité. Le logarithme étant une fonction

monotone et croissante, elle n'influence pas la décision. Nous avons :

$$\begin{aligned} h^* &= \underset{j \in N}{\operatorname{argmax}}\ P(X^l|C_j)P(C_j) \\ &= \underset{j \in N}{\operatorname{argmax}}\ d_j(X^l) \end{aligned}$$

$d_j(X^l)$ représente la valeur de décision de choisir la classe C_j. Compte tenu de notre hypothèse sur la densité de probabilité,

$$\begin{aligned} d_j(X^l) &= ln(P(X^l|C_j)P(C_j)) && (2.2) \\ &= ln(P(X^l|C_j)) + ln(P(C_j)) && (2.3) \\ &= ln(P(C_j)) - \frac{n}{2}ln(2\pi) - \frac{1}{2}ln(|\Sigma_j|) - \frac{1}{2}\left[(X^l - m_j)^T\Sigma_j^{-1}(X^l - m_j)\right] && (2.4) \end{aligned}$$

Le terme constant peut être supprimé. La règle du *MAP* lorsque nous supposons que les caractéristiques de la forme suivent une distribution gaussienne s'écrit :

$$\boxed{d_j(X^l) = ln(P(C_j)) - \frac{n}{2}ln(|\Sigma_j|) - \frac{1}{2}\left[(X^l - m_j)^T\Sigma_j^{-1}(X^l - m_j)\right]} \quad (2.5)$$

Définition 15 *et l'expression* $(X^l - m_j)^T\Sigma_j^{-1}(X^l - m_j)$ *est appelée distance de Mahalanobis.*

Cette règle de décision est du second ordre, par conséquent l'espace des caractéristiques est découpé par des hyperquadratiques. En dimension 2, nous avons des ellipses, des hyperboles, des paraboles ainsi que des droites qui découpent l'espace. Étudions les trois cas possibles :

Cas 1 : $\Sigma_i = \sigma^2\mathbf{I}$ Dans ce premier exemple, nous supposons les matrices de covariance proportionnelles à la matrice unité. Cela représente le cas où les caractéristiques sont indépendantes et de mêmes variances. Les points sont englobés par des hypersphères avec une séparation entre deux classes (i, j) définies par $d_j(X^t) = d_i(X^t)$. Cette courbe de démarcation entre les classes est en fait un hyperplan, orthogonal à la ligne joignant les moyennes (voir illustration de la figure 2.3).

Cas 2 : $\Sigma_i = \Sigma$ Lorsque les caractéristiques ne sont pas indépendantes, mais que cette variation est la même pour chaque classe, alors $\Sigma_i = \Sigma$. Les points sont ainsi englobés dans des ellipses et non des sphères. La décision d'appartenance à une classe est donc :

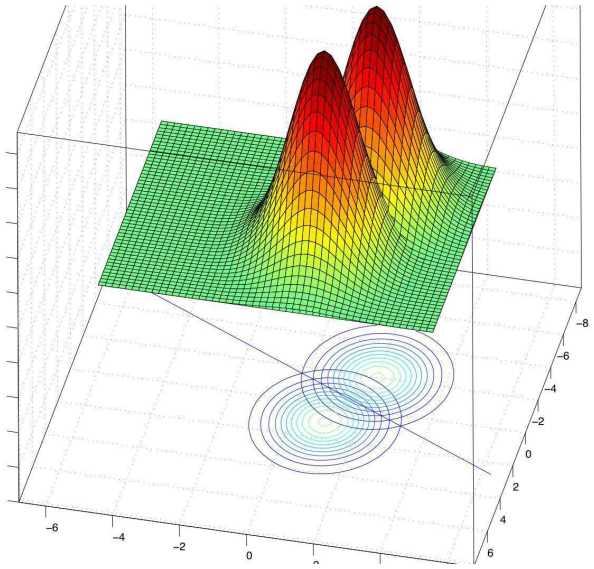


FIG. 2.3: Exemple de deux distributions gaussiennes, dont les moyennes sont $\mu_1 = [1; 1]$, $\mu_2 = [-2, 2]$ et dont les matrices de covariance sont diagonales $\Sigma 1 = \Sigma_2 = [2, 0; 0, 1]$. Les gaussiennes décrivent au sol des cercles. La règle de décision est donnée par l'équation de la droite $\frac{3}{2}x - y + \frac{9}{4} = 0$ qui est perpendiculaire à l'axe reliant les moyennes.

$$\boxed{d_j(X^l) = -\frac{1}{2}\left[(X^l - m_j)^T \Sigma_j^{-1}(X^l - m_j)\right] + ln(P(C_j))} \tag{2.6}$$

La séparation entre deux classes est également un hyperplan, qui à la différence du cas 1, n'est généralement pas orthogonal à la ligne joignant les moyennes.

Cas 3 : Σ quelconque C'est le cas le plus général pour une distribution gaussienne. Les points des classes sont englobés dans des hyperellipses. L'équation permettant la classification entre deux classes i et j est donnée, comme dans les cas précédents par $d_i = d_j$. Aucune simplification n'est possible, excepté si les probabilités a priori sont égales. L'espace des caractéristiques est partitionné pour deux classes i et j par les différentes hyperquadratiques possibles.

Géométriquement, le principe de la classification bayésienne est déterminé par la position du vecteur X^t dans l'espace des caractéristiques partitionné par les courbes d'iso-décisions. D'un point de vue calculatoire, ce choix revient à prendre la valeur maximale des fonctions de décisions. Ce principe de classification est bien optimal dans le sens qu'il minimise le risque de mauvaise classification. Cependant, il est intéressant d'utiliser l'équation 2.5, non pas pour obtenir une réponse binaire définie par le maximum, mais pour exploiter l'ensemble des valeurs de $d(X^l)$ conduisant à cette décision. En effet, plus un point est proche d'une frontière, plus l'incertitude est grande entre les deux classes. De

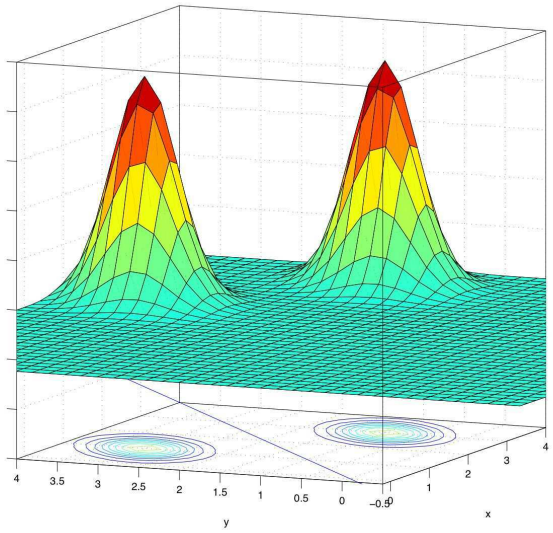


FIG. 2.4: Les paramètres des deux gaussiennes sont $\mu_1 = [3; 1]$, $\mu_2 = [1; 3]$ et $\Sigma_1 = \Sigma_2 = \frac{1}{16}[3, 1; 1, 3]$. Le fait que Σ ne soit pas diagonale implique que les gaussiennes décrivent au sol non plus des cercles, mais des ellipses. La règle de décision est une droite d'équation : $x - y = 0$

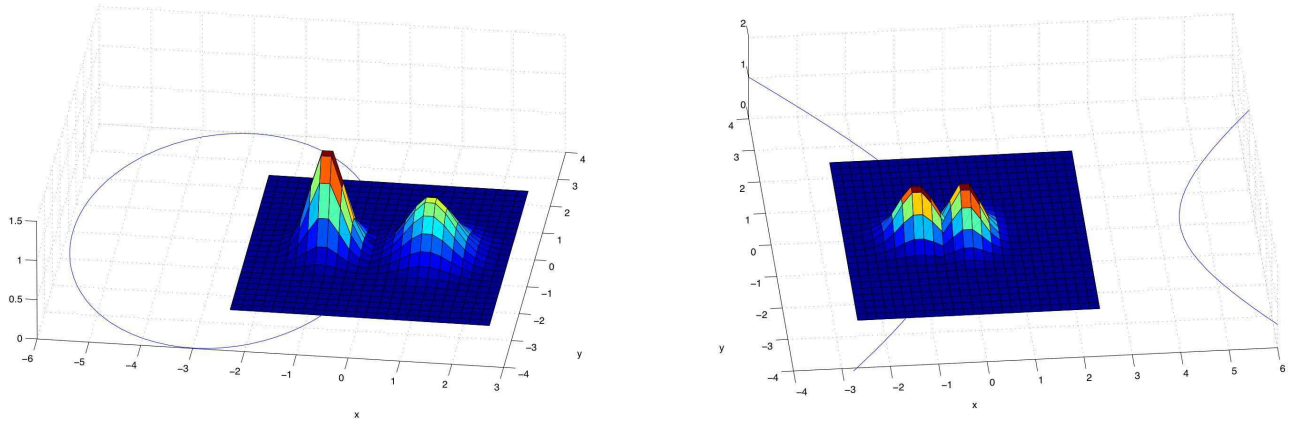


FIG. 2.5: Illustration des courbes de décision lorsque les matrices de covariance changent. Dans les deux cas, les moyennes sont identiques ($\mu_1 = [0; 0]$ et $\mu_2 = [-1; 0]$), à gauche les matrices de covariance sont $\Sigma_1 = [0.1, 0.0; 0.0, 0.15]$ et $\Sigma_2 = [0.2, 0; 0, 0.25]$, la courbe est alors une ellipse d'équation $-\frac{5}{2}x^2 - \frac{4}{3}y^2 - 15x - 1.296 = 0$. À droite, $\Sigma_1 = [0.1, 0; 0, 0.15]$ et $\Sigma_2 = [0.15, 0; 0, 0.1]$ l'équation devient une hyperbole ($-x^2 + y^2 + 4x + 2 = 0$).

la même façon, plus un point est proche du centre des caractéristiques de la forme, plus la certitude de bonne classification est forte. La fonction $d_j(X^l)$ donne par conséquent, une indication sur la valeur de ressemblance avec la classe C_j.

Utilisons l'éventail des valeurs de la fonction de décision comme critère de similarité avec la forme que nous recherchons. La règle de décision pour une forme C de moyenne m et de covariance Σ est :

$$\boxed{d(X^l) = ln(P(C)) - \frac{n}{2}ln(|\Sigma|) - \frac{1}{2}\left[(X^l - m)^T\Sigma^{-1}(X^l - m)\right]} \tag{2.7}$$

La probabilité $P(C)$ ainsi que la constante étant les mêmes pour tous les points, la mesure de similarité par rapport à la forme C est bien la distance de Mahalanobis.

$$\boxed{d(X^l) = (X^l - m)^T\Sigma^{-1}(X^l - m)} \tag{2.8}$$

Plus le vecteur X^l est proche des statistiques de l'échantillon de la forme, plus sa distance est faible. Cette notion est très intuitive : plus une forme est proche d'une autre, plus nous pouvons imaginer une distance faible entre les deux.

Une conséquence également très importante de cette équation est que nous passons d'un vecteur de dimension $l \geq 1$ à une valeur dans $\mathbb{R}$. Il y a par conséquent un double intérêt d'utiliser cette distance : premièrement, nous obtenons une mesure de similarité et deuxièmement nous avons une réduction du nombre de dimensions.

Quelques exemples de cartes de similarités calculées sur des images couleur sont présentées en figures 2.6 et 2.7. Les cartes des similarités sont les négatifs des cartes des distances. La figure 2.8 présente les cartes de distance obtenues à partir de l'image couleur de la figure 2.6. La carte obtenue par calcul de la distance de Mahalanobis est comparée à celle obtenue par calcul de la distance euclidienne. Dans le calcul de la distance euclidienne, l'origine est le point 0. La référence statistique pour le calcul de la distance de Mahalanobis est un carré centré en 0 (le noir pur) sur la mire des couleurs : voir figure 2.8. Lorsque la variance de l'échantillon augmente, la dynamique de la carte des distances décroît, mais dans tous les cas, les deux modes de l'image (les pâquerettes et l'arrière-plan) sont mieux séparés par la distance de Mahalanobis que par la distance euclidienne.

2.1.2 Construction d'images de rang

Les distances (quelles soient statistiques ou euclidiennes) ne permettent pas d'ordonner totalement les vecteurs dans les images multivaluées. Dans le cas de la couleur, deux vecteurs correspondant à deux couleurs différentes peuvent tout à fait être équidistants de la référence. Du fait de la complexité des images et du bruit, on peut penser que le nombre de vecteurs distincts et équidistants est faible. Il n'en est rien. Quelques tests réalisés sur des images différentes ont montré que 20 à 30% des points sont

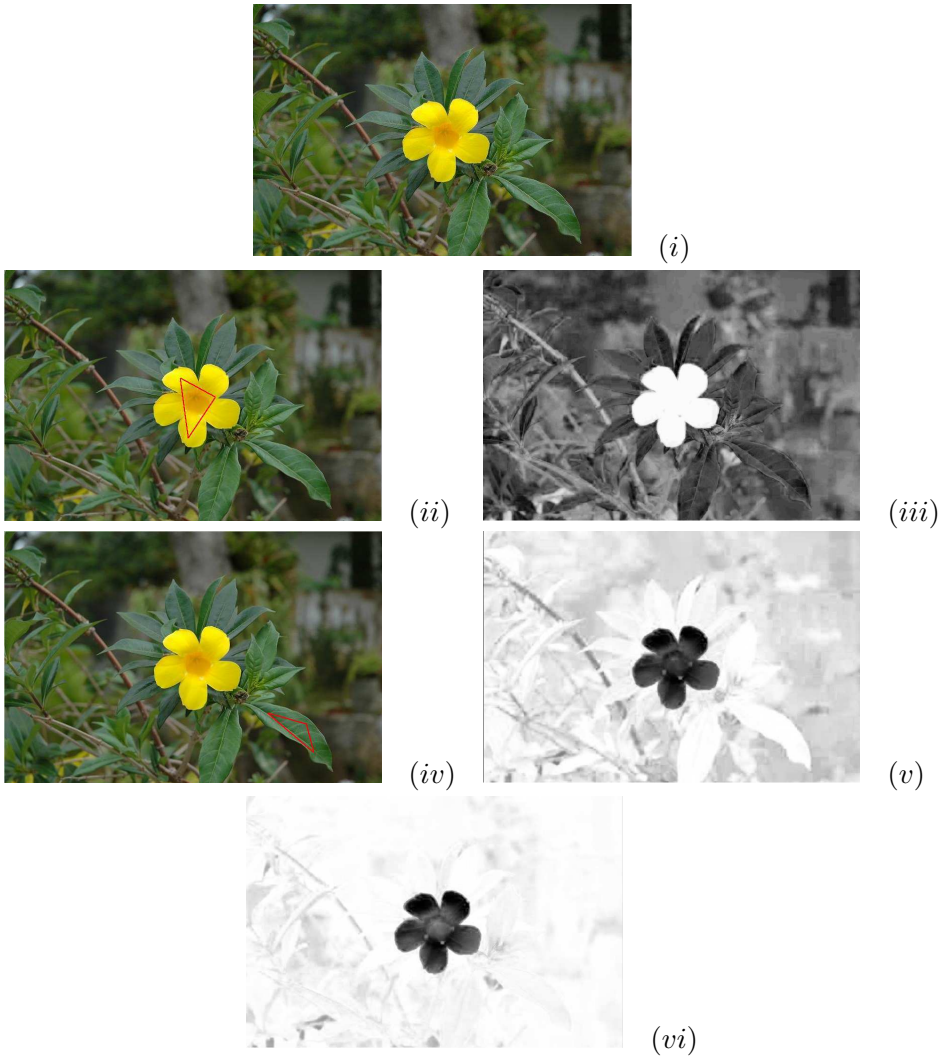


FIG. 2.6: Cartes de similarité associées à une image couleur calculée à partir de la distance de Mahalanobis en utilisant différents échantillons statistiques de référence. Sur les images de similarité, les points de forte luminance localisent les formes statistiquement similaires à l'échantillon de référence tandis que les points de faible luminance sont des points peu ou pas similaires à l'échantillon. (i) est l'image originale. (iii) est l'image des distances associée à la référence (ii) (l'échantillon est la région délimitée en rouge sur l'image (ii)). (iv) est la carte des distances associée à la référence (v) et (vi) est la carte des distances obtenue en prenant comme référence un blanc le blanc (après ajout d'un bruit gaussien). Sur les cartes des distances, les structures de l'image sont ordonnées selon leur similarité à l'échantillon de référence.

FIG. 2.7: Cartes de similarité calculées à partir de la distance de Mahalanobis. La référence statistique est l'image (v).

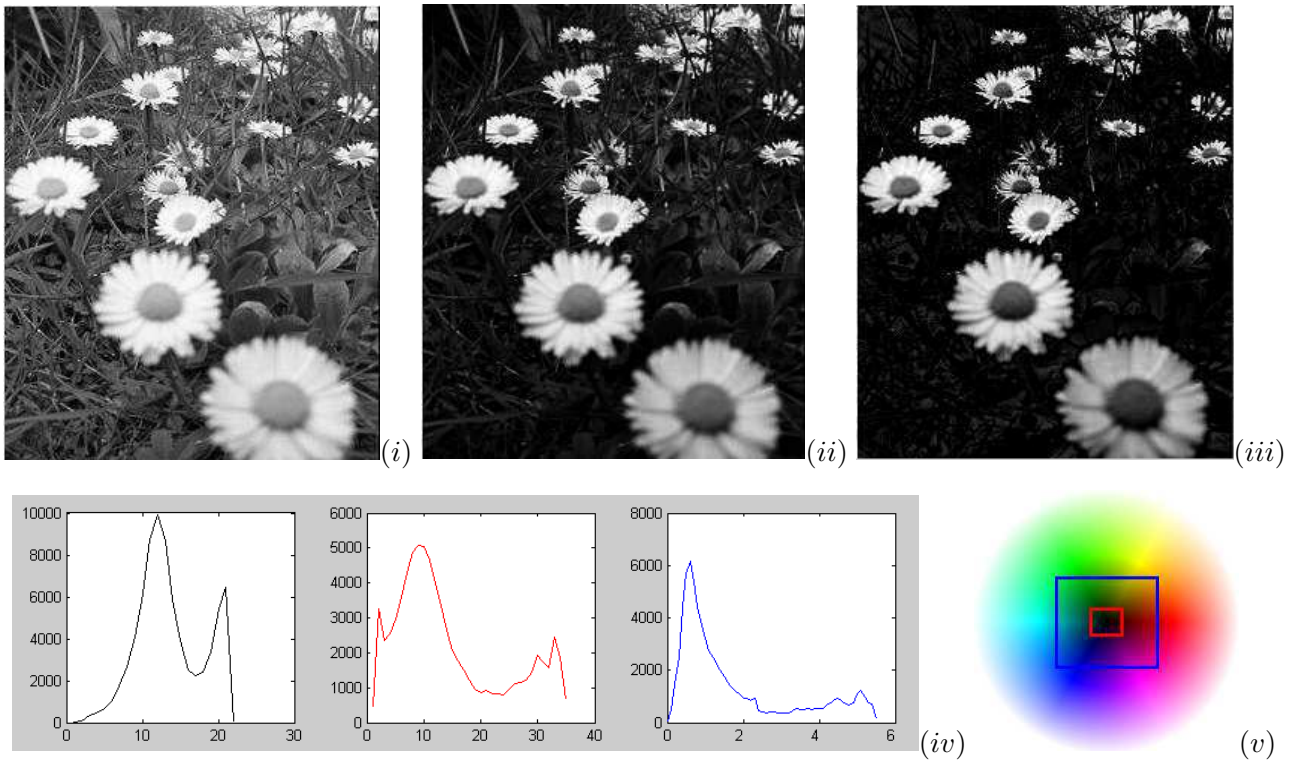


FIG. 2.8: Commparaison des distances d'Euclide (i) et de Mahalanobis ((ii) et (iii)) pour des échantillon de référence de variance croissante. L'origine pour le calcul de la distance euclidienne est le point 0. Les échantillons de référence pour le calcul de la distance de Mahalanobis sont indiqués en rouge (cas (ii)), bleu (cas (iii)) sur la figure (v).

dans ce cas de figure. De fait, on ne peut ignorer ce problème, car cela provoquerait des dysfonctionnements importants sur les opérateurs que nous allons définir. L'expérience a été menée en programmant une érosion sur les vecteurs à partir de l'ordre partiel déduit de la distance de Mahalanobis. Il y a deux conséquences graves au fait que l'ordre ne soit pas total : premièrement, le programme n'est pas reproductible (dans le cas où deux vecteurs différents définissent tous les deux un infimum, l'un ou l'autre peut être choisi sans aucune maîtrise de l'aléatoire lié à ce choix). Deuxièmement, comme conséquence de ce premier point, au sein d'une même transformation, des vecteurs différents peuvent être choisis, provoquant une sorte de bruit sur l'image résultat. La combinaison des opérateurs morphologiques (pour la définition de filtres) ne ferait qu'amplifier ce désordre... Si théoriquement, un ordre partiel suffit, la mise en oeuvre pratique requiert quant à elle un ordre total.

Dans le cas présent, l'objectif étant d'obtenir une classification en accord avec l'échantillon de référence, rien a priori ne peut justifier le choix d'un ordre ou d'un autre pour départager les vecteurs équidistants. La solution la plus simple est donc retenue. Elle consiste à ordonner les vecteurs équidistants selon l'ordre lexicographique (voir section 1.4.2). Dans le cas de la couleur, la cascade lexicographique peut être appliquée dans l'espace RGB ou dans tout autre espace de représentation de la couleur (espaces LAB, HSV...voir section 1.3.2). Pour les illustrations que nous présentons dans ce document, l'espace RGB est utilisé. Encore une fois, ce n'est pas la perception de la couleur qui est en question ici, mais l'ordre induit par l'analyse statistique. L'ordre lexicographique est juste un outil mathématique permettant une manipulation aisée des données.

La distance de Mahalanobis (ou plutôt son inverse sous forme de carte de similarité) complétée de l'ordre lexicographique conduit à la formation d'une image scalaire qui sera appelée *image de rang*. En tout point de l'image et sur tout voisinage du point, il est non seulement possible de définir un plus grand et un plus petit voisin, mais il est même possible de ranger tous les voisins du point. L'ordre est total.

L'option choisie est de construire une image de rang faite d'entiers consécutifs. Une approche similaire avait été proposé par [127] utilisant la notion de table de hachage entre les vecteurs et leur ordre. Nous avons poursuivi dans cette direction, en associant à chaque vecteur de l'image d'origine un unique rang entier et en décidant d'occuper tous les rangs. Si N est le nombre de couleurs dans l'image d'origine (plus généralement le nombre de vecteurs distincts dans l'image d'origine), alors le rang maximal dans l'image de rang est $N-1$: le vecteur de plus forte similarité est associé au rang $N-1$, la couleur de plus faible similarité étant associée au rang 0.

Notons que si nous traçons une courbe passant par les sommets de l'image de rang nous retrouvons la « space filling curve » présentée par [26]. L'image de rang est une autre représentation de l'ordre induit.

La transformation de rang est la transformation qui associe à chaque vecteur V de l'image F son rang n : $T_{rank}(V) = n$. Soit $\mathcal{V}_{\mathcal{F}}$ l'ensemble de tous les vecteur de l'image F, T_{rank} est une bijection de $\mathcal{V}_{\mathcal{F}}$ vers $[0, N-1]$ (voir figure 2.9).

T_{rank} étant bijective, F peut être reconstruite à partir de l'image de rang qui lui est associée $f = T_{rank}(F)$ par la transformation inverse

$$F = T_{rank}^{-1}(f)$$

mais deux images F et G ne peuvent pas être comparées entre elles via leurs images de rang puisque les valeurs des rangs dans les images de rang sont parfaitement liées à l'image : à chaque calcul, la transformation T_{rank} n'est pas définie pour toutes les couleurs de la table des couleurs mais uniquement pour les couleurs de l'image courante. Par conséquent, cette transformation ne permet pas d'ordonner les images entre elles, sauf cas particulier de deux images composées ayant toutes deux leurs valeurs dans l'espace $\mathcal{V}_{\mathcal{F}}$. Pour munir l'ensemble des images multivaluées d'une relation d'ordre cohérente avec la démarche entreprise, il suffit de revenir à la distance de Mahalanobis. Si d_M représente la distance de Mahalanobis par rapport à un échantillon de référence fixé une fois pour toutes, alors

$$F \leq_M G \Leftrightarrow \forall p \in E,\ \ d_M(F(p)) \leq d_M(G(p))$$

Le treillis des images multivaluées est donc muni de l'ordre induit par la distance de Mahalanobis. L'ordre sur les images n'est pas total. Le treillis ainsi construit est bien complet. Le plus petit élément est l'image de teinte uniforme correspondant à la couleur de la table des couleurs la plus éloignée de l'échantillon de référence au sens de la distance de Mahalanobis. Le plus grand élément est l'image teinte uniforme dont la couleur correspond à la couleur moyenne de l'échantillon de référence.

L'ordre sur les images est partiel mais l'ordre sur les valeurs des images est total et les deux ordres sont cohérents puisqu'ils émergent tous les deux de la même distance (la distance de Mahalanobis). Toute la stratégie décrite au chapitre 1 reste valide. Les opérateurs morphologiques sur les images multivaluées seront donc définis via leurs homologues sur fonctions scalaires par transformation de l'image de rang puis retour dans l'espace des valeurs multivaluées :

$$T = T_{rank}^{-1} \circ t \circ T_{rank}$$

L'image de rang est une image constituée d'entiers consécutifs. Bien sûr, il serait tout à fait possible de travailler sur une carte des similarités à valeurs réelles (tout en assurant un ordre total sur les vecteurs). Le choix des entiers consécutifs peut paraître peu judicieux dans la mesure où la dynamique de la carte de similarité est perdue : deux vecteurs à des distances très différentes de l'échantillon peuvent très bien avoir des rangs voisins. Ce choix est motivé par des considérations algorithmiques et calculatoires.

La conception de la plupart des logiciels développés en morphologie multivaluée se heurte à de sérieuses difficultés de mise en oeuvre dues au fait qu'il faut manipuler des types de données non déterminés de manière stricte (float, int...) [43]. Dans le cas présent, le logiciel repose sur une simple mémorisation de la table des correspondances couleurs-rangs. Cette opération est simple, rapide et ne nécessite que très peu de développements informatiques (on revient sur ce point dans le chapitre 3).

La deuxième raison de ce choix, qui est la raison principale, est que le fait de travailler sur des images d'entiers consécutifs fournit d'emblée une solution au problème de la soustraction entre vecteurs.

Typiquement, comme nous l'avons mentionné au chapitre 1, les résidus morphologiques sur images multivaluées (tels que les gradients morphologiques ou les chapeaux haut de forme) sont définis soit en utilisant la soustraction vectorielle usuelle (ce qui introduit de faux effets de couleur), soit en privilégiant une composante vectorielle particulière (dans le cas de la couleur, on choisit l'espace HLS pour ne travailler que sur la composante H des luminances). Ensuite, les autres composantes sont réinjectées afin de reconstruire un vecteur à partir de la seule composante traitée [101, 59, 7]. Dans le cas présent, les rangs étant consécutifs, toute opération de soustraction entre les rangs (sous réserve que le résultat soit positif) est encore un rang valide (c'est-à-dire un rang défini dans l'image de rang). Par conséquent, toute soustraction entre deux vecteurs V_1 et V_2 d'une même image F peut être définie comme la différence de leurs rangs respectifs $(r_1 - r_2)$ et le retour dans l'espace des couleurs est possible : $T_{rank}^{-1}(r_1 - r_2)$ est bien défini.

Les opérateurs morphologiques n'introduisent pas de nouvelle luminance dans les images (ils sont soit extensifs, soit anti-extensifs ou bien ils résultent de la composition d'opérateurs de ces deux types). Non seulement, le résultat de la transformation est bien comparable à l'original dans l'espace multivalué (via l'ordre induit par la distance de Mahalanobis) mais il est possible de définir le résidu de la transfor-

mation puisque les soustractions mises en jeu ne font intervenir que des couleurs recensées dans la table de correspondance construite via la transformation de rang.

Prenons l'exemple du gradient morphologique, défini dans le cas des images à valeurs scalaires comme la différence entre la dilatation et l'érosion :

$$g_B = \frac{\delta_B - \epsilon_B}{2r}$$

si B est un disque de rayon r. En pratique, la normalisation par r n'est pas indispensable, elle permet juste d'assurer que g_B est bien la norme du gradient de l'image lorsque $r \to 0$ et en différenciant f au sens des distributions.

Pratiquement, la définition suivante (qui est la plus couramment utilisée) a tout à fait du sens :

$$g_B = \delta_B - \epsilon_B$$

D'après ce que l'on a dit, le gradient morphologique vectoriel peut être défini à partir du gradient calculé sur l'image des rangs :

$$G_B = T_{rank}^{-1} \circ g_B \circ T_{rank} = T_{rank}^{-1}[\delta - \epsilon]T_{rank} = T_{rank}^{-1}[\delta T_{rank} - \epsilon T_{rank}]$$

L'opération $T_{rank}^{-1}[g_B \circ T_{rank}]$ est possible car les valeurs dans δT_{rank}, ϵT_{rank} et $[\delta - \epsilon]T_{rank}$ peuvent toutes être interpétées comme des couleurs existantes dans l'espace des couleurs de l'image d'origine.

Des exemples de gradients couleur ainsi calculés sont montrés en figure 2.33. Bien entendu, tout ce qu'on vient de dire vaut également pour les transformations de type chapeau haut de forme qui sont les résidus des ouvertures ou fermetures.

2.2 Expoitation directe de la carte des distances : le cas particulier de la segmentation

La segmentation des images, nous l'avons dit en introduction du chapitre 1, consiste à partitionner l'espace en régions connexes et disjointes. L'issue des algorithmes de segmentation est une image binaire ou labellisée où chaque point de l'espace est affecté à une région (éventuellement, la segmentation peut

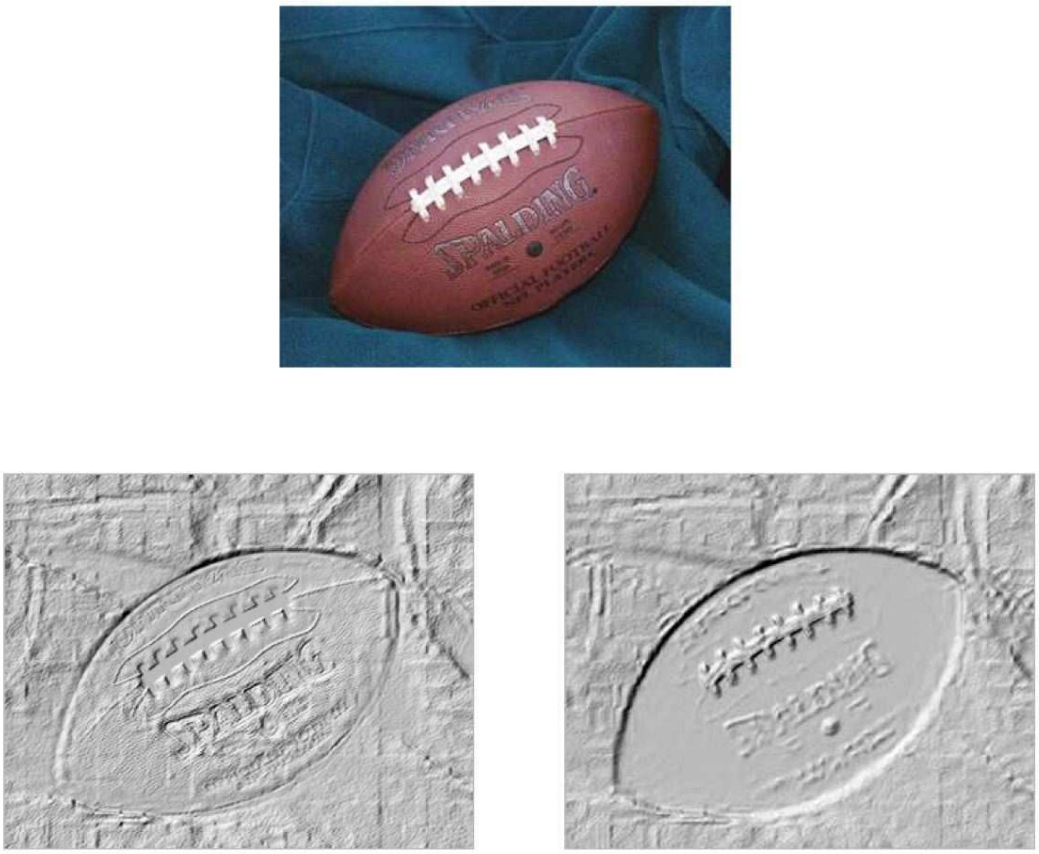


FIG. 2.9: Une image couleur et image de rang associées. Les images de rang rendent compte de la similarité de chaque point de l'image à la référence (ici une partie du blanc de la corde du ballon puis un échantillon de rouge du ballon) tout en assurant un ordre total sur les vecteurs et une manipulation facilitée des données. Le choix de l'échantillon conditionne fondamentalement l'ordre sur les couleurs.

se réduire à l'extraction d'un unique objet dans l'image et dans ce cas, l'espace est partitionné en deux régions correspondant à l'intérieur et l'extérieur de l'objet).

La segmentation est un cas particulier puisque l'issue n'est pas une image de même nature que la première : c'est une image binaire, le retour dans l'espace des valeurs multivaluées n'est pas nécessaire (sauf contexte spécifique de construction d'images mosaïques par exemple).

Un simple seuillage de la carte des distances est toujours envisageable. Ceci revient à effectuer une classification statistique classique, la décision s'effectuant point à point en fonction de la similarité locale sans qu'aucune information d'ensemble (régionale) ne soit intégrée. Pour tenir compte de la topologie de l'image et permettre par exemple une segmentation correcte dans le cas où les classes d'appartenance ne sont pas clairement distinctes, des paradigmes de segmentation plus évolués peuvent être utilisés.

De nombreux paradigmes de segmentation ont été proposés dans la littérature ces vingt ou trente dernières années fondés sur des formalismes très variés parmi lesquels les méthodes statistiques, les classifications floues, les champs de Markov, les réseaux de neurones, les approches variationnelles, les modèles Bayesiens, les contours actifs, les fonctionnelles d'énergie et notamment la célèbre fonctionnelle de Mumford et Shah, les opérateurs de Canny and Deriche...etc...

De très nombreux modèles et parmi les plus intéressants (si l'on en juge par les applications qui y font appel) s'inspirent d'analogies avec les phénomènes physiques : c'est le cas de nombreuses méthodes fondées sur les minimisations d'énergie comme le modèle ballon de Cohen [30], les méthodes de calcul de chemins minimaux qui sont une application du principe de Fermat [40, 31], les modèles électrostatiques qui exploitent la loi de Coulomb [55], le paradigme morphologique qui s'inspire de la notion topographique de ligne de partage des eaux [19]...

Parmi les travaux pionniers dans le domaine de la segmentation, deux paradigmes ont émergé très tôt de l'école morphologique sous l'impulsion de Matheron et Serra, qui font encore référence dans le domaine : la partition en zones plates et la partition en bassins versants.

La présentation qui suit a deux objectifs. Premièrement, illustrer sur des exemples d'images réelles et sur la base des segmentations obtenues, la puissance sémantique des cartes de distance ou des cartes de similarité construites grâce à la distance de Mahalanobis. Deuxièmement, rappeler deux points de vue différents de la segmentation (qui correspondent aux deux paradigmes de la segmentation morphologique). Le premier point de vue qui est illustré ici distingue fondamentalement le problème du filtrage ou de la simplification du problème de la segmentation : c'est le premier paradigme morphologique fondé sur la ligne de partage des eaux. Le deuxième point de vue formule au contraire la segmentation d'image comme un problème de simplification, le problème revenant à construire les filtres adéquats : c'est le second paradigme morphologique fondé sur la partition en zones plates. Nous illustrerons dans cette partie le premier paradigme. Le second sera largement discuté dans la suite de ce chapitre (sections 2.3.3 et 2.4). Afin de clarifier notre propos, nous commençons par présenter chacun des paradigmes.

2.2.1 Deux paradigmes morphologiques pour la segmentation

Segmenter une image consiste à partitionner le plan image en un ensemble de régions non vides, connexes et disjointes. La définition suivante est due à Serra [121].

Définition 16 *Une partition D est une application extensive et disjonctive associant au support A de l'image un ensemble de sous-ensembles de A.*

D extensive signifie que pour tout point x du support de l'image, $x \in D(x)$ et par conséquent $D(x) \neq \emptyset$ et $A = \bigcup D(x)$. D disjonctive signifie que $x \neq y \Rightarrow D(x) \cap D(y) = \emptyset$ ou $D(x) = D(y)$ [121].

En notant $(R_i)_{i \in I}$ la famille des sous-ensembles définis par la partition D, les propriétés suivantes sont satisfaites

- $\forall i \; R_i \neq \emptyset$
- $\forall i, j \; ; \; i \neg j \; R_i \cap R_j = \emptyset$
- $A = \underset{i}{\cup} R_i$

et nous retrouvons la définition proposée par Coquerez et Philipp [35] que nous avions mentionnée au chapitre 1.

Ces définitions sont simples, mais elles ne sont pas opérationnelles : à toute image de support A correspond une infinité de partitions. C'est là un des points importants mis en avant dans le chapitre 1 de ce document : la question centrale dans tout problème de segmentation consiste à choisir la partition qui ait du sens pour l'image et pour l'utilisation que l'on veut faire de la segmentation. Pour choisir parmi toutes les partitions possibles, il va falloir imposer de nouvelles contraintes.

La partition en zones plates La partition en zone plate est la partition la plus immédiate qui dérive de la définition 16 si on lui ajoute comme contrainte supplémentaire l'homogénéité des régions $D(x)$: dans le cas d'une image scalaire, les zones les plus homogènes sont bien les zones plates, c'est-à-dire les composantes connexes de points de même luminance : alors, $D(x)$ est la plus large composante connexe contenant x et constituée de points ayant la même luminance que x sur l'image [111]. En d'autres termes, en regardant l'image comme un relief topographique et en interprétant les luminances comme des altitudes, chaque zone plate correspond à un plateau de l'image : voir figure 2.10.

Pourquoi la partition en zones plates est-elle si intéressante ? Parce qu'elle est non paramétrique et parfaitement inhérente à l'image et parce qu'elle conduit naturellement à une description multiéchelle de la scène [111] une fois associée à des familles de filtres permettant de simplifier l'image de manière à élargir la partition en zones plates : ces filtres sont les opérateurs dits connexes [111]. A titre d'exemple, les zones plates et zones quasi-plates d'une image à teintes de gris sont présentées en figure 2.11 ainsi qu'un filtre connexe produisant une image simplifiée dont les zones plates correspondent aux zones

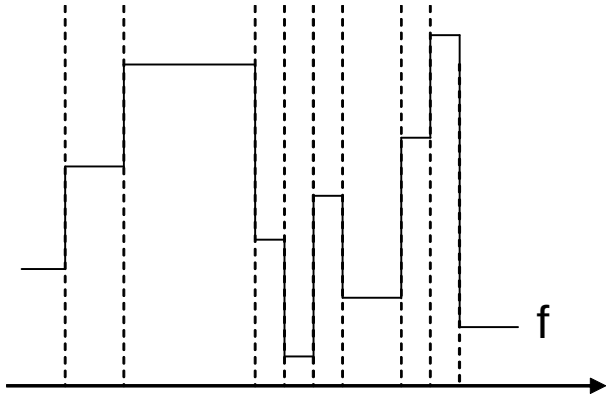


FIG. 2.10: Une fonction constante par morceaux et sa partition en zones plates.

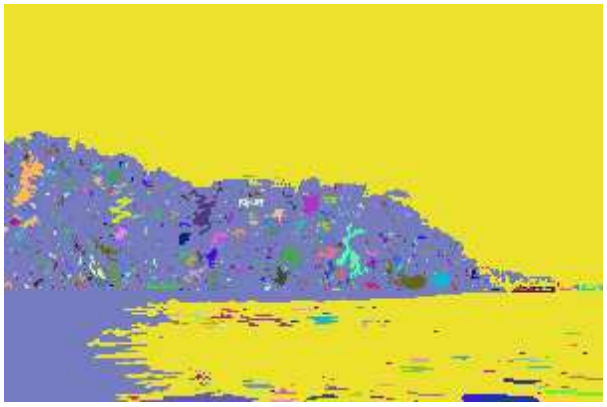
quasi-plates de l'originale : voir figure 2.12. Les zones plates ayant des luminances proches sont fusionnées. L'image simplifiée est obtenue en attribuant aux zones ainsi obtenues, valeur moyenne de l'image originale sous la zone.

La partition en bassins versants Les partitions en zones plates sont bien adaptées lorsque les transitions de luminance (ou de couleur) dans les images sont brutales (images de type *cartoon*) ; elles le sont moins dans le cas d'images aux variations de luminance (ou de couleur) lentes. Dans le cas d'images de texture, elles sont parfaitement inadaptées, à moins de travailler sur une image plus élaborée que la texture originale... Dans le cas des images à variations lentes (images naturelles), les contours ne coïncident pas avec les lieux de transition de luminance, mais avec les lieux de transition *maximale* de luminance. Si f est une image scalaire et si $\|\nabla f\|$ désigne la norme du gradient de f, alors, les régions homogènes coincident non pas avec les zones plates de f mais avec les bassins versants de f : voir figure 2.13.

Que représentent les bassins versants ? En topographie, un bassin versant correspond à la zone géographique drainée par une rivière ou un réservoir. Si une goutte de pluie tombe en un point du bassin versant, elle s'écoulera le long du relief jusqu'à rejoindre finalement la rivière ou le réservoir en question. En morphologie, nous préférons parler de source d'inondation, car nous nous référons plus souvent à une notion duale des bassins versants qui est celle de ligne de partage des eaux. La ligne de partage des eaux topographique est l'ensemble des points qui divisent des zones géographiques drainées par différentes sources d'eau (rivières ou réservoirs). L'idée fondamentale qui, en traitement d'image, a permis d'aboutir à la partition en bassins versants est l'analogie entre la fonction numérique (qui le plus souvent est choisie comme la norme du gradient de l'image pour les raisons évoquées précédemment)

(i)

(ii)



(iii)

FIG. 2.11: Zones plates et quasi-zones plates d'une image à teintes de gris. (i) Image originale (image de taille 200×300 pixels $\times 8$ bits) (ii) les 27988 zones plates de (i) (en fausses couleurs). (iii) les 700 zones quasi-plates de (i) : les différences de luminance inférieures à 38 sont négligées.

et le relief topographique [32, 104]. Nous avons déjà évoqué cette analogie au chapitre 2. En regardant l'image gradient comme un relief, les contours des objets dans l'image correspondent aux lignes crête du gradient (points de variation de luminance maximale), c'est-à-dire aux frontières des bassins versants qui correspondent exactement aux points de partage des eaux (voir figure2.14).

Cette idée proposée dans les années 80 par Lantuéjoul et Beucher [71, 19, 18] a conduit au paradigme morphologique bien connu sous le nom de ligne de partage des eaux morphologique. Les travaux qui suivirent furent très nombreux et riches. Ce paradigme fut utilisé avec succès pour résoudre bon nombre d'applications avec les techniques d'imagerie les plus diverses. En effet, comme la partition en zones plates, la partition en bassins versant est une partition sans paramètre parfaitement intrinsèque à l'image. D'ailleurs, on peut remarquer très aisément que les deux paradigmes coïncident parfaitement pour les

(i) (ii)

FIG. 2.12: Effets des opérateurs connexes sur les fonctions numériques. (i) est l'image originale. (ii) est obtenueà partir de (i) par fusion des zones plates ayant des niveaux de gris similaires : les différences plus faibles que 38 sont négligées. Quand deux zones plates fusionnent, la luminance moyenne est attribuée à la zone plate résultante. Les partitions en zones plates associées à ces opérateurs sont celles de la figure 2.11.

fonctions en escalier, c'est-à-dire constantes par morceaux (en les différentiant au sens des distributions). Disposer d'un paradigme sans paramètre est bien, mais ne suffit pas. Encore faut-il pouvoir aisément contraindre l'algorithme à produire la segmentation désirée.

La première étape après l'« invention de la *LPE* » a consisté à exprimer le paradigme en termes de squelette par zone d'influence [18] afin de pouvoir la programmer et la définir à partir d'un ensemble de sources quelconques disposées sur l'image (dans le schéma le plus simple, les sources d'inondation sont les minima régionaux de la fonction). Permettre à l'utilisateur de choisir les sources d'inondation c'est lui donner la capacité de contraindre l'algorithme afin que ne soient segmentées que les régions désirées (sous réserve d'être capable de les localiser).

Considérant un ensemble de sources $(x_i)_{i\in I}$, la zone d'influence associée à x_i est consituée de l'ensemble des points du plan (support de l'image) plus proche de x_i que de toute autre source x_j :

$$IZ(x_i) = \{x \in E, \forall j \neq i,\, d(x, x_i) \leq d(x, x_j)\} \tag{2.9}$$

Le squelette par zone d'influence (noté SKIZ) est l'ensemble des points du plan équidistants de deux sources (au moins). Si d est la distance euclidienne, le squelette par zone d'influence correspond à la partition de Voronoi. Si d est la distance topographique (encore appelée distance image), alors le

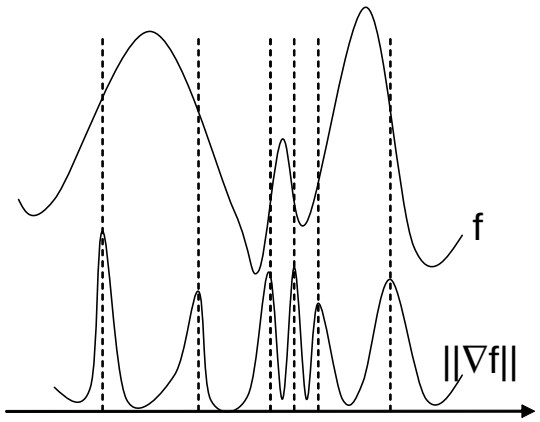


FIG. 2.13: Fonction à variations lentes et partition en bassins versants.

squelette par zone d'influence correspond à la ligne de partage des eaux morphologique [95, 96]. In [96], la distance topographique calculée sur l'image f entre les points x et y est définie par

$$d_f(x,y) = \inf_{\gamma \in \Gamma_{xy}} \int \|\nabla f(\gamma(s))\| ds$$

Boomgaard a par ailleurs formalisé la ligne de partage des eaux comme la solution d'une minimisation d'énergie [41].

Mais la description la plus célèbre de la ligne de partage des eaux morphologique est certainement celle qui l'identifie comme un processus de croissance de région (et dont les algorithmes les plus classiques s'inspirent).

L'image est à nouveau vue comme un relief et des sources d'inondation sont disposées de part et d'autre du relief. Une inondation progressive du relief est simulée. L'eau pénètre à niveau constant à partir des sources d'inondation. Ainsi, des lacs se forment inondant les bassins versants de l'image. Lorsque des eaux provenant de deux sources différentes se rencontrent, un barrage est érigé : voir figure 2.14. Le processus continue jusqu'à ce que le relief soit complètement inondé. La ligne de partage des eaux correspond à l'ensemble des barrages finalement érigés. Comme ce processus l'indique clairement, le relief et les sources d'inondations conditionnent complètement la ligne de partage des eaux finalement obtenue [129].

La segmentation par ligne de partage des eaux combine les avantages des approches région et des approches contour. L'approche région consiste à envahir l'objet en partant d'une zone initiale (une source). La croissance s'effectue à l'aide de critères plus ou moins élaborés sur l'homogénéité de l'objet.

Dans l'approche frontière, on procède de manière duale en recherchant les frontières de l'objet et non son contenu. L'hypothèse soutenue montre que les fortes variations lumineuses mises en évidence par l'image gradient délimitent l'objet. À partir de cette image gradient, un simple seuillage peut suffire à retrouver des contours simples. La segmentation par ligne de partage des eaux réalise un compromis particulièrement intéressant puisqu'elle travaille par croissance de région sur l'image gradient.

Le paradigme de segmentation par ligne de partage des eaux peut donc se résumer comme la simulation d'une inondation d'un relief à partir de sources. Tel quel, le processus est entièrement non paramétrique. Pour introduire des contraintes supplémentaires et satisfaire les objectifs de la segmentation (et de ses utilisations futures), différentes solutions sont possibles. On peut agir sur les sources, le relief ou encore la nature de l'inondation. Ces questions ont fait l'objet de nombreuses recherches depuis les travaux fondateurs de Lantuéjoul et Beucher. Trouver les bonnes sources d'inondation signifie avoir localisé (même grossièrement) les régions à segmenter dans l'image. L'analyse granulométrique [81], l'étude des extrema des images numériques [117, 56, 57, 129, 130, 134], les filtres connexes [110] et les nivellements morphologiques [82, 39, 84, 120] sont autant d'outils ayant permis une bonne maîtrise de ce problème dans des applications particulièrement délicates. La question du choix du relief à inonder est également un problème important. Dans le processus d'inondation, ce choix indique avec quelle priorité les points du relief sont agrégés aux sources. Nous reviendrons sur ce point dans le paragraphe suivant, car c'est à ce niveau que nous nous proposons d'intervenir. Enfin, le choix de l'inondation a permis de contraindre la géométrie des régions segmentées afin d'obtenir des contours plus réguliers (à la manière des snakes par exemple) et d'augmenter la robustesse du processus de segmentation sans pour autant perdre en précision. Ceci a donné naissance à la ligne de partage des eaux morphologique visqueuse [131][1].

2.2.2 Ligne de Partage des Eaux et distance de Mahalanobis

Lorsqu'il est possible de travailler à partir d'un unique échantillon de référence (ce qui est souvent le cas en médical car on focalise souvent sur un seul organe ou sur une seule pathologie), alors le calcul de la distance de Mahalanobis fournit une carte topographique suffisante pour pouvoir segmenter les images multivaluées. Le paradigme de segmentation est donc très simple. Il se fait en deux temps : un, le calcul de la carte des distances ; deux, le calcul de la *LPE* sur la carte des distances.

[1] Ce point sera abordé en section 2.4

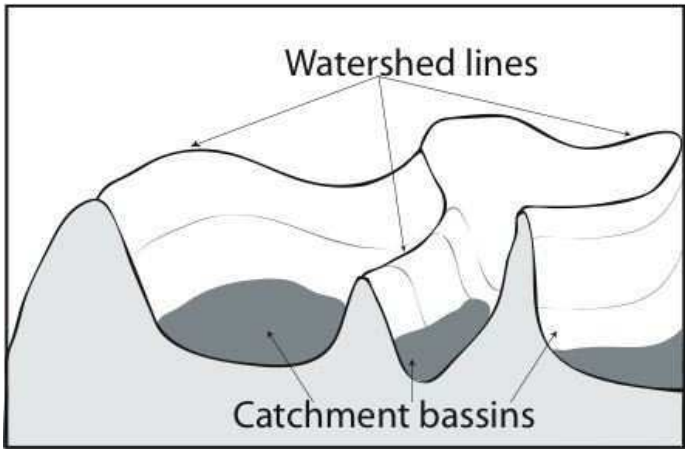


FIG. 2.14: Représentation de l'inondation des bassins versants. La Ligne de partage des eaux est l'ensemble des points de séparation des eaux provenant de différents bassins. Elle est localisée sur la crête du relief. La stratégie combine une approche région et une approche frontière : la croissance s'effectue autour des minimums régionaux (ou plus généralement de sources d'inondation prédéfinies) et la rencontre à lieu sur les lignes de crête de l'image gradient.

A titre d'exemple, nous illustrons maintenant le fonctionnement et les performances de cet algorithme dans le cas d'images couleur. Nous aborderons en fin de chapitre les applications médicales.

Quelques exemples couleur Comme il a été mentionné au chapitre 1, le choix de l'espace colorimétrique est un point important pour l'analyse des images couleur. Cependant, en ce qui nous concerne, ce choix rentre dans un problème plus général. En effet, nous avons montré au paragraphe 2.1.1 que les classes doivent être contenues dans des hyperellipses afin de pouvoir être séparées par la distance de Mahalanobis. Par conséquent, nous devons regarder la représentation de l'image dans l'espace des caractéristiques afin de mesurer la complexité du problème (figure 2.16).

Si les classes sont vraiment mélangées, il faut étudier d'autres alternatives comme le choix d'un autre système colorimétrique ou le fait d'ajouter d'autres critères (nous illustrerons les possibilités en la matière dans le paragraphe suivant dans le cas de la segmentation d'images texturées).

L'exemple de la figure 2.15, montre que malgré le passage à la carte des distances, la segmentation n'est pas immédiate. L'objet est trop complexe. La nécessité d'une régularisation est évidente, car sans cette étape, les contours obtenus directement à partir de la carte des distances seraient particulièrement fluctuants. Ainsi, avant de segmenter l'image, la carte des distances est filtrée (en l'occurrence par fermeture morphologique) de telle sorte que la ligne de partage des eaux produite soit plus régulière ; le processus de segmentation sera plus également plus robuste.

Le choix de la fermeture morphologique est un choix très classique, car cet opérateur permet de fermer des contours qui ne l'étaient pas et pour lesquels la ligne de partage des eaux donne des résultats extrêmement imprécis. Des choix plus judicieux qu'une simple fermeture peuvent être imaginés (et notamment le recours à une fermeture visqueuse : confer section 2.4). Dans le cas présent, notre souci n'est pas la régularisation de la segmentation, mais la définition et l'expérimentation d'un paradigme.

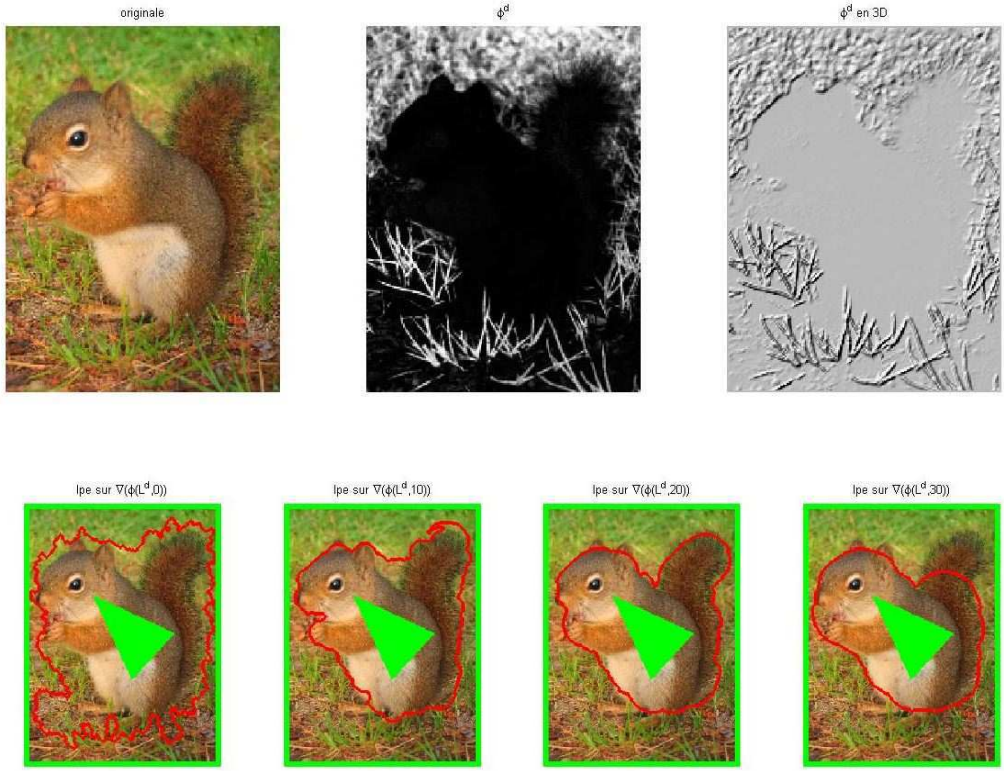


FIG. 2.15: Segmentation par calcul de la ligne de partage des eaux sur la carte des distances. En haut, l'image originale, la carte des distances et sa représentation $3D$. L'échantillon de référence est l'écureuil (zone triangulaire verte représentée en superposition sur l'image). Cette zone sert également de source d'inondation lors du calcul de la ligne de partage des eaux. En bas, trois résultats de la segmentation par *LPE* sur le gradient de la carte des distances. En fonction des différentes fermetures (disque de taille $0, 10, 20, 30$), le résultat est plus ou moins grossier. Une fermeture avec un disque de taille 20 (troisième image) permet d'approcher au mieux la forme recherchée.

L'exemple présenté en figure 2.15 illustre bien les performances de la segmentation obtenue via la carte des distances. Rapportées à une représentation scalaire par simple projection (sans prise int{egration d'information sur la nature statistique des régions), ces images sont extrêmement difficiles à segmenter : la plupart des algorithmes classiques (sans sur-paramètrage) échouent. On peut à titre d'exemple comparer les segmentations obtenues via la carte des distances et via une analyse par composante : voir les figures 2.16 et 2.17.

Les exemples des figures 2.18 et 2.19 montrent l'impact du choix de la métrique sur les résultats. La distance de Mahalanobis pour les raisons que nous avons détaillées semble de loin une excellente métrique. À titre d'illustration, nous présentons les résultats en utilisant deux métriques et l'image en niveaux de gris. Dans tous les cas, la même stratégie de segmentation est utilisée.

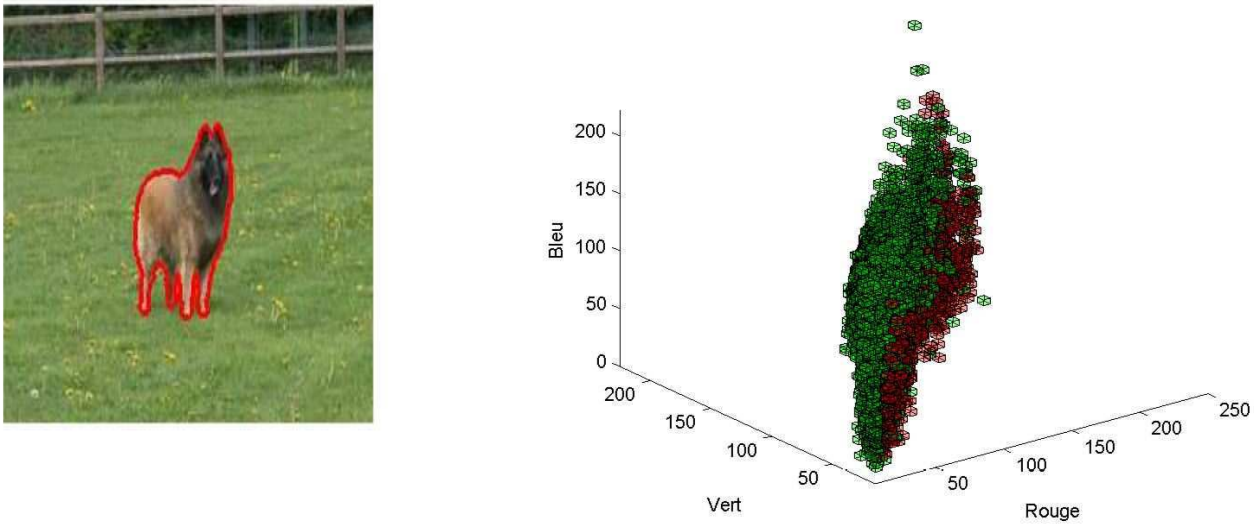
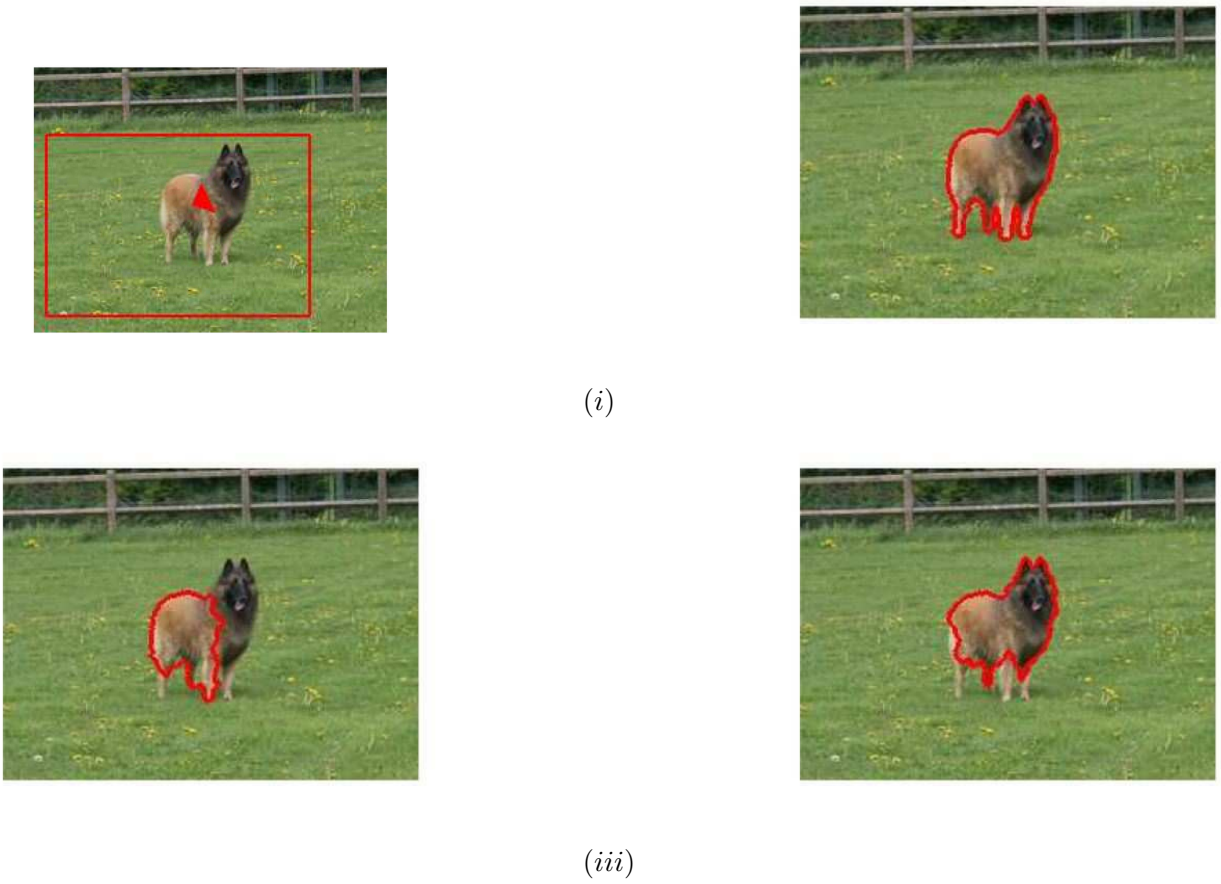


FIG. 2.16: Afin de comprendre la complexité du problème, il faut transposer l'image dans l'espace des caractéristiques. Dans cet exemple, deux classes sont représentées dans l'espace RGB. Les deux classes sont correctement séparées et peuvent être incluses dans des hyperellipses. Par conséquent, la carte des distances constituera un support favorable à une segmentation.

FIG. 2.17: Segmentation obtenue sur chaque composante vectorielle (dans la représentation RGB). Sur cet exemple, les mêmes marqueurs que ceux utilisés en figure 2.16 sont retenus. Dans chaque cas, la ligne de partage des eaux est calculée sur le gradient de l'image.

Maintenant, dans le cas où plusieurs objets sont recherchés dans l'image, le problème est un peu plus complexe puisque plusieurs références statistiques doivent être simultanément prises en compte. Dans

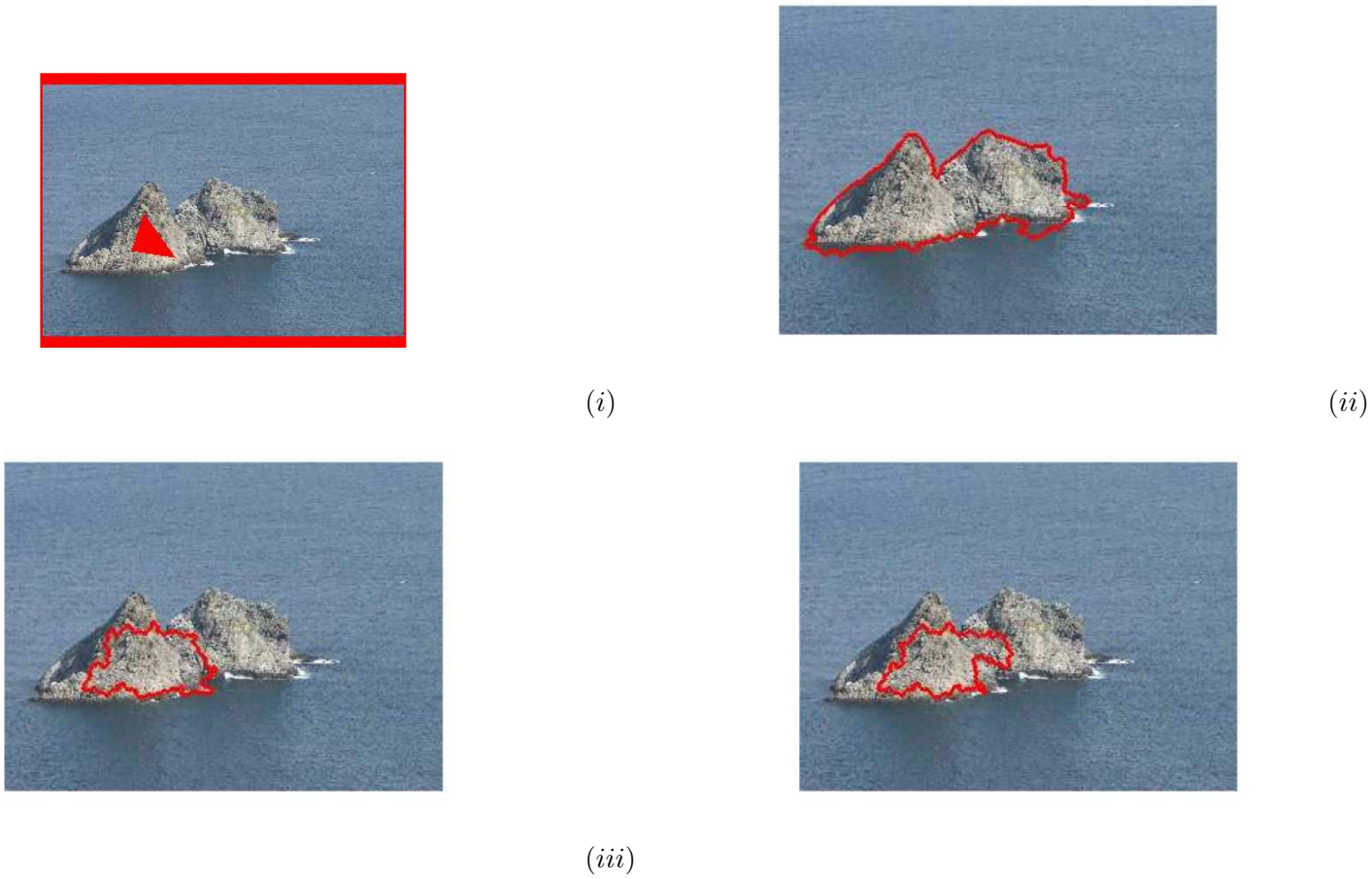


(i) (ii)

(iii) (iv)

FIG. 2.18: Comparaison des segmentations obtenues via différentes cartes de distance. (i) est l'image originale, les marqueurs intérieurs et extérieurs sont superposés en rouge à l'image. (ii) est le résultat obtenu via la distance de Mahalanobis (l'échantillon de référence est le marqueur rouge sur le chien). (iii) est le résultat de la segmentation via la distance euclidienne. (iv) est la segmentation obtenue par segmentation d'une unique composante (meilleur résultat de la figure 2.17).

ce cas, le paradigme de segmentation par ligne de partage des eaux doit être reformulé sur le modèle des approches de type *SRG*, de l'anglais « Seeded Region Growing » initialement introduites par Adams et Bischof en 1996 [1]. Tout comme dans le calcul de la ligne de partage des eaux, des sources pointant sur les zones à segmenter sont définies, les points libres de l'image sont agrégés aux sources dont ils sont les plus proches. A l'origine, Adams et Bischof proposent une distance au sens de la valeur moyenne. Contrairement aux cas précédents, chaque source définit une référence statistique pour le calcul des distances : les sources sont ainsi mises en compétition (voir figure 2.20). Les références statistiques sont fixées à l'initialisation et ne croissent pas avec le temps : cela n'aurait comme conséquence que de niveler les statistiques et donc de rendre plus difficile la décision pour la croissance. Cet algorithme de



(i) (ii)

(iii) (iv)

FIG. 2.19: Segmentation obtenue par calcul de la ligne de partage des eaux sur l'image gradient de la carte des distances de Mahalanobis ((ii)), euclidiennes ((iii)), ou encore en ne retenant qu'une seule composante de l'image couleur ((iv)). (i) est l'image originale ; les marqueurs sont superposés en rouge à l'image.

segmentation est illustré en figures 2.21, 2.22, 2.23 et 2.24]. La distance de Mahalanobis propose une normalisation par la variance ce qui induit une bonne robustesse au bruit.

Cas d'une image multivaluée formée par analyse granulométrique Comme il a été dit en introduction, les images couleur ne sont pas les seuls exemples d'images multivaluées. A titre d'illustration, nous présentons ici un exemple de segmentation de texture recourant à une analyse granulométrique et statistique. La technique utilisée ici pour l'analyse de la texture a été proposée par Dougherty [42]. Elle prend le nom de *local pattern spectrum*. Elle est fondée sur une analyse granulométrique de l'image combinée à un calcul des moments statistiques locaux. L'analyse granulométrique consiste à appliquer sur l'image des ouvertures de taille croissante ce qui permet de représenter l'image pour différentes échelles de taille. Plus la taille de l'ouverture augmente, plus la représentation est grossière. Le « pattern

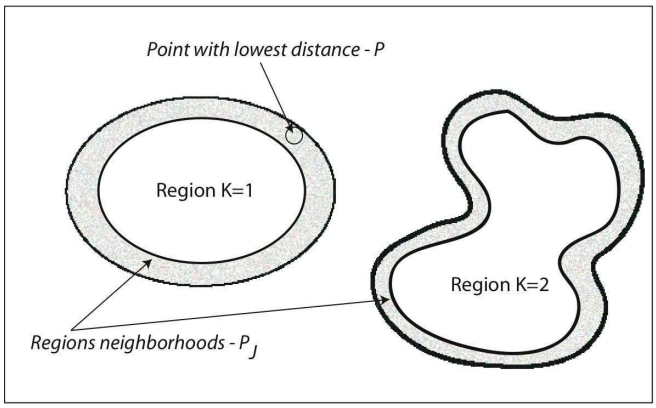


FIG. 2.20: Modification du protocole de croissance de région dans le cas où chaque source définit une référence statistique. Chaque région grossit à un rythme fixé par la distance statistique des points libre à la source.

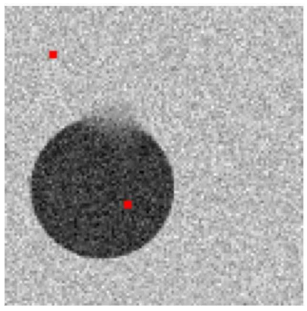


FIG. 2.21: Image originale avec les deux marqueurs initiaux. Un bruit gaussien est ajouté à l'intérieur et à l'extérieur. Les algorithmes de segmentation classiques peu robustes au bruit échouent.

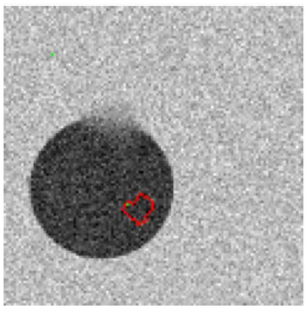


FIG. 2.22: Résultat obtenu par calcul directe de la ligne de partage des eaux. Sans pré-traitement particulier, l'algorithme échoue.

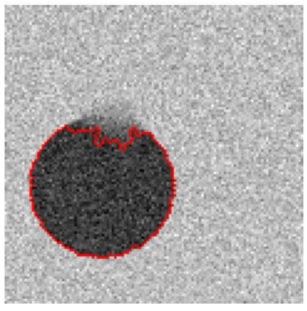


FIG. 2.23: Résultat obtenu par *SRG* classique (utilisant la distance à la moyenne comme critère d'évolution).

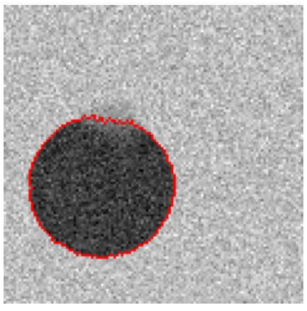


FIG. 2.24: Résultat obtenu en conditionnant la croissance de région par la distance de Mahalanobis à chaque source.

spectrum » [78] consiste à construire une courbe granulométrique associant à chaque échelle de représentation (chaque taille d'ouverture) le ratio $1 - \frac{A_\lambda}{A_0}$ où A_λ est l'aire de l'image à l'échelle λ et A_0 l'aire de l'image originale. La dérivée de cette distribution est une densité de probabilité. Le *local pattern*

spectrum est construit sur le même principe, mais l'analyse est rendue locale en considérant une fenêtre centrée sur chaque de l'image. Ensuite les trois premiers moments statistiques sont calculés. Ainsi, pour chaque résolution et en chaque point de l'image, un vecteur de dimension 3 est obtenu qui caractérise la texture à laquelle le point appartient.

Le *local pattern spectrum* conduit donc à la formation d'une pile d'images à valeurs vectorielles, les vecteurs étant de dimension $(3 \times N)$ (3 moments calculés pour N tailles d'ouvertures différentes pour l'analyse granulométrique). Sur l'exemple de la figure 2.25 [2], $N = 5$. Cette image multivaluée est ensuite segmentée par calcul de la LPE sur la carte des distances, trois sources (trois références statistiques étant mises en compétition), une dans chaque texture. Le résultat obtenu est présenté en figure 2.25.

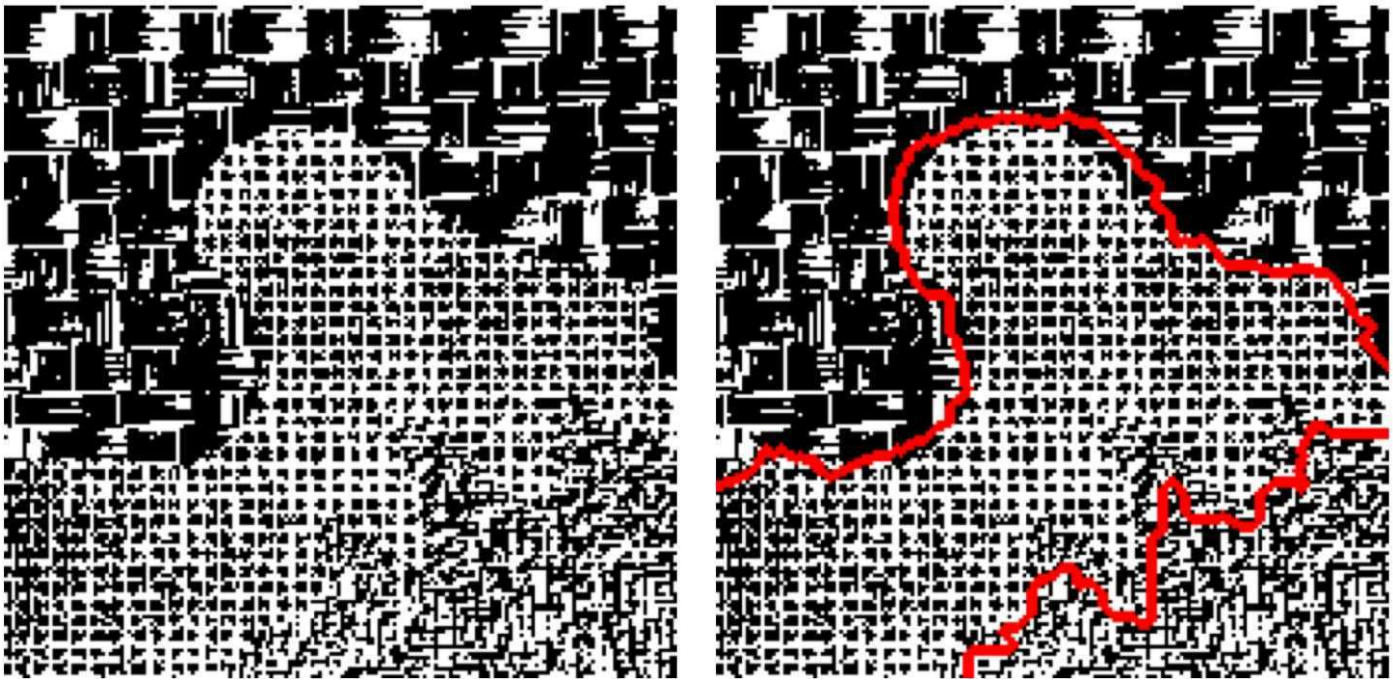


FIG. 2.25: Segmentation d'une texture après décomposition en plusieurs résolutions selon la méthodologie proposée par [42]. L'image segmentée est constituée de vecteurs de dimension 15. Trois marqueurs locaux (quelques pixels) sont positionnés dans les 3 zones d'intérêt.

2.2.3 Conclusion

Les exemples présentés dans cette partie montrent clairement que la distance de Mahalanobis est une mesure de similarité robuste ; les images que nous avons présentées correspondent à des problèmes rela-

[2] Nous remercions Lucas Tamarit et Michel Kocher de l'école d'ingénieur de Genève (EIG) pour nous avoir fourni cette image.

tivement complexes de segmentation pour lesquels les cartes de similarité se révèlent fort intéressantes. Comme nous l'avons dit en début de section, l'exploitation des distances statistiques et de la distance de Mahalanobis pour la segmentation est maintenant bien connue. La question du filtrage des images par contre n'a pas fait l'objet d'études approfondies. Notamment, considérant que les cartes de similarités ont une forte valeur sémantique puisque chaque point de l'image est ainsi hiérarchisé selon son importance dans la scène et pour l'application, il semble naturel de construire des filtres susceptibles de traiter différemment les points de l'image selon leur niveau de similarité, autrement dit des filtres auto-adaptatifs. Nous reviendrons ensuite sur des questions de segmentation, mais via le premier paradigme morphologique, la partition en zones plates, et nous verrons comment il est possible de définir des opérateurs connexes adaptatifs permettant une prise en compte plus importante de l'a priori statistique dans toute la chaîne d'analyse.

2.3 Filtrage morphologique des images multivaluées via la transformation de rang

Les applications de filtrage, contrairement aux applications de segmentation, requièrent un retour dans l'espace des valeurs vectorielles. Nous avons dit déjà qu'une manière simple de procéder consiste à se munir d'un ordre total sur les vecteurs. Il a été proposé en 2.1.2, la construction d'une image de rang constituée d'entiers consécutifs permettant d'ordonner totalement les vecteurs en combinant la distance de Mahalanobis et la cascade lexicographique. Cette image de rang définie, tous les opérateurs morphologiques développés pour les images à teintes de gris peuvent être étendus aux images multivaluées. Pour des raisons de simplicité, la majorité de nos illustrations porteront sur des images couleur.

Dans cette partie, les opérateurs morphologiques classiques étendus aux images multivaluées sont étudiés. La section suivante exploitera les conclusions de cette étude et engagera la construction d'opérateurs plus exotiques.

2.3.1 Action des opérateurs morphologiques fondamentaux

Puisque l'image de rang scalaire est une représentation sans perte de l'image vectorielle F (voir section 2.1.2), le traitement morphologique de l'image vectorielle revient à transformer l'image de rang puis à revenir dans l'espace des valeurs vectorielles :

$$T = T_{rank}^{-1} \circ t \circ T_{rank}$$

où t représente l'opérateur agissant sur les fonctions scalaires et T l'opérateur équivalent agissant sur les fonctions vectorielles.

Nous allons maintenant expérimenter les filtres morphologiques les plus classiques et nous servir de ces exemples pour expliquer comment la distance de Mahalanobis ordonne les différents éléments qui composent la scène et interpréter l'action des différents opérateurs. Différents cas de figure sont examinés. Le premier est le plus simple : la référence statistique est composée d'une seule couleur (mais contient des teintes plus ou moins contrastées et du bruit). L'ordre induit sur les vecteurs est relativement intuitif. Dans le second cas, la référence statistique est un objet plus complexe composé de plusieurs couleurs. Comme nous allons le voir, la distance de Mahalanobis produit des effets de filtrage moins intuitifs (l'ordre sur les vecteurs ne correspond plus à une échelle de perception des couleurs) car les différentes nuances de couleur constituant la référence sont mélangées. Dans le troisième exemple, la référence statistique est une texture ce qui correspond à l'utilisation la plus adaptée de la distance de Mahalanobis.

Cas d'une référence statistique d'une seule couleur Ce cas correspond aux figures 2.26 à 2.28. Dans l'exemple de la figure 2.26, un échantillon de rouge du ballon est choisi comme ensemble de référence pour le calcul de la distance de Mahalanobis. Les points du ballon ont des fortes valeurs de similarité, les points à l'arrière-plan et le logo blanc sont des points de faible similarité. Souvenons-nous que l'image de rang est une image de similarité. Par dilatation, les points rouges sont dilatés et les points de faible similarité sont recouverts par les points de forte similarité avoisinants. Ainsi, le logo blanc du ballon disparaît, le ballon rouge grossit sensiblement.

Sur l'exemple des figures 2.27 et 2.28 maintenant, la référence statistique est encore extraite de l'image : c'est la fleur jaune au premier plan de l'image. Les vecteurs couleur de l'image originale sont ordonnés dans le sens de la référence. Les teintes de jaune sont les teintes dominantes (de plus forte similarité), les teintes de vert puis de gris sont des teintes de plus faible similarité. L'ordre induit par la distance de Mahalanobis est présenté sous forme de mire en figure 2.27, (ii). On remarque que certaines teintes de jaune s'intercalent parmi les teintes de vert et de gris ce qui correspond à une mauvaise classification. Comme nous l'avons dit plus haut, une hypothèse fondamentale de l'utilisation de la

distance de Mahalanobis comme mesure de similarité est que l'échantillon de référence soit gaussien. Dans le cas présent, cette hypothèse est clairement non satisfaite. Les effets de la mauvaise classification sont visibles lorsqu'on calcule les transformations morphologiques par exemple sur la sous-figure $(xiii)$ de la figure 2.27 : une structure jaune apparaît sous une structure verte.

FIG. 2.26: Dilatation morphologique d'une image couleur définie par dilatation de l'image de rang. A gauche, l'image originale. A droite, le résultat de la dilatation par un disque euclidien. La référence statistique est ici un échantillon de rouge du ballon. L'érosion permet d'enlever l'ensemble des inscriptions sur la face du ballon.

Cas d'une référence statistique inhomogène Ce cas est illustré en figure 2.29. L'échantillon de référence est une pâquerette faite de nombreuses teintes, les couleurs majoritaires étant le jaune et le blanc. Comme illustré sur la figure 2.30, ordonner les couleurs dans ce cas n'est pas un problème trivial. La référence statistique résulte de l'union de deux nuages de points distincts (relatifs aux deux couleurs majoritaires, jaune et blanc dans notre cas). La distance de Mahalanobis est une distance euclidienne pondérée au centre du nuage. D'autres options (d'autres distances) pourraient être envisagées afin de donner la prédominance à un seul des deux nuages par exemple. Ceci n'a pas fait l'objet d'expérimentation, mais cet exemple ouvre des perspectives sur la question du choix de la distance statitisque selon la nature de l'échantillon considéré. La mire de classification des couleurs obtenue par calcul de la distance de Mahalanobis indique clairement que les teintes de jaune et blanc s'intercalent ; la couleur prédominante est tantôt le jaune, tantôt le blanc ce qui a des effets notables par filtrage (voir notamment les résultats obtenus par les filtres alternés séquentiels (sur la figure 2.29, voir les images (ix) et (x)).

Il faut cependant préciser que notre but ici n'est pas d'expérimenter les filtres morphologiques usuels. Nous nous orientons vers des opérateurs de simplification des images susceptibles de trier les structures

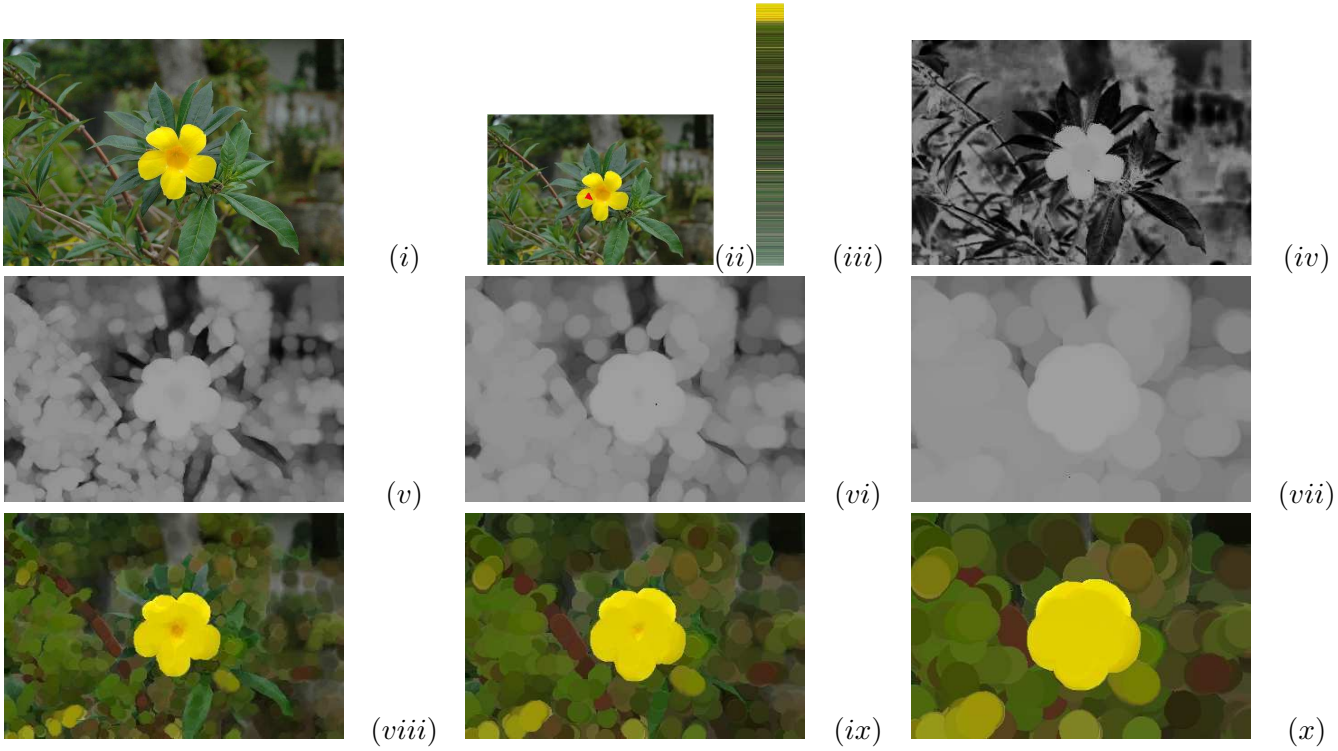


FIG. 2.27: Dilatation d'une image couleur par dilatation de son image de rang. L'élément structurant est un disque euclidien. (i) est l'image originale, (ii) est l'échantillon de référence, (iii) est la mire de classification des couleurs de l'image originale (la classification est obtenue par calcul de la distance de Mahalanobis combiné avec une cascade lexicographique), (iv) est l'image de rang. C'est une carte de similarité. $(v - vii)$ sont les résultats de dilatations morphologiques par des disques de taille 5, 10 et 20 de l'image de rang et $(viii - x)$ sont les résultats des mêmes dilatations avec retour dans l'espace des couleurs.

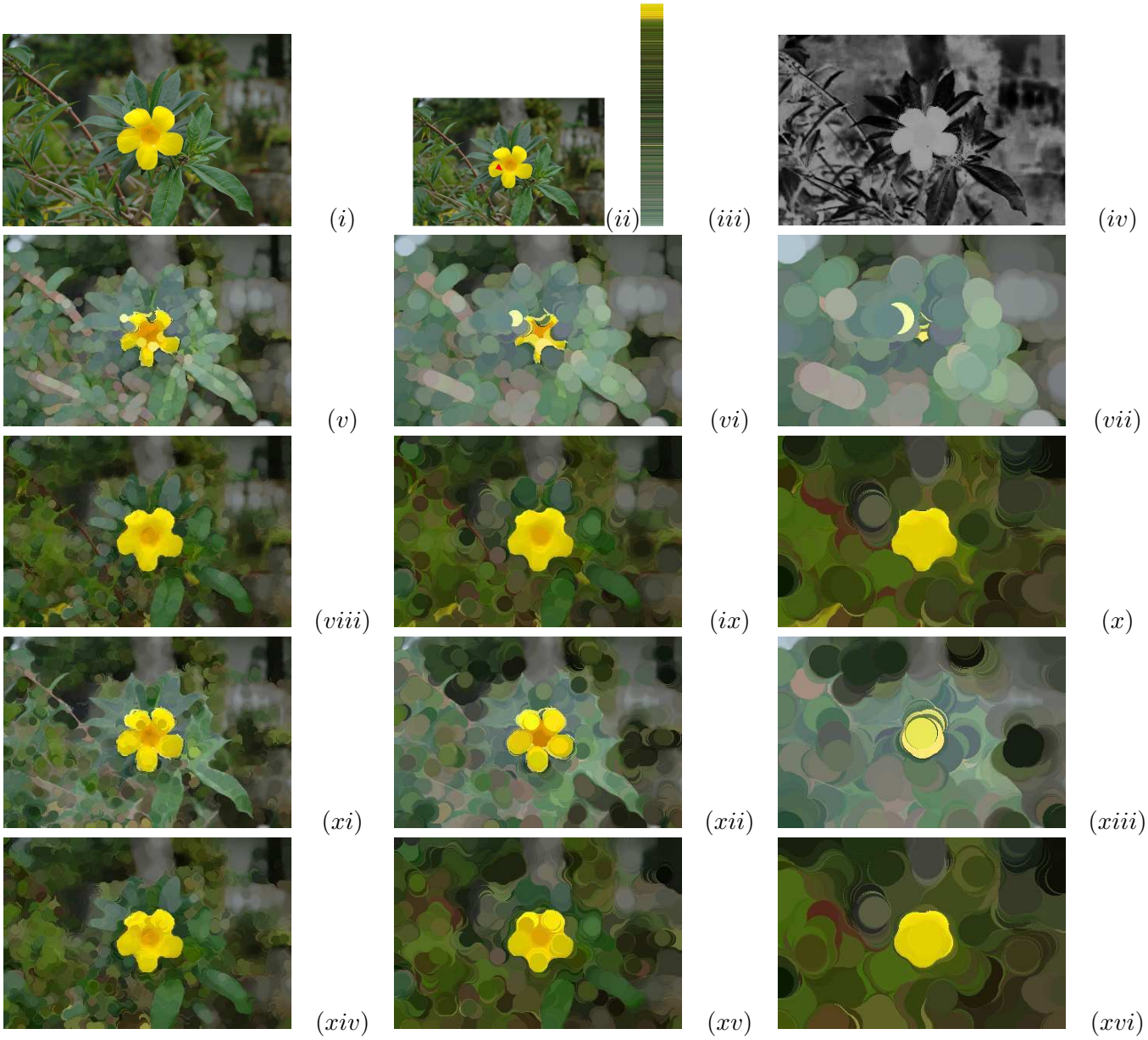


FIG. 2.28: Erosions, ouvertures et fermetures d'une image couleur (même principe que pour la figure 2.27). $(v - vii)$ sont les résultats d'érosions morphologiques par des disques de taille 5, 10 et 20. $(viii - x)$ sont les résultats de fermetures morphologiques de taille 5, 10 et 20. (xi-xiii) sont les résultats d'ouvertures morphologiques de taille 5, 10 et 20. (xiv-xvi) sont les résultats de filtres alternés séquentiels (FAS) de taille 5, 10 et 20.

dans l'image selon leur ressemblance au modèle statistique de référence. L'organisation des couleurs dans la mire n'a pas lieu de suivre la perception visuelle.

Cas où la référence est une texture Cet exemple est illustré en figure 2.31. La distance de Mahalanobis est bien adaptée aux échantillons texturés pour lesquels le modèle gaussien est généralement valide. Dans l'exemple de la figure 2.31, la référence est un échantillon de la pelure du chien. Comme cela est confirmé par la mire de classification des couleurs, la distance de Mahalanobis est très performante et les opérateurs morphologiques sur les images multivaluées donnent des résultats très cohérents.

Par ouverture, toutes les structures de petite taille de l'image qui ont une plus forte similarité que leur voisinage sont éliminées ; les grosses (comme le chien) sont partiellement préservées. Par érosion, les fleurs ressortent (elles sont dilatées) car elles sont à la base de la mire : le jaune est la couleur la moins similaire au chien. L'action des filtres est complètement déterminée par le choix de la référence statistique dont découle l'ordre sur les couleurs. Sur l'exemple de la figure 2.32, la référence est un échantillon d'herbe ce qui modifie complètement la carte de similarité et cette fois-ci, par dilatation, le chien qui est peu similaire à l'herbe est érodé.

Revenons à l'exemple de la figure 2.31. Par fermeture, l'effet dual est obtenu. Le chien est lissé tandis que les fleurs dans l'herbe sont rasées. Par ouverture, fermeture ou FAS, la texture du chien est lissée puisqu'à l'intérieur du chien, il y a de petites variations en termes de similarité. Pour ces images et pour ce type de traitements où l'a priori statistique joue un rôle fondamental, il semble judicieux de ne pas filtrer les points de forte similarité comme les points de faible similarité. Autrement dit, si l'on veut transformer les images *dans le sens* de la référence statistique, il faut définir des filtres adaptatifs permettant d'éliminer l'information peu significative (les points de faible similarité) tout en préservant les points de plus forte similarité.

2.3.2 Les résidus morphologiques

Comme nous l'avons dit en section 2.1.2, les résidus morphologiques sont les transformations définies par des différences. Les résidus de dilatations et érosions définissent des gradients ; les résidus d'ouvertures et fermetures définissent des chapeaux haut de forme. Les deux points forts de la transformation de rang sont un, son caractère bijectif, deux, la production d'images de rang constituées d'entiers consécutifs dont on a vu le potentiel pour la définition de soustractions entre vecteurs. Dans cette partie, nous illustrons sur des exemples, les images obtenues par calcul des résidus morphologiques : voir

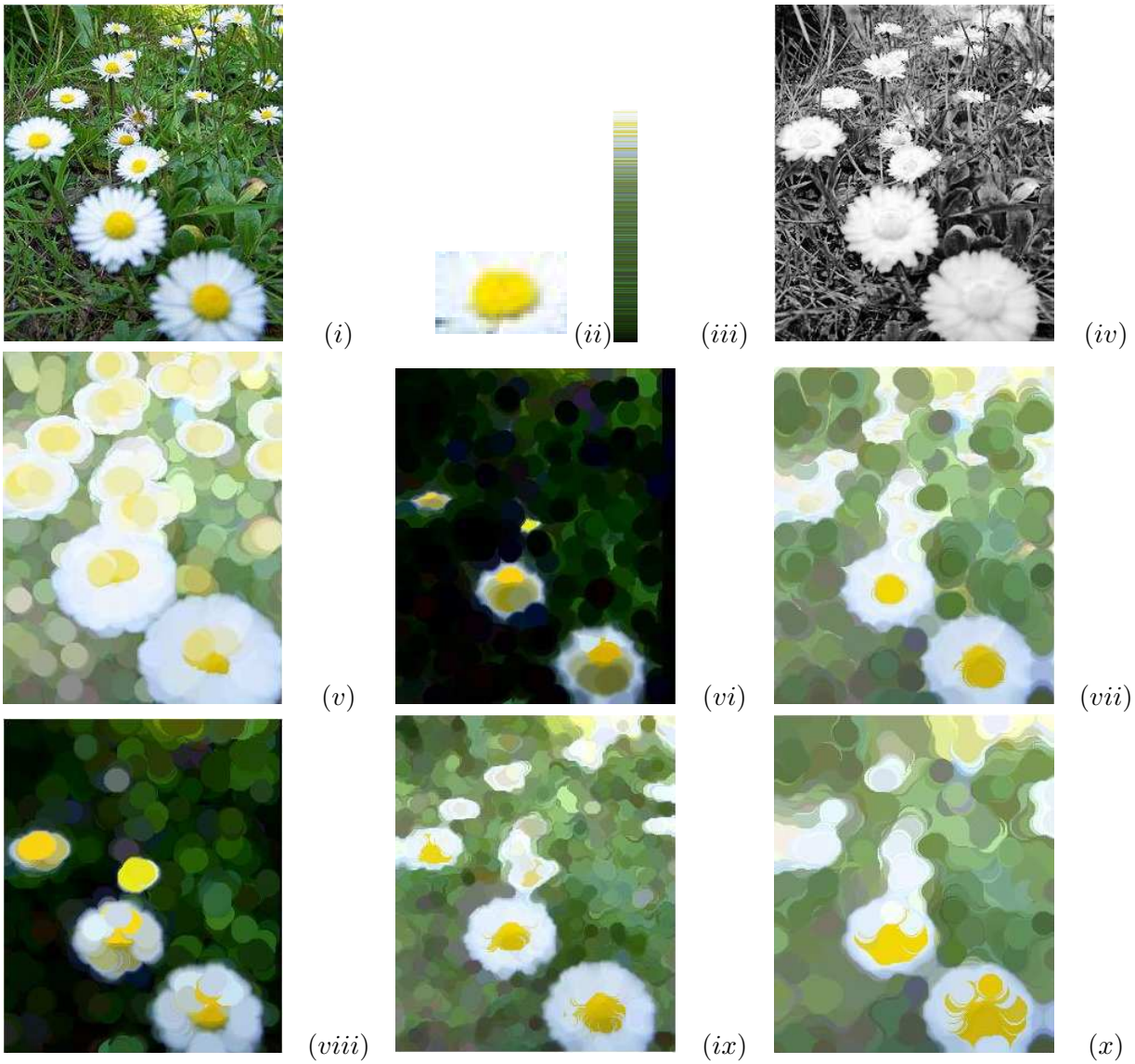


FIG. 2.29: Cas d'une référence composée de plusieurs couleurs. (i) est l'image originale (de taille $250 \times 300 \times 3$). (ii) et (iii) sont l'échantillon de référence et la mire de classification des couleurs. (iv) est l'image de rang. $(v - viii)$ sont les résultats d'une dilation, érosion, fermeture, ouverture de taille 10. (ix) et (x) résultent d'un FAS de taille 10 puis 20 pixels.

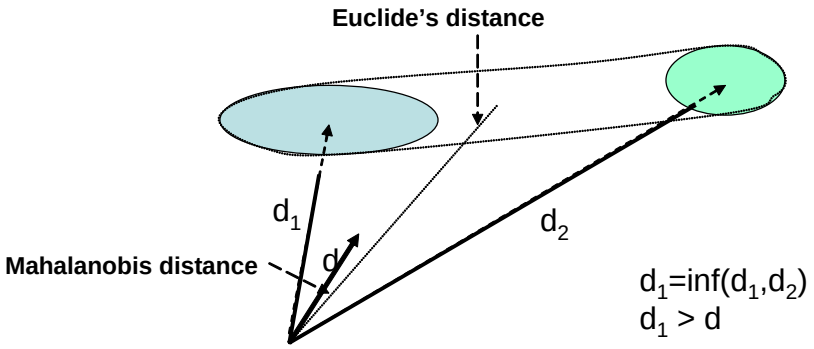


FIG. 2.30: Si l'échantillon de référence est une combinaison de plusieurs structures de couleurs différentes, la distance de Mahalanobis ordonne correctement les points du plan mais les couleurs dans l'échantillon peuvent s'intercaler.

figures 2.33 à 2.37. Le calcul n'introduit aucune nouvelle couleur qui n'existait déjà dans l'image originale, par contre, il n'y a aucune interprétation particulière des couleurs possibles sinon que les variations fortes sont représentées par des couleurs de haut rang et que les variations faibles sont représentées par des couleurs de faible rang.

Sur l'exemple de la figure 2.37, les grains de beauté sont extraits par gabarit de taille en calculant des chapeaux haut de forme par fermeture. Cependant, la représentation en fausses couleurs ne présente que peu d'intérêt si le problème est la détection et la classification (par la taille) des grains de beauté. Dans ce cas, le chapeau haut de forme calculé sur l'image de rang est une information tout aussi consistante et plus aisément exploitable.

2.3.3 Simplification des images multivaluées par filtrage morphologique connexe

Le but du filtrage est de produire une version simplifiée d'une scène. Dans de nombreuses applications, la qualité visuelle de l'image produite importe. Notamment, au niveau perceptuel, la préservation des transitions de luminance ou de couleur est de première importance. C'est le cas par exemple lorsqu'on cherche à produire des images plus simples en vue de leur codage. Ce problème rejoint la question de la segmentation telle qu'elle formulée par exemple par Mumford et Shah [92] : produire une image constante par morceaux la plus simple possible tout en restant au plus proche de l'original.

En morphologie mathématique, cette question est d'abord formulée comme un problème de segmentation. Il s'agit de produire des images simplifiées qui élargissent la partition en zones plates associée à l'image originale (nous avons défini la partition en zones plates en section 2.2.1). Ainsi, aucune nouvelle transition ne peut être introduite par filtrage ; il y a élargissement des zones de gradient nul. Les

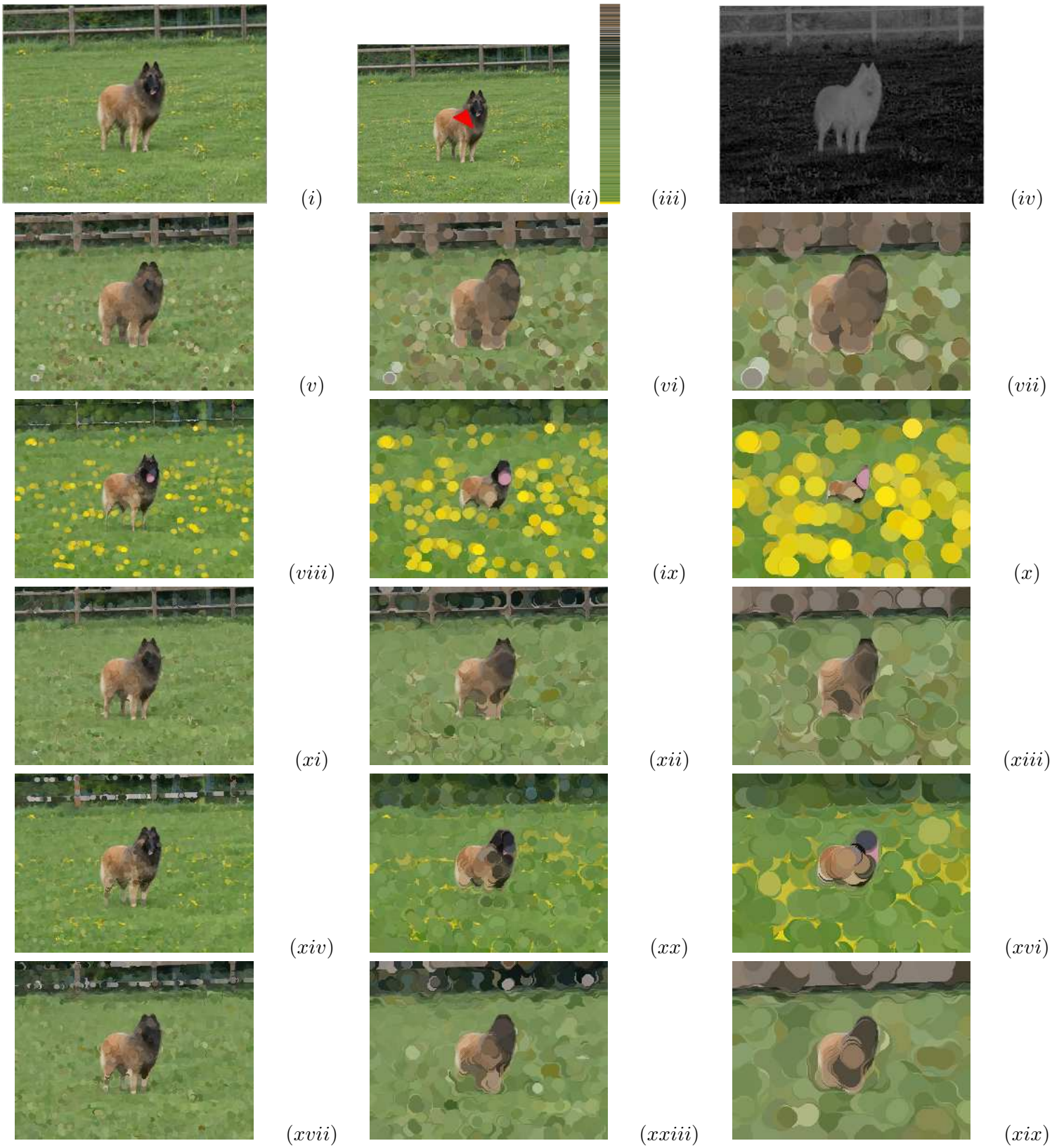


FIG. 2.31: Cas d'un échantillon texturé. (i) est l'image originale, (ii) indique la référence (iii) est la mire de classification des couleurs, (iv) est l'image de rang. Résultats de dilatations $(v-vii)$, érosions $(viii-x)$, fermetures $(xi-xiii)$, ouvertures $(xiv-xvi)$ et FAS (xvii-xix) pour des éléments structurants de taille 2, 5 et 10.

FIG. 2.32: Changement de référence : cette fois, l'échantillon de référence est prélevé dans l'herbe. À gauche, l'image originale. À droite, le résultat de la dilatation morphologique par un disque de taille 10 pixels.

(i)

(ii)

FIG. 2.33: Résidus morphologiques couleur. (i) gradient mophologique. (ii) chapeau haut de forme par ouverture.

FIG. 2.34: De gauche à droite : image originale, gradient externe ($\delta - Id$) et gradient interne épais (l'élément structurant est un disque de taille 5). La référence pour le calcul de la distance est l'herbe.

opérateurs morphologiques qui ont de telles propriétés sont les opérateurs dits connexes. Ces opérateurs ont été formalisés par Serra et Salembier dans les années 95 [112] et utilisés avec succès dans les applications de codage [109]. Parmi les opérateurs connexes, on trouve toute une panoplie de filtres morphologiques comme les reconstructions [119] et les nivellements [82, 39, 84, 120]. Les nivellements sont des opérateurs connexes qui non seulement n'introduisent aucune nouvelle transition dans les images,

FIG. 2.35: De gauche à droite : image originale, gradient interne épais obtenu avec un disque de taille 5. La référence pour le calcul de la distance est le rouge du ballon.

mais qui en plus préservent le sens des transitions dans les images. Ainsi, l'organisation des luminances ou des couleurs est préservée.

Nous avons vu que les opérateurs morphologiques classiques multivalués lorsqu'ils sont définis via l'image de rang et donc via l'ordre induit par la distance de Mahalanobis peuvent produire des agencements de couleur perceptuellement inhabituels. Dans le cas des filtres connexes, chaque zone plate de l'image couleur (chaque zone constituée de points de couleur uniforme) sera soit préservée, soit élargie par absorption d'une ou plusieurs zones plates adjacentes. En effet, une propriété remarquable de la transformation de rang est de préserver les zones plates de l'image couleur. Deux vecteurs identiques de l'image multivaluée ont forcément le même rang. Partitionner l'image de rang en zones plates revient donc à partitionner l'image multivaluée en zones plates. Par les nivellements morphologiques, les agencements de couleur seront préservés. Si un objet est éliminé, il y a création à sa place d'une zone de couleur uniforme. Par conséquent, les nivellements ne vont pas se heurter au problèmes des agencements de couleur comme les filtres morphologiques plus standards. Ceci est illustré par les exemples des figures 2.38 à 2.40.

2.3.4 Conclusion

Les filtres morphologiques usuels peuvent être étendus aux images multivaluées via la distance de Mahalanobis et l'image de rang qui s'en déduit. Cependant, en mettant une référence statistique au centre de la construction de l'ordre sur les vecteurs, toute l'information est hiérarchisée dans le sens de

cette référence. Les possibilités ne se réduisent pas à la seule analyse morphologique conventionnelle. Il est tout à fait possible d'utiliser la hiérarchie introduite via la distance statistique de telle sorte à produire des images simplifiées qui préservent parfaitement les formes statistiquement proches de la référence tout en simplifiant les structures peu similaires à la référence (ou bien l'inverse dans le cas où l'on désire éliminer les structures prises comme référence sans toucher au reste de l'image). Sur nos exemples précédents, la question se formule ainsi : est-il possible de simplifier les images sans abîmer la fleur jaune de la figure 2.38, les pâquerettes de la figure 2.39 ou encore le chien de la figure 2.40 ? Les filtres morphologiques conventionnels ne permettent pas ce type de simplification puisqu'ils agissent en tout point de l'image d'une même manière que ces points soient de forte ou faible similarité. Ainsi en voulant simplifier l'arrière-plan des images, les plus petites fleurs sont perdues, la texture du chien est gommée.

Pour orienter l' analyse dans le sens de la référence statistique, il faut réintroduire l'a priori statistique dans la formulation des filtres, soit encore recourir à des opérateurs adaptatifs pour lesquels l'activité de filtrage en chaque point de l'image sera déterminée en fonction de la similarité du point à la référence. Nous allons définir de tels opérateurs dans la section qui suit.

FIG. 2.36: De haut en bas : image originale, image fermée avec un disque de taille 3, chapeau haut de forme. La référence pour le calcul de la distance de Mahalanobis est de la peau saine. Par fermeture, les plus petits grains de beauté sont éliminés (ils ont une forte dissimilarité donc de faibles rangs). Par chapeau haut de forme, ils sont extraits.

FIG. 2.37: Suite de chapeaux haut de forme par fermeture pour des disques de taille croissante : $2, 3, 4$ et 5.

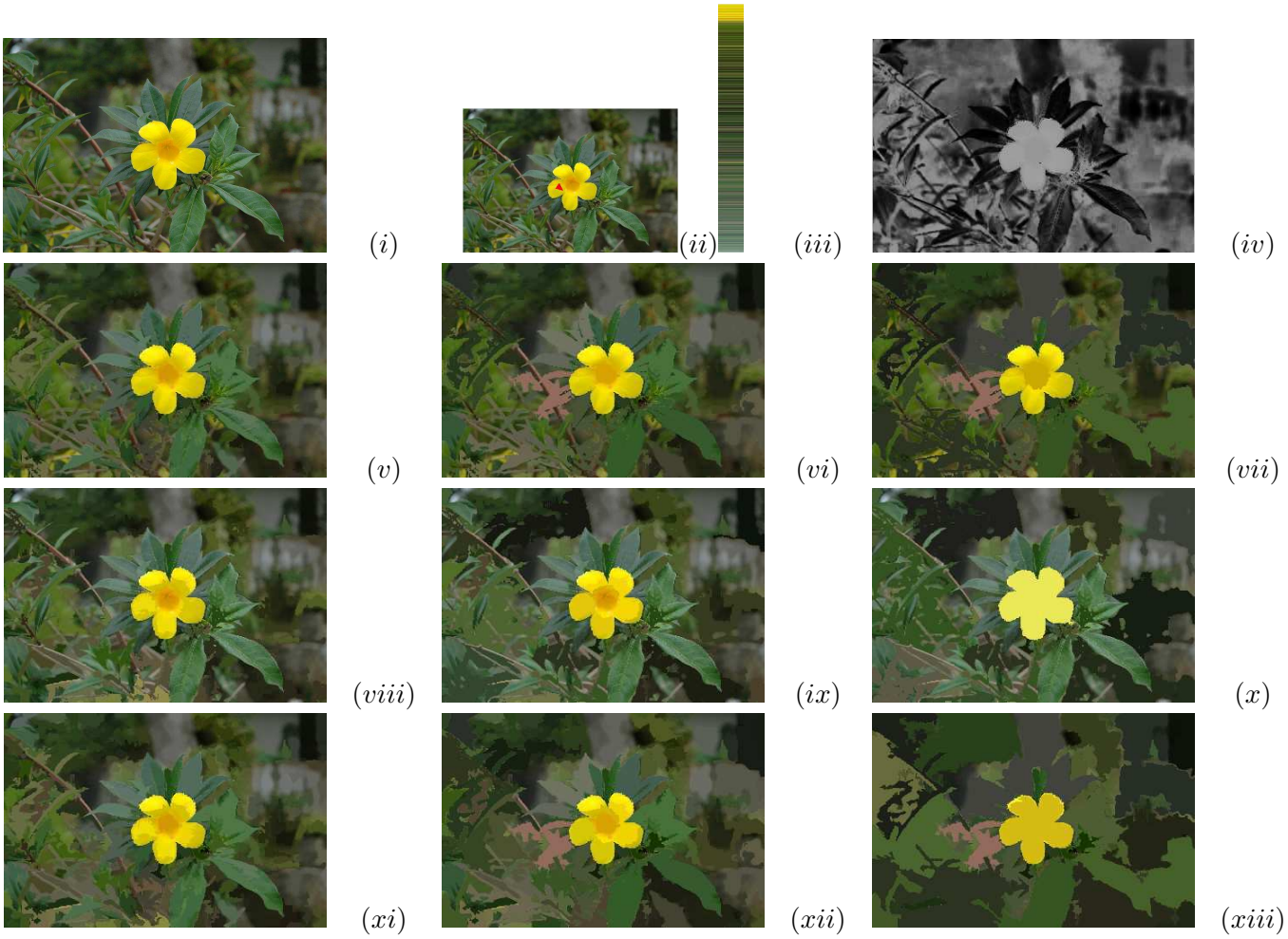


FIG. 2.38: Quelques filtres morphologiques par reconstruction. $(v - vii)$ Fermetures par reconstruction. L'élément structurant est un disque de rayon 5, 10 puis 20 $(viii - x)$ Même chose avec des ouvertures par reconstruction puis avec des FAS par reconstruction : $(xi - xiii)$.

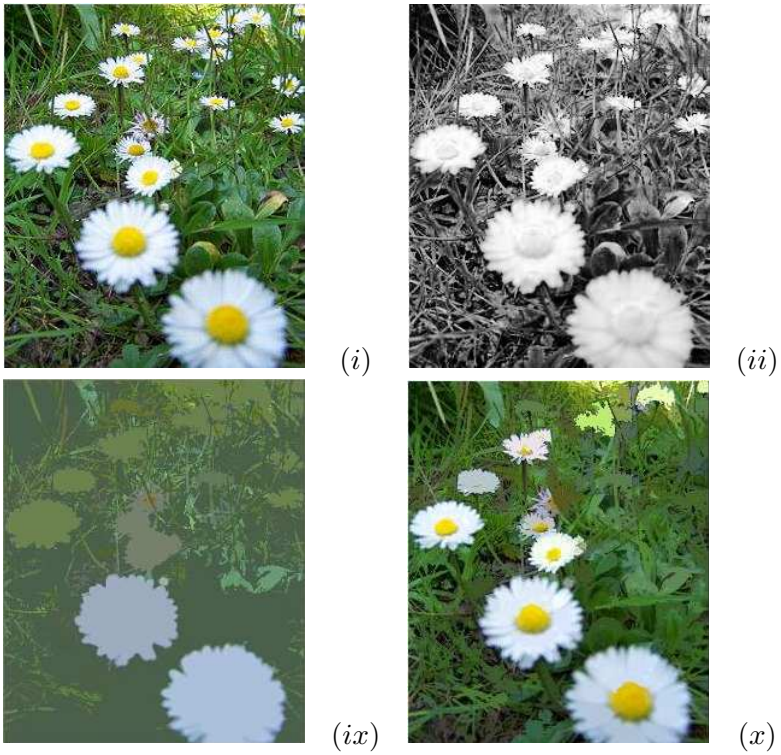


FIG. 2.39: (iii) Exemple d'un filtre alterné séquentiel par reconstruction : chaque ensemble de niveaux de l'image de rang ou de son inverse est soit intégralement préservé soit intégralement éliminé. (iv) Effet d'un nivellement surfacique [135] : tout ensemble de niveaux de l'image de rang ou de son inverse de surface inférieure à 500 pixels est éliminé.

FIG. 2.40: Effet des filtres morphologiques par reconstruction. Cas des ouvertures morphologiques par reconstruction $(v - vii)$, des fermetures morphologiques par reconstruction $(viii - x)$, des FAS par reconstruction $(xi - xiii)$ pour des disques de taille 5, 10 puis 20.

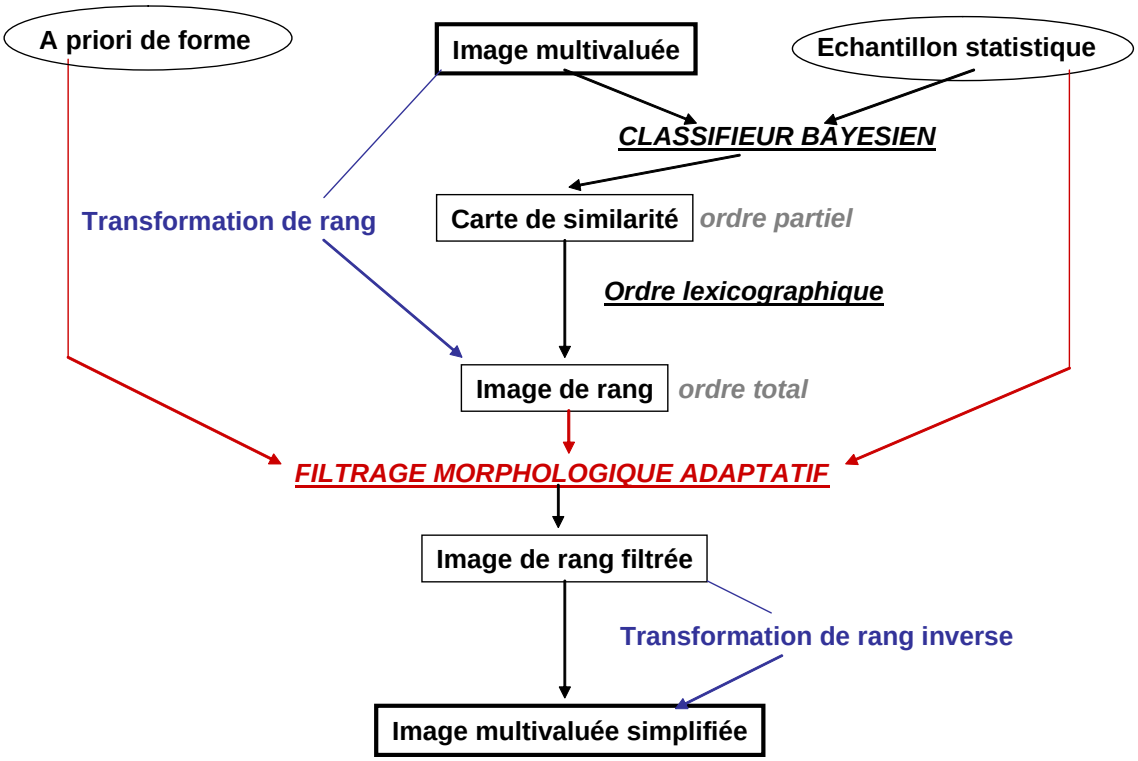


FIG. 2.41: Comment tenir compte de la référence statistique dans l'analyse morphologique d'une image multivaluée.

2.4 Définition d'opérateurs morphologiques auto-adaptatifs

Les opérateurs morphologiques standards agissent sur l'image conformément à un critère de taille et de forme prédéfini par l'intermédiaire de l'élément structurant. Les illustrations que nous avons montrées dans la section précédente ont été effectuées à partir d'éléments structurants dont la taille était fixée une fois pour toutes, et ceci, pour tous les points de l'image indépendamment de leur similarité à l'échantillon. Si maintenant on veut tenir compte de l'a priori statistique dans l'analyse morphologique, une solution simple consiste à exploiter l'information de l'image de rang. Sur cette image, les luminances ont une haute valeur sémantique ; elles indiquent si oui ou non un point est similaire à l'échantillon. En combinant résolument l'analyse morphologique à l'analyse statistique, il est alors possible de prendre en compte simultanément les informations dont on dispose sur la nature statistique et sur la géométrie des objets d'intérêt : ceci est résumé par l'organigramme de la figure 2.41.

L'idée que nous allons mettre en oeuvre est de simplifier beaucoup les objets de faible similarité tout en préservant les objets fortement similaires à la référence. La figure 2.42 illustre ce principe sur un profil de luminance de l'image de rang et sur l'exemple d'une fermeture adaptative. Aux seuils

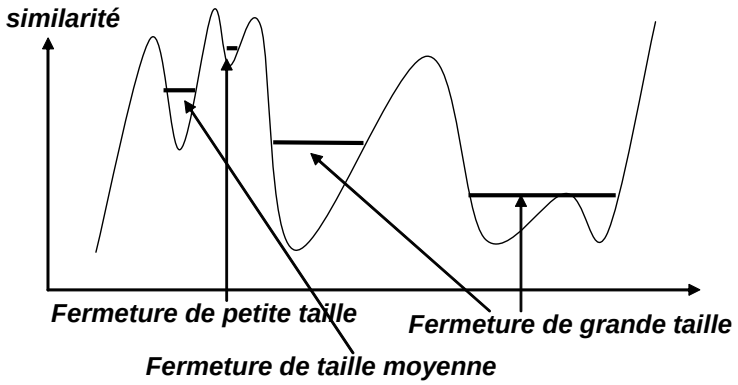


FIG. 2.42: Comment il est possible de définir des opérateurs adaptatifs sur l'exemple d'un processus de fermeture. La fonction représente un profil type d'une image de rang. Aux points de haut rang (forte similarité), on applique une fermeture de petite taille. Aux points de faible rang (faible similarité), on applique une fermeture de grande taille. La scène est simplifiée, mais les structures proches de la référence sont préservées.

correspondant à des hautes similarités sont associés des filtres de forte activité (l'élément structurant est de grande taille), tandis qu'aux seuils correspondant à de faibles similarités sont associés des filtres de faible activité (l'élément structurant est de petite taille).Bien sûr il est tout à fait possible d'inverser le processus afin de filtrer beaucoup les points de forte similarité et peu les points de faible similarité.

De tels opérateurs ont été introduits et formalisés dans le cas des images à teintes de gris par Meyer et Vachier, à l'origine dans le but de régulariser les contours extraits par calcul de la ligne de partage des eaux [131, 85, 132]. Deux modèles de transformations visqueuses ont été proposés. Le premier modèle (le modèle huile) permet de définir des transformations adaptées à la luminance locale sur l'image. Le deuxième modèle (le modèle mercure) permet d'adapter le traitement localement en fonction du contraste du point considéré. Dans le cas présent, ce sont les transformations de type huile qui nous intéressent puisque, sur les images de rang, c'est justement la luminance qui est significative. Notons enfin que jusqu'à présent les opérateurs visqueux n'ont pas été adaptés au cas des images multivaluées ou couleur.

2.4.1 Formulation classique dans le cas des images à valeurs scalaires

La définition des opérateurs visqueux est très simple lorsqu'on décompose les fonctions (scalaires) en ensembles de niveaux.

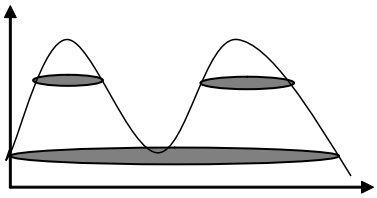


FIG. 2.43: Décomposition d'une fonction à valeurs scalaires en ensembles de niveau.

L'image numérique à valeurs scalaires étant notée f, nous notons $X_h(f)$ son seuil de niveau h

$$X_h(f) = \{x \in E,\ f(x) \geq h\}$$

et $\chi_h(f)$ l'indicatrice associée :

$$\forall x \in E,\ \chi_h(x) = \begin{cases} 1 & \text{if } x \in X_h \\ 0 & \text{else} \end{cases}$$

Toute fonction numérique f, sous réserve de bonnes propriétés de semi-continuité supérieure, peut être décomposée en ensembles de niveau et reconstituée à partir de ses ensembles de niveau (voir figure 2.43) :

$$f = \bigvee_h h\ \chi_h(f)$$

Les opérateurs morphologiques standards lorsqu'ils sont associés à des éléments structurants plans commutent avec les seuils. En d'autres termes, une dilatation par exemple de taille λ satisfait

$$\delta_\lambda(f) = \bigvee_h h\ \delta_\lambda[\chi_h(f)]$$

Ici, δ_λ désigne la dilatation par un disque euclidien de rayon λ.

La dilatation morphologique visqueuse par contre agit différemment sur les points de forte et de faible luminance [132] :

$$\delta^v_{R_0}(f) = \bigvee_h h\ \delta_{r(h)}[\chi_h(f)] \tag{2.10}$$

$\delta^v_{R_0}$ est la dilatation visqueuse de taille R_0. Cela signifie qu'*au maximum*, les points de l'image sont dilatés par un élément structurant de taille R_0. Théoriquement, $r(h)$ peut être n'importe quelle fonction

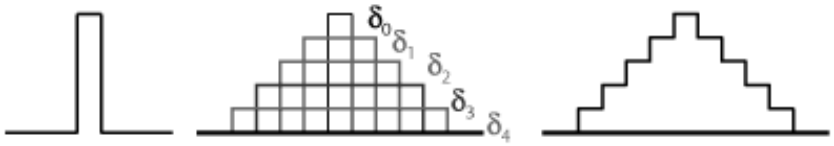


FIG. 2.44: Comment les dilatations visqueuses agissent sur les fonctions à valeurs scalaires. Ici les faibles luminances sont fortement dilatées et les plus hauts niveaux de luminance sont laissés inchangés. Sur ce shéma, la taille de l'élément structurant est une fonction linéaire de la luminance : $r(h) = N - h$

croissante ou décroissante où à variations variables, de la luminance h. En pratique, jusqu'ici, seules les fonctions croissantes ou décroissantes et de plus linéaires ont été expérimentées [132]. Par exemple, si $r(h) = R_0 \frac{M-h}{M}$ où M désigne la valeur maximale de la fonction $f : M = sup_{x \in E}\{f(x)\}$, alors les points de faible luminance sont fortement dilatés ($r(0) = R_0$) et les points de plus forte luminance sont laissés inchangés ($r(M) = 0$). Inversement, pour laisser inchangés les points de faible luminance et dilater fortement les points de forte luminance, il suffit de prendre la variation inverse : $r(h) = R_0 \frac{h}{M}$. Le choix de la fonction $r(h)$ est une question qui dépend fortement de l'image et du contexte. Ce point sera discuté au moment des expérimentations.

Maintenant, les érosions visqueuses sont définies de manière duale :

$$\epsilon^v_{R_0}(f) = \bigvee_h h \ \epsilon_{r(h)}[\chi_h(f)]$$

Si la loi $r(h)$ est défini une fois pour toutes, alors le couple $(\delta^v_{R_0}, \epsilon^v_{R_0})$ définit une adjonction morphologique, mais, attention, les deux opérateurs ne sont pas duaux [133]. Pour obtenir une dualité dilatation-érosion, il faut inverser les lois de variation du rayon $r(h)$ mais dans ce cas, on perd l'adjonction [133]. Quoiqu'il en soit, les ouvertures et fermetures visqueuses peuvent être définies par le couple formant une adjonction, c'est-à-dire par composition des dilations et érosions selon les cascades [133] :

$$\gamma^v_{R_0} = \delta^v_{R_0} \circ \epsilon^v_{R_0} \quad \text{et} \quad \varphi^v_{R_0} = \epsilon^v_{R_0} \circ \delta^v_{R_0}$$

ou encore et c'est équivalent, en revenant à une définition semblable à celle donnée dans le cas des dilatations et érosions visqueuses :

$$\gamma^v_{R_0}(f) = \bigvee_h h \ \gamma_{r(h)}[\chi_h(f)] \quad \text{and} \quad \varphi^v_{R_0}(f) = \bigvee_h h \ \varphi_{r(h)}[\chi_h(f)] \tag{2.11}$$

(i) (ii) (iii)

FIG. 2.45: Effet de l'ouverture morphologique visqueuse sur une image à teintes de gris. (i) est l'image couleur originale (ii) est son image de rang (calculée à partir de la distance de Mahalanobis voir section 2.1.2). (iii) est le résultat de l'ouverture visqueuse de l'image (ii) avec $r(h) = R_0 \left[\frac{M-h}{M}\right]$, $R_0 = 10$ et $M = sup_{p\in E}\{f(p)\}$.

Comme c'était le cas pour les dilatations et érosions visqueuses, si $r(h)$ est une fonction décroissante de l'intensité h, alors les structures de faible luminance sont fortement ouvertes ou fermées tandis que les structures de forte luminance sont laissées inchangées.

Bien sûr, il est possible de décliner ainsi tous les filtres morphologiques classiques, pour définir par exemple des filtres visqueux surfaciques, des filtres visqueux par reconstruction ou encore toute la gamme des filtres alternés séquentiels ; il suffit de remplacer la dilatation par l'opérateur choisi dans la définition 2.10. Par exemple, l'ouverture par reconstruction visqueuse se définit comme suit :

$$\gamma_{R_0}^{v,rec}(f) = \bigvee_h h\ \gamma_{r(h)}^{rec}[\chi_h(f)] \quad \text{and} \quad \varphi_{R_0}^{v,rec}(f) = \bigvee_h h\ \varphi_{r(h)}^{rec}[\chi_h(f)]$$

Les effets des opérateurs visqueux sur les images à teintes de gris sont illustrés sur les figures 2.45 et 2.46. Ici, l'image à teintes de gris est l'image de rang associée à l'image couleur présentée. Les pâquerettes ont de hautes valeurs dans l'image de rang ; elles sont préservées par l'ouverture visqueuse tandis que les autres structures (les herbes) sont gommées et le gommage est d'autant plus important que les structures sont de faible similarité.

(i) (ii)

FIG. 2.46: Effet de l'ouverture morphologique visqueuse par reconstruction sur une image à teintes de gris. (i) est l'image originale (même image que dans la figure 2.45). (ii) est le résultat de l'ouverture visqueuse par reconstruction de l'image (i) avec les même paramètre que dans la figure 2.45. Les pâquerettes (qui correspondent à la référence statistique) sont préservées tandis que les herbes sont gommées.

2.4.2 Extension au cas des images multivaluées avec prise en compte d'a priori statistiques

Nous arrivons maintenant au coeur de la construction des opérateurs adaptatifs. Sur les images de rang, les luminances s'interprètent comme des niveaux de similarité, par conséquent les opérateurs adaptatifs pour les images multivalués peuvent être définis très simplement par le biais des opérateurs visqueux appliqués aux images de rang. La transformation de rang étant bijective, l'extension de ces opérateurs définis dans le cas des images à valeurs scalaires au cas des images multivaluées est immédiate. Si $\Delta^v_{R_0}$ et $\Sigma^v_{R_0}$ représentent les dilatations et érosions visqueuses sur les images multivaluées, elles se déduisent de leur équivalent scalaire selon les relations 2.12 et 2.13.

$$\Delta^v_{R_0}(F) = T^{-1}_{rank} \circ \delta^v_{R_0} \circ T_{rank}(F) = T^{-1}_{rank}[\bigvee_h h\delta_{r(h)}\chi_h(f)] \tag{2.12}$$

$$\Sigma^v_{R_0}(F) = T^{-1}_{rank} \circ \epsilon^v_{R_0} \circ T_{rank}(F) = T^{-1}_{rank}[\bigvee_h h\epsilon_{r(h)}\chi_h(f)] \tag{2.13}$$

Dans ces définitions, F est la fonction multivaluée et f la fonction à valeurs scalaires.

Bien évidemment, cette construction vaut également pour l'ensemble des transformations visqueuses (ouvertures, fermetures, filtres alternés connexes ou non).

Le shéma ci-dessous résume le principe de construction des opérateurs visqueux agissant sur les images multivaluées, sur l'exemple de la dilatation visqueuse.

$$F \xrightarrow{T_{rank}} f \begin{array}{ccccccc} \nearrow & & \chi_0(f) & \rightarrow & \delta_{r(0)}[\chi_1(f)] & \searrow & \\ \rightarrow & & \chi_h(f) & \rightarrow & \delta_{r(h)}[\chi_h(f)] & \rightarrow & \delta^v_{R_0}(f) \xrightarrow{T^{-1}_{rank}} \Delta^v_{R_0}(F) \\ \searrow & & \chi_{N-1}(f) & \rightarrow & \delta_{r(N-1)}[\chi_{N-1}(f)] & \nearrow & \end{array}$$

Décomposition en ensembles de niveau de l'image de rang

2.4.3 Le calcul effectif des transformations visqueuses

D'un point de vue calculatoire, le schéma précédent fondé sur une analyse seuil par seuil des images de rang aboutit à un algorithme très coûteux en temps de calcul. Les images de rang ordonnant totalement l'ensemble des vecteurs des images multivaluées, leur dynamique est extrêmement importante. Typiquement, sur les exemples montrés dans ce document, quelques 10 à 20000 rangs sont recensés. On peut imaginer la lenteur d'un processus calculatoire fondé sur une analyse seuil à seuil. Des méthodes de calcul par équations aux dérivées partielles viennent très récemment d'être proposées par Maragos et Vachier [79]. L'avantage majeur de la formulation par EDP est de conduire à des temps de calcul proportionnels à la taille de la transformation (R_0) et parfaitement indépendants de la dynamique de l'image. Sans entrer dans les détails, l'EDP des dilatations visqueuses est une variante de l'équation dévolution classique des dilatations, la vitesse d'évolution étant simplement non plus constante comme dans le cas des dilatations simples, mais fonction de la luminance en chaque point de l'espace. La formulation que nous allons présenter est due à Maragos et Vachier [79].

Dans la formulation par EDP, les dilatations et érosions d'une image f par un élément structurant convexe et plan tB sont décrites comme des opérateurs multi-échelles, t représentant l'échelle.

$$\delta(x, y, t) = (f \oplus tB)(x, y), \quad \epsilon(x, y, t) = (f \ominus tB)(x, y)$$

avec $\delta(x, y, 0) = \epsilon(x, y, 0) = f(x, y)$.

$\oplus$ et $\ominus$ sont respectivement l'addition et la soustraction de Minkowski.

Dans le cas où l'élément structurant B est un disque euclidien unitaire ($B_p = \{(x, y) : \|(x, y)\| \leq 1\}$, les EDP qui généralisent les dilatations et érosions planes de f par B sont

$$\partial_t \delta = \|\nabla \delta\|, \quad \partial_t \epsilon = -\|\nabla \epsilon\|$$

Pour les signaux 1D, les disques unitaires deviennent des intervalles $[-1, 1]$ et la norme $\|\nabla u\|$ devient simplement $|\partial_x u|$.

Maintenant, les EDPs qui gouvernent les dilatations et érosions visqueuses (toujours dans le cas où l'élément structurant est un disque euclidien) sont obtenues en modulant les vitesses d'évolution selon les valeurs de la fonction en chaque point. Par exemple, pour une loi d'évolution $r(h)$ linéaire $r(h) = M - h$, où M est la valeur maximale de f, on obtient l'EDP 2.14 qui correspond à un cas particulier des EDPs de type Hamilton-Jacobi.

$$\begin{aligned} \partial_t \delta^v(x, y, t) &= (f_{max} - \delta^v(x, y, t))\|\nabla \delta^v\| \\ \partial_t \epsilon^v(x, y, t) &= -(f_{max} - \epsilon^v(x, y, t))\|\nabla \epsilon^v\| \end{aligned} \tag{2.14}$$

Le schéma de simulation numérique est décrit en détail dans [79].

2.4.4 Expérimentations

Dans toutes les expérimentations qui sont présentées ici, le but du filtrage est une simplification des images qui vise à préserver intégralement les régions définies par le modèle statistique choisi comme référence. Les structures d'intérêt correspondent donc aux points de forte luminance dans les images de rang. Par conséquent, les filtres visqueux sont définis de manière telle que les transformations les plus importantes affectent les points de faible luminance. La taille $r(h)$ de la transformation du seuil de niveau h est donc une fonction décroissante de h. Les fonctions suivantes ont été expérimentées : voir figure 2.47.

$$r(h) = R\left[1 - \frac{h}{N-1}\right] \quad \text{or} \quad r(h) = R\left[1 - \frac{h}{N-1}\right]^{\frac{5}{2}}\left[1 + \frac{h}{N-1}\right]^{\frac{5}{2}}$$

Dans le premier cas, $r(h)$ est simplement une fonction linéaire de h. Dans le second cas, $r(h)$ est une approximation de la sigmoïde. La seconde fonction est plus appropriée à nos expérimentations, car elle

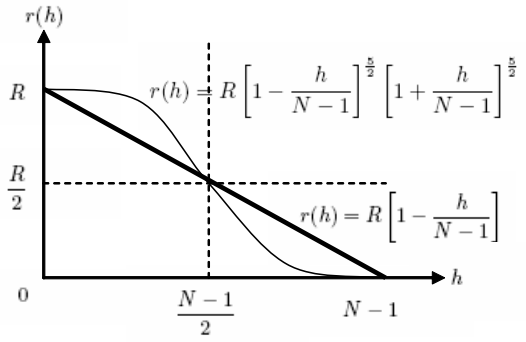


FIG. 2.47: Deux exemples de relation de dépendence entre l'activité du filtre $r(h)$ et le rang (ou similarité) h.

permet de filtrer intensivement les points de faible similarité tout en laissant une certaine latitude quant à leur rang ; de même, la sigmoïde permet de ne pas filtrer les points de forte similarité tout en laissant une certaine marge quant à la valeur de leur rang. Entre les deux extrêmes, l'évolution est quasi-linéaire.

Plusieurs filtres visqueux ont été testés sur différents types d'images couleur : voir les figures 2.48 à 2.54. Ces résultats peuvent être comparés à ceux obtenus via les filtres non visqueux (voir les exemples de la section 2.3.1). A titre d'exemple, en figure 2.50, les images obtenues par des filtres alternés séquentiels par reconstruction standard et visqueuse sont comparées. La fleur qui correspond à la référence statistique n'est pas altérée par la simplification visqueuse alors que l'arrière-plan est notablement simplifié. Les gradients calculés sur ces deux images confirment cette observation. Certaines régions ont conservé tous leurs détails tandis que d'autres sont devenues de larges zones de couleur uniforme. Une autre illustration très parlante est la figure 2.51 : les pâquerettes, quelle que soit leur taille, ne sont pas altérées par le processus de simplification comme c'était le cas dans les processus de filtrage non visqueux (voir figure 2.29).

Les résultats obtenus via les deux fonctions d'évolution du rayon (courbes de la figure 2.47) sont comparés dans les figures 2.48 et 2.52. Dans les deux cas, l'évolution lente de type sigmoïde donne de meilleurs résultats visuels.

Enfin, en figure 2.54, le résultat obtenu par nos filtres adaptatifs est testé de deux manières. Tout d'abord, la sensibilité à l'échantillon est éprouvée ; ensuite, le résultat obtenu est comparé à celui que fournirait un simple seuillage de l'image de rang (avec retour dans l'espace des valeurs multivaluées). On note une certaine sensibilité aux variations d'échantillon statistique. Quant au seuillage, il fournit une image très simple, mais de moins bonne qualité visuelle que le filtre adaptatif. Il faut ajouter à cette

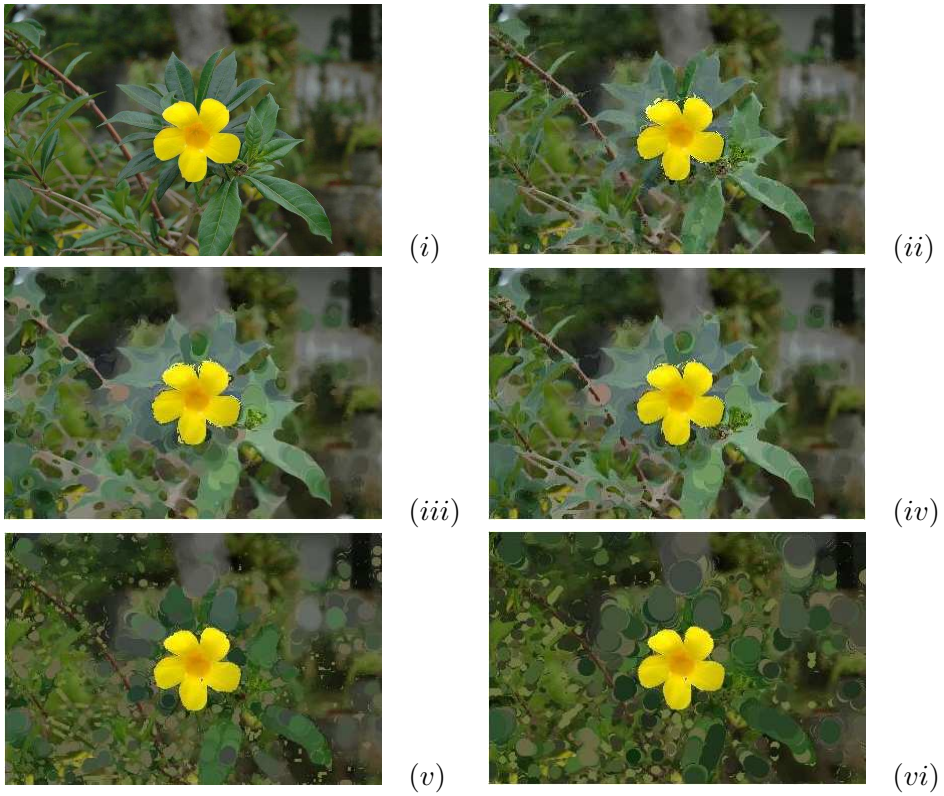


FIG. 2.48: Définition de filtres morphologiques adaptatifs : ici des ouvertures (cas (ii), (iii) et (iv)) et fermetures visqueuses (cas (v) et (vi)). (i) is the original image. L'échantillon de référence est celui de la figure 2.27. L'élement structurant est un disque et $R_0 = 5$ (ii) puis 10 (iii) (iv) et (v) et 20 (vi). Dans les cas (ii) et (iii), l'activité de filtrage $r(h)$ est une fonction linéaire du rang h : $r(h) = R\left[1 - \frac{h}{N-1}\right]$. Dans les cas (iv), (v) et (vi), l'évolution suit une loi en forme de sigmoïde : $r(h) = R\left[1 - \frac{h}{N-1}\right]^{\frac{5}{2}}\left[1 + \frac{h}{N-1}\right]^{\frac{5}{2}}$.

remarque que le choix du seuil reste toujours un point sensible (ici, il a été effectué manuellement jusqu'à obtenir l'image désirée). Dans le cas des opérateurs visqueux, il faut choisir une taille maximale de filtrage. Ce choix est, dans le contexte de simplification décrit, en rapport avec le niveau de simplification désiré non avec la taille supposée des structures à préserver. On peut donc prétendre que ce paramètre est moins sensible que le choix d'un seuil.

2.5 Conclusion

Comme en témoignent les opérateurs adaptatifs que nous venons de présenter, la structure choisie pour l'analyse des images multivaluées ouvre des perspectives intéressantes. Plusieurs questions demandent cependant à être étudiées. Tout d'abord, même si nombre d'applications tombent dans ce contexte, on ne dispose pas toujours d'un échantillon statistique des objets d'intérêt dans l'image. Dans

FIG. 2.49: Filtres visqueux par reconstruction. L'élément structurant (au niveau d'activité maximale) est un disque de rayon $R = 20$. La loi d'évolution du rayon est : $r(h) = R\left[1 - \frac{h}{N-1}\right]^{\frac{5}{2}}\left[1 + \frac{h}{N-1}\right]^{\frac{5}{2}}$. (i) est le résultat d'une ouverture visqueuse par reconstruction. (ii) est le résultat d'une fermeture visqueuse par reconstruction. (iii) est le résultat d'un FAS par reconstruction.

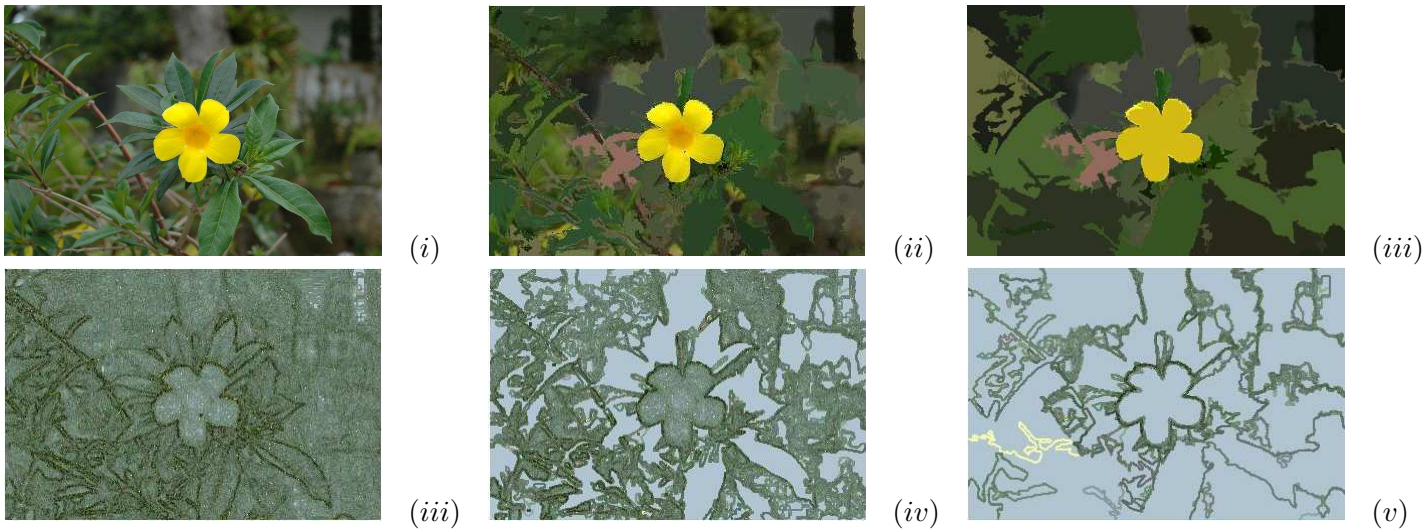


FIG. 2.50: (i) et (iv) Image originale et son gradient morphologique. (ii) et (iv) Image simplifiée par un FAS visqueux de taille $R = 20$ par reconstruction et son gradient morphologique. (iii) et (v) Image simplifiée par un FAS standard de taille 20 par reconstruction et son gradient morphologique.

ce cas, comment choisir la référence ? Est-il possible d'automatiser ce choix ? Le second point porte sur le choix de la distance. Dans le cas où les structures ne satisfont pas l'hypothèse gaussienne, une autre distance que la distance de Mahalanobis peut se révéler mieux adaptée. Notamment, la distance de Mahalanobis est une sorte de distance euclidienne pondérée ; comment cette pondération peut-elle être modifiée afin de mieux convenir à des échantillons très différents (qui peuvent être bruités ou non, simples ou complexes). Troisièmement, l'utilisation d'une image de rang à valeurs entières à la place d'une carte des similarités à valeurs réelles présente des avantages de mise en oeuvre algorithmique, mais a également un défaut majeur : en passant à l'image de rang, la dynamique de la carte des similarités est perdue ce qui peut introduire de mauvais comportements lors du calcul des filtres adaptatifs. Ce point ne pose pas de difficulté théoriques importantes mais demande tout de même à être analysé.

FIG. 2.51: (i) Image originale, (ii) Résultat d'une ouverture visqueuse par reconstruction, (iii) Résultat d'un FAS visqueux par reconstruction. La référence est une pâquerette (voir figure 2.7). L'activité de filtrage est une fonction décroissante de la similarité.

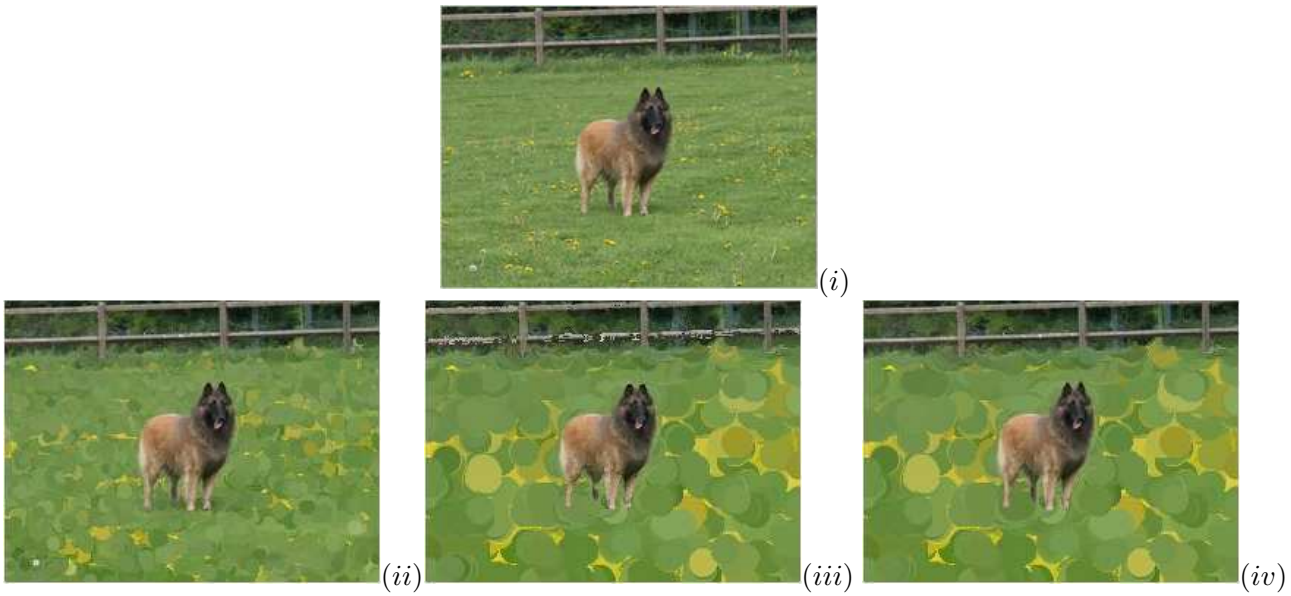


FIG. 2.52: (i) est l'image originale. L'échantillon de référence est pris dans le chien (voir figure 2.40). (ii) et (iii) sont les résultats d'ouvertures visqueuses de taille 5 et 10, l'activité $r(h)$ est une fonction linéaire de h : $r(h) = R\left[1 - \frac{h}{N-1}\right]$). (iv) est le résultat obtenu par ouverture visqueuse de taille 10, l'activité étant cette fois une fonction à variations plus lentes pour les très fortes et très faibles similarités : $r(h) = R\left[1 - \frac{h}{N-1}\right]^{\frac{5}{2}}\left[1 + \frac{h}{N-1}\right]^{\frac{5}{2}}$

Enfin, le calcul rapide des filtres adaptatifs est un point crucial pour que ces opérateurs puissent être effectivement exploités. Dans la configuration actuelle, leur calcul est absolument prohibitif. Les solutions fondées sur les schémas numériques des EDPs que nous avons présentés sont en cours d'évaluation [79].

FIG. 2.53: Cette fois-ci, c'est l'herbe qui sert de référence et une dilatation visqueuse a été appliquée.

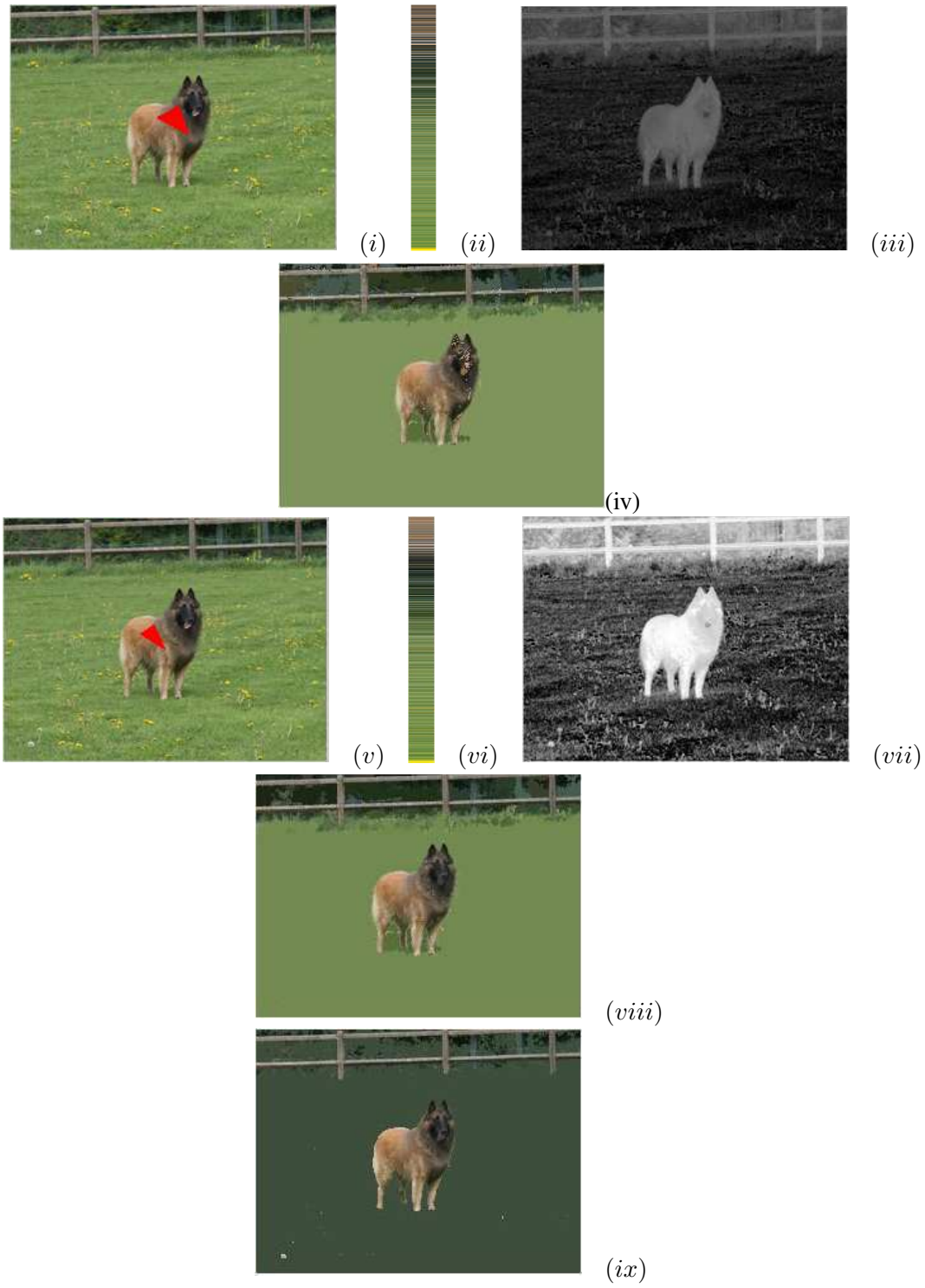


FIG. 2.54: Tests avec FAS visqueux par reconstruction : (iv) et $(viii)$. L'élément structurant est un disque euclidien de taille (maximale) 20. L'activité suit une loi de type sigmoïde. Les résultats ont été obtenus pour deux échantillons de référence différents (voir (i) et (v)). (ix) a été obtenue par simple seuillage de l'image de rang.

Chapitre 3

Développements logiciels pour les applications cliniques

Dans cette partie nous présentons les principaux développements logiciels à destination des médecins des Hôpitaux Universitaires de Genève réalisés pendant la thèse. Ces travaux sont en rapport direct avec les recherches présentées dans les deux premiers chapitres de la thèse.

Premièrement, des outils d'aide au diagnostic ont été développés pour le dépistage du mélanome de la peau et l'analyse des IRMs cardiaques. Deuxièmement, un ensemble de fonctionnalisés de traitement des images ont été incorporés au logiciel d'imagerie médicale OsiriX développé par les ingénieurs des HUG. Enfin, dans le souci de transmettre les algorithmes proposés dans cette thèse, une librairie libre (open source) *vmorph* a été élaborée qui comprend toutes les fonctionnalités proposées dans cette thèse, de la formation des cartes de distance au calcul des filtres morphologiques visqueux sur images vectorielles.

Concernant les applications médicales, nous ne prétendons pas ici présenter des solutions validées des problèmes d'imagerie présentés. Les algorithmes proposés visent plutôt à ouvrir le champ des investigations dans le domaine. Sans validation clinique, leur valeur est strictement expérimentale.

Enfin, des travaux réalisés spécifiquement avec l'équipe médicale des HUG sont présentés en annexe sous la forme des publications dont ils ont fait l'objet.

3.1 Dépistage du mélanome de la peau : le projet Naeview

Le projet Naeview dirigé par Francis Klumb de l'Hôpital Universitaire de Genève (Unité d'Imagerie Numérique) vise à élaborer une solution logicielle d'aide au dépistage du mélanome par une analyse photographique de l'ensemble du corps.

Les motivations du projet sont dues à une croissance du nombre de cancers de la peau ces dernières années. Par exemple, sur le territoire helvétique, nous constatons une évolution de 1500 nouveaux cas par an. La prévention par le contrôle visuel des grains de beauté est essentielle. Les principaux projets ont pour but l'aide au diagnostic, afin de détecter la malignité d'un grain de beauté [44, 115], mais très peu à notre connaissance font de véritables cartographies du corps. Ce problème est d'autant plus important lorsque le nombre de grains de beauté est important. L'examen clinique est long, et il est difficile de mesurer l'évolution entre deux visites. La quantité des données à analyser est gigantesque et représente une charge très importante pour le clinicien. L'aide au diagnostic est sur ce thème (comme dans de nombreux domaines médicaux où des dépistages de masse sont entrepris) devenue urgente.

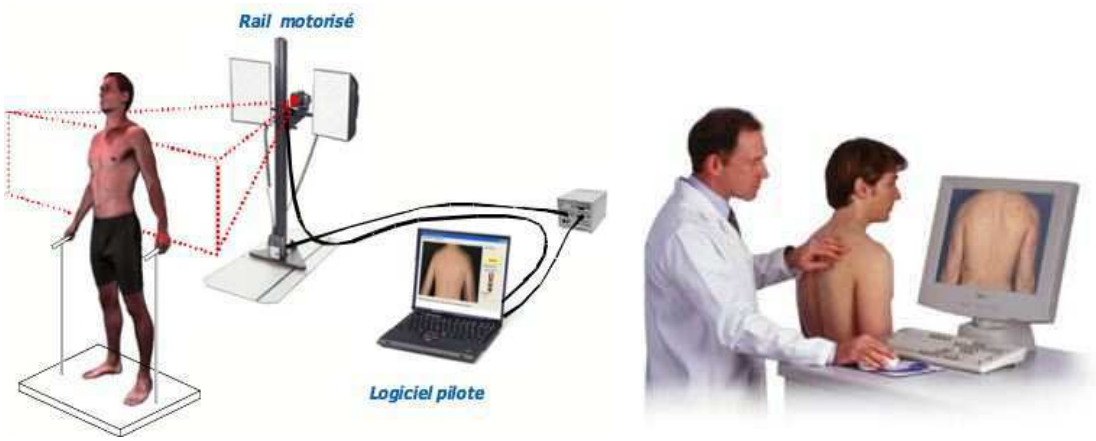


FIG. 3.1: Robot d'acquisition et logiciel de pilotage.

Le travail de l'équipe de Naeview a d'abord consisté à élaborer une stratégie d'acquisition des données. Un robot extrêmement évolué, couplé avec un logiciel, pilote un appareil photo haute résolution (figure 3.1). Une *webcam* est même embarquée afin d'effectuer un premier recalage visuel entre deux acquisitions, directement depuis le logiciel Naeview. Le dispositif est maintenant opérationnel et utilisé en routine clinique. Le principe est d'effectuer différentes acquisitions (figure 3.2) pour obtenir une cartographie des grains de beauté.

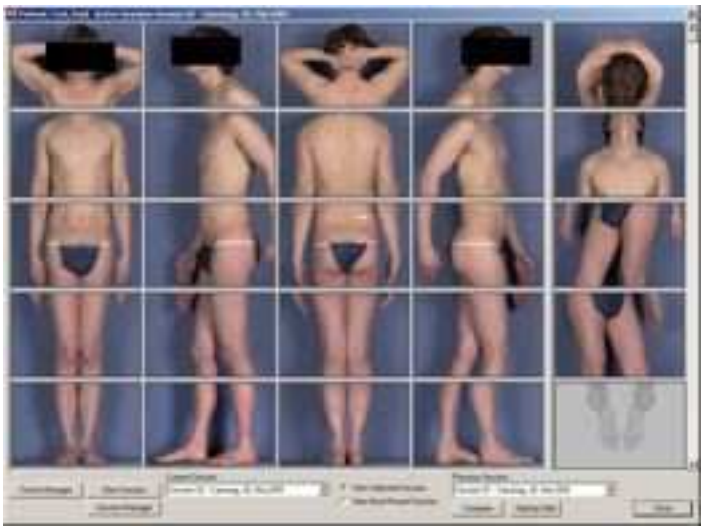


FIG. 3.2: Le patient prend différentes positions afin d'obtenir une cartographie complète du corps.

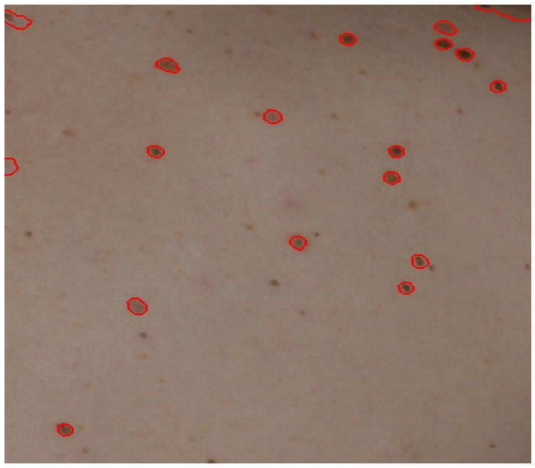
Une hypothèse est prise en accord avec les cliniciens pour limiter la recherche aux grains de beauté supérieurs à $5mm$. L'accent est mis sur la comparaison entre les examens cliniques. En effet, les grains de beauté malins croissent très rapidement. Ceci justifie l'importance du suivi médical.

La complexité de la problématique de segmentation des grains de beauté a plusieurs sources liées aux différences de couleur de peau, à la variabilité de la pilosité, aux ombrages (creux du dos, omoplate), aux variations de teinte entre les visites successives... Actuellement, ce projet est dans une phase d'évaluation de différents algorithmes de segmentation, des premiers résultats sont disponibles [94, 93].

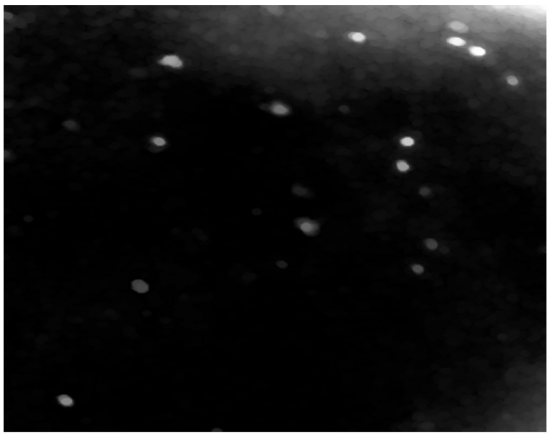
L'extraction des grains de beauté est donc un problème de détection et de segmentation sur images couleur avec a priori statistique sur peau saine (évidemment, l'échantillon est redéfini pour chaque patient). Le classifieur bayésien (via la distance de Mahalanobis) une fois initialisé avec une zone de peau saine génère une image de rang où chaque grain de beauté est rehaussé. Ici, un seul échantillon de peau est nécessaire et une segmentation par calcul direct de la *LPE* sur la carte des distances est possible. La figure 3.3 résume les grandes étapes de l'algorithme. Les figures 3.4 à 3.11 présente la routine mise en oeuvre en situation clinique et les résultats obtenus sur plusieurs échantillons de peau.

C'est la présence d'ombres sur les images qui semble poser le plus de difficultés. Une solution consiste à appréhender le problème par bandes successives afin d'éviter les variations d'intensité lumineuse entre les zones. Les bords du corps sont également un challenge. Cependant, les différentes acquisitions doivent permettre de ne pas avoir à analyser ces zones.

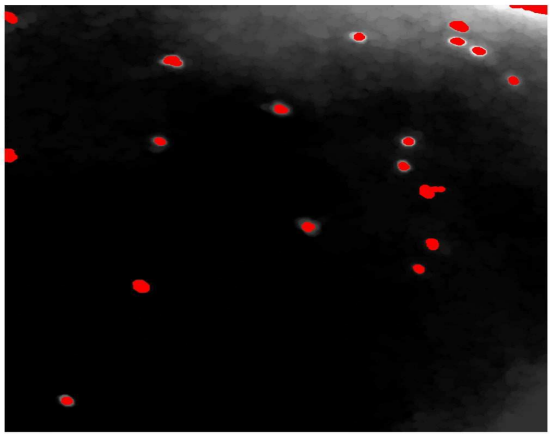
1)



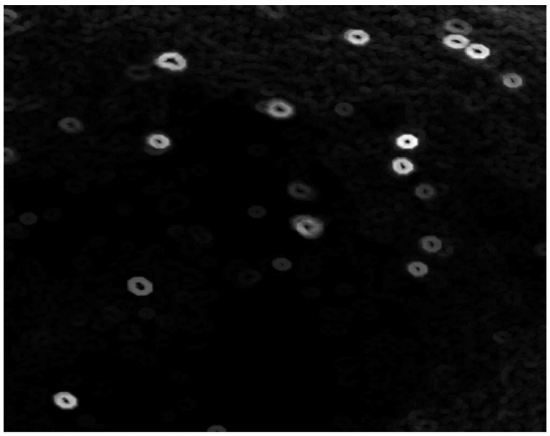
6)



2)



5)



3)

4)

FIG. 3.3: Les différentes étapes permettant de segmenter les images de peau : 1) l'image originale avec la zone servant de référence. 2) Ouverture morphologique de la carte de similarité (carte des distances de Mahalanobis) avec un disque de rayon 5 pixels. 3) Gradient épais (disque de rayon 5 pixels) afin d'accentuer les bords. 4) Calcul de la dynamique des maxima [56, 57, 134]. 5) Maxima régionaux de dynamique supérieure à 30 (ils marquent les grains de beauté). 6) Résultat de la *LPE* sur l'image gradient avec les marqueurs pré-cités (clic sur peau saine et maxima de forte dynamique)

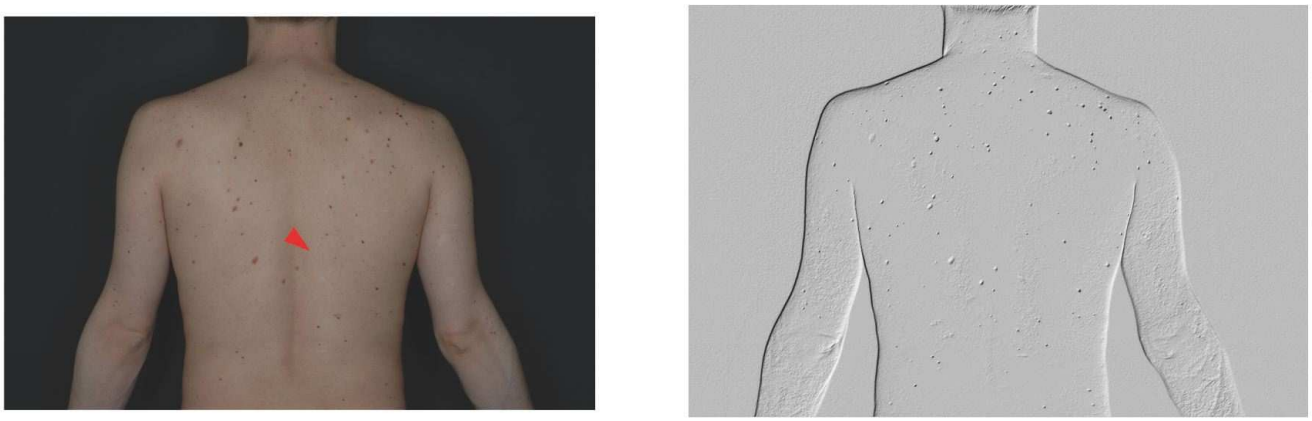


FIG. 3.4: À gauche, l'image originale sur laquelle nous prenons comme référence une partie de peau sans grain de beauté. À droite, nous représentons en relief l'image dans $\mathcal{L}^d$.

FIG. 3.5: Focus sur zones sélectionnées.

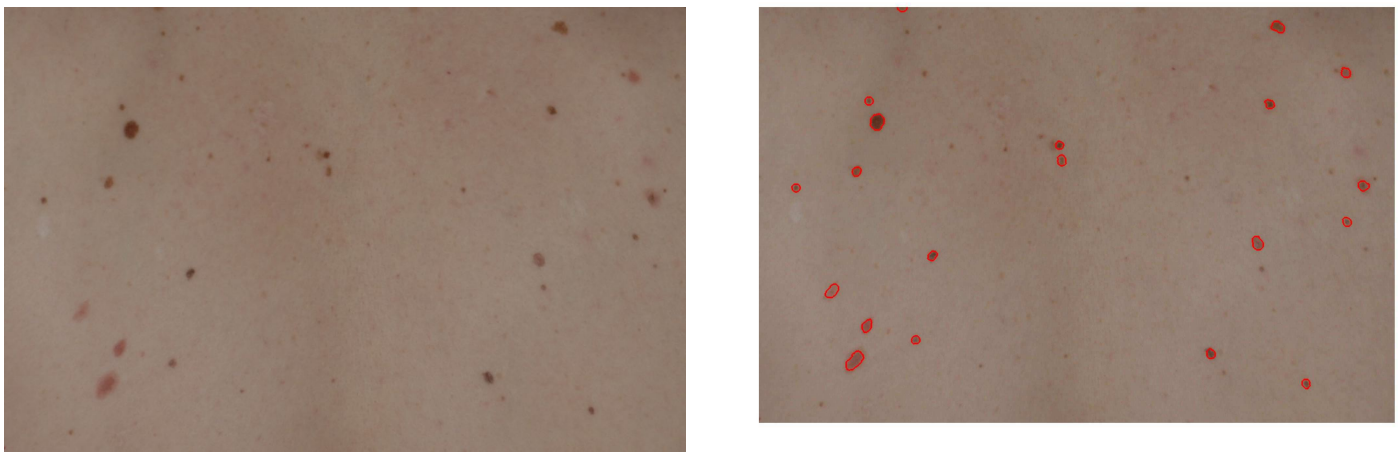


FIG. 3.6: Résultats obtenus pour la zone gauche de la figure 3.5.

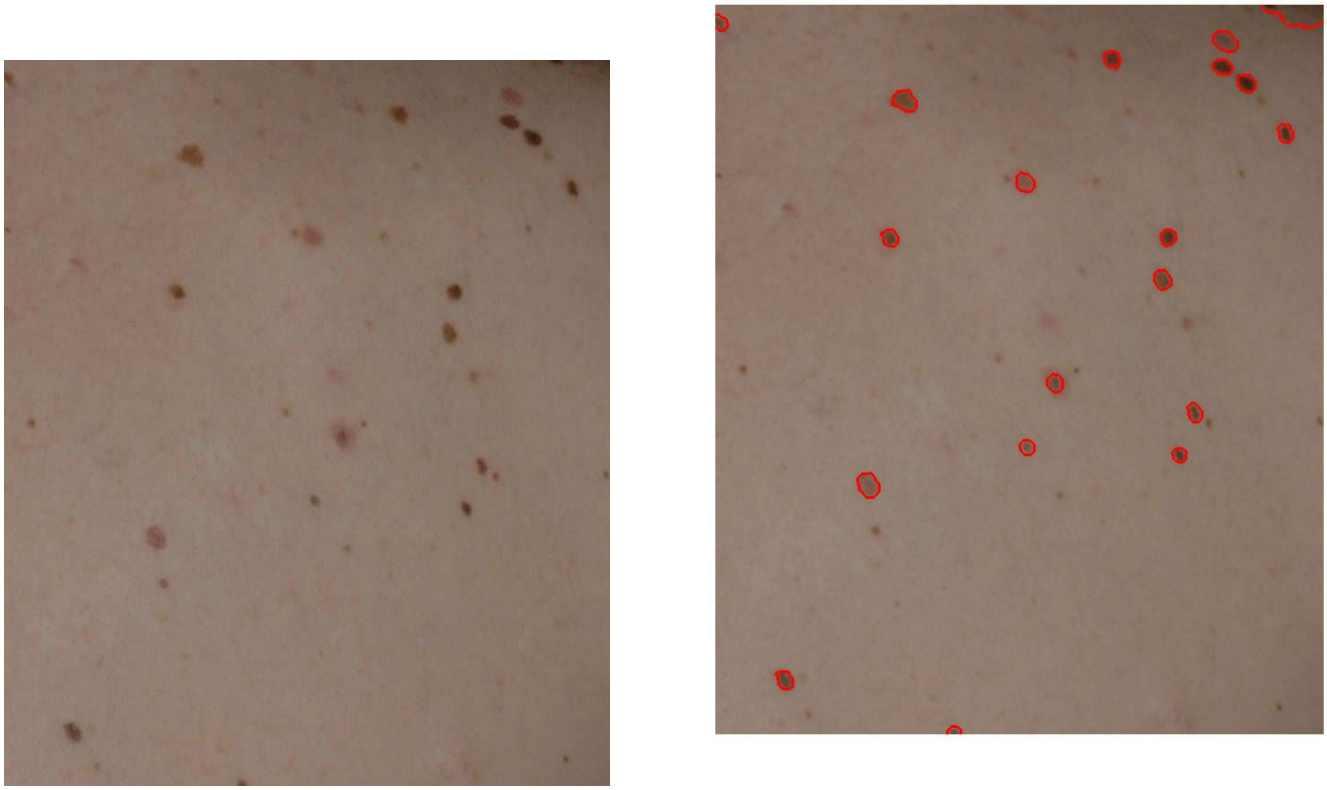


FIG. 3.7: Résultats obtenus pour la zone droite de la figure 3.5.

FIG. 3.8: Focus sur 2 nouvelles zones.

Les premiers résultats sont encourageants. La méthode est simple et robuste. Les routines de développement sont sous forme de filtres ITK[1], l'équipe de Naeview peut par conséquent utiliser immédiatement les algorithmes. Une validation clinique est maintenant nécessaire pour confronter les premiers

[1]ITK (Insight Toolkit), projet open source offrant une librairie de traitement des images. La librairie est spécialisée pour l'imagerie médicale, elle est disponible à l'adresse `http://www.itk.org`

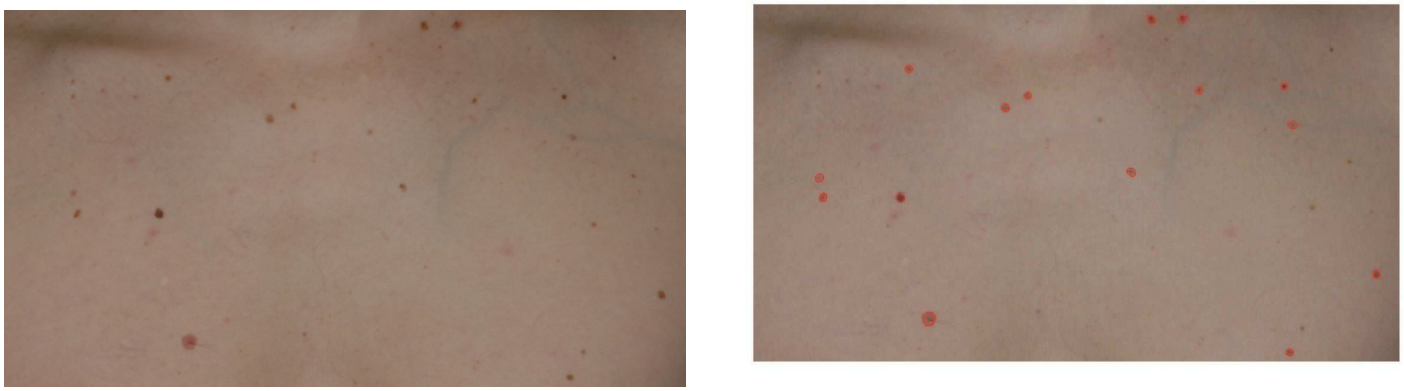


FIG. 3.9: Résultats obtenus pour la zone gauche de la figure 3.8

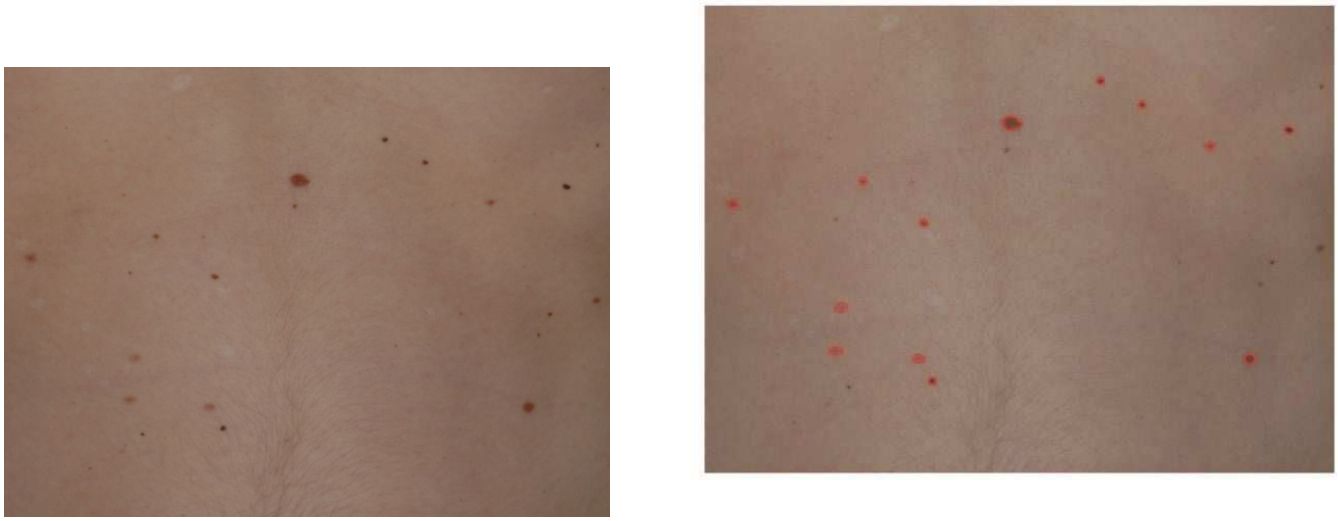


FIG. 3.10: Résultats obtenus pour la zone droite de la figure 3.8

résultats obtenus à un très grand nombre d'images, condition pour qu'un logiciel utilisable en milieu clinique puisse voir le jour. Nous espérons en tout cas que nos travaux avec l'équipe de Naeview se poursuivront.

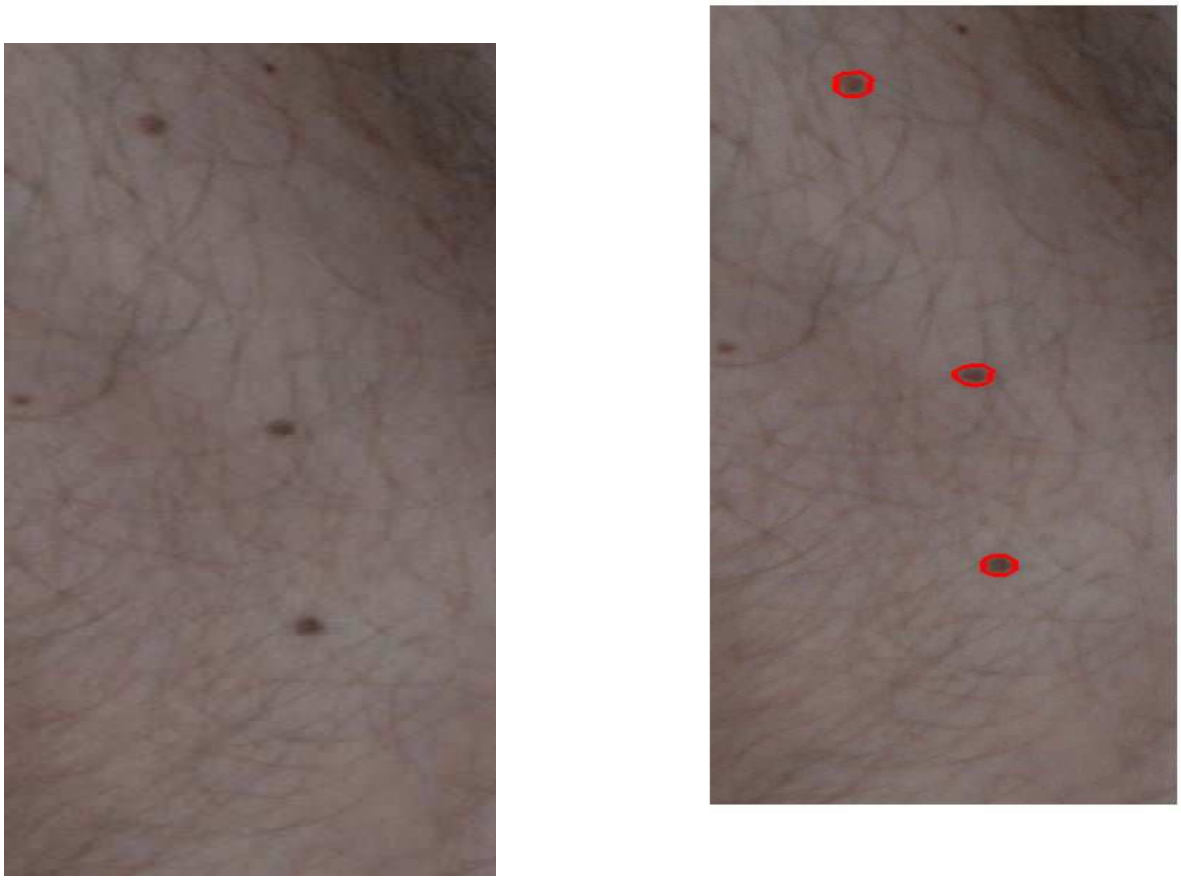


FIG. 3.11: Résultats sur une zone à forte pilosité.

3.2 Segmentation d'IRM cardiaques

L'IRM est devenue une méthode d'imagerie incontournable en raison de la qualité des images fournies ainsi que de son aspect non invasif.

3.2.1 L'IRM cardiaque

Le principe d'obtention de ces images repose sur la résonance magnétique nucléaire (RMN) des protons de l'eau. Le phénomène de RMN est dû aux découvertes indépendantes de Bloch [20] et Purcell [105] en 1946 dont ils obtiendront le prix Nobel en 1956. Le principe repose sur la rotation des électrons de l'atome qui induit un moment angulaire intrinsèque ou spin. À chaque moment angulaire est associé un moment magnétique en raison de la charge électrique (figure 3.12). Dans le cas d'atomes ayant un

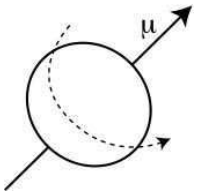


FIG. 3.12: Représentation du moment magnétique associé à un atome. L'atome se comporte comme un petit aimant lorsqu'il est plongé dans le champ magnétique de l'IRM.

nombre impair de protons, l'équilibre magnétique est rompu, permettant ainsi d'obtenir un noyau se comportant comme un petit aimant. Sans champ électromagnétique externe, les protons sont orientés aléatoirement produisant localement un champ magnétique nul (figure 3.13).

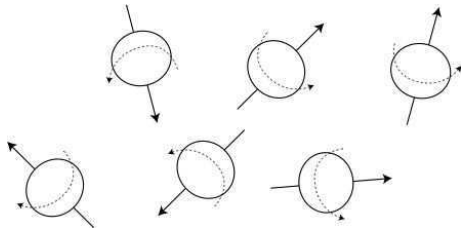


FIG. 3.13: Sans champ magnétique externe, les protons se positionnent aléatoirement, le champ magnétique local résultant de cette organisation est nul.

Cependant, si un champ externe est activé les protons s'alignent sur ce champ et se mettent à tourner à une fréquence w_0, dite fréquence de Larmor. On constate deux configurations fonction des deux spins possibles du proton (figure 3.14). Ces configurations ne sont pas équivalentes énergétiquement, la

position parallèle nécessite moins d'énergie entraînant toujours plus de protons dans ce sens. C'est ce décalage énergétique entre les deux positions qui va permettre la détection du signal par l'IRM.

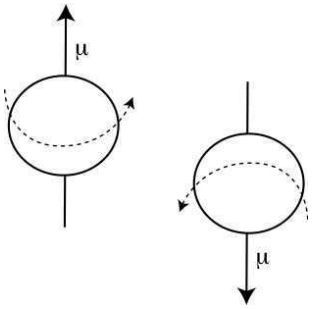


FIG. 3.14: Plongés dans un champ magnétique les protons s'alignent dans la direction du champ magnétique. En fonction de leur spin, les protons ont alors deux configurations possibles : parallèle (Sud Nord) et antiparallèle (Nord Sud).

Le principe consiste maintenant à émettre des impulsions électromagnétiques (radio fréquence ou RF) équivalentes à un autre champ magnétique perpendiculaire B_1. La RF émet des ondes d'une fréquence égale à w_0, fréquence de résonance permettant ainsi le basculement de l'axe de précession d'un angle α. L'angle varie en fonction du temps d'émission des ondes et peut aller jusqu'à l'inversion totale (180˚). En règle générale les angles les plus courants sont 90˚et 180˚.

L'arrêt du champ B_1 provoque un retour à la position initiale sous la forme d'une spirale en libérant l'énergie emmagasinée (figure 3.15). Un courant induit en forme de sinusoïde amortie est alors récupéré par des bobines. On peut décomposer cette phase de relaxation en deux parties. En effet, le vecteur de magnétisation M se décompose en une composante longitudinale (M_z) et une composante transversale M_{xy}. La composante longitudinale correspond à une phase de remagnétisation durant laquelle de nombreux échanges énergétiques se font avec l'environnement. Le temps de retour à l'équilibre est défini par le temps $T1$. La composante transversale quant à elle met en évidence le temps nécessaire, temps $T2$, de retour du déphasage des spins. Ces deux temps $T1$ et $T2$ reflètent les propriétés physiques des tissus graisseux, musculaires, ...

La construction de l'image consiste à utiliser la relation entre la fréquence de Larmor (w_0) et le champ magnétique B_0. L'idée consiste à créer un gradient magnétique linéaire suivant l'axe (OZ) afin de localiser la position de la coupe. En utilisant la même idée sur les axes OY puis OX, la position spatiale peut ensuite être décodée en utilisant les méthodes usuelles de traitement du signal.

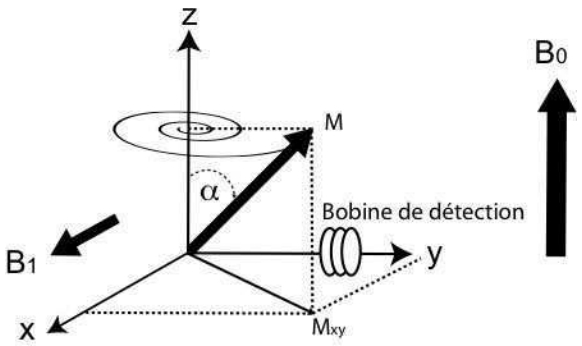


FIG. 3.15: Phase de relaxation : La RF à w_0 équivalente à un champ magnétique B_1 augmente l'énergie des protons et déplace l'axe de précession d'un angle α. L'arrêt de la RF permet un retour à l'état initial d'énergie suivant un parcours en spirale. Durant la phase de retour au calme, l'énergie précédemment emmagasinée est libérée ce qui provoque un courant induit récupéré par une bobine de l'IRM.

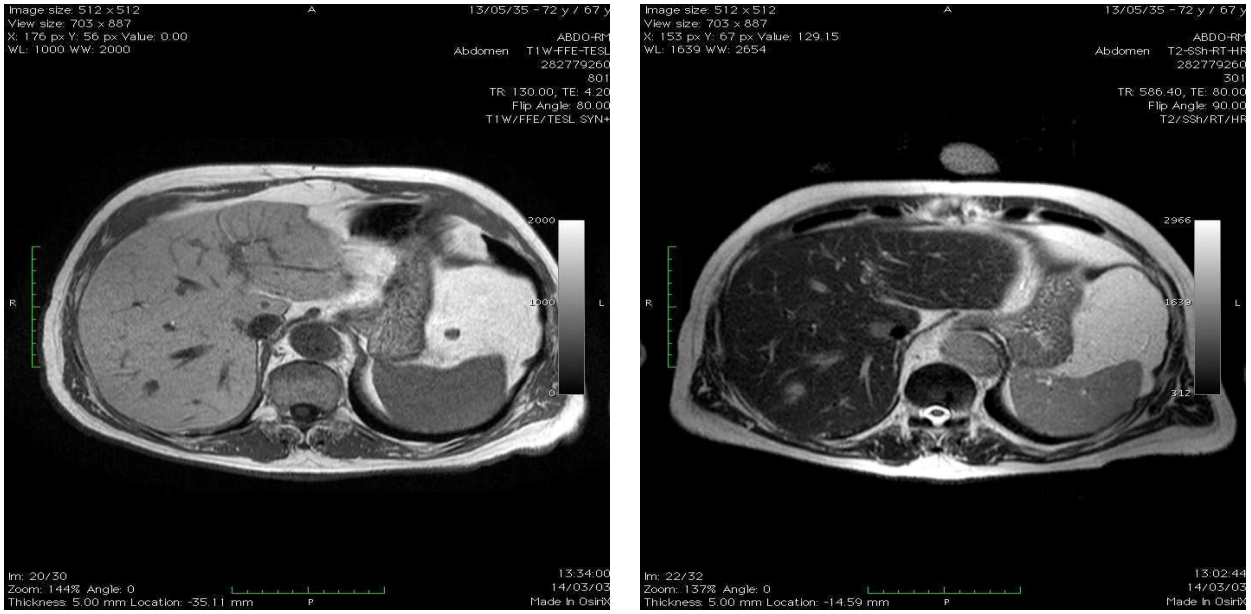


FIG. 3.16: Mise en évidence des différents contrastes entre les deux temps T1 et T2. IRM hépatique T1 à gauche et T2 à droite.

3.2.2 L'activité cardiaque

La paroi du coeur est composée de trois types de tissus : le péricarde, le myocarde et l'endocarde. Le péricarde constitue l'enveloppe externe du coeur, c'est un tissu fibreux dont le rôle est de protéger le coeur. En allant vers l'intérieur, nous trouvons le myocarde formé de cellules musculaires cardiaques. C'est le muscle responsable de la contraction et qui représente la principale masse du coeur. Pour finir, une fine couche forme l'endocarde avant de laisser place aux cavités.

Quatre cavités permettent d'effectuer la circulation sanguine 3.17 : deux oreillettes et deux ventricules. Le sang arrive par les oreillettes, l'oreillette gauche collectant le sang venant des poumons et la droite, le sang provenant du reste de l'organisme. Les ventricules permettent ensuite d'éjecter le sang comme de véritables petites pompes. Le ventricule droit éjecte le sang usé vers les poumons tandis que le ventricule gauche beaucoup plus gros et par conséquent plus puissant éjecte le sang oxygéné dans la circulation sanguine par l'aorte. Dans ce mécanisme, les muscles papillaires permettent de bloquer les valves des oreillettes au moment de l'éjection. Nous soulignons ce point puisque dans le cadre de la segmentation du myocarde les muscles papillaires augmentent la difficulté puisqu'ils apparaissent dans la cavité (figure 3.18).

Nous limitons notre étude à la segmentation du ventricule gauche, puisqu'il est le principal indicateur du bon fonctionnement de la fonction cardiaque. En effet, comme nous l'avons précisé les zones infarcies diminuent localement la contractibilité du coeur ce qui se traduit par une éjection plus faible et des troubles de la contraction. Afin de mesurer ces effets, nous proposons de travailler sur les coupes en courts axes (figure 3.18) puisqu'elles permettent d'effectuer cette quantification.

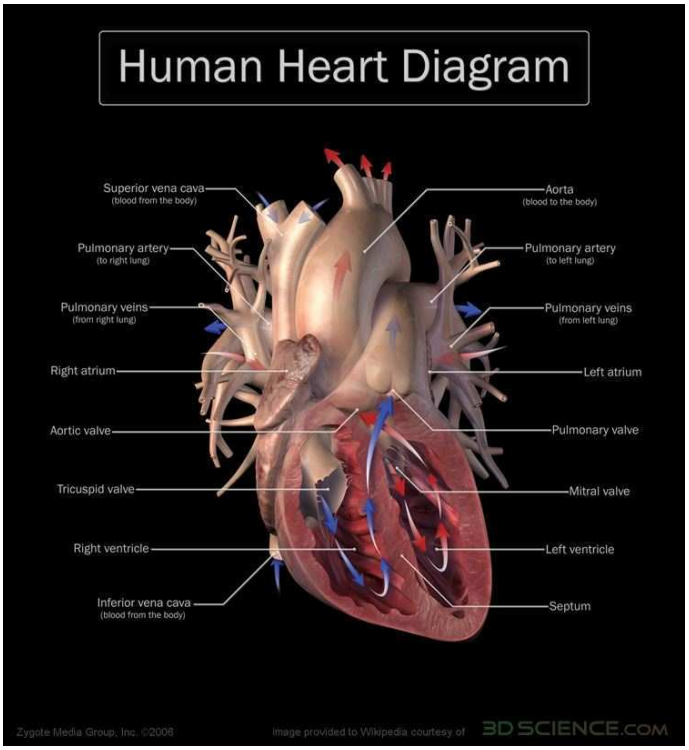


FIG. 3.17: Image faite par Zygote Media Group du site http ://www.3DScience.com

Il est important de distinguer deux phases dans la révolution cardiaque : la phase de diastole correspondant à la dilatation et la phase de systole pour la contraction. Durant ces périodes, les phases

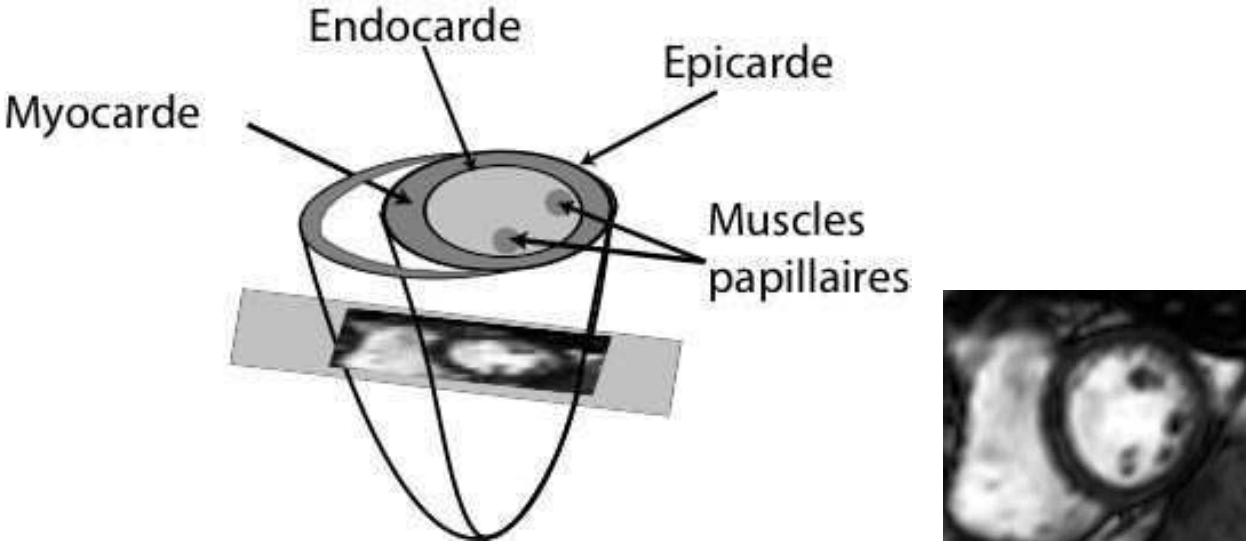


FIG. 3.18: À gauche, schéma du coeur où nous mettons en évidence le ventricule gauche avec ces principaux repères anatomiques et le plan de coupe axial. Les IRM en court axe sont obtenues à partir de ce plan et permettent de visualiser le myocarde. La segmentation va consister à délimiter l'endocarde et l'épicarde. Une difficulté supplémentaire est due aux muscles papillaires en raison de leur couleur semblable au myocarde.

extrêmes de contraction maximale (télésystole) ou bien de dilation maximale (télédiastole) sont primordiales puisqu'elles permettent de calculer la fraction d'éjection. Ce paramètre est calculé à partir du volume systolique (volume en télédiastole (VTD) moins le volume en télésystole VTS), normalisé par le VTD.

La fraction d'éjection permet d'obtenir un premier indicateur objectif sur l'état global de la fonction cardiaque. Son calcul implique uniquement la segmentation de la cavité du ventricule gauche en télédiastole et en télésystole puisque même si les variations de pression sont cinq fois plus faibles dans le ventricule droit, le volume de sang pompé dans les deux ventricules est identique. Cependant, cet indicateur global ne permet pas de mettre en évidence les problèmes locaux de contraction. Par conséquent, il est nécessaire d'offrir une segmentation totale du myocarde en segmentant l'endocarde et l'épicarde.

3.3 La segmentation du myocarde : focus sur deux stratégies classiques

La segmentation du ventricule gauche du coeur est un sujet très abordé dans la littérature en raison de son intérêt évident en terme de santé publique. Un très grand nombre de techniques de segmentation d'image ont été testées sur ce problème. Une revue des différentes méthodes est proposée dans [126, 76]. Les méthodes fondées sur les contours actifs (*snakes*) [67] ou les *level set* [97] sont parmi les plus

performantes, car elles permettent d'inclure dans le processus une forte dose de modélisation. Parmi les plus récents travaux dans ce domaine, on pourra consulter les communications de la liste non exhaustive suivante [38, 65, 37, 29].

L'image IRM est difficile à analyser, les contours du myocarde étant parfois occultés ou noyés dans les structures organiques attenantes. Par conséquent, la phase de modélisation est très importante afin de mener à bien la segmentation. Malgré l'abondance des recherche menées sur le sujet, le clinicien reste dépourvu de logiciels opérationnels et le traitement des images cardiaques reste un problème fondamental notamment pour les recherches médicales menées sur les maladies cardiaques. Avant de décrire l'algorithme qui a été développé pour l'équipe médicale des HUG, nous allons donner les principaux fondements des solutions les plus couramment expérimentées dans ce domaine : les modèles déformables et les atlas.

3.3.1 Les modèles déformables

Les modèles déformables consistent à faire évoluer une courbe dans le temps afin de converger vers la solution. Différentes implémentations sont possibles (contour actif, level-set). Ces approches sont très souvent utilisées pour la segmentation cardiaque [83, 87], puisque le myocarde peut se représenter deux formes plus ou moins elliptiques [103].

L'approche par contour actif consiste à faire évoluer une courbe dans le plan de l'image en fonction de différentes forces internes et externes : les forces internes contrôlent l'élasticité ou la rigidité de la courbe tandis que les forces externes poussent la courbe vers les frontières des objets. Le succès des contours actifs repose en partie sur le contour initial et la bonne convergence vers les frontières [113]. Dans la pratique, cela implique l'utilisation de forces externes [30] permettant d'éviter de tomber dans des minima locaux, augmentant par la même occasion le nombre de paramètres.

L'approche par *level set* consiste à encapsuler le contour 2D initial dans une fonction de distance signée 3D ϕ (figure 3.20) avant de faire évoluer la fonction vers un minimum d'énergie. L'intersection du plan de niveau 0 avec la fonction de distance signée donne le résultat de la segmentation (zéro *level set*). L'intérieur de la courbe (négatif) et l'extérieur de la courbe (positif) permettent aisément de prendre en compte des paramètres régions afin d'avoir une combinaison possible entre régions et frontières : on pourra par exemple se référer à [99].

Ces approches sont basées sur la définition initiale d'une énergie que l'on minimise ensuite. Les *level set* comme les *snakes* sont par conséquent sensibles aux problèmes des minima locaux. Cependant,

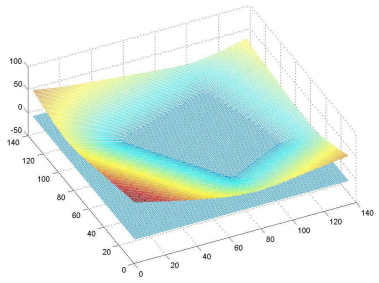


FIG. 3.19: Représentation de la fonction ϕ avec son intersection avec le plan de niveau 0. Nous visualisons parfaitement le contour initial C

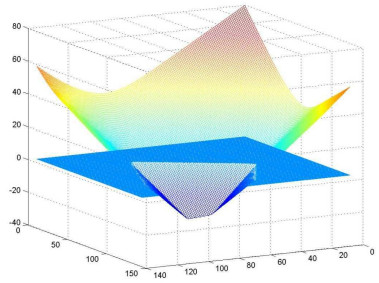


FIG. 3.20: Visualisation de ϕ sous un autre angle. Sous cet angle, le contour n'est pas visible, mais nous visualisons correctement la fonction de distance signée. Notons que $|\nabla\phi| = 1$

nous pouvons expliquer le succès des *level set* face au *snakes* en raison d'une meilleure résistance au bruit et surtout par la prise en compte des changements de topologie sans reparamétrisation.

Cependant, malgré l'engouement des chercheurs sur ce problème et un gros effort d'évaluation de l'ensemble des algorithmes de segmentation en milieu clinique, la question de la segmentation du myocarde sur IRM est encore d'actualité. La raison principale tient au fait que le problème est doublement complexe. Premièrement, les images d'IRM bien qu'ayant le meilleur contraste par rapport aux autres types d'imageries (scanner, échographie, etc.) sont quand même fortement bruitées. Le bruit peut être très important lorsque le patient souffre et ventile anormalement. Deuxièmement, au-delà de la qualité des images, l'épicardium est particulièrement difficile à délimiter, puisque bien souvent il touche d'autres structures de même contraste. Enfin, l'algorithme doit être robuste vu le grand nombre d'images à traiter et leur variabilité.

3.3.2 L'utilisation d'atlas

Sans information au niveau de la frontière, les techniques de régularisation qui ont participé au succès des contours actifs et des *level set* ne sont pas suffisantes. Pour pallier à ce problème, une idée

consiste à intégrer un a priori de forme[69, 75]. Pluempitiwiriyawej [103] dans une approche par *level set* intègre dans l'énergie une forme elliptique. La fonction ϕ évolue en fonction de quatre critères : la détection des gradients, l'homogénéité des régions, une régularisation et un a priori géométrique. Ces critères confèrent à la méthode une bonne robustesse. Une autre approche consiste à ne pas utiliser une forme géométrique fixe, mais un atlas. La construction de l'atlas nécessite de disposer d'un grand jeu de données et d'effectuer un recalage entre les données afin d'obtenir les variations autour de la forme moyenne. Le problème du recalage est plus délicat. La construction d'un atlas s'effectue généralement en utilisant des distances signées, car cela s'intègre aisément dans un cadre variationnel [108, 98]. Bien souvent, afin de restreindre les nombres de dimensions, une analyse en composante principale est faite afin de déterminer les principaux modes et bâtir l'atlas [128]. Les auteurs précédemment cités utilisent tous les atlas sur la problématique cardiaque.

L'étude de l'évolution des techniques de segmentation du ventricule gauche semble confirmer le besoin d'ajouter un critère supplémentaire aux traditionnelles approches frontière/région. L'ajout d'une forme géométrique permet de combler les faibles informations dont nous disposons. Nous suivrons ces choix en ajoutant une forme géométrique. Cependant, à la différence des auteurs cités nous n'intégrons pas de forme pure ou bien d'atlas. Notre a priori de forme est basé sur la segmentation de l'endocarde. En effet, la segmentation de l'endocarde est relativement simple, les frontières étant bien définies, la seule difficulté résident autour des muscles papillaires. Une fois segmenté, l'endocarde nous semble le meilleur a priori géométrique, pour évaluer la forme du myocarde. Une forme géométrique ou même un atlas ne sont que des formes moyennes alors que l'endocarde est directement en relation avec l'épicarde du patient. Analysons maintenant les étapes dans le détail de la méthode proposée.

3.4 Algorithme proposé

3.4.1 Augmenter les dimensions

Au-delà des techniques utilisées, l'évolution du problème de la segmentation cardiaque tend vers une extension des dimensions. Les physiciens produisent de plus en plus couramment des séquences $3D$. La segmentation $3D$ [62] ou $2D + T$ [62, 77] apporte bien évidement des informations supplémentaires qui permettent de paliers aux bruits ou pertes de contrastes sur un volume ou une coupe. L'exploration $4D$ ($3D + t$) est plus rare [74] et va certainement devenir un standard d'ici quelques années. Les travaux récents [38, 65, 37, 29] vont dans ce sens.

Concernant notre travail, l'objectif est non pas de résoudre la problématique cardiaque, qui est en soi une thématique de thèse entière, mais d'illustrer comment intégrer un a priori géométrique afin de contre-balancer la perte de contraste. L'ajout d'un atlas et plus généralement d'a priori n'est généralement pas simple. Dans notre cas, il suffit d'empiler les critères (image multivaluée) et d'utiliser la transformation de rang pour se ramener à un cas $2D$.

3.4.2 Initialisation

L'algorithme est initialisé par une détection grossière de l'extérieur et de l'intérieur du myocarde : un point est positionné manuellement à l'intérieur de la cavité cardiaque ; un point est mis à l'extérieur. Les marqueurs externes sont construits à partir de ces deux positions comme indiqué sur la figure 3.21.

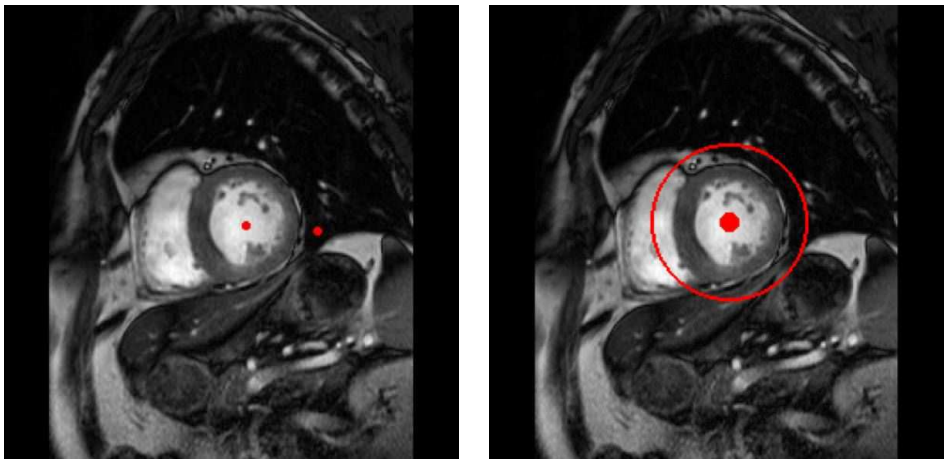


FIG. 3.21: Initialisation : marquage manuel pour la segmentation de l'endocarde.

3.4.3 Segmentation de l'endocarde

La segmentation de l'endocarde ne pose pas de difficulté majeure (c'est l'épicarde qui pose problème). Elle échoue rarement même avec des techniques simples de segmentation. Ici, nous proposons une segmentation par calcul de la *LPE* sur une image gradient (le gradient morphologique [18] est utilisé). Les muscles papillaires ne sont logiquement pas pris puisqu'ils sont de la même couleur que le fond. Cependant, l'étude de nombreuses images montre qu'ils sont soit représentés par des points isolés, soit collés au myocarde. Lorsqu'ils sont isolés, la *LPE* règle immédiatement le problème puisqu'elle encercle le point avant de l'inonder. Le deuxième cas est plus difficile puisque la *LPE* s'arrête au bord des muscles, comme dans l'exemple de la figure 3.22a. La création de ces concavités ou trous sur les bords du résultat est très simplement contournée par une fermeture comme nous le montrons sur la figure 3.22b.

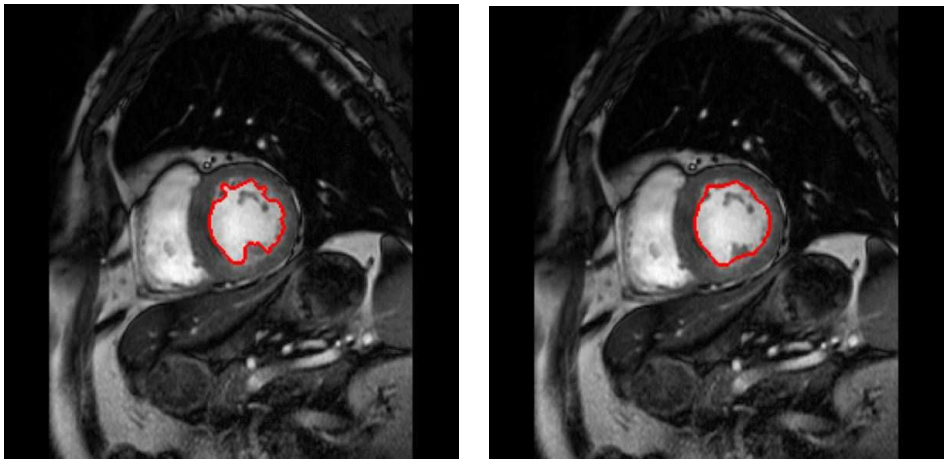


FIG. 3.22: À gauche la figure montre le résultat de la segmentation par *LPE*. Les muscles papillaires internes sont absorbés par la *LPE* mais pas ceux contre le myocarde. La fermeture (image de droite) enlève les muscles papillaires récalcitrants.

3.4.4 Construction du critère de forme

La segmentation de l'épicarde nécessite d'ajouter un critère supplémentaire. L'idée consiste à utiliser le résultat précédent et faire l'hypothèse que l'épicarde est presque un homothétique de l'endocarde en d'autres termes, nous allons utiliser la distance à l'endocarde pour construire un indicateur de confiance et positionner l'épicarde. Une carte statistique de position est ainsi élaborée : les points à distance constante de l'endocarde (obtenus par simple dilatation morphologique de l'endocarde) forment une courbe fermée, simple qui est valuée avec sa valeur moyenne sous le l'image. La carte est obte-

nue en considérant quelques dilatations de l'endocarde (fixée en fonction de l'épaisseur présumée du myocarde). Elle met en évidence les zones homogènes (le myocarde) créant la famille dont nous avons besoin pour segmenter le myocarde. L'image ainsi obtenue est présentée par un exemple en figure 3.23. La segmentation de cette image seule ne permet bien sûr pas d'obtenir le myocarde comme nous le montrons sur la figure 3.24. Il est par conséquent nécessaire de travailler sur deux cartes simultanément, l'image et la carte statistique, afin de bénéficier de toutes les informations disponibles.

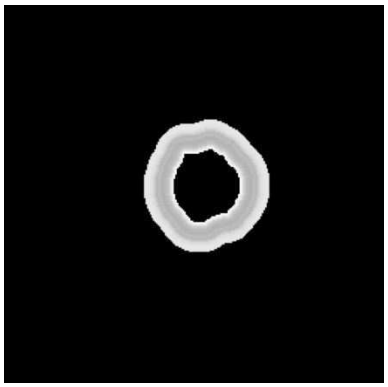


FIG. 3.23: Carte statistique obtenue en calculant la moyenne sur le contour de l'endocarde dilaté.

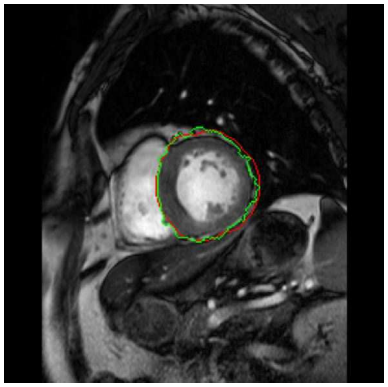


FIG. 3.24: En rouge le résultat de la segmentation en utilisant la *LPE* avec uniquement l'image statistique. L'épicarde est très largement surévalué. En vert, le résultat en combinant les deux critères. L'irrégularité est gommée par une ouverture.

3.4.5 Segmentation dans l'espace des similarités

La segmentation de cette image bivaluée est effectuée en utilisant la transformation de rang et la *LPE*. Les échantillons sont collectés en partant de l'endocarde. Deux ou trois dilatations garantissent d'être dans le myocarde et d'obtenir un masque permettant de constituer l'échantillon.

Les marqueurs sont simplement deux cercles, l'un à l'intérieur du myocarde et l'autre à l'extérieur. On remarque lors des expérimentations que le contour obtenu peut déborder légèrement du myocarde provoquant des petits cisaillements sur les bords de l'épicarde (trait vert de la figure 3.24). Ceci est dû au fait que nous n'avons pas imposé de contrainte de régularisation lors de la segmentation. Une simple ouverture de la zone trouvée correspondant à l'intérieur du myocarde suffit à garantir une segmentation correcte.

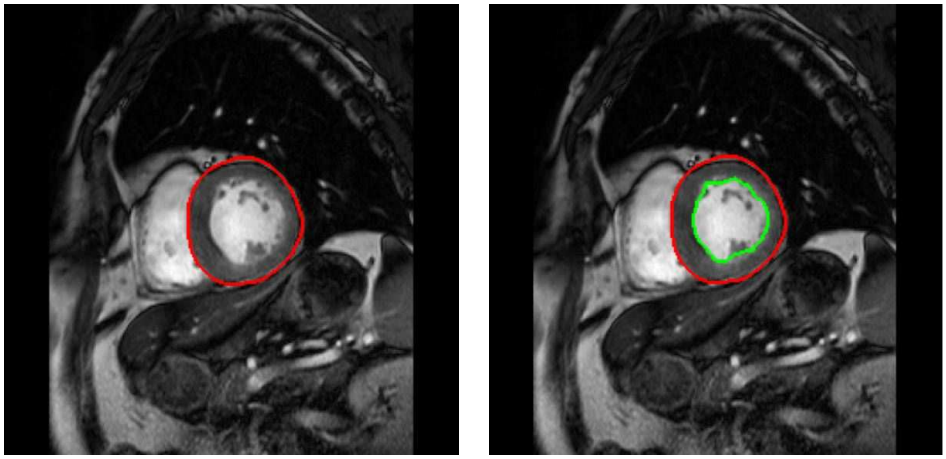


FIG. 3.25: Segmentation de l'épicarde. À gauche, l'endocarde et l'épicarde.

3.5 Conclusion : validation et segmentation 3D

Il faut bien évidemment de nombreux tests ainsi qu'une validation à grande échelle afin de pouvoir prétendre avoir résolu ce problème de segmentation. Cependant, les premiers résultats sont très encourageants (voir figures 3.26 à 3.28). L'initialisation est simple et les résultats sont reproductibles malgré l'action de l'utilisateur. Une démonstration complète de notre algorithme est faite sur internet à l'adresse suivante : `http://www.dailymotion.com/arnaud74130/video/x2ky8p_demo-segmentation-cardiaque_tech`. La démonstration est faite sur des IRM du coeur d'un rat. La qualité de l'image est nettement moins bonne et les contrastes sur les bords sont vraiment faibles. Malgré cela le résultat est correct (figure 3.28).

L'étape suivante consiste à segmenter l'ensemble du coeur. L'idée est d'utiliser le même algorithme, mais en utilisant des volumes. Les acquisitions 3D du coeur sont pratiquement finalisées. Concernant notre algorithme, le passage en 3D ne demande que peu de modifications.

Au-delà de l'intérêt médical, l'objectif était de montrer comment intégrer simplement un critère supplémentaire lorsque les frontières sont peu marquées. L'ajout d'a priori géométriques dans les processus de segmentation est une question d'intérêt dans nombre d'applications ...

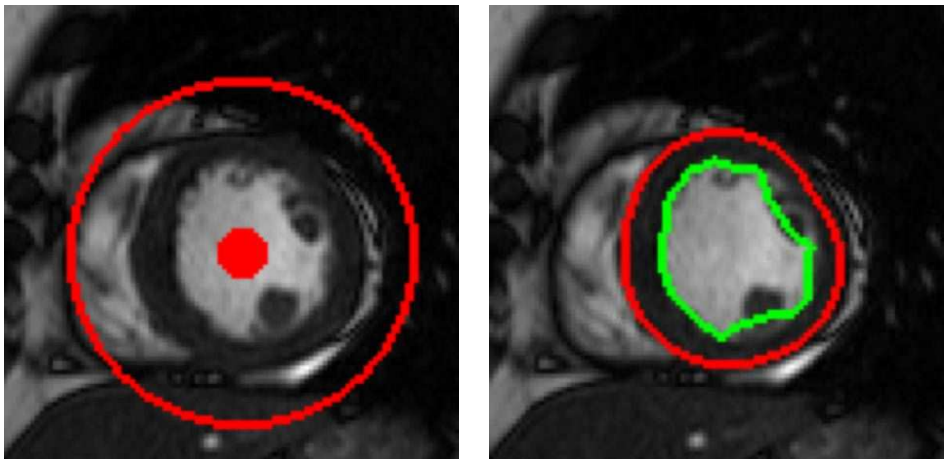


FIG. 3.26: Patient 1. De gauche à droite : les marqueurs, le résultat. Les marqueurs sont positionnés uniquement par deux points que nous dilatons afin d'obtenir la configuration de l'image de gauche. Les muscles papillaires particulièrement gros sur cette coupe provoquent une sous-évaluation de l'endocarde.

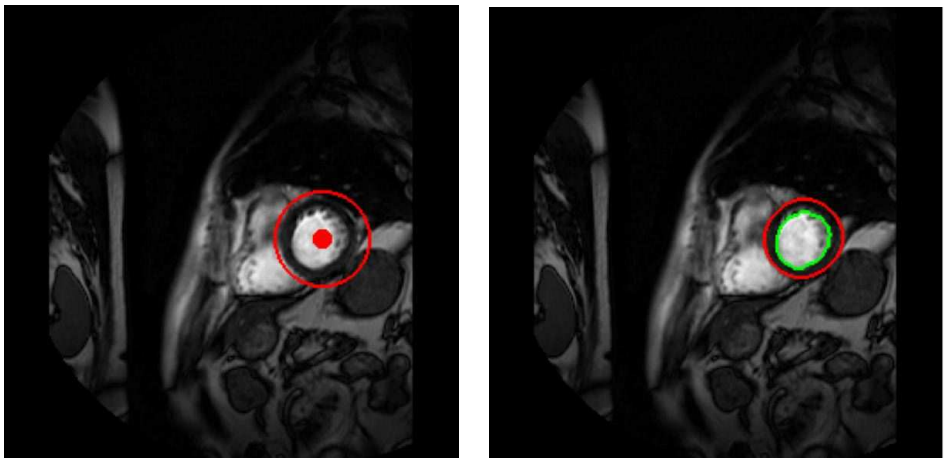


FIG. 3.27: Patient 2.

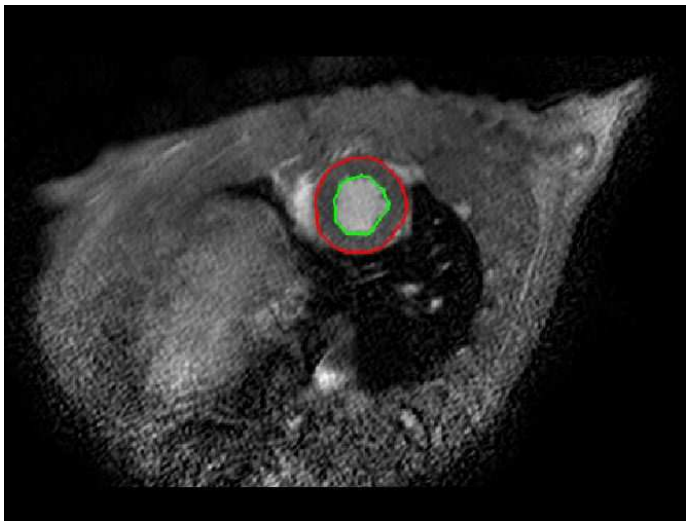


FIG. 3.28: IRM de rat. La qualité des images ainsi que les contrastes sur les bords sont nettement moins bons que chez l'homme, malgré cela le résultat est correct.

3.6 Osirix, un outil d'aide à l'analyse des images médicales

Le logiciel OsiriX [107] permet de visualiser les images au format médical (DICOM[2]). Ce logiciel écrit par Antoine Rosset est libre. La gratuité et les qualités du logiciel ont permis un déploiement massif dans le monde. En raison de son succès et compte tenu du fait que l'équipe de développement d'OsiriX était à Genève, nous avons porté notre algorithme sous OsiriX.

Le portage du *MSRG* sous OsiriX était nécessaire pour évaluer l'algorithme sur des cas concrets. En effet, les prototypes développés sous Matlab ne permettent pas de manipuler des volumes importants. Lancer le *MSRG* sur un scanner de crâne de 800 coupes avec deux types d'acquisitions (IRM T1 et

[2] `http://medical.nema.org/`

T2) implique un développement optimisé ainsi qu'une interface pour l'utilisateur. Le médecin étant spécialiste, il peut renseigner la machine sur le contexte en fournissant notamment des échantillons pertinents des objets recherchés dans l'image. Ici, une illustration effectué sur le Visible human[3] eyes 3.29.

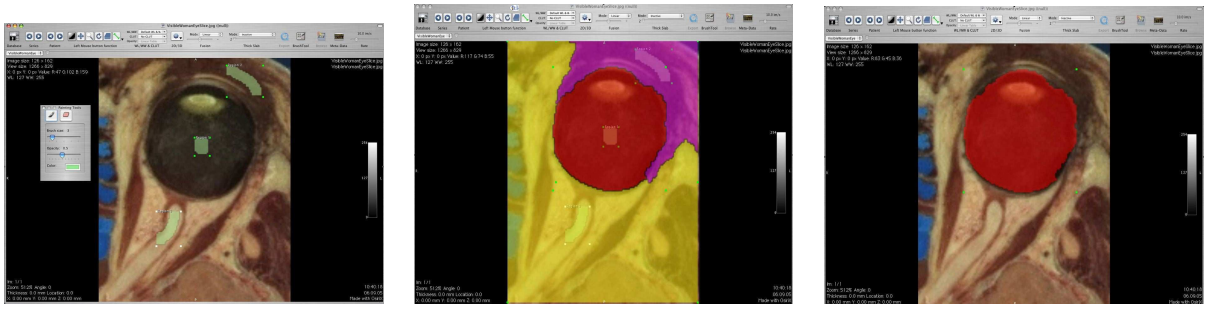


FIG. 3.29: De gauche à droite. L'image directement chargée dans OsiriX, nous pouvons placer les marqueurs avec l'aide du pinceau réalisé pour ce type d'algorithme. Une fois les marqueurs positionnés, l'utilisateur lance alors le *MSRG*. La segmentation effectuée, le médecin peut ensuite à l'aide d'un gestionnaire de région d'intérêt, supprimer les régions des autres tissus et obtenir des statistiques diverses (aire, volume)

Afin de tester aisément le *MSRG*, nous avons par conséquent développé avec l'équipe d'OsiriX une interface complète pour les algorithmes de segmentation. Cette étape coûteuse en temps et éloignée des problèmes de traitement de l'image est pourtant fondamentale pour effectuer des tests. Ce point est très important et explique la création et la direction stratégique de la librairie *vmorph* (confer section 3.7) capable d'être utilisée en mode recherche ou en mode production avec des fonctions identiques.

L'interface proposée dans OsiriX permet de positionner ou d'enlever les marqueurs, de limiter la zone de traitement et de définir l'évolution du *MSRG* en 2D ou 3D. Nous offrons également la possibilité de sélectionner plusieurs examens avant de lancer l'algorithme. Il est alors possible de prendre deux volumes 3D (T1,T2 par exemple) ou deux types d'imagerie (PET,CT).

Un *ROIManager* donne ensuite les informations des différentes régions, (aires, volumes).

[3] `http://www.nlm.nih.gov/research/visible/visible_human.html`

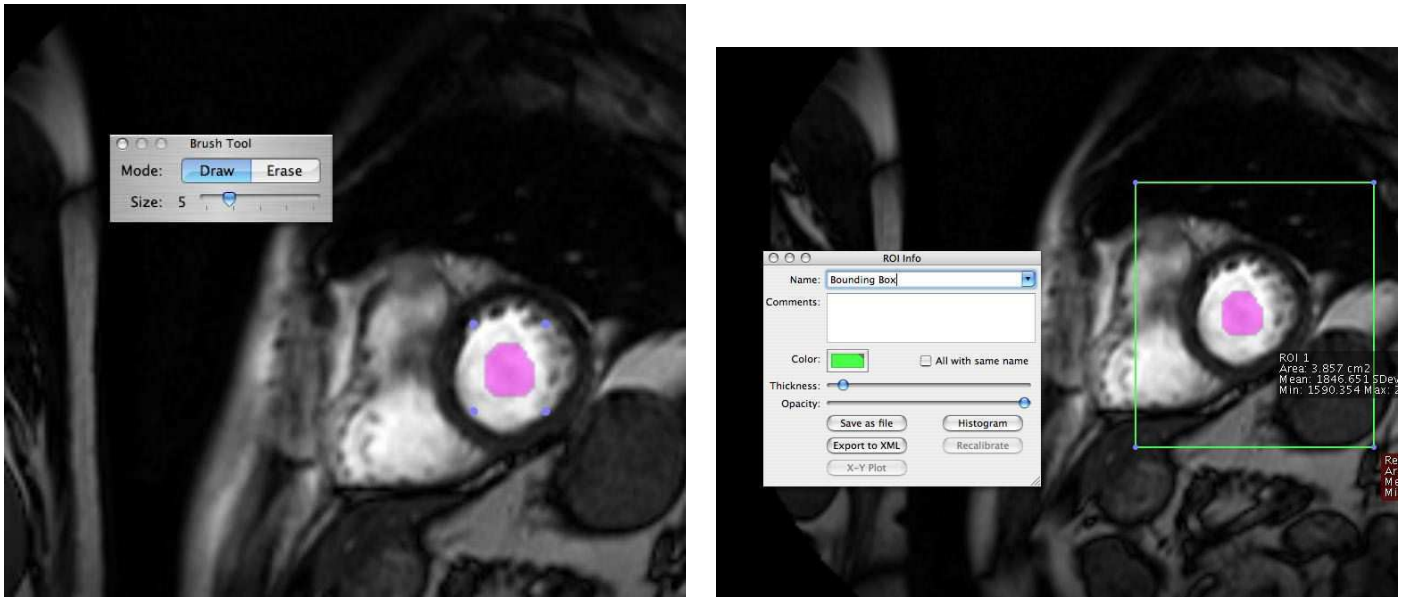


FIG. 3.30: Positionnement des marqueurs

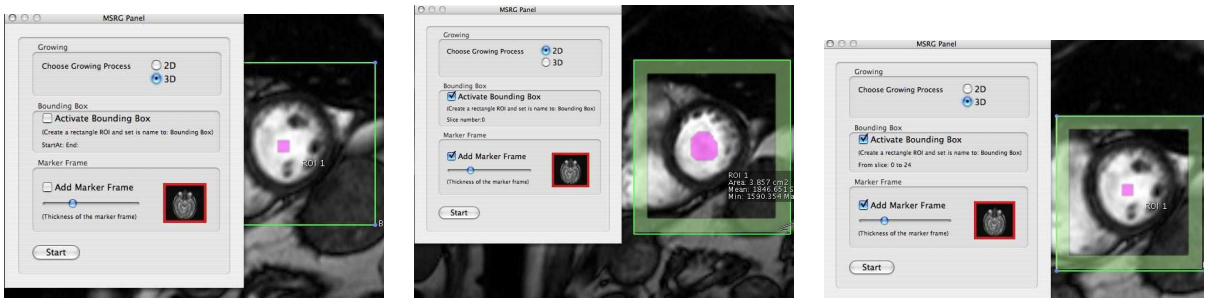


FIG. 3.31: Paramétrage du lancement de l'algorithme.

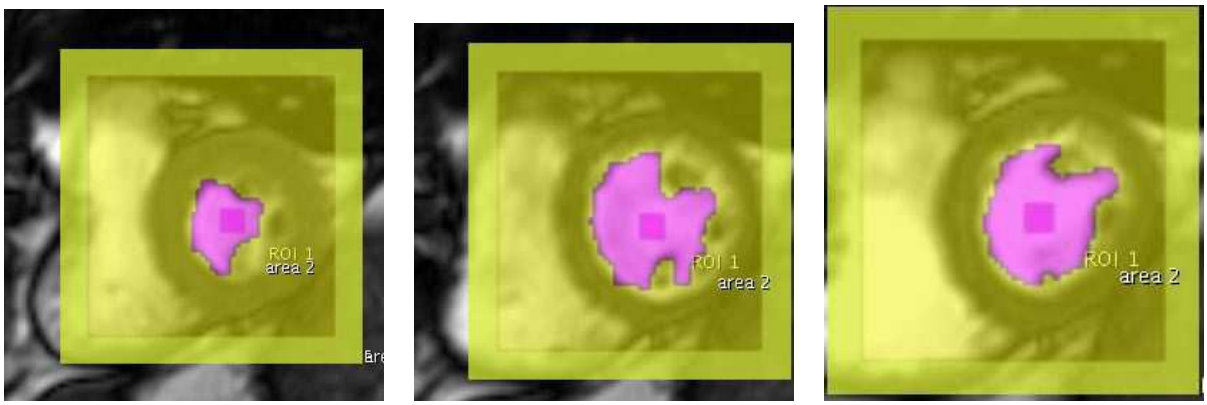


FIG. 3.32: Résultats de la segmentation de l'endocarde à différentes coupes.

3.7 vmorph

La librairie *vmorph*[4] est une librairie opensource écrite dans le langage JAVA permettant le traitement de la morphologie mathématique aussi bien sur des scalaires que sur des vecteurs. Cette librairie est disponible sur l'ensemble des plateformes (linux, macosx, windows, solaris). Une intégration de *vmorph* est également faite avec le logiciel Matlab. Cette intégration permet l'utilisation de toutes les fonctions de *vmorph* dans l'environnement Matlab. L'utilisation de Matlab et *vmorph* offre un moyen efficace et puissant pour la construction des algorithmes. L'intégration avec Scilab et Octave est en cours d'évaluation.

La librairie *vmorph* est construite à partir des recherches faites dans ce travail. Les réflexions sur l'extension aux vecteurs en morphologie mathématique ont été traduites informatiquement afin de confronter la théorie et la pratique. L'extension des opérateurs de morphologie mathématique se fait par conséquent en utilisant l'image de rang (2.1.2). Cette initialisation est indispensable afin de définir l'ordre dans le treillis et de permettre la généralisation des opérateurs. L'opération de soustraction nécessaire pour les transformations du type chapeau haut de forme ou pour le calcul des gradients suit également les principes théoriques que nous avons proposés (2.3.2).

3.7.1 Architecture

Nous montrons les grands principes de construction de la librairie *vmorph*. L'objectif n'est pas de présenter l'implémentation de tous les algorithmes, mais de mettre en avant les principes architecturaux que nous avons retenus.

3.7.2 Principes

Nous utilisons, le principe *DRY*, « Don't repeat yourself »[64] très présent dans le développement Web. Bien qu'apparemment évident la mise en application l'est nettement moins. En effet, la création des opérateurs morphologiques peut rapidement se résumer en une somme de copier/coller. Ce problème est encore plus frappant lorsque nous étendons les opérateurs aux vecteurs puisque la redondance de code est plus importante. Érosion et dilation sont d'un point vue algorithmique très proche aussi bien pour les scalaires que pour les vecteurs. Il faut par conséquent réussir à centraliser le code commun.

[4] Le nom de vmorph vient de vectorial morphology.

La recopie du code augmente le nombre d'erreurs de façon impressionnant. En effet, une erreur sur le traitement des bords par exemple doit être recopiée environ une dizaine de fois pour l'ensemble de la librairie, alors que tous doivent gérer le débordement. Le principe *DRY* permet de concentrer les parties sensibles afin que la correction se répercute sur l'ensemble des routines. Cependant, cela impose une organisation et une réflexion plus importante lors de la construction des algorithmes.

La librairie *vmorph* est également construite en utilisant la méthodologie provenant de l'eXtreme programming [16]. L'idée est de proposer une architecture évolutive afin de prendre en compte plus facilement les changements. Le code est par conséquent flexible et ouvert aux modifications. Afin de garantir cette souplesse, à chaque objet sont associées des classes de tests. Encore une fois, ce principe bien que faisant appel au bon sens est souvent négligé dans les phases de développement. Les développeurs se focalisent souvent sur la construction de l'algorithme et enchaînent les fonctionnalités avant de stabiliser les algorithmes avec des programmes de tests. La mise en place d'une véritable stratégie de tests évite la régression fonctionnelle et permet de garantir les évolutions rapides. Ces modifications et restructurations du code en fonction des besoins sont appelées *refactoring* [46]. La librairie *vmorph* dispose d'une batterie de tests permettant d'effectuer du *refactoring* et donc de répondre aux futures évolutions sereinement.

3.7.3 Implémentation

Pour répondre au principe *DRY*, nous utilisons des itérateurs[5] et des classes abstraites[6]. La figure 3.33 montre le diagramme de classe suivant le formalisme UML [23] pour les opérateurs standards. Sans rentrer dans le détail, l'implémentation consiste à concentrer toute la logique métier dans la classe abstraite. L'érosion et la dilatation ne font qu'instancier la classe *MorphoOperator* en précisant simplement l'opérateur permettant d'ordonner les éléments. Le code est ainsi minimal.

```
1 public class Dilatation extends MorphoOperator  {
2     public Dilatation(ImageDouble in, SE se) {
3         super(in, se, DILATATION);
4     }
5 }
```

[5] Wikipedia : En génie logiciel, l'itérateur est un patron de conception (design pattern) comportemental. Un itérateur est un objet qui permet de parcourir tous les éléments contenus dans un autre objet

[6] Wikipedia : En programmation orientée objet (POO), une classe abstraite est une classe dont l'implémentation n'est pas complète et qui n'est pas instanciable. Elle sert de base à d'autres classes dérivées (héritées)

Nous avons montré que l'érosion et la dilatation reposaient sur la même classe *MorphoOperator*. Nous pouvons garder cette classe pour l'extension aux vecteurs en utilisant les itérateurs. La classe *MorphoOperator* implémente la classe *Iterator*. Par conséquent, à chaque opération (ligne 16), c'est-à-dire à chaque point il est possible de prendre la main et de mettre à jour l'image vectorielle (ligne 18).

```
1  public abstract class VectOperator implements VOperator {
2      public TicToc chrono = null;
3      private VectData in = null;
4      private MorphoOperator operator = null;
5
6      public VectOperator(VectData in, MorphoOperator operator) {
7          chrono = new TicToc();
8          this.in = in;
9          this.operator = operator;
10
11     }
12
13     public VectData process() {
14         chrono.tic();
15         VectData out = VectData.emptyDmDataFromSrc(in);
16         while (operator.hasNext()) {
17             Value point = (Value) operator.next();
18             out.getRank().poke(point.xp, point.yp, point.value);
19             out.getStack().setVectorAtRowAndCol(point.yp, point.xp,
20                     in.getStack().getVectorAtRowAndCol(point.y, point.x));
21         }
22         chrono.toc();
23         return out;
24     }
25 }
```

Nous avons montré comment nous pouvons mettre en place le principe *DRY* pour une implémentation d'opérateurs en morphologie mathématique. Cette concentration du code dans une seule classe permet de limiter les sources d'erreurs et d'étendre très simplement les principes aux autres types de données.

3.7.4 Intégration dans Matlab

Afin de garantir une intégration totale avec Matlab, *vmorph* utilise des *doubles*, ce qui est loin d'être optimal pour la mémoire. Différents types seront supportés par *vmorph*, mais dans un premier temps il est capital d'offrir cette passerelle avec Matlab. L'intégration est quasi immédiate puisque Matlab supporte les types de JAVA. Il est à noter qu'il faut faire extrêmement attention lors de la manipulation des tableaux multidimensionnels qui diffèrent au niveau des indices entre JAVA et Matlab. *Vmorph* offre

les fonctions de transfert entre Matlab à JAVA. Ci-dessous un exemple de fonction Matlab permettant l'appel d'une fonction de *vmorph* pour le calcul d'une érosion vectorielle.

```
function [rank,stack]=vmerovect(f,rank,size)
    % Compute the vectorial erosion
    %   INPUT: f: multivalued image, rank:  rank Image, size: of the disk
    %   OUTPUT: rank: rank image updated, stack :  eroded multivalued image
    import com.vmorph.image.*;
    import com.vmorph.image.morpho.vect.*;
    import com.vmorph.image.morpho.se.*;

    jStack=StackDouble.fromMatlab(double(f));
    jRank=ImageDouble(double(rank));

    jSe=SEDisk(size);
    jVectData=VectData(jRank,jStack);

    ero=VectEro(jVectData,jSe);
    jOut=ero.process();

    rank=jOut.getRank().getBuff();
    stack=jOut.getStack().toMatlab();
```

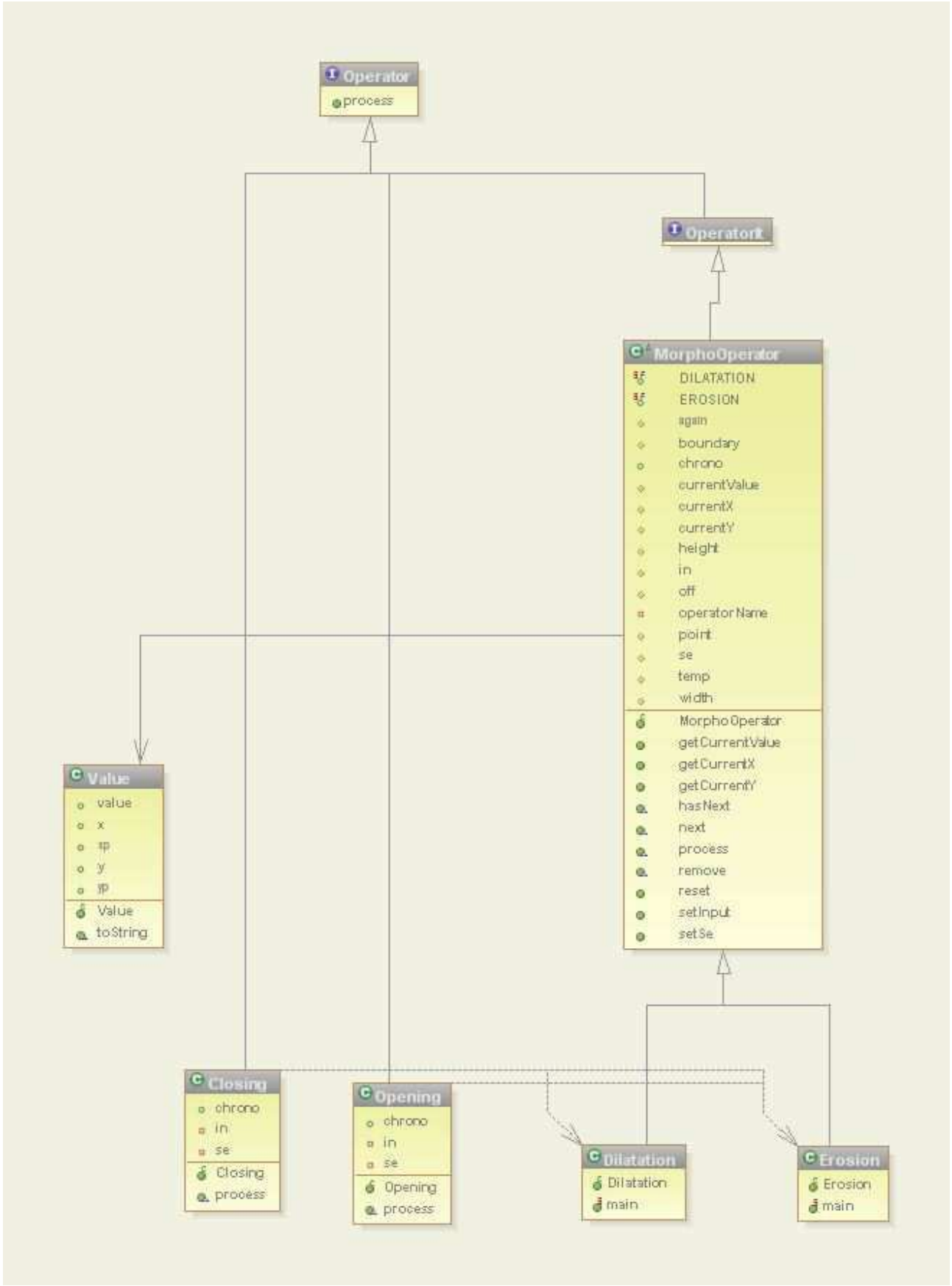


FIG. 3.33: Package contenant la construction des opérateurs morphologiques pour les scalaires. Érosion et dilation dérivent tous les deux d'une classe abstraite contenant toute la logique métier. La classe de base *MorphoOperator* implémente également la classe *Iterator* afin d'offrir un moyen simple et efficace pour rendre la main à chaque traitement.

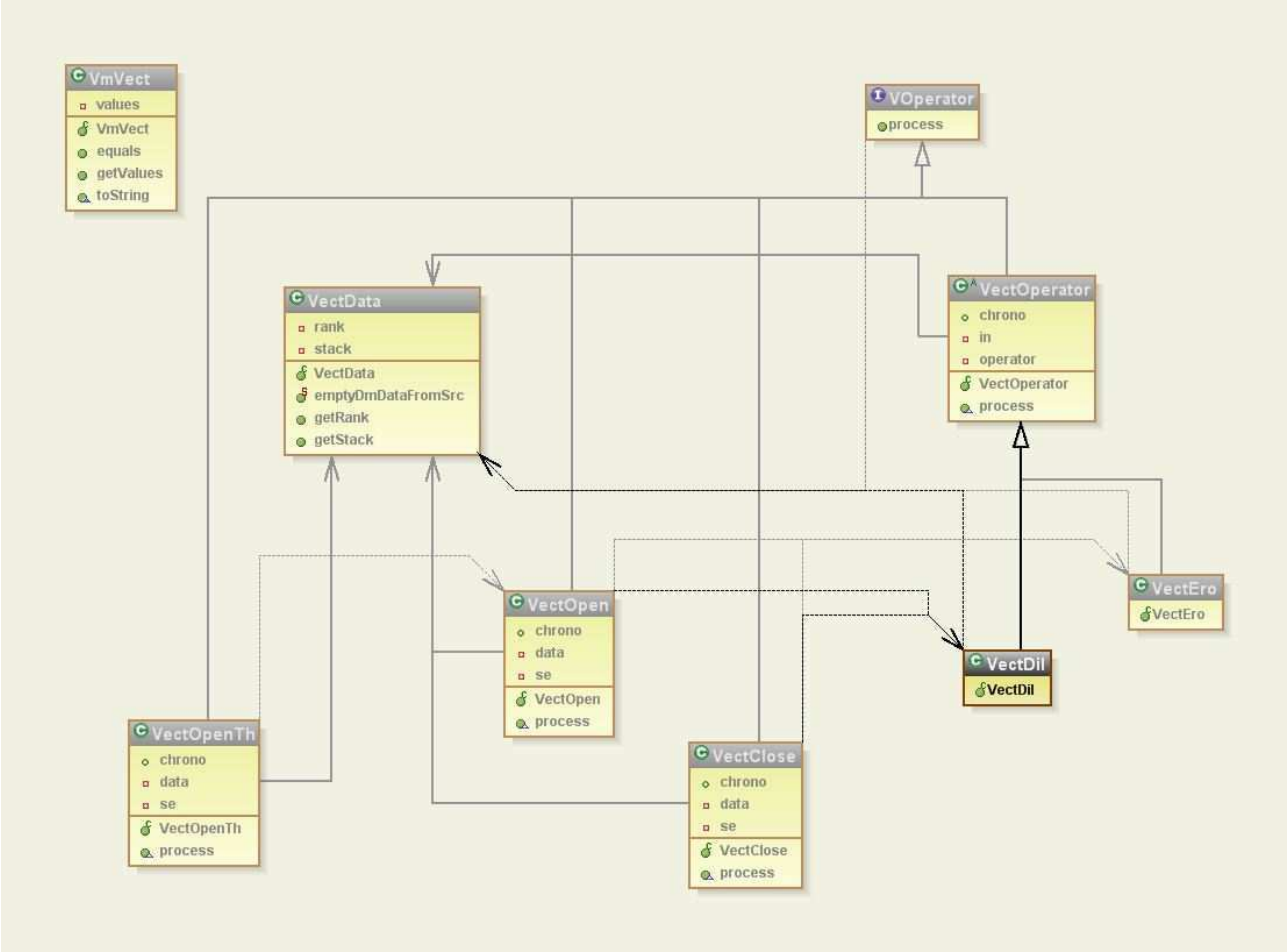


FIG. 3.34: Package contenant la construction des opérateurs vectoriels. Érosion et dilation dérivent également d'une classe abstraite. Cette classe utilise la logique métier de *MorphoOperator* que nous utilisons pour le traitement des scalaires.

Chapitre 4

Conclusion et perspectives

Face à la difficulté pour une machine de réaliser les mêmes tâches que l'homme, les algorithmes doivent être conçus de telle sorte à exploiter au mieux l'information intrinsèque à l'image et la connaissance susceptible de lui être transmise par l'opérateur. La solution que nous avons présentée, testée et exploitée sur des applications médicales concrètes dans cette thèse est construite sur une analyse statistique et morphologique de l'image. L'analyse statistique permet de confronter le contenu de l'image à un échantillon de référence. Dans le cas des images multivaluées, elle réduit la complexité de l'analyse de l'image en modifiant la représentation des données : on peut passer de données vectorielles à des données scalaires sans perte d'information. L'analyse morphologique vient compléter l'analyse statistique et permet la prise en compte des propriétés géométriques de l'image.

La stratégie proposée permet de définir très aisément l'intégralité des opérateurs morphologiques sur les images multivaluées puisque l'on passe sans perte des données vectorielles aux données scalaires.

Des illustrations sur images couleur ont été montrées ; elles permettent de visualiser le comportement d'un opérateur et éventuellement de le comparer aux résultats obtenus par des opérateurs plus conventionnels ou de le confronter aux requêtes de l'utilisateur.

Une innovation importante concerne la définition des opérateurs adaptatifs pour les images multivaluées. Les expérimentations effectuées donnent à penser que la voie est prometteuse. Ces opérateurs permettent de prendre en compte les a priori statistiques et morphologiques pour l'analyse des images, et dans le cas des images couleur, permettent de simplifier les images d'une manière très intéressante pour nombre d'applications. Parmi celles-ci on peut penser aux applications de codage et aux outils de retouche d'images par exemple.

Sur le plan des applications médicales, des algorithmes originaux ont été présentés pour l'étude des mélanomes dans le cadre du projet Naeview et la segmentation d'IRM cardiaques. Dans ce type de projets, la validation des algorithmes est indispensable pour passer de l'étude de faisabilité au logiciel clinique. Les équipes des HUG ont commencé ce travail qui demande un investissement très important. À ce jour nous n'avons pas encore les résultats, mais les premières expérimentations sont encourageantes.

Par ailleurs, cette thèse a fait l'objet de développements informatiques importants. J'ai livré et intégré dans le logiciel OsiriX toute une boîte à outils pour la segmentation interactive des images médicales. Cette intégration a été réalisée en collaboration avec le groupe de développement d'OsiriX pendant mon séjour à Genève. Dans un deuxième temps, j'ai développé une librairie permettant l'utilisation de tous les développements proposés dans cette thèse. Cette librairie gratuite et open source fait l'objet de développements continus. *Vmorph* permet de travailler aussi bien en mode recherche avec des logiciels comme Matlab qu'en mode production pour une utilisation en routine clinique. J'ai toujours essayé de confronter la théorie au domaine appliqué. La construction de *vmorph* s'inscrit dans cette démarche.

Sur le plan théorique, les réflexions amorcées dès le début de ma thèse sur les vecteurs et la morphologie mathématique semblent toujours plus d'actualité. La technologie permet plus facilement le maniement des vecteurs, les appareils physiques génèrent de plus en plus de données, l'imagerie PET-CT met bien en évidence l'arrivée de nouvelles images. Il y a donc fort à parier que le développement des imageries multivaluées va se développer. Concernant le travail qui a été entrepris pendant ma thèse, les recherches sur la manipulation des vecteurs et la construction des opérateurs pour les images multivaluées me semblent mériter des travaux plus approfondis. Par exemple, les premiers résultats obtenus à partir des opérateurs visqueux sont surprenants. Plusieurs questions se posent et notamment : comment choisir les références dans le cas de l'analyse d'images couleur génériques (lorsqu'aucun a priori sur le contenu des images n'est disponible) ? Est-il possible de mettre en place un processus de choix automatique des références ?

Un autre axe de recherche est bien sûr celui de la validation clinique des algorithmes proposés. J'espère pouvoir poursuivre ma collaboration sur ce point avec les équipes de recherche des Hôpitaux Universitaires de Genève. Il est d'ores et déjà prévu de mettre l'accent sur la problématique cardiaque.

Annexe A

Quelques communications

La liste des publications qui suit n'est pas exhaustive. Seules les communications dont le contenu scientifique n'est pas mentionné dans le corps du document sont présentées. Les autres communications sont référencées dans la bibliographie. On peut découper les travaux en trois grandes parties. Une première partie sur la morphologie mathématique (segmentation et filtrage), dont ce mémoire fait une large synthèse[2, 49, 47, 48]. Une deuxième partie sur la recherche de contenu (Content base retrieval) [90, 91, 89]. Enfin, une dernière partie, plus liée au système d'information[14] et au problématique de distribution des calculs[88].

A.1 Benefits from content-based visual data access in radiology

H. Müller, A. Rosset, A. Garcia, J.-P. Vallée, and A. Geissbuhler.

in radiology. Radiographics, 25 (3) :849–858, 2005.

Introduction to visual information retrieval

Content-based visual data access without the use of textual descriptions is a very active research topic in computer vision and image processing. Many applications exist in the research domain as well as commercial systems. Medical applications are often cited as one of the principal areas where content-based visual queries can be beneficial. Still, only very few visual image retrieval systems have been used in clinical routine, most remain prototypes.

More on general image retrieval systems for non-medical applications such as journalists' image archives or on trademark retrieval can be found in [1,2,3]. Most current systems use the approach to formulate queries using example images, called QBE (**Query by example(s)**). This requires having a proper starting image to formulate the query. Other systems allow formulating queries by selecting regions from pre-segmented images [4] or by text and images combined [5]. In general, the images are represented in the databases by automatically extracted visual features that are supposed to correspond to the visual image content or the way we perceive it. Mainly used for image retrieval are:

- grey levels and color descriptors, in a local or global fashion,
- texture descriptors,
- shapes of segmented objects.

Grey levels in the image and their distributions or layout throughout the image are often represented through histograms that can be compared with a simple intersection or a Euclidean distance. Local grey level descriptors can be represented by the most frequent grey in a certain area or by local grey level histograms. T**extures** can be described by wavelet filter responses [6] that measure the changes of the grey levels in various directions and scales throughout the image, or features derived from co-occurrence matrices that count the frequency of neighboring grey levels in various directions and distances to describe a texture. This allows describing the scale of a

texture, the principal directions and whether the changes are very quick or rather gradual. Texture descriptors make mainly sense when they are extracted from a region that is homogenous in texture. S**hape** features can be used to characterize identifiable or segmented objects. Used features are mathematical moments of the shape but also features that describe the roundness of the form or the number of changes between convex and concave segments of the contour. Often, the goal is to extract features that are invariant with respect to the size of the object and with respect to rotations. By comparing the features of two images, we can calculate a similarity score between the two. Different distance measures for comparisons exist such as the simple Euclidean or the city block distance.

In general, all the features are on a fairly low semantic level in contrast to text that might come with the images. These features are also often in contrast to the high-level semantic concepts that users are mainly looking for like an object such as a tumor or a certain texture representing a disease such as emphysema. This semantic difference between image representation and the image content is called the **semantic gap**. Another gap or information loss is the **sensory gap** that is already due to the information loss from the original physical structure to the digital image, for example, due to projection where three-dimensional structures are represented by two-dimensional images (e.g. Chest x-rays). The always-limited resolution of digital images is another reason for a sensory gap between images and reality.

Most published articles on content-based medical image retrieval seem to be either written in medical departments where a clear need for image retrieval systems exists and is often defined [7,8,9] or in computer science departments where medical data sets are used but no link with the clinical routine exists [10,11]. Only few active research projects with clear clinical goals and running prototypes exist currently. The ASSERT project (Automatic Search and Selection Engine with Retrieval Tools) of Purdue University is a rather active one. It focuses on the analysis of

textures in high-resolution CTs of the lung [12]. A clinical test using the system as diagnostic aid shows an improvement in diagnostic quality, especially for less experienced practitioners [13]. 11 persons had to diagnose the same set of images once with the help of ASSERT and once without any help. The tests were performed with several weeks in between so the persons did not remember the previous results. The percentage of correct diagnoses improved from 29% to 62% with aid from ASSERT. The improvement was higher for general radiologists than for chest specialists. No group decreased its diagnostic performance. The IRMA project is also very active [14]. It works on visual similarity retrieval and automatic image classification. In IRMA, a multi-dimensional code was created to annotate image databases [15] with axes for modality, body orientations, body region examined, and the biological system under control. An image database is currently being annotated containing 10'000 images from clinical routine, mainly conventional radiographies. The medGIFT retrieval system described in this paper differs from IRMA and ASSERT. It uses a very large feature space and is based on techniques well known from text retrieval. Relevance feedback and user interaction are very important components. Another difference is the fact that it does not need classification and a-priori knowledge for retrieval. The features itself are supposed to model the visual similarity of the images. More references and a detailed review of content-based medical image retrieval systems can be found in [16].

Sometimes, medical images are also retrieved with text, only [9,17]. This cannot really be called content-based retrieval but rather context-based retrieval as the text describes the context in which the image was taken or evaluated and rarely the visual content. As text, the radiology report or the text supplied by the teaching file can be taken, if available with the image in digital form. These texts are generally treated to remove very frequent so-called stop words (like "the", "a"). Then, stemming removes the unimportant end part ("contained", "containing" both become "contain") before the remaining words can be indexed and used for retrieval. The imageCLEF competition

(http://ir.shef.ac.uk/imageclef2004/) shows that both text and visual features have an important influence on retrieval quality. Best results can be obtained when combining the two. Whereas text has the advantage to cover semantics, it has the disadvantage of being task and user dependent, and even the same person will annotate the same image in slightly different words when performing the same task again. When a new image is produced, there is no text available and the radiologist will need to formulate queries. Automatically extracted visual features are "objective" for one image and can be obtained without additional work.

Materials and Methods used in this article

This article gives an overview of the literature available on applications in medical visual image retrieval. From the available literature, important application fields are identified and scenarios for the potential benefits of image retrieval in radiology are presented with visual retrieval examples using the retrieval system medGIFT.

The image retrieval engine medGIFT is described in more detail. It is based on the open source software tool GIFT (GNU Image Finding Tool). Our teaching file system casimage offers integration into the visual retrieval framework and gives us access to large teaching files. CasImage is an in-house development and medGIFT is equally being developed and specialized for the use with medical images in the medical informatics service, Geneva. medGIFT retrieves images based on local and global grey level and texture similarities. medGIFT was evaluated for medical image retrieval in the context of the imageCLEF competition [18,19]. Some of the evaluation results concern the number of grey levels that deliver best retrieval results. This number is surprisingly low for optimal image retrieval. The system was used with several grey level quantizations and the performance was evaluated against a gold standard generated by a radiologist.

Applications for visual data access

Teaching

Teaching can be the first domain to really profit from content-based access methods [20]. Many teaching files such as casimage [21] or the online system myPACS (http://www.mypacs.net/) exist. The systems are supposed to give a maximum flexibility to the practitioner entering cases, being as much integrated into the clinical workflow as possible. Like this, interesting or typical cases can be exported directly from the PACS or the viewing station without the necessity of complex transformations. Inclusion of images into presentations and texts should equally be easy by drag and drop. Such an easy-to-use system gives flexibility to practitioners but also prohibits a strict control of the entered data for validity. As a consequence the data are often of mediocre quality, containing spelling errors and non-standardized abbreviations. The records stored in the casimage database also contain multi-lingual entries posing more problems. Sometimes, single records are multi-lingual as data was copy/pasted from a French document with a translation for a web demonstration being started but never finished. Content-based search can be an easy option to complement the text-based or hierarchical access methods to the data. This can allow students to browse the available data themselves in an easy and straightforward fashion by clicking “show me similar images”. This can stimulate self-learning and a comparison of similar cases and their particularities. On the other hand, lecturers can find optimal cases for teaching even in parts of the database that they did not generate themselves, which might be annotated differently from their cases or simply in another language. This can also include visually similar cases with a different diagnosis, which can be important for teaching. Good starting images can still be found using text-based or hierarchical search. Especially very large databases with more than 60,000 images (such as casimage) can instantly profit from a new way of browsing with automatically extracted visual

features. Evaluation of retrieval based on the exactly same diagnoses is not necessary for navigation in teaching files and thus current quality is sufficient for this kind of research.

The RSNA already created the **MIRC** (Medical Imaging Resource Centre, http://mirc.rsna.org/) standard for sharing image data for teaching. Currently, several large databases can be queried via a web page by textual queries. It would be very interesting to index all these images in a visual form and extend the MIRC standard to allow visual queries with example images. This is a very useful extension of the MIRC protocol.

Research

In research, the situation is very similar to teaching. The quality of retrieval does not always have to be on a diagnosis level and a little time for browsing can be spent to optimize query results. Content-based methods can be used in a variety of applications to complement text. They are an option for the retrieval of images of a certain kind to be included into a study. Visual access can also be used as a quality control to find images that might have been misclassified. Images of newly discovered diseases can attempt to be found in old databases even when it was not clear how exactly they were indexed textually. Visual features can be included directly into medical studies. What are the visual features that patients of certain diseases in a certain stadium of the disease have? Data mining in the visual features can be used to find potentially important visual characteristics of diseases or visual differences between diseases. One goal is a real **visual knowledge management**, where images and associated textual data can be analyzed together. Multimedia data mining can lead to unknown links between visual features and diagnoses or other patient demographics. The implicit information that is stored in an image, the textual case description, treatment and outcome can consequently be used to improve future, similar cases.

Diagnostic aid

Most systems that are currently described in the literature are tools destined for diagnostic aid. Visual features have been used to aid lung diagnostics [12], to classify pathology slides [22], and melanoma images [23], and for many more applications. Figure 1 shows an example for an application of image retrieval as a diagnostic aid for lung CT images with a typical result.

(Figure 1)

An image retrieval system can help in places where the diagnosis depends strongly on direct visual properties of images in the context of evidence-based medicine [24] or case-based reasoning [25]. Main problem is the evaluation of systems for diagnostic aid. Most often, only a very small database is extracted and systems are optimized on this database and then evaluated. This cannot lead to good results as the algorithms need to have a much larger base for the optimization, otherwise algorithms will not work on other images as they are too specialized. Another "problem" is the advancement of medical imaging techniques. New techniques deliver other, often better images. For automatic retrieval and analysis of images this means that the algorithms might not work with the new images in the same way they did on the old ones.

There is a clear need for tools that can easily be adapted for various fields of applications and that can learn the features based on a new group of images and imaging techniques. Systems need to be frameworks of reusable components where each component can easily be replaced. The basis for evaluation is the availability of image databases and ground truth for various tasks. Initiatives for such evaluation databases are underway in various organizations [26]. The need for standard data sets cannot be underestimated. In areas such as text retrieval, standard test sets and databases have lead to a significant improvement in retrieval quality. Problem and importance of reference databases have already been discussed in the 1970s [27]. Research in content-based medical image

retrieval can profit largely from such datasets. Ground truth needs to be available to advance research through a comparison of techniques on the same basis.

PACS and electronic patient record

Of course, the goal of image retrieval has to be an integration of content-based access into various clinical applications such as the PACS and the electronic patient record. This has already been proposed several times [10,11,17]. Still, the main problem of integration into PACS is the sheer amount of data that is being produced in hospitals. Without a proper selection algorithm for cases and slices, the indexed data will quickly become unmanageable, especially when using modern multi-slice devices that produce hundreds or thousands of images in a single series. Often, these problems are neglected in the literature.

Integration into the electronic patient record and access to all cases via content-based retrieval as well as textual retrieval would of course be an ultimate solution to be able to use all implicit knowledge being stored in the images and their accompanying textual information. Still, for such a scenario, multiple problems will need to be solved and appropriate retrieval algorithms for all sorts of images need to be implemented. Problems will also include privacy of the patients as their treatment data is used to improve the treatment of new cases.

Limits of automatic visual retrieval

Many possibilities of image retrieval have already been discussed but there are also several limits and problems. Most problems are linked to the low-level visual features being used. The system does not know its limits and search is not semantic but based on broad visual appearance, only.

Figure 2 is an example with limited retrieval quality. Although the first retrieved image is relevant, among the next images shown on screen are several that are not at all scintigraphic images. This is due to the rather unsharp lines and the light grey background color. Other images with the same grey background and similarly unsharp objects are found.

[Figure 2]

Only with relevance feedback, marking several images as relevant or non-relevant, the system can refine the search and find only scintigraphic images The system does not know which part is a-priori the most important part of the image for the user. Only when feeding back several images, the system can adapt to the user needs.

The medGIFT retrieval system

MedGIFT was developed to work together with casimage (http://www.casimage.com/), a radiological teaching file that has been in daily routine use for several years now [21,28]. More than 60,000 images from more than 10,000 medical cases have been indexed, so far. The database is available on the Intranet of the hospital, with a smaller database being publicly available via Internet and MIRC. MedGIFT itself is an image retrieval engine [29]. It is based on the open source system GIFT (http://www.gnu.org/software/gift/), outcome of the viper project of the University of Geneva (http://viper.unige.ch/). This system offers components for content-based indexing and retrieval of images such as feature extraction algorithms, feature indexing structures and a communication interface called MRML (Multimedia Retrieval Mark-up Language, http://www.mrml.net/). The interface allows for an easy integration into various applications such as a teaching file, document management systems or tools for diagnostic aid. GIFT uses techniques from text retrieval such as frequency-based feature weights, inverted file indexing structures and relevance feedback mechanisms [30]. Frequency-based feature weights mean that the importance of visual features is determined by their frequency in the image and by the frequency in the entire collection of all images, similar to the weighting of words in text retrieval engines. Rare words contain more information and are more important than frequent words. The inverted file structure is also commonly used in textual search engines such as google. Inverted file means that the index is not based on documents that refer to features (or words) but on the features that point to the

documents in which they appear. This is in analogy to google where an index of all words exists and for each word a list of web pages that contain it. As visual features to represent images, four feature groups are chosen:

- local and global texture features based on responses of Gabor filters;
- color/grey scale characteristics on a global image scale and locally within image regions.

Gabor filters measure the change in the image in a certain direction and scale. This means that it describes a texture with respect to its directions as well as with respect to the size and strength. Small or slow changes can easily be distinguished from quick and large changes. Local features are obtained by successively dividing the image into four regions of the same size and extracting the mode color of each region. This creates a multi-scale representation of the image. Figure 3 shows an example of such an image representation with several blocks at various scales.

[Figure 3]

Local Gabor filters allow determining in which region which shapes or textures occur. The potential feature space is very large (85,000 possible visual features). Each image contains roughly 1'000-2'000 features. Frequencies of visual features are similar to frequencies of words in text. The weighting scheme consequently weights rare features higher than frequent features in analogy with text search. More details on the GIFT technology can be found in [30].

To improve results with medical images that are primarily in grey scale, the number of grey levels was increased from the 4 grey levels of GIFT. For color photographs, the grey levels are shown to be unimportant for retrieval as the human visual system is less sensitive to them than to colors. The number of texture descriptors based on Gabor filter responses was equally raised as textures are expected to be more important in medical images than in color photography. Best overall results in first tests were obtained when using ~4-16 grey levels which is a surprisingly small number. To fine-tune the number of grey levels, we simply index the database in various ways and then

evaluated the results for each quantization against a gold standard defined by a radiologist [31]. More tests are needed to define an optimal number of grey levels for each query task. A much larger number of grey levels seems to create too specific queries and miss relevant images. This is far from the 256 gray levels that JPEG offers and even further from the resolution of CR or DR images in DICOM. Still, low-level features for retrieval work better when this information is less specific. A new user interface based on php (http://www.php.org/), a scripting language to create web-based interactive applications, was developed showing the diagnosis of retrieved images under the image thumbnail and is linked with the teaching file as can be seen in Figure 4.

[Figure 4]

The retrieval engine allows submitting an unlimited number of images combined as query and also images as negative examples or negative feedback to focus and refine the research further. On screen, the images are sorted by visual similarity to the query image(s) and the similarity score is displayed under the image. The diagnosis and the level of similarity are also shown. When clicking on one of the images shown, the case database is opened with the corresponding case, including a textual description and further images in full resolution. The system is an interactive tool, which means that response times need to be below 1 second [32]. On a current Pentium 4 computer with 2.8 GHz using a database of 9'000 images the response times for one-image-queries are always below 0.5 seconds.

Discussion

The examples and the description of the casimage/medGIFT combination have shown that content-based image retrieval can be used for the management of medical image data. Still, there are several open questions and problems to solve. One important question is the evaluation of medical image retrieval systems using textual as well as visual retrieval. A benchmarking event for image retrieval system comparison has been established at the CLEF conference (Cross Language

Evaluation Forum). A medical image retrieval task was added in 2004 and 11 research groups from Europe, North America and Asia participated. A first evaluation of our system shows that when using one step of relevance feedback, on average 14 of the first 20 images are relevant which shows that the technology can be used in non-critical domains such as the search of interesting cases for teaching.

Current applications are often extremely specialized for a very small application domain and hard to adapt to new requirements and new types of images, or they are extremely general without the possibility to use them for diagnosis-based (specialized) retrieval. New image retrieval projects will need to be based on common platforms to allow on the one hand the important specialization for clinical domains using as much a-priori information as possible and on the other hand very general retrieval in PACS-like databases or teaching files with a large variety of images. Not only a single research group should share such a platform. Also between several research groups such a sharing of source code should be done so new technologies and features can easily be implemented and also be compared. Reimplementation of basic functionalities needs to be avoided. Specialization is very important to get applications working and into clinical routine for tasks where the radiologists can profit from the help. Extremely important is the evaluation of algorithms on real-life data. Evaluation databases will need to be generated for specialized retrieval, including ground truth for the task being evaluated. The importance of evaluation cannot be underestimated. Projects for the identification of interesting medical imaging problems and on the generation of reference image data sets are underway in the US and Europe [26]. One important factor for image retrieval research is also the evaluation of the user behavior with a retrieval system. It is important to adapt systems to user needs using interfaces that the users accept [33].

With respect to content-based data access it is important to explain the technology, its potential and problems so expectations are realistic and users are not promised perfect retrieval results. System

improvements are only possible through several loops of feedback to include as much medical knowledge as possible into the retrieval engine for a certain task. Close cooperation between radiologists and computer scientists will be necessary for successful projects.

Conclusion

Image retrieval systems in the medical domain are in the process of getting into first applications to complement the conventional text-based search. They allow accessing and navigating in extremely large visual archives and extract hidden information without the high cost of manual annotation and codification of databases. Still, visual access to databases will stay a complement to text-based search and will at least in the foreseeable future not replace it. It is important that the two are developed closely together. Still, to get acceptance in the clinical domain, there is a clear need for real clinical applications that use content-based access mechanisms. Only working clinical applications will help to get acceptance in the medical community for more than "playing" with a retrieval system. To achieve this, systems will need to include as much medical knowledge as possible. A very close cooperation between medical practitioners and medical computer scientists will be necessary to achieve this goal. Promising applications will need to be identified and implemented based on a framework of components for image retrieval, so redevelopments of software are avoided and easy adaptation of the software will be possible.

Acknowledgements

Thanks to Julien Vignali for finishing our user interface in time for the RSNA presentation.

References

[1] Smeulders AWM, Worring M, Santini S, Gupta A, Jain R, Content-Based Image Retrieval at the End of the Early Years, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (12):1349-1380, 2000.

[2] Enser PGB, Pictorial information retrieval. (Progress in Documentation) Journal of Documentation 51(2), 1995, 126-170.

[3] Flickner M, Sawhney H, Niblack W et al. Query by Image and Video Content: The QBIC system. IEEE Computer, 28(9):23-32, September 1995.

[4] Carson C, Thomas M, Belongie S, Hellerstein JM, Malik J, Blobworld: A System for Region-Based Image Indexing and Retrieval, Conference on Visual Information Systems, Amsterdam, The Netherlands, pp. 509-516, 1999

[5] Westerveld T, Image Retrieval: Content versus Context, In Content-Based Multimedia Information Access, RIAO 2000, pp 276-284, Paris, France, 2000.

[6] Vetterli M, Kovacevic J, Wavelets and subband coding, Prentice Hall, 2000.

[7] Lowe HJ, Antipov I, Hersh W, Arnott Smith C. Towards knowledge-based retrieval of medical images. the role of semantic indexing, image content representation and knowledge-based retrieval. In Proceedings of the Annual Symposium of the American Society for Medical Informatics (AMIA), pages882-886, Nashville, TN, USA, October 1998.

[8] Tagare HD, Jaffe C, Duncan J. Medical image databases: A content-based retrieval approach. Journal of the American Medical Informatics Association, 4(3):184- 198, 1997.

[9] El-Kwae E, Xu H, and Kabuka MR. Content-based retrieval in picture archiving and communication systems. Journal of Digital Imaging, 13(2):70-81, 2000.

[10] Orphanoudakis S, Chronaki C, Vamvaka D. I2C net: Content-based Similarity Search in Geographically Distributed Repositories of Medical Images. Computerized Medical Imaging and Graphics, 20 (4): 193-207, 1996.

[11] Ogiela MR, Tadeusiewicz R, Semantic-oriented syntactic algorithms for content recognition and understanding of images in medical databases, IEEE international conference on Multimedia and Expo, Tokyo, Japan, 2001.

[12] Shyu CR, Brodley CE, Kak AC, Kosaka A, Aisen AM, Broderick LS. ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases. Computer Vision and Image Understanding, 75(1/2):111-132, July/August 1999.

[13] Aisen AM, Broderick LS, Winer-Muram H, Brodley CE, Kak AC, Pavlopoulou C, Dy J, Shyu CR, and Marchiori A. Automated storage and retrieval of thin-section CT images to assist diagnosis: System description and preliminary assessment. Radiology, 228:265-270, 2003.

[14] Lehmann TM, Güld MO, Thies C, Fischer B, Spitzer K, Keysers D, Ney H, Kohnen M, Schubert H, Wein BB, Content-based image retrieval in medical applications, Methods of Information in Medicine – to appear.

[15] Lehmann TM, Schubert H, Keysers D, Kohnen M, Wein BB. The IRMA code for unique classification of medical images. Proceedings SPIE 5033, 2003.

[16] Müller H, Michoux N, Bandon D, Geissbuhler A, A review of content-based image retrieval applications–clinical benefits and future directions, International Journal of Medical Informatics 73:1-23, 2004.

[17] Le Bozec C, Jaulent MC, Zapletal E, Degoulet P, Unified Modeling Language and Design of a Case-Based Retrieval System in Medical Imaging, AMIA Annual Symposium 1998, Nashville, TN, USA, 1998.

[18] Müller H, Geissbuhler A, Ruch P, Report on the CLEF Experiment: Combining Image and Multi-Lingual Search for Medical Image Retrieval, Proceedings of the 2004 Cross Language Evaluation Forum, Bath, England, 2004.

[19] Clough P, Sanderson M, Müller H, The CLEF Cross Language Image Retrieval Track (ImageCLEF) 2004, Springer Lecture Notes in Computer Science 3115, pages 243-251, 2004.

[20] Bucci G, Cagnoni S, de Dominicis R, Integrating content-based image retrieval in a medical reference database, Computerized Medical Imaging and Graphics, 20(4):231-241, 1996.

[21] Müller H, Rosset A, Vallée JP, Geissbuhler A. Integrating content-based visual access methods into a medical case database. In Proceedings of the Medical Informatics Europe Conference (MIE 2003), St. Malo, France, May 2003.

[22] Tang LH, Hanka R, Ip HHS, Lam R, Extraction of semantic features of histological images for content-based retrieval of images, IEEE Symposium on computer-based medical systems (CBMS), Houston, TX, USA, 2000.

[23] Schmid-Saugeon P, Guillod J, Thiran JP, Towards a computer-aided diagnosis system for pigmented skin lesions, Computerized Medical Imaging and Graphics 27:65-78, 2003.

[24] Bui AAT, Taira RK, Dioniso JD, Aberle DR, El-Saden S, Kangarloo H, Evidence-based radiology – requirements for electronic access, Academic Radiology 9:662-669, 2002.

[25] Abidi SSA, Manickam S, Leveraging XML-based medical records to extract experimental clinical knowledge – an automated approach to generate cases for medical case-based reasoning systems, International Journal of Medical Informatics 68:187-203, 2002.

[26] Horsch A, Prinz M, Schneider S, Sipilä O, Spinnler K, Vallee JP, Verdonck-de Leeuw I, Vogl R, Wittenberg T, Zahlmann G, Establishing an International Reference Image Database for Research and Development in Medical Image Processing, Methods of Information in Medicine 43, pp. 409-415, 2004.

[27] Sparck Jones K, van Rijsbergen CJ, Report on the need for and the provision of an ideal information retrieval test collection, British Library Research and Development Report 5266, Computer Laboratory, University of Cambridge, 1975.

[33] Rosset A, Ratib O, Geissbuhler A, Vallée, JP. Integration of a multimedia teaching and reference database in a PACS environment. RadioGraphics, 22(6):1567-1577, 2002.

[28] Rosset A, Müller H, Martins M, Dfouni N, Vallée JP, Ratib O, Casimage Project–a digital teaching files authoring environment, Journal of Thoracic Imaging 19(2):1-6, 2004.

[30] Squire DMG, Müller H, Müller W, Marchand-Maillet S, Pun T, Design and Evaluation of a Content-based Image Retrieval System, In Syed M Rahman ed., Design and Management of Multimedia Information Systems: Opportunities and Challenges, 7, pp. 125-151, Idea Group Publishing, 2001.

[31] Müller H, Rosset A, Vallée JP, Terrier F, Geissbuhler A, A reference data set for the evaluation of medical image retrieval systems, International Journal on Computerized Medical Imaging and Graphics, 28, pp. 295-305, 2004.

[32] Nielsen J, Usability Engineering, Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 1993.

[33] Kushniruk AW, Patel VL, Cognitive and usability engineering methods for the evaluation of clinical information systems, Journal of Biomedical Informatics 37(1):56-76, 2004.

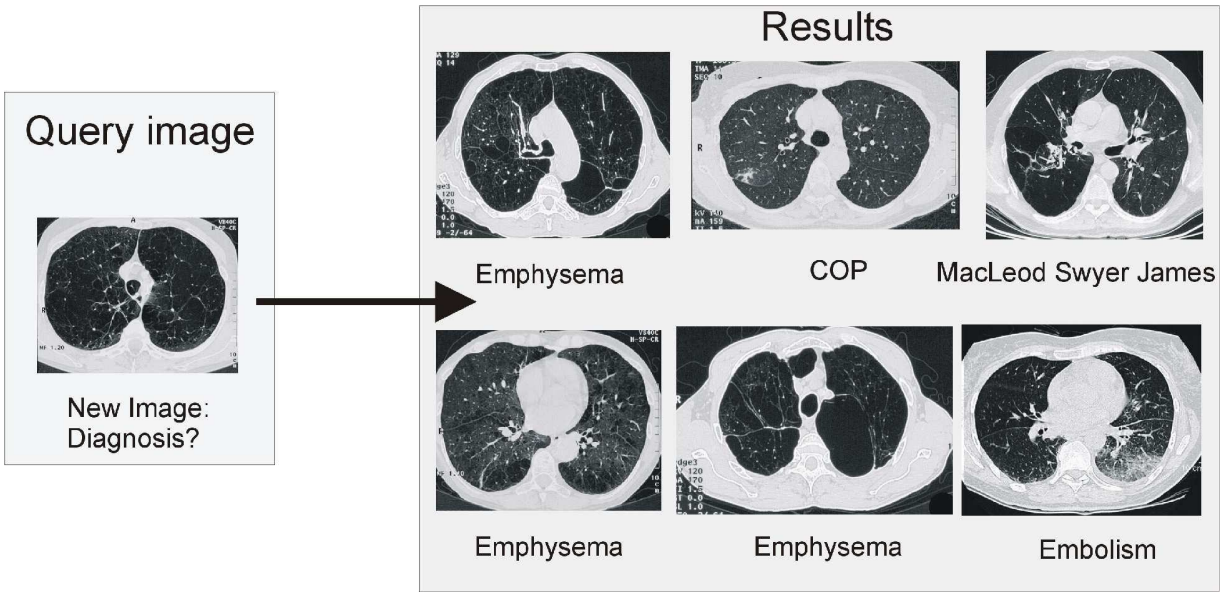


Figure 1: Content-based image retrieval as a diagnostic aid using MedGift and the database Casimage. A new image is presented and the results are shown on the right screen with their diagnoses and a link to the complete case description. The query image on the left shows emphysematous lesions with multiple, confluent, centrilobular and paraseptal areas of low attenuation without visible walls. The resulting search proposes 5 cases of Emphysema including 1 case of an unilateral emphysema (MacLeod Swyer Syndrome) and 2 cases of small area of consolidation in the pulmonary parenchyma (COP and pulmonary embolism). The typical pattern of pulmonary parenchyma destruction observed in the 5 cases of emphysema strongly suggests the diagnosis of emphysema for the query image. COP= ryptogenic Organized Pneumonia.

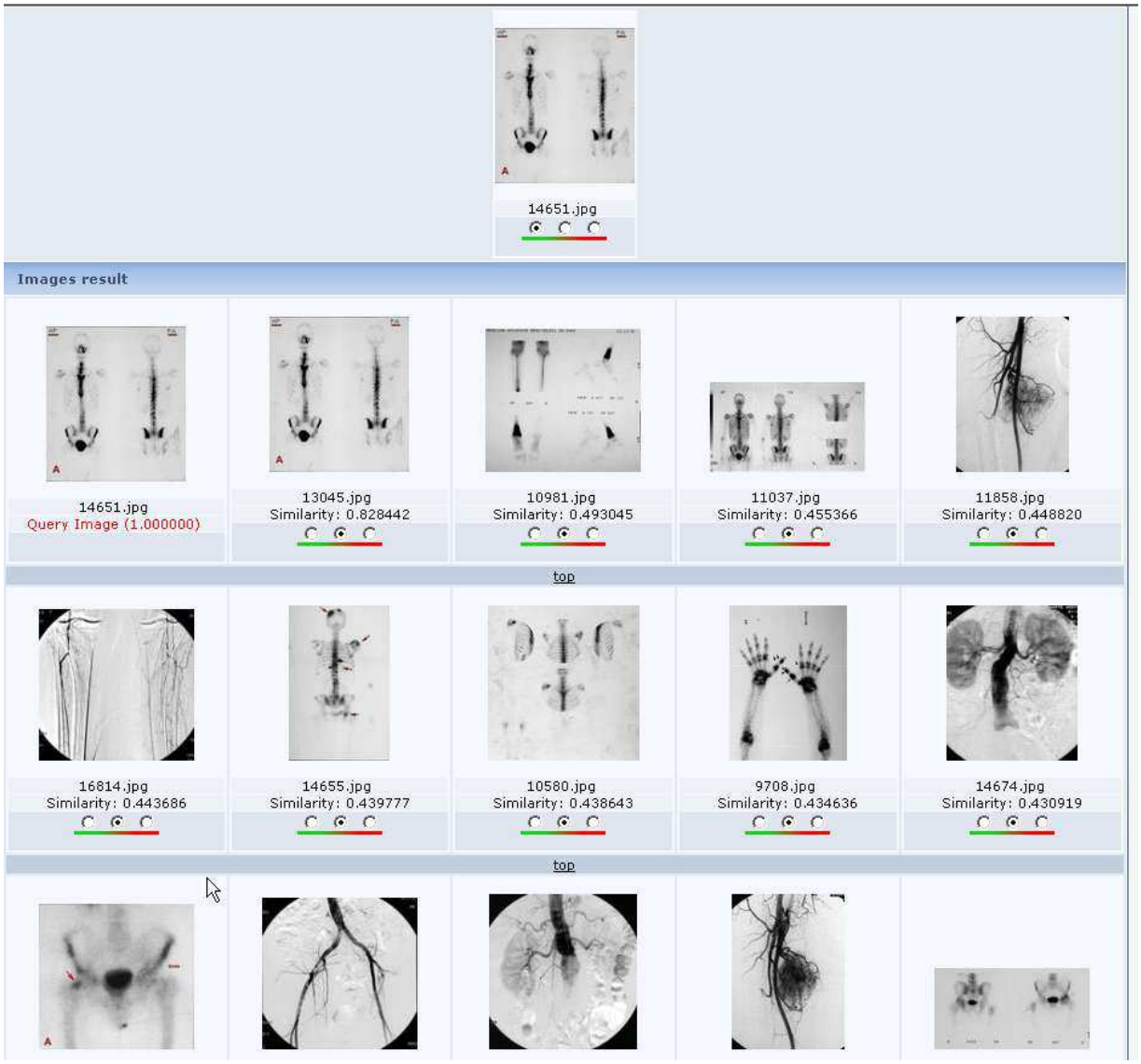


Figure 2: A retrieval result where the retrieval partly fails because the query does not contain much information with respect to varying grey level changes or strong textures. There is no sharply lined object in the image, which would ease retrieval.

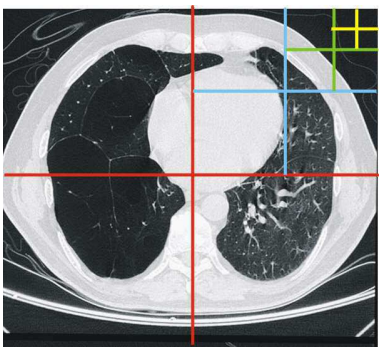


Figure 3: The image is first partitioned into four equally sized regions and this is repeated for each of the sub regions to extract local image characteristics.

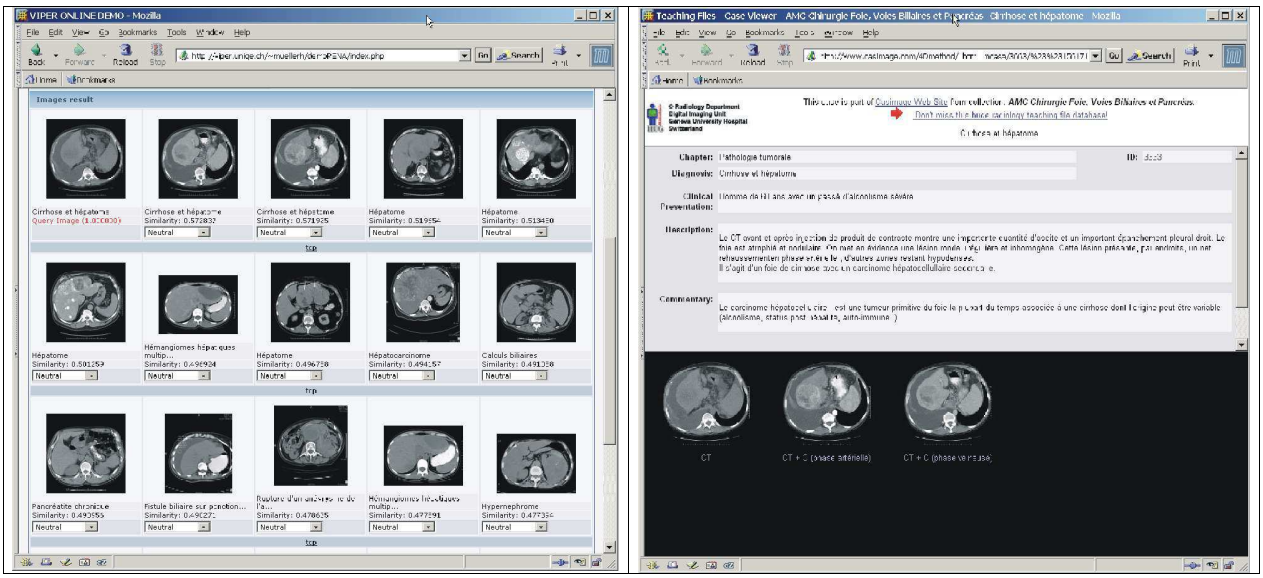


Figure 4: The interface of medGIFT and the corresponding textual CasImage case description when clicking on an image in the medGIFT interface.

Listing of Figure legends:

Figure 1: Content-based image retrieval as a diagnostic aid using MedGift and the database Casimage. A new image is presented and the results are shown on the right screen with their diagnoses and a link to the complete case description. The query image on the left shows emphysematous lesions with multiple, confluent, centrilobular and paraseptal areas of low attenuation without visible walls. The resulting search proposes 5 cases of Emphysema including 1 case of an unilateral emphysema (MacLeod Swyer Syndrome) and 2 cases of small area of consolidation in the pulmonary parenchyma (COP and pulmonary embolism). The typical pattern of pulmonary parenchyma destruction observed in the 5 cases of emphysema strongly suggests the diagnosis of emphysema for the query image. COP = Cryptogenic Organized Pneumonia.

Figure 2: A retrieval result where the retrieval partly fails because the query does not contain much information with respect to varying grey level changes or strong textures. There is no sharply lined object in the image, which would ease retrieval.

Figure 3: The image is first partitioned into four equally sized regions and this is repeated for each of the sub regions to extract local image characteristics.

Figure 4: The interface of medGIFT and the corresponding textual CasImage case description when clicking on an image in the medGIFT interface.

A.2 Building an Enterprise-wide PACS for all diagnostic images

D. Bandon, P. Trolliard, A. Garcia, C. Lovis, A. Geissbühler, and J.-P. Vallée.

In CARS, pages 279–284, 2004.

Building an Enterprise-wide PACS for all diagnostic images

David Bandon[a,*], Patrice Trolliard[a], Arnaud Garcia[a], Christian Lovis[a], Antoine Geissbühler[a], Jean-Paul Vallée[a]

[a] *Service of Medical Informatics, University Hospitals of Geneva, Geneva, Switzerland*

Abstract. PACS has the vocation to manage all diagnostic images acquired within the hospital. To address the various situations encountered in the imaging services, the traditional architecture used for radiological image has to evolve. We present in this paper our preliminarily results towards an enterprise wide PACS. Our solution is based on an existing radiological PACS system from which images are distributed through an electronic medical record to the entire hospital. This platform is enriched with a flexible integration framework supporting DICOM and DICOM-XML formats. In addition, a generic workflow engine highly customisable is used to define the work processes. Cardiology and haematology are the first implemented applications. The global strategy for further development is then described based on three possible architectures for an enterprise-PACS.

Keywords: PACS, Diagnostic Images, Image distribution, EMR, Workflow engine

1. Introduction

PACS is nowadays well accepted by the radiologist community and is widely implemented. The PACS industry is mature and offers archiving solutions and reading stations that fulfill the needs of the users. Moreover, IHE provides a solid framework to ensure a good integration between the different information systems (HIS, RIS, EMR). The issue at stake is now outside the radiology.

With the progress of the diagnosis and treatments, the patient is taken in charge by a continuously increasing number of different medical specialties and as a result the quantity and variety of the information generated explodes. The distribution of the patient information between departments is therefore a critical point and sets a new challenge for the information systems. We can expect that better information in terms of quality and quantity will give new perspectives to the physicians and will ultimately

* Corresponding author. *E-mail address*: david.bandon@sim.hcuge.ch.

improve the outcome for the patient.

The expectation for information systems is to address all the hospital wide imaging acquisitions. Indeed there are many specialties beyond radiology that produce body images. Those specialties understand the benefit from the digital storage of their images within a unified system. The PACS can provide help for the patient follow-ups and the creation of teaching materials. We have spotted in our institution the following specialties as targets for this extended PACS: cardiology, dermatology, ophthalmology, surgery, oto rhino laryngology, hematology pathology and gastroenterology. So far those specialties have only had at their disposal very specialized solutions that can manage images only locally without any real perspective for distribution inter-departments. For example, in the pathology segment, an imaging system we have found the need to connect a network of microscopes in order to provide report tools, image displaying and processing facilities.

In place of the currently large number of information systems addressing the specific needs of each specialty we envisage a unified PACS storing all diagnostic images acquired within the hospital and a unique distribution point via the EMR application. The challenge is to develop a common system that is flexible enough to address the specificities of each specialties. The majority of the commercial PACS have been developed for radiology and require some structural evolution to fulfill this ambition. The PACS has definitely vocation to extend beyond the radiological images to become the common platform for all diagnostic images distribution hospital wide. We will depict here initial practical achievements in that direction.

2. Methods

We take advantage of a running PACS [1] and a home-made Electronic Medical Record (EMR).

2.1 A hospital-wide PACS for radiological images

Our core imaging facilities rely on a commercial PACS system (supplier: [2]) deployed within the radiology department in 1999 and on a hospital wide distribution through our home-made EMR application (Electronic Medical Record). Currently, this system stores 15'000 images daily delivered by 33 modalities on a three-tier hierarchical storage: a 700 gigabyte magnetic disk cache, a 5-terabyte near-line archive based on optical disk and a 17-terabyte tape backup. As a result, 95% of image production is digital and effectively stored on the PACS.

2.2 A federative EMR

We logically promote our EMR solution called DPI (*Dossier Patient Intégré,* [3]), as a federated access to all medical information. Image access is provided since 2002 with other clinical information and lab results in an seamless integrated way. A Java applet image viewer is directly launched and fully embedded in the EMR user interface. Technically speaking, the image retrieval mechanism is performed in two steps:

- use of informational services to document the radiological record content. The supplied imaging list is displayed to user for study selection within the EMR;
- launching of a java applet viewer implemented by our PACS supplier. This java viewer loads images from an image repository acting as a web server.

To transform our current PACS in an enterprise-PACS, we introduced an open integration framework and a generic workflow engine.

2.3 An integration framework relying on DICOM and a DICOM-XML format

We mainly rely on the DICOM standard in radiology. However, we adopted a lighter format based on XML for other specialties since many imaging suppliers lack the support for DICOM. The rationale is that a DICOM-XML like format is more easily accepted because it is easier to implement. Moreover it is compliant to the federative middleware applied within the institution [4]. Basically this format represents a simplified DICOM information model: each XML tag matches a corresponding DICOM data element but we only retain the most relevant information (see fig. 1). Images are sent enclosed as JPEG or TIFF formats. We implemented some DICOM-XML to DICOM converter brokers to communicate with our PACS. We intend to support motion images in the future through DICOM Visible Light supplement.

2.4 A generic workflow engine to fit the specialties needs

Since the medical specialties adopts various ways of working with images, we needed to implement different workflows. To reach that goal, we acquired a Business Process Management System ([5], supplier: [6]) acting as a generic workflow engine highly customisable. Each specialty corresponds to its own imaging workflow. A workflow process specifies the tasks to be done by the involved actors – technicians, physicians - as well as the notification to be sent to the different systems.

A new workflow is implemented in a loop cycle: process definition graphically designed by a functional expert (see fig. 2), implementation (integration connector, database connections), process deployment through the workflow engine. The last step, workflow statistical analysis, is used to highlight the potential bottlenecks to be adjusted in the process design.

```
<DICOM_XML>
    <APPLICATION>Application name</APPLICATION>
    <PROFILE>Imaging service: haematology</PROFILE>
    <PATIENT>
        <DE GRP="10" ELT="10">Last Name^First Name</DE>
        <DE GRP="10" ELT="20">Patient Id</DE>
        <STUDY>
            <DE GRP="8" ELT="20">Study date</DE>
            <DE GRP="8" ELT="50">Accession Number</DE>
            <DE GRP="8" ELT="1030">Study desc. Hemogram</DE>
            <SERIE>
                <DE GRP="8" ELT="60">Modality OT</DE>
                <DE GRP="8" ELT="21">Series date</DE>
                <DE GRP="8" ELT="103e">Series desc.:color blue</DE>
                <DE GRP="20" ELT="11">Series number: 1</DE>
                <IMAGE>
                    <DE GRP="8" ELT="22">Acquistion date</DE>
                    <DE GRP="7fe0" ELT="10">Path to image file</DE>
                </IMAGE>
```

Fig. 1. Simplified example of the DICOM-XML structure used.

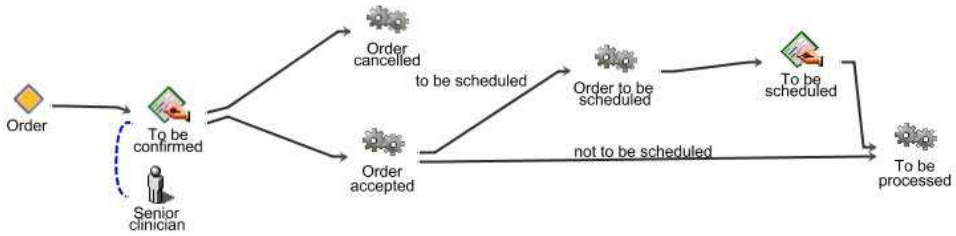


Fig 2. Example of a workflow design

3. Results

3.1 Cardiology

In 2003, the ultrasound unit acquired a dedicated mini PACS fully customized to their needs including all predefined measurement macros for cardiac imaging. This system has been interfaced to the workflow engine to enable a complete workflow cycle: order by clinician in the EMR, validation by cardiologists, scheduling, image loop acquisition on ultrasounds, report and image review, image transfer to the enterprise-PACS (see fig. 3). Since the workflow engine tracks all steps, the clinician is informed in real time of the examination state.

3.2 Haematology

We implemented a system allowing hematologists to generate JPEG images from a digital microscope, to open a patient record within the EMR and finally to link these images to a corresponding hemogram (graphic record of a differential blood count). Images are automatically uploaded, converted to DICOM (modality Secondary Capture) and pushed to the PACS. They finally become available for display through the EMR.

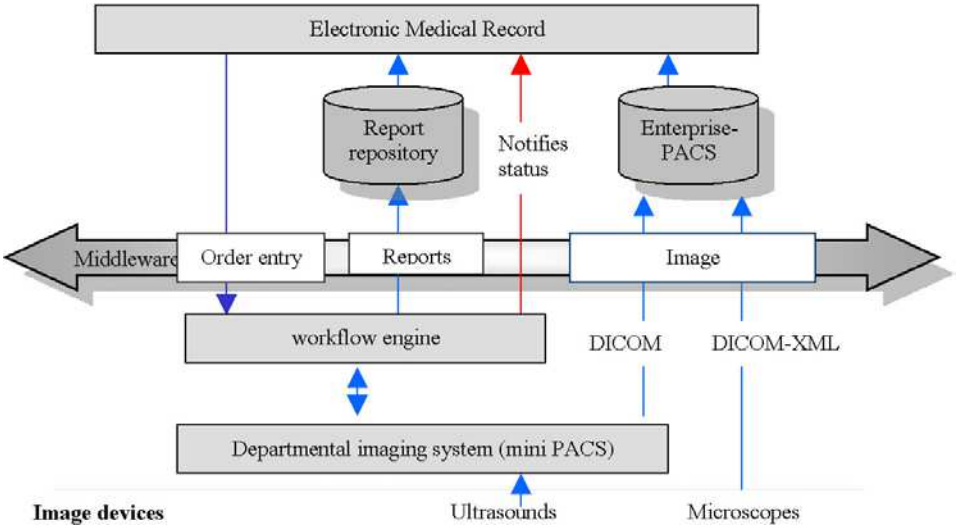


Fig. 3. Architecture used for cardiology (left) and haematology (right).

4. Discussion

An enterprise-PACS could be designed following three different architectures:

- A unique PACS system: this system manages all images produced within the hospital by directly connecting the image producers.
- A hierarchical model coupling multiple mini PACS to a master PACS: each mini-PACS fulfills the needs of its specialty and stores the most recent images while the master PACS archives in the long-term.
- Multiple PACS systems: each department owns its self-standing PACS. A search engine capable to query all storage units gathers the information pertaining to a patient from the different departments.

The unique system approach is the simplest architecture since it avoids the redundancy of storage and minimizes the integration tasks. However there is a major threat for robustness and performance issues due to the software complexity. The second architecture allows flexibility and specialization since each system matches a specific workflow. In the other hand, it burdens the integration and communication. Moreover, it is financially the least economic since it multiplies the number of different systems to be acquired. Finally, the last architecture directs the complexity towards the search engine.

So far we have implemented the second approach for pragmatic reason in order to cope with the heterogeneity of the installed basis in our institution. However our strategy in the future is an hybrid of the second and third architectures. We foresee a new architecture based on two enterprise-PACS systems, one dedicated to radiology and the second to the other specialties. This distribution on two systems will allow us to balance loads. A search engine – the portal as named by O.Ratib [7] - will be implemented to query the two systems. This solution will be interfaced to a workflow engine to

encompass the workflow specificities (scheduling, image management, report and diagnosis). Finally a unique system will be used to centrally manage the various access rights.

Additionally it seems important to us to elect the EMR as the main software piece to access the diagnostic image folder in a unified way. The EMR should supply a generic viewer software able to display all image types or motion image. To encompass the various needs, this viewer should call customized processing tools or display protocols via plugins.

Finally the question arises whether images have to be stored systematically in the DICOM format. DICOM will be the most appropriate format once the two following requirements will be fulfilled. Firstly DICOM should continue to evolve to support the video compression codecs [8] (MPEG, MJPEG) used for endoscopy. Secondly a lobbying pressure has to be exerted to get suppliers outside radiology effectively supporting DICOM. In the mean time a workaround based on DICOM converters can be conceived. A second alternative could be that the enterprise-PACS supports DICOM as well as the most frequent multimedia formats [9]. The data consistency would be ensured by keeping the DICOM arborescence (patient, study and series).

5. Conclusion

An enterprise-PACS offers real-added value for clinical activity since it allows a multidisciplinary image review. With image distribution, it reinforces the patient-centered EMR as the ideal platform to access the medical records in a unified and exhaustive ways. In a more global perspective we envisage to enlarge the scope of this enterprise-PACS. Physicians would have the possibility to associate any diagnostic images to a medical context (diagnosis, therapy planning, follow-up), some documents or to a predefined workflow stage.

Acknowledgments

The authors are grateful to people who strongly contributed to this project success: Andrew Parrott as system engineer, Michael Dehouck, Lydia Dorsaz, Cyrille Duret, and Marcelo Echeverria, for the implementation part.

References

[1] Bandon D, Lovis C., Jean-Paul Vallée J-P, Geissbuhler A, Terrier F , The global Image Management Strategy within the HUG, Swiss Medical Informatics (SMI), 2003, 51:37-42.
[2] Image Devices GmbH (a CERNER company). http://www.imagedev.com.
[3] Lovis C, Baud RH, Rassinoux AM, Scherrer JR. Value to add to the patient record when making the EPR. Iwg, editor. Eprimp, Rotterdam; 1998. 17:225-227.
[4] Geissbuhler A, Lovis C, Lamb A, Sphani S: Experience with an XML/HTTP-based federative approach to develop a hospital-wide clinical information system. Medinfo 2001; 10(Pt 1):735-9.
[5] *Fischer L.,* The Workflow Handbook 2004. Published by Future Strategies Inc., ISBN 0-9703509-6-1
[6] AKAZI. http://www.akazi.com
[7] Ratib O, Swiernik M, McCoy, From PACS to integrated EMR, Computerized Medical Imaging and Graphics, 2003, 27:207-215
[8] Moving Picture Experts Group. MPEG home page: *http://www.mpeg.org*
[9] The Graphics File Format Page. http://www.dcs.ed.ac.uk/home/mxr/gfx/2d-hi.html. Accessed March 10, 2004.

A.3 Grid Computing at the University Hospitals of Geneva

H. Müller, A. Garcia, J.-P. Vallée, and A. Geissbühler. Grid computing at the university hospitals of geneva.

In healthGrid Conference, 2003.

Abstract

Although no implementation with respect to grid technologies has taken place as of yet, the University Hospitals of Geneva plan to apply these technologies in several pilot studies. The implementations will first take place in research projects but grid networks are also seen as a strategic long-term technology. Grid-enabled applications can change the way the resources in a hospital are used. Computation on demand can soften performance peaks and allow new diagnostic tools. Storage grids can ease the exchange of information in a secure way, where all the partners have their view of the data that they are allowed to access and where distributed resources are seen as one single, large hard disk.
The fact that there are more than 5,400 computers in the University Hospitals of Geneva with most of them being used at most 5%-10% of the time shows how much potential there is to use the resources in a more efficient way. This article describes our vision for grids and the tools that need to be developed to use this technique in a hospital environment. The extension of this concept to the emerging community health information network is also considered.

Keywords: Grid computing, grid usage, image retrieval, text mining, hospital infrastructure

1. Introduction

In medical hospitals, the amount of information and especially digital information produced and used is rising strongly. In the Radiology Department of the University hospitals of Geneva, for example, the number of images produced daily (in 2002) is at more than 12,000 which results in more than 4 million images per year accounting to roughly 2 Terabytes of image data.
Other departments might not have quite as much data but it all still amounts to large quantities that have to be stored and also made accessible for example in the electronic patient records. Data comes from various sources in varying formats and is distributed using various information systems and protocols. The same is valid for the exchange of information with other institutions. Data have to be transmitted to other medical practitioners, to other hospitals and anonymized data also needs to be transferred to the public offices for statistics and to insurance companies for billing.
While this is mainly a data storage and communication problem, in several research fields as well as in imaging diagnostics, there is a need for powerful computational resources, regularly. In general, medical departments do not have the computing power that particle physics institutes such as the CERN[1] have as these can cost millions of Swiss Francs, which are not available for a hospital research computer infrastructure.
Grid networks [1,2] have the potential to help with these two problems. They are forecasted to be for computation what the Internet was for communication. A large set of computers can, in principle, be used as one large resource for the storage of data and for computation. With authentification of the users, access to computational and storage capacities can be granted or denied. Like this, virtual organisations can be created that share, depending on projects, the same resource for computing and storage. This can be the case on several levels, *i.e.* on a hospital level, a regional health care provider level and also on a project level, where resources from centres such as the CERN might be made available for certain computational tasks. Such a computation on demand can also open up new diagnostic methods and tools when speeding up procedures.
These virtual organisations to share data and computational power can also be imagined within research projects between several institutes.
Of course, the data also needs to be in a well-defined format such as XML to make the exchange possible and easy. Sharing of data alone does not necessarily help much. Proprietary applications need to be replaced by open ones.

[1] http://www.cern.ch/

2. Computer architectures in hospitals and the storage of data

Basically all hospitals have computerized information systems to store and access all sorts of data, from the administrative data such as goods purchase in the hospital information system to treatment data in clinical information systems and images in the PACS (Picture Archiving and Communication System). Various departmental information systems exist (RIS-Radiology, LIS- Laboratory).
The sources of the data are extremely varied and there exist information systems at various levels such as the systems for departments. Often, the various information systems for departments and on a global level are connected with communication interfaces using standard communication protocols such as HL7[2] for patient or diagnostic data or DICOM[3] [3] for image data. Although there exist standards describing these communication protocols, interfacing still remains a problem and many interfaces rest proprietary because the standards do not cover all the needed functionalities.
In a large hospital there are thousands of computers to access the stored data when and where they are needed. To have all the data accessible from all these computers, large integrated systems are needed to administer the integrated, electronic patient record.

2.1 University Hospitals of Geneva

The University Hospitals of Geneva[4] (HUG) have been pioneering the development of hospital information systems since the 1970s. The PACS system for image storage was one of the first of its kind in the early 1990s. By now, the entire image producing, diagnostics and storing process is digital.

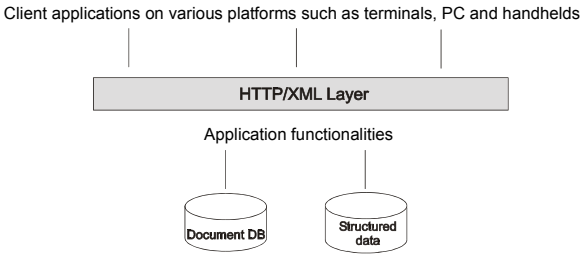


Figure 1: The model of the hospital information system at the University Hospitals of Geneva.

The information system has recently been redeveloped using a component architecture, with many technologies being based on web standards such as XML for describing data and HTTP for communication [4]. Figure 1 demonstrates this model when clients access functionalities via HTTP/XML

[2] http://www.hl7.org/
[3] http://medical.nema.org/
[4] http://www.hcuge.ch/

that have access to the various databases containing documents and structured patient data.
Such an architecture allows having a set of standard interfaces and eases the reuse of functionalities and the data interchange. Functionalities can be used by several subsystems in exactly the same way.
A tool that is used for many research projects internally is the image case database system Casimage[5] [5]. This system is routinely used by Medical Doctors (MDs) in the hospital to store interesting or typical cases with the corresponding images. Its main use is for research and for teaching but it can as well be used for diagnostics in domains such as evidence-based medicine or case-based reasoning. For this, new search facilities might be useful or even necessary.

3. Architecture for research in the medical environment

Departments of Medical Informatics in large University Hospitals often have a significant number of employees, in the case of Geneva, more than 50 people. Still, most activities are not for research but for service with respect to the hospital infrastructure such as the electronic patient record or the PACS system to store images. Often, the two are combined to transfer results from research into routine use.
Research groups in medical informatics work on a large variety of different projects from image processing to database technologies, text mining and security aspects or mobile computing. Many of these activities demand the storage and analysis of very large quantities of data and the computational need can as well be very high. Still, rarely, very powerful computer are accessible by these research departments as the focus of the departments work is often on the service of the important hospital infrastructure for diagnostics, data storage and data access rather than research.
For new research areas such as the integration of grids into the hospital information structure the means are often even more restricted. At the University Hospitals of Geneva, currently two persons are working to obtain knowledge on the implementation and use of grid technologies while developing pilot applications. All this is an additional workload to the usual charges and no dedicated person for applying these technologies exists. Other hospitals will most likely not have more resources for similar technologies and projects.

4. A vision of grids

Some of the ideas on grid networks in this section might yet be a bit naïve but they are not from a network expert's point of view but from a user standpoint and might omit some of the domain problems.
The big idea of coupling several single resources to create a larger one sounds very appealing. Especially the integration of strong authentification routines into this concept is very important for the medical field. Such an infrastructure can allow having a physical structure almost as it is at the moment, and at the same time the logical structure can be fairly different. Virtual organisations can be created in research projects or in other cooperations that share storage and/or computational resources. Figure 3 demonstrates this idea, where different institutions like the CERN, the University Hospitals of Geneva (HUG), statistics institutes and also other MDs can create a single grid network where depending on the access rights of a certain person or institution, different resources can be used and different data can be accessed. Normally, the logical structure is larger than the physical one as resources are shared between partners.
This can help in medical research projects, where the computational power of the CERN can be used for short periods of time and also the CERN that might need test users for their technologies, can profit from these tests. Statistics institutes can directly access the data in the hospital information system that they are allowed to use and the information exchange can get much easier as data do not need to be transmitted and be stored in different versions and with changes being made regularly. MDs can also access actual data which can help to avoid errors due to a bad information flow which are frequent in medicine [6,7] and according to several studies cause thousands of deaths per year in the US alone.

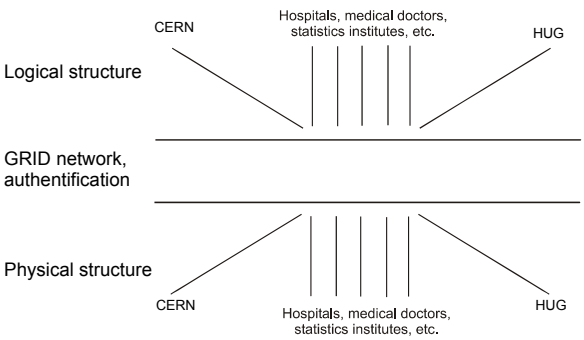


Figure 2: A grid network where the same players have a physical infrastructure and a larger logical structure including the access to all resources that they are permitted to use.

Such a common infrastructure can also ease other development projects such as a common Geneva health information system [8]. This can not only improve the quality of the data but also reduce development and data interchange costs and accelerate the development. Health professionals could efficiently share patient-related information, and decision makers would have easy access to up-to-date anonymized data.

[5] http://www.casimage.com/

Still, not all problems will be solved that easily. The various different data formats and many proprietary applications will stay to be used for a while and format converters will continue to be useful. But it is at least a first step to create an infrastructure for easy communication and sharing of computational capacities.

5. Useful applications of grid technologies in the University Hospitals of Geneva

This section compiles a few test applications that are planned to be "griddified" in a first step. These applications are chosen because of their relevance to current practice and the expected benefit that the grid can have for these specific applications. The applications already exist and they are relatively easy to parallelize. They can as well be tested with anonymized data so it is not necessary to have a grid cluster within the hospital but resources can be used elsewhere for some first tests.

5.1 Content-Based Image Retrieval

Content-based image retrieval is an extremely active research domain that is trying to develop tools and methods to manage the large amount of visual and multimedia data that are produced and made accessible via the Internet. In a hospital environment, the generation of multimedia data for diagnostics, teaching and therapy planning is also on a rise. These methods normally use automatically extracted visual features such as colours, contours and textures to describe and search images as contrast to other search methods, which are usually text-based. A good overview of the field with many references can be found in [9].

Figure 3: The interface of the medGIFT.

With the availability of the open source image retrieval system GIFT[6] (GNU Image Finding Tool) a tool is available to easily build content-based retrieval applications for several domains. This project was founded from the University of Geneva's Viper[7] project [10,11]. An overview of the architecture of the system can be found in [12]. The system is component-based and to access the retrieval engine, a communication protocol has been developed to separate interface and search engine and also to allow access for other applications such as meta search engines. The language for the communication between the interface/programs and the search engine is called MRML[8] (Multimedia Retrieval Markup Language). Figure 3 shows the interface of the program with the result of an example query.

The system can profit from both main aspects of grids, the storing of a large amount of visual data and the need for computing power to quickly return calculated retrieval results to the user. Most of the algorithms are fairly easy to parallelize which has already been implemented using PVM (Parallel Virtual Machine) in [13].

Main applications for the use in grids are the computationally intense extraction of visual features and as well the execution of queries that access the large amounts of visual features that are stored in an inverted file index for efficient searching. Our current image database contains more than 25,000 images which results in a feature index of more than 1 Gbyte.

The GIFT system is in the process of being integrated into the CasImage system, which is a medical case database system that includes image data and corresponding annotations. This will insure the routine use of this search technology and will create a need for fast indexing and quick responses to visual queries. CasImage is in routine, daily use in the hospital since more than two years.

5.2 Text Mining

Text mining in the medical environment poses different problems, as it is more a question of the large amount of data that needs to be analysed. Text mining techniques can be used on various medical data, mostly on the electronic patient record that contains large amounts of more or less structured data. The quality of the data is not always extremely high as much of the content is intended for internal use only, anyways. Thus algorithms to correct these data are one first step before further analysis and retrieval can take place [14].

These technologies can profit from the use of storage grids where large amounts of data can be accessed quickly and in a distributed way.

Text mining techniques can be used for text categorisation where concepts are automatically assigned to pieces of text. This often implies the use

[6] http://www.gnu.org/software/gift/

[7] http://viper.unige.ch/

[8] http://www.mrml.net/

of extremely large, controlled vocabularies that could also be used efficiently in storage grids.
Text categorisation is often used in medical informatics: for billing or decision-making purposes or for cross lingual information retrieval (UMLS, [15]). Most advanced categorisation approaches to handle large sets of categories (between 20,000 (MeSH) to 2,000,000 (UMLS)) cannot use machine-learning strategies (k-nearest neighbours, SVM, Linear Least Square Fit) for computational reasons. Therefore, the use of a GRID-like architecture could help in order to scale up such data-intensive approaches.

5.3 3D Image Processing

For 3D image processing or other image processing algorithms such as fusion of images from various modalities, large computational power is needed. Especially the idea to make these algorithms interactive needs access to fast computing resources.
To integrate these techniques even better into the diagnostic process, it is necessary to make them fast. Having much shorter reaction time for image fusion or functional MRI (fMRI) can also open up new diagnostic possibilities as images can be created and processed during an operation and not necessarily a day in advance. This can help to recognize the shift of organs in the body after opening and allow easier recognition and navigation, for example.
Many of these algorithms are relatively easy to parallelize. Segmentation algorithms can be executed on several machines slice-by-slice and only little communication is needed for optimisation. The same holds for 3D reconstruction where nodes can work on parts of the data and communicate the results for display.

6. Needs for grid technologies and steps towards an implementation

Although several projects on grid technologies exist and a large amount of software has been developed in these projects, there is not yet a free "out-of-the-box-solution". Most projects such as the Globus toolkit or the DIET software [16] are fairly difficult to install and work only in a very strictly defined environment on one operating system. Creating local miniGrids is thus fairly difficult and demands very experienced system administrators.
Current technologies are also not really compatible among each other. They are developed in projects and it is important for choosing one of the packages that they are continued to be supported in the long term. Otherwise, a project can be hindered strongly by a wrong choice of platform or middle ware.
The usage of running grids is not extremely easy, either. Normally, no easy-to-use application programming interface (API) exists for the current middle ware applications.

6.1 Needs for grid software to make it usable in a medical environment

There are a number of projects that produce software that can help to install and use grids. This includes the Globus[9] toolkit that is part of the European dataGrid[10] project.
The Unicore[11] forum is another option that allows downloading software to ease the development of grids.
Besides these developments what is really needed is an easy to use *application programming interface* (API) that allows the development of applications that use the resources a grid provides.
This would perfectly be usable on *different platforms* such as PCs under Windows and Linux, Apple computers and Unix systems from various vendors, so that existing, heterogeneous computer infrastructures with a large number of PCs (as they exist in the hospital) can be reused. Tools for *load balancing* are also indispensable to be able to distribute the charge on the machines depending on their computational power and the needed bandwidth to transmit data.
Extremely *fast communication links* are needed if a lot of data needs to be transmitted by the used algorithms
The *parallelisation* of existing algorithms also needs to be made as easy as possible so existing software can easily be transferred into efficient grid-applications. Partly this might be possible within compilers but for compilers there will definitely be limits and some manual intervention and program planning is for sure necessary.
The software to install and *manage grid nodes* also needs to be as easy as possible to install. A good option would also be to not only use specific grid computers for inclusion into a grid node but to be able to demand further resources among the PCs that are, for example, not used at a certain point. This can be somewhat similar to the SETI@home[12] project but would most likely require a much more complex structure to protect the data, for example.
Maybe the most important points to make grids usable in a medical environment are *security issues*! This goes beyond simple access right to data and encryption of the communication. In first steps, it is clear that only anonymized data can be used outside of the hospital environment to test applications on grids, even if all security mechanisms to protect the data do exist. Only after some confidence has been build up, it might be possible to use clinically relevant and personalized files for griddified

[9] http://www.globus.org/
[10] http://eu-datagrid.web.cern.ch/eu-datagrid/
[11] http://www.unicore.org/
[12] http://setiathome.ssl.berkeley.edu/

applications. Tests will need to be performed regularly to insure the security of grid systems.
For running large-scale grid applications these security mechanisms need to be well in place and tested.

6.2 Our further steps for implementation

Although our first goal was originally to build a miniGrid at the interior of the hospital to test technologies and couple a number of PCs for larger calculations, this idea has been omitted for now due to the large effort that would be necessary to maintain such a network.
Our first steps will thus be the *use of resources* from the CERN, which offered us this possibility. This will include the comparison of execution times of programs on normal PCs and in a grid environment. Later on, a *miniGrid* in the hospital seems reasonable that can be used for data storage and exchange and sharing of computational load in a couple of projects for research or for data exchange with statistical offices.
A first application to test in a grid environment is the pilot application for *content-based image retrieval*. This has several reasons. (1) The application is available in open source and thus relatively easy to change and adapt to grid specifications. (2) The application has two different parts that can both be parallelized separately, fairly easily. (3) The feature extraction and index generation phase can be extremely parallelized as the features are extracted from every single image in the database. The Casimage database currently contains 25,000 images that could be sent to 25,000 computers, each to extract the features of one image, in an extreme scenario for feature extraction. This could reduce the visual feature calculation phase from 14 hours (2 seconds per image) to a moderate 2 seconds plus a few seconds for all the data transmission.
The query phase is not as easy to parallelize but several computer can do part of the evaluation, and then, the results can be merged in the end. The gain for this will not be as high as the cost for merging is significantly high. Evaluations have to verify the gain or loss.
After this first test phase, the other pilot applications need to be implemented and also evaluated. Also the integration of the University Hospitals into a Geneva-wide medical information system to gather epidemiological data about the health system is an option, as proper data interchange between various partners can be eased by such a Geneva-wide grid network.

7. Conclusions

Grid technologies have a large potential in the medical domain as well as in others with respect to better use of existing resources but also for obtaining computation and storage on demand to make applications faster and data access easier.
This can not only be the case for medical research departments as it might be in a first phase but also in clinical routines where the availability of diagnostic tools in real time can help the decision making process and open up new possibilities. The storage of data within the hospital can also profit from the tools for secure storage and access to data and the exchange of information can be eased.
The idea to use all the 4,500 computers available for computationally intensive tasks when the computers are not used otherwise sounds very logical and can remove the need to spend much money on specialized high performance computers to realize quick response times. By simply using the existing resources more efficiently this money can be used elsewhere.
Still, much work needs to be done before grids can be programmed as easily as single computer systems at the moment. The middleware needs to be much easier to install grids and use the grid services within programs. A simple-to-use API needs to be created for such an easy access.
For the use in storage grids, a security infrastructure needs to be created that can subsequently be applied to the hospital so only anonymized data can get to the outside and only authorized personnel can access personal patient data.
Such a security and computing infrastructure can make the exchange of data within research projects, to official statistics offices and to other partners much easier than it is at the moment as this can create virtual organisations that share part of the data and/or part of the computing resources depending on projects or institutional links.

8. References

[1] I. Foster, C. Kesselman, Steven Tuecke, The Anatomy of the Grid – Enabling scalable virtual organisations, *International Journal on Supercomputer Applications*, 2001.

[2] F. Gagliardi, B. Jones, M. Reale, S. Burke, European Datagrid project: Experiences of deploying a large scale testbed for e-science applications, *Performance 2002*, pp. 480-500, 2002.

[3] B. Revet, DICOM Cookbook for implementations in modalities, *technical report Philips medical systems*, 1997.

[4] A. Geissbuhler, C. Lovis, A. Lamb, S. Spahni, Experience with an XML/http-based federative approach to develop a hospital-wide clinical information system, *Medinfo 2001*, IOS Press, Amsterdam, 2001.

[5] A. Rosset, O. Ratib, A. Geissbuhler and J-P. Vallée, Integration of a Multimedia Teaching and Reference Database in a PACS Environment, *Radiographics*, 22(6):1567-1577, 2002

[6] L. L. Leape, Error in medicine, *JAMA, The Journal of the American Medical Association* 272(23), pp. 1851-1861, 1994.

[7] L. T. Kohn, J. M. Corrigan, M. S. Donaldson, editors, To err is human – Building a Safer Health System, *National Academic Press*, Washington D.C., 1999.

[8] D. Hausser, Analyse stratégique de l'implantation du système de santé à Genève, *technical report*, IDHEAP, June 2002.

[9] AWM. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain, Content-Based Image Retrieval at the End of the Early Years, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(12):1349-1380, 2000.

[10] DMcG. Squire, W. Müller, H. Müller and J. Raki, Content-Based Query of Image Databases, Inspirations from Text Retrieval: Inverted Files, Frequency-Based Weights and Relevance Feedback, *Scandinavian Conference on Image Analysis (SCIA)*, pp. 143-149, Kangerlussuaq, Greenland, 1999

[11] D. McG. Squire, H. Müller, W. Müller, S. Marchand-Maillet and T. Pun, Design and evaluation of a content-based image retrieval system, *Chapter 7, Design and Management of Multimedia Information Systems: Opportunities and Challenges, Idea Group Publishing*, S. M. Rahman, editor, pp. 125-151, 2001.

[12] H. Müller, W. Müller, DMcG. Squire, Z. Pecenovic, S. Marchand-Maillet and T. Pun, An Open Framework for Distributed Multimedia Retrieval, *Computer Assisted Information Retrieval (RIAO)*, volume 1, pp. 701-712, Paris, 2000.

[13] H. Müller, D. McG. Squire, W. Müller, T. Pun, Efficient Access methods for content-based image retrieval with inverted files, *Multimedia Storage and Archival Systems IV, SPIE proceedings volume 3846,* Boston, Massachussetts, 1999.

[14] P. Ruch, R. Baud, A. Geissbuhler, Evaluating and reducing the effect of data corruption when applying bag of words approaches to medical records, *International Journal of Medical Informatics* 67, pp. 75-83, 2002.

[15] D. Eichmann, M. Ruiz, P. Srinivasan, Cross-Language Information Retrieval with the UMLS Metathesaurus, *Research and Development in Information Retrieval*, pp. 72-80, 1998.

[16] F. Desprez, DIET: A distributed interactive engineering toolbox for client-server applications in a grid environment, *ERCIM News*, No. 45, April 2001.

Bibliographie

[1] R. Adams and L. Bischof. Seeded region growing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(6) :641–647, 1994.

[2] A.Garcia, N.Boch, M. Kocher, N. Michoux, C.Constantin, F. Terrier, and J-P. Vallée. 3d watershed approach to volume estimation of hepatic tumors on mri. In *annual congress of the Swiss society of radiology*. Interlaken, Switzerland, May 2004.

[3] H. M. Al-Otum. Morphological operators for color image processing based on Mahalanobis distance measure. *Optical Engineering*, 42 :2595–2606, September 2003.

[4] J. Angulo and J. Chanussot. *Morphologie mathématique : approches non déterministes, mise en oeuvre et applications, Chapter 5, Couleurs et Images Multivariées*. L. Najman et H. Talbot, 2008.

[5] J. Angulo and J. Serra. Color segmentation by ordered mergings. In *ICIP03*, pages II : 125–128, 2003.

[6] J. Angulo and J. Serra. Morphological coding of color images by vector connected filters. In *In IEEE Proc. of the Seventh International Symposium on Signal Processing and Its Applications (ISSPA'2003)*, volume 1, pages 69–72, July 2003.

[7] J. Angulo and J. Serra. Modelling and segmentation of colour images in polar representations. *IVC*, 25(4) :475–495, April 2007.

[8] Jesús Angulo. *Morphologie mathématique et indexation d'images couleur. Application à la microscopie en biomédecine*. PhD thesis, Centre de Morphologie Mathématique, Ecole des Mines de Paris, France, December 2003.

[9] Jesús Angulo. Morphological colour operators in totally ordered lattices based on distances : Application to image filtering, enhancement and analysis. *Comput. Vis. Image Underst.*, 107(1-2) :56–73, 2007.

[10] E. Aptoula and S. Lefevre. A comparative study on multivariate mathematical morphology. *Pattern Recognition*, 2007. to appear.

[11] E. Aptoula, S. Lefevre, and C. Collet. Mathematical morphology applied to the segmentation and classification of galaxies in multispectral images. In *European Signal Processing Conference : EUSIPCO'06, September 4-8, Florence, Italy, 2006.*, 2006.

[12] Erchan Aptoula. Segmentation d'images astronomiques multi-bandes par morphologie mathématique. Master's thesis, Université Louis-Pasteur, Strasbourg, 2004/2005.

[13] Guy Auliac, christiane Cocozza-thivent, sophie Mercier, and Michel Rossignol. *Mathématiques - Intégration et probabilités - 3ème année*. EdiScience, 2005.

[14] David Bandon, Patrice Trolliard, Arnaud Garcia, Christian Lovis, Antoine Geissbühler, and Jean-Paul Vallée. Building an enterprise-wide pacs for all diagnostic images. In *CARS*, pages 279–284, 2004.

[15] V. Barnett. The ordering of multivariate data. *Journal of the Royal Statistical Society. Series A (General)*, 139(3) :318–355, 1976.

[16] Kent Beck. *Extreme Programming Explained*. Addison-Wesley Professional, 1999.

[17] J.A. Benediktsson, J.A. Palmason, and J.R. Sveinsson. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Trans. Geoscience and Remote Sensing*, 43(3) :480–491, March 2005.

[18] S. Beucher. *Segmentation d'images et morphologie mathématique*. PhD thesis, Ecole des Mines de Paris, Juin 1990.

[19] S. Beucher and C. Lantuéjoul. Use of watersheds in contour detection. *Proc. Int. Workshop on image processing, Rennes(France)*, pages 17–21, Sept. 1979.

[20] F. Bloch. Nuclear introduction. *Phys. Rev*, 70 :460–474, 1946.

[21] I. Bloch. Mathematical Morphology on Bipolar Fuzzy Sets. In *International Symposium on Mathematical Morphology (ISMM 2007)*, Rio de Janeiro, Brazil, 2007.

[22] I. Bloch and H. Maître. Constructing a fuzzy mathematical morphology : alternative ways. Technical report, Télécom-Paris 92C002, September 1992.

[23] Grady Booch, James Rumbaugh, and Ivar Jacobson. *Le guide de l'utilisateur UML*. Eyrolles, 2000.

[24] Yves Bouchard. *Le holisme épistémiologique dans la Critique de la raison pure de Kant*. PhD thesis, Université de Montréal, 1997.

[25] Jocelyn Chanussot. *Approches vectorielles ou marginales pour le traitement d'images multi-composantes*. PhD thesis, Université de Savoie (ESIA), 1998.

[26] Jocelyn Chanussot and Patrick Lambert. Total ordering based on space filling curves for multivalued morphology. In *ISMM '98 : Proceedings of the fourth international symposium on Mathematical morphology and its applications to image and signal processing*, pages 51–58, Norwell, MA, USA, 1998. Kluwer Academic Publishers.

[27] M. Chapron. A new chromatic edge detector used for color image segmentation. *IEEE International Conference on Pattern Recognition*, A :311–314, 1992.

[28] H. D. Cheng, X. H. Jiang, Y. Sun, and Jingli Wang. Color image segmentation : advances and prospects. *Pattern Recognition*, 34 :2259–2281, 2001.

[29] S. Clément-Guinaudeau, T. Goissen, J. Cousty, J.-L. Dubois-Randé, A. Rahmouni, P. Guéret, J. Garot, and L. Najman. Fully automated segmentation of 4d cardiac cine-mri dataset in humans. In *Journées Européennes de la Société Française de Cardiologie*, 2007.

[30] L. D. Cohen. On active contour models and balloons. *CVGIP :Image Understanding*, 53(2) :211–218, March 1991.

[31] L.D. Cohen. Minimal paths and deformable models for image analysis. *Actes des journées d'études SEE : le traitement d'image à l'aube du XXIième siècle*, Mars 2002.

[32] S.H. Collins. Terrain parameters directly from a digital terrain model. *The Canadian Surveyor*, 29 :507–518, 1975.

[33] M. Comer and E. Delp. Morphological operations for color image processing. *Journal of Eletronic Imaging*, 8(3) :279–289, 1999.

[34] Mary L. Comer and Edward J. Delp. Empirical study of morphological operators in color image enhancement. In Sullivan J. R., Dawson B. M., and Rabbani M., editors, *Proc. SPIE Vol. 1657, p. 314-325, Image Processing Algorithms and Techniques III, James R. Sullivan ; Benjamin M. Dawson ; Majid Rabbani ; Eds.*, volume 1657 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 314–325, may 1992.

[35] J.-P Coquerez and S. Philipp. *Analyse d'images : filtrage et segmentation*. MASSON, 1995.

[36] Antoine Cornuéjols and Laurent Miclet. *Apprentissage artificiel : Concepts et algorithmes*. Eyrolles, 2002.

[37] J. Cousty, L. Najman, and M. Couprie. Segmentation of 4d cardiac mri : automated method based on watershed cuts. *to be published*, 2008.

[38] Jean Cousty. *Segmentation et analyse d'images cardiaques IRM*. PhD thesis, Laboratoire A2SI, Groupe ESIEE, 2007.

[39] José Crespo and Ronald W. Schafer. Locality and adjacency stability constraints for morphological connected operators. *Journal of Mathematical Imaging and Vision*, 7 :85–102, 1997.

[40] E. Deléchelle and J. Lemoine. La trajectoire déformable : un modèle optique des contours géodésiques fondé sur le principe de fermat. *VI'99, Trois-Rivières, Québec, Canada*, Mai 1999.

[41] R. V. den Boomgaard and R.W̃. Smeudlers. The morphological structure of images : the differential equations of morphological scale-space. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 16 :1,101–1,113, 1994.

[42] E.R. Dougherty, J.T. Newell, and J.B. Pelz. Morphological texture-based maximum-likelihood pixel classification based on local granulometric moments. *Pattern Recognition*, 25(10) :1181–1198, October 1992.

[43] Raffi Enficiaud. *Algorithmes multidimensionnels et multispectraux en Morphologie Mathématique : Approche par méta-programmation*. PhD thesis, Ecole des Mines de Paris, 2007.

[44] S. Fischer, P. Schmid-Saugeon, and J. Guillod. Analysis of skin lesions with pigmented networks. In *Proceedings of International Conference on Image Processing ICIP'96*, volume 1 of *Journal of Structural Biology*, pages 323–326. IEEE, 1996.

[45] Florence Forbes and Andrian E. Raftery. Bayesian morphology : Fast unsupervised bayesian image analysis. Technical Report 3374, INRIA, Mars 1998.

[46] Martin Fowler and Kent Beck. *Refactoring : Improving the Design of Existing Code*. Addison-Wesley Professional, 1999.

[47] A. Garcia, C. Vachier, and J.-P. Vallée. Multivariate mathematical morphology and Bayesian classifier application to colour and medical images. In *Image Processing : Algorithms and Systems VI. Edited by Astola, Jaakko T. ; Egiazarian, Karen O. ; Dougherty, Edward R. Proceedings of the SPIE, Volume 6812, pp. 681203-681203-11 (2008).*, volume 6812 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, March 2008.

[48] Arnaud Garcia and Corinne Vachier. Self-adaptive morphological operators on color images based on the use of statistical patterns. *to be submitted (JMIV)*, 2008.

[49] Arnaud Garcia, Corinne Vachier, Antoine Rosset, and Jean-Paul Vallée. Multi-criteria seeded region growing for multi-contrast mri. In *ISMRM06*, 2006.

[50] J. Gauch and Chi-Wan Hasia. A comparison of threee color image segmentation algorithm in four color spaces. *SPIE Visual Communication and Image Processing*, 1818 :1168–1181, 1992.

[51] R.C. Gonzalez and R.E. Woods. *Digital Image Processing, Second Edition*. Prentice Hall, 2002.

[52] J. Goutsias, H.J.A.M. Heijmans, and Sivakumar K. Morphological operators for image sequences. *Report - Department of Operations Research, Statistics, and System Theory ISSN 0924-0659*, 19 :1–33, 1994.

[53] J. Goutsias, H.J.A.M. Heijmans, and K. Sivakumar. Morphological operators for image sequences. *Computer Vision and Image Understanding*, 62(3) :326–346, November 1995.

[54] J. I. Goutsias. Morphological transformations of image sequences : a lattice theory approach. In P. D. Gader, E. R. Dougherty, and J. C. Serra, editors, *Proc. SPIE Vol. 1769, p. 306-317, Image Algebra and Morphological Image Processing III, Paul D. Gader ; Edward R. Dougherty ; Jean C. Serra ; Eds.*, volume 1769 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 306–317, June 1992.

[55] T. Grigorishin and Y.H. Yang. Image segmentation : an electrostatic field based approach. *Vision interface'98, Vancouver*, pages 279–286, 1998.

[56] M. Grimaud. *La géodésie numérique en morphologie mathématique. Application à la détection automatique de microcalcifications en mammographie numérique*. PhD thesis, Ecole des Mines de Paris, Dec. 1991.

[57] M. Grimaud. New measure of contrast : dynamics. *Image Algebra and Morphological Processing III, San Diego CA, Proc. SPIE*, 1992.

[58] C. Gu. *Mathematical Morphology and Its Applications to Image and Signal Processing*, chapter Multivalued Morphology and its Application in Moving Object Segmentation and Tracking, pages 345–352. Springer-Verlag New York, LLC, 2002.

[59] A.G. Hanbury and J. Serra. Morphological operators on the unit circle. *Image Processing, IEEE Transactions on*, 10(12) :1842–1850, 2001.

[60] Allan Hanbury. *Morphologie Mathématique sur le Cercle Unité*. PhD thesis, Ecole des Mines de Paris, 2002.

[61] Allan Hanbury and Jean Serra. Mathematical morphology in the l*a*b* color space. Technical report, Ecole des Mines de Paris, 2001.

[62] Alejandro F. Frangi Hans C. van Assen, Mikhail G. Danilouchkine, Sebastián Ordás, Jos J.M. Westenberg, Johan H.C. Reiber, and Boudewijn P.F. Lelieveldt. Spasm : A 3d-asm for segmentation of sparse and arbitrarily oriented cardiac mri data. *Medical Image Analysis*, 10(2) :286–303, 2006.

[63] H.J.A.M. Heijmans and C. Ronse. The algebraic basis of mathematical morphology. i. dilations and erosions. *Comput. Vision Graph. Image Process.*, 50(3) :245–295, 1990.

[64] Andrew Hunt and David Thomas. *he Pragmatic Programmer : From Journeyman to Master*. Addison-Wesley Professional, 1999.

[65] L. Najman J. Cousty, M. Couprie, S. Clément-Guinaudeau, T. Goissen, and J. Garot. Automated, accurate and fast segmentation of 4d cardiac cine-mr images based on discrete mathematical morphology. *to be published*, 2008.

[66] Emmanuel Kant. *La critique de la raison pure*. PUF, 1944.

[67] M. Kass, A. Witkin, and D. Terzopoulos. Snakes : Active contour models. *International Journal of Computer Vision*, 1(4) :321–331, 1988.

[68] A. Katartzis, V. Pizurica, and H. Sahli. Application of mathematical morphology and markov random eld theory to the automatic extraction of linear features in airborne images, 2000.

[69] Michael R Kaus, Jens von Berg, Jurgen Weese, Wiro Niessen, and Vladimir Pekar. Automated segmentation of the left ventricle in cardiac mri. *Med Image Anal*, 8(3) :245–254, 2004.

[70] P. Lambert and J. Chanussot. Extending mathematical morphology to color image processing. In *CGIP'00 - 1st International Conference on Color in Graphics and Image Processing*, pages 158–163, october 2000.

[71] C. Lantuéjoul. *La squelettisation et son application aux mesures topologiques des mosaïques polycristallines*. PhD thesis, Ecole des Mines de Paris, 1978.

[72] Hyunsook Lee. A convex hull peeling depth approach to nonparametric massive multivariate data analysis with applications. Technical report, Department of Statistics, The Pennsylvania State University, 1997.

[73] J. Li and Y. Li. Multivariate mathematical morphology based on principal component analysis : Initial results in building extraction. *international archives of photogrammetry remote sensing and spatial information sciences*, 35 :1168–1173, 2004.

[74] Maria Lorenzo-Valdes, Gerardo I. Sanchez-Ortiz, Andrew G. Elkington, Raad H. Mohiaddin, and Daniel Rueckert. Segmentation of 4d cardiac mr images using a probabilistic atlas and the em algorithm. *Medical Image Analysis*, 8(3) :255–265, july 2004.

[75] J. Lotjonen, S. Kivisto, J. Koikkalainen, D. Smutek, and K. Lauerma. Statistical shape model of atria, ventricles and epicardium from short- and long-axis mr images. *Medical Image Analysis*, 8(3) :371–386, 2004.

[76] M. Lynch, O. Ghita, and P. F. Whelan. Automatic segmentation of the left ventricle cavity and myocardium in mri data. *Comput Biol Med*, 36(4) :389–407, Apr 2006.

[77] Uzümcü M, van der Geest RJ, Swingen C, Reiber JH, and Lelieveldt BP. Time continuous tracking and segmentation of cardiovascular magnetic resonance images using multidimensional dynamic programming. *Investigative Radiology*, 41(1) :52–62, 2006.

[78] P. Maragos. Pattern spectrum and multiscale shape representation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 11(7) :701–716, July 1989.

[79] Petros Maragos and Corinne Vachier. A pde formulation for viscous morphological operators with extensions to intensity-adaptive operators. *to be publisched in proceedings of International Conference on Image Processing*, 2008.

[80] G. Matheron. *Eléments pour une Théorie des Milieux Poreux*. Masson, Paris, 1967.

[81] G. Matheron. *Random Sets and Integral Geometry*. John Wiley and Sons, New York, 1975.

[82] G. Matheron. Les nivellements. Technical report, Technical report, Centre de Morphologie Mathématique, 1997.

[83] T. McInerney and D. Terzopoulos. A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4d image analysis. *Computerized Medical Imaging and Graphics*, 19(1) :69–83, 1995.

[84] F. Meyer. From connected operators to levelings. *Proceedings of the fourth international symposium on Mathematical morphology and its applications to image and signal processing*, pages 191–199, 1998.

[85] F. Meyer and C. Vachier. *On the Regularization of The Watershed Transform*. Advances in imaging and electron physics (AIEP)", Hawkes, P.W. (ed.), Academic Press, 2007.

[86] M. Mignotte, C. Collet, P. Perez, and P. Bouthemy. Markov random field model and fuzzy formalism-based data modeling for the sea-floor classification. In F. Preteux, A. Mohammad-Djafari, and E. R. Dougherty, editors, *Proc. SPIE Vol. 3816, p. 229-240, Mathematical Modeling, Bayesian Estimation, and Inverse Problems, Francoise Preteux ; Ali Mohammad-Djafari ; Edward R. Dougherty ; Eds.*, pages 229–240, June 1999.

[87] S. C. Mitchell, B. P. F. Lelieveldt, R. J. van der Geest, H. G. Bosch, J. H. C. Reiber, and M. Sonka. Multistage hybrid active appearance model matching : Segmentation of left and right ventricles in cardiac mr images. *IEEEE trans on Medical Imaging*, 20 :415–423, 2001.

[88] Henning Müller, Arnaud Garcia, J-P. Vallée, and Antoine Geissbühler. Grid computing at the university hospitals of geneva. In *healthGrid Conference*, 2003.

[89] Henning Müller, Arnaud Garcia, Jean-Paul Vallée, Antoine Rosset, and Antoine Geissbuhler. Potential benefits from content-based visual data access in radiology. In *RSNA InfoRad Demonstration*. Chicago, IL, USA, December 2003.

[90] Henning Müller, Antoine Rosset, Arnaud Garcia, Jean-Paul Vallée, and Antoine Geissbuhler. Benefits of content-based visual data access in radiology. *Radiographics*, 25 (3) :849–858, 2005.

[91] Henning Müller, Julien Vignali, Arnaud Garcia, Jean-Paul Vallée, and Antoine Geissbuhler. Allowing content-based visual data access to a radiologic case database. In *annual congress of the Swiss society of radiology*. Interlaken, Switzerland, May 2004.

[92] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and variational problems. *Communications on Pure and Applied Mathematics*, XLII(5) :577–685, 1988.

[93] B. Naegel and N. Boch. Un système d'aide au diagnostic dermatologique par cartographie et suivi de grains de beauté. In *Proceedings of the french conference ORASIS 2007*, 2007.

[94] Benoît Naegel, Nicolas Passat, Nicolas Boch, and Michel Kocher. Segmentation using vector-attribute filters : methodology and application to dermatological imaging. In Gerald Jean Francis Banon, Junior Barrera, Ulisses de Mendonça Braga-Neto, and Nina Sumiko Tomita Hirata, editors, *Proceedings*, volume 1, pages 239–250, São José dos Campos, October 10–13, 2007 2007. Universidade de São Paulo (USP), Instituto Nacional de Pesquisas Espaciais (INPE).

[95] L. Najman. *Morphologie mathématique : de la segmentation d'images à l'analyse multivoque*. PhD thesis, Paris Dauphine, Avr. 1994.

[96] L. Najman and M. Schmitt. A dynamic hierarchical segmentation algorithm. *ISMM'94 : Mathematical Morphology and its Applications to Image Processing - Poster Contributions*, pages 13–14, Sept. 1994.

[97] S. Osher and J. Sethian. Fronts propagating with curvature dependent speed : Algorithms based on hamilton-jacobi formulations. *J. Comput. Phys.*, 79 :12–49, 1988.

[98] N. Paragios, Mikael Rousson, and Visvanathan Ramesh. Matching distance functions : A shape-to-area variational approach for global-to-local registration. *European Conference on Computer Vision*, pages II :775–790, 2002.

[99] Nikos Paragios. A variational approach for the segmentation of the left ventricle in mr cardiac images. In *VLSM '01 : Proceedings of the IEEE Workshop on Variational and Level Set Methods (VLSM'01)*, page 153, Washington, DC, USA, 2001. IEEE Computer Society.

[100] Claire Parnet. *Dialogues Gilles Deleuze*. Flammarion, 1977.

[101] R. A. Peters. Mathematical morphology for angle-valued images. In *Proc. SPIE*, volume 3026, 1997.

[102] A. Plaza, P. Martinez, R. Perez, and J. Plaza. Spatial/spectral endmember extraction by multidimensional morphological operations. *Geoscience and Remote Sensing, IEEE Transactions on*, 40(9) :2025–2041, 2002.

[103] Charnchai Pluempitiwiriyawej. *Cardiac MR Image Segmentation : STACS, A New Active Contour*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA 15213, 2003.

[104] T.H. Puecker and D.H. Douglas. Detection of surface specific points by local parallel processing of discrete terrain elevation data. *Computer Graphics and Image Processing*, 4 :375–387, 1975.

[105] E.M. Purcell, H.C. Torrey, and R.V. Pound. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev*, 69 :37, 1946.

[106] Christian P. Robert. *Le choix bayésien : principes et pratique*. Springer - Statistical Theory and Methods, 2006.

[107] Antoine Rosset, Luca Spadola, and Osman Ratib. Osirix : An open-source software for navigating in multidimensional dicom images. *Journal of Digital Imaging*, 17(3) :205–216, September 2004.

[108] Mikael Rousson and Nikos Paragios. Shape priors for level set representations. In *ECCV*, pages 78–92, Copenhangen, Denmark, 2002.

[109] P. Salembier and J.R. Hidalgo. Connected operators based on reconstruction process for size and motion simplification. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, Orlando, USA, May 2002.

[110] P. Salembier and J. Serra. Morphological multiscale image segmentation. *Visual Communication and Image Processing, SPIE*, 1818 :620–631, 1992.

[111] P. Salembier and J. Serra. Flat zones filtering, connected operators, and filters by reconstruction. *IEEE Transactions on Image Processing*, 4(8) :1153–1160, 1995.

[112] Philippe Salembier and Jean Serra. Flat zones filtering, connected operators, and filters by reconstruction. *IEEE Transactions on Image Processing*, 4(8) :1153–1160, 1995.

[113] M. F. Santarelli, V. Positano, C. Michelassi, M. Lombardi, and L. Landini. Automated cardiac mr image segmentation : theory and measurement evaluation. *Med Eng Phys*, 25(2) :149–59, Mar 2003.

[114] L. J. Sartor and A. R. Weeks. Morphological operations on color images. *Journal of Electronic Imaging*, 10 :548–559, apr 2001.

[115] P. Schmid-Saugeon, J. Guillod, and J. Thiran. Towards a computer-aided diagnosis system for pigmented skin lesions. *Computerized Medical Imaging and Graphics*, 27(1) :65–78, 2003.

[116] M. Schmitt and J. Mattioli. *Morphologie Mathématique*. Masson, Dec. 1993.

[117] M. Schmitt and F. Prêteux. Un nouvel algorithme en morphologie mathématique : les r-h maxima et r-h minima. *Proc. 2ieme Semaine Internationale de l'Image Electronique*, pages 469–475, Avr. 1986.

[118] J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, 1982.

[119] J. Serra. *Image Analysis and Mathematical Morphology. Theoretical Advances*. Academic Press, 1988.

[120] J. Serra. Quelques propriétés des nivellements. Technical report, Technical report, Centre de Morphologie Mathématique 30/98/MM, 1998.

[121] J. Serra. Connexions et segmentation d'image. *Traitement du Signal, Numéro spécial : "Le traitement du signal à l'aube du XXIe siècle"*, 20(3) :243–254, 2003.

[122] J. C. Serra. Anamorphoses and function lattices. In E. R. Dougherty, P. D. Gader, and J. C. Serra, editors, *Proc. SPIE Vol. 2030, p. 2-11, Image Algebra and Morphological Image Processing IV, Edward R. Dougherty ; Paul D. Gader ; Jean C. Serra ; Eds.*, volume 2030 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 2–11, June 1993.

[123] Jean Serra. *Image Analysis and mathematical Morphology*. Ac. Press, 1982.

[124] P. Soille. *Morphological Image Analysis : Principles and Applications*. Springer-Verlag, 1999.

[125] P. Soille. Constrained connectivity and transitions regions. *to be publisched in proceedings of International Conference on Image Processing*, 2008.

[126] Jasjit S. Suri, S. Kamaledin Setarehdan, and Sameer Singh, editors. *Advanced algorithmic approaches to medical image segmentation : state-of-the-art application in cardiology, neurology, mammography and pathology*. Springer-Verlag New York, Inc., New York, NY, USA, 2002.

[127] H. Talbot, C. Evans, and R. Jones. Complete ordering and multivariate mathematical morphology. In *ISMM '98 : Proceedings of the fourth international symposium on Mathematical morphology and its applications to image and signal processing*, pages 27–34, Norwell, MA, USA, 1998. Kluwer Academic Publishers.

[128] A. Tsai, Jr. Yezzi A., W. Wells, C. Tempany, D. Tucker, A. Fan, W.E. Grimson, and A. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *Medical Imaging, IEEE Transactions on*, 22(2) :137–154, 2003.

[129] C. Vachier. *Extraction de caractéristiques, segmentation d'image et morphologie mathématique*. PhD thesis, Ecole des Mines de Paris, Dec 1995.

[130] C. Vachier and F. Meyer. Extinction value : a new measurement of persistence. *Proc. of 1995 IEEE Workshop on Nonlinear Signal and Image Processing*, pages 254–257(Vol.I), Juin 1995.

[131] C. Vachier and F. Meyer. The viscous watershed transform. *Journal of Mathematical Imaging and Vision*, 22 :251–267, 2005.

[132] C. Vachier and F. Meyer. News from viscous land. In G-J-F Banon and J. Barrera, editors, *Mathematical Morphology and its Applications to Signal and Image Processing, proceedings of the 8th International Symposium on Mathematical Morphology*, volume 8, October 2007.

[133] C. Vachier and F. Meyer. The viscous morphological operators. *work in progress - script available at http ://www.cmla.ens-cachan.fr/Membres/corinne-vachier-mammar.html*, 2008.

[134] C. Vachier and L. Vincent. Valuation of image extrema using alternating filters by reconstruction. *Image Algebra and Morphological Processing, San Diego CA, Proc. SPIE*, Juil. 1995.

[135] L. Vincent. Morphological area openings and closings for grayscale images. *Shape in Picture, NATO Workshop, Driebergen*, Sept. 1992.

[136] Stephen S. Wilson. Theory of matrix morphology. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(6) :636–652, 1992.