DEA
Mathématiques,
Vision,
Apprentissage.

Deuxième année de
Magistère MMFAI

Projet ÉPIDAURE

Approche statistique pour la segmentation d'images tridimensionnelles du foie

Nils RAYNAUD

Rapport de stage de Diplôme d'Etudes Approfondies, septembre 2000

stage dirigé par Nicholas Ayache

encadré par Xavier Pennec
et Hervé Delingette

Table des matières

Chapitre 1

Introduction

La visualisation tridimensionnelle d'un organe à partir de l'information donnée par des images en coupe est une opération mentale bien difficile, même pour un spécialiste. Une image en trois dimensions est toujours plus explicite et plus facile à analyser. Or, lors de la visualisation d'une image tridimensionnelle, on est obligé de procéder coupe par coupe, ce qui nuit à la compréhension. Il serait donc particulièrement intéressant de parvenir à reconstruire en trois dimensions, à partir d'une telle image, les organes qui intéressent le praticien, afin de pouvoir les isoler les uns des autres, et de pouvoir les regarder sous tous les angles possibles.
Cela suppose que l'on soit capable d'extraire de la série de coupes constituant une image en trois dimensions, la surface constituant le contour de l'objet que l'on veut isoler. Cette opération est appelée une «segmentation» de l'objet en question.
Il existe de nombreuses méthodes pour segmenter des images, celles-ci étant plus ou moins automatisées. Dans le cas d'images médicales, les images sont très bruitées, peu contrastées, et les erreurs ne sont pas envisageables. Les segmentations automatiques classiques sont donc à exclure, car elles ne font pas la différence entre le contour d'un organe donné et celui d'un autre organe. Mais une segmentation entièrement manuelle réalisée par un spécialiste, bien qu'elle soit optimale, est longue et fastidieuse. Un compromis efficace entre ces deux extrêmes n'est donc pas dénué d'intérêt. Il faudrait donc incorporer dans notre méthode de segmentation une forme d'intelligence, qui permette de faire la différence entre ce qui fait partie de l'organe à segmenter et ce qui correspond à un autre organe. Pour cela, on introduira une forme a priori de l'organe, qui subira ensuite des déformations. Ces déformations ne seront acceptées que dans la mesure où notre modèle conserve une forme qui soit caractéristique de cet organe. Ceci sera fait par l'intermédiaire d'une étude statistique menée sur un ensemble d'apprentissage représentatif de l'organe. On bâtit ainsi un modèle extrêmement contraint, qui ne fonctionne que pour l'organe pour lequel on a fait des statistiques. Mais c'est aussi là que réside la force de cette méthode : en introduisant dans notre modèle la notion de forme caractéristique, on empêche toute erreur grossière consistant à segmenter des bouts d'autres organes, et l'on améliore considérablement la segmentation là où les contours de l'organe ne sont pas suffisamment contrastés pour être repérés par un algorithme classique de segmentation. Il ne sera pas inutile, ceci fait, de compléter la segmentation ainsi obtenue par une segmentation automatique classique, afin d'affiner le résultat.
Nous appliquerons notre méthode à une ensemble d'images de foies, provenant d'acquisitions fournies par l'IRCAD.
Nous commencerons, après quelques généralités sur le foie, par exposer la méthode de représentation surfacique, par maillages simplexes, que nous avons adoptée. Nous montrerons ensuite comment obtenir une forme moyenne de foie, puis nous exposerons une première méthode pour calculer des déformations caractéristiques, appelée analyse en composantes principales. Ceci fait, nous introduirons une autre méthode moins classique et plus en vogue, l'analyse en composantes indépendantes.

Nous montrerons alors les résultats obtenus par des segmentations utilisant l'analyse en composantes principale, et nous les comparerons avec des segmentations semi-automatiques.
Nous verrons enfin dans quelle mesure nous pouvons améliorer notre segmentation contrainte par les statistiques en imposant les coordonnées de certains sommets de notre maillage.

Ce stage s'est déroulé du 10 avril 2000 au 9 septembre 2000 à l'Institut National de Recherche en Informatique et Automatique, au sein du projet ÉPIDAURE. Il constitue la conclusion de mon année de DEA à l'ENS Cachan, dans le cadre du DEA «Mathématiques, Vision, Apprentissage», dirigé par Jean-Michel Morel. Ce stage était dirigé par Nicholas Ayache, directeur du projet ÉPIDAURE, encadré par Xavier Pennec et Hervé Delingette, chercheurs du projet ÉPIDAURE, et placé sous la responsabilité de Laurent Younes au niveau du DEA.

1.1 Le projet ÉPIDAURE

L'objectif du projet ÉPIDAURE est de développer des outils de traitement d'image médicales, avec de nombreuses applications : construction d'atlas anatomiques, simulations d'opérations chirurgicales, recalage d'images volumiques, aide au diagnostic pré-opératoire, réalité virtuelle et augmentée, etc.
Le projet ÉPIDAURE regroupe quatre chercheurs, le Dr. Nicholas Ayache, son directeur, Xavier Pennec, Hervé Delingette et Grégoire Malandain, ainsi qu'Éric Bardinet, ingénieur expert, Janet Bertot, ingénieur système, Francoise Pezé, chargée de l'administration, une dizaine de thésards, de nombreux collaborateurs, tant dans le milieu médical que dans celui de la recherche scientifique, et bien sûr des stagiaires.

1.2 Remerciements

Je tiens à remercier Nicholas Ayache pour son enthousiasme inébranlable et ses encouragements, Xavier Pennec et Hervé Delingette pour leur patiente et l'aide précieuse qu'ils m'ont fournie, sans oublier Luc Soler et Jacques Marescaux de l'IRCAD (Institut de Recherche contre le Cancer de l'Appareil Digestif) qui nous ont procuré les images abdominales. Merci à Johan Montagnat pour ses explications sur ses travaux antérieurs, à Yves Chau pour son expertise, Éric Bardinet pour l'intérêt qu'il a porté à mon travail et pour sa disponibilité, à Grégoire Malandain et aux thésards de l'équipe, Jonathan Stoeckel, Sylvain Prima, Guillaume Picinbono, David Rey, Alexis Roche, Sébastien Granger, Clément Forest, Pascal Cachier, Sébastien Ourselin, Maxime Sermesant, pour leur sympathie et l'aide occasionnelle qu'ils ne m'ont jamais refusé, à Gérald Bianchi et Guillaume Flandin pour leurs nombreux conseils de programmation. Merci enfin à Françoise Pezé pour l'efficacité de son travail, et bien entendu à Janet Bertot qui est toujours là pour résoudre les problèmes informatiques.

1.3 Sujet de stage

«Construction de modèles statistiques de formes hépatiques (stage de DEA)

Ce stage porte sur la conception de modèles statistiques de formes permettant de définir une représentation moyenne de la forme d'un organe, et une description de sa variabilité. Après une étude de l'état de l'art sur ce sujet, le stage consistera à proposer une méthode, et à l'appliquer à une base de donnée d'images scanner du foie qui comprend près de 40 organes. Le stage s'appuiera sur des travaux théoriques et pratiques de l'équipe dans ce domaine. On pourra utiliser un outil interactif de segmentation qui permet d'extraire la géométrie des organes dans les images scanner très efficacement. »

1.4 Motivations

La segmentation du foie dans une image scanner de l'abdomen d'un patient s'inscrit dans un processus visant à réaliser un simulateur chirurgical, afin d'effectuer un meilleur diagnostic préopératoire. Ce projet dans lequel l'INRIA est impliqué, est dirigé par l'IRCAD. Les techniques proposées ici pourront être étendues à la segmentation de bien d'autres organes, en vue de simuler l'intégralité des organes de la zone à opérer.
Il est évident que des applications peuvent être trouvées dans d'autres milieux que la chirurgie.

1.5 État de l'art

Outre les techniques classiques de segmentation automatique, qui ne tiennent pas compte de la spécificité des images à traiter, et ne sont donc pas bien adaptées à la segmentation d'acquisitions médicales, en raison du bruit important et du faible contraste qui caractérise ce type de données, plusieurs méthodes plus fines ont été élaborées. Ainsi, dans [ELF00], les auteurs représentent un contour par les zéros d'une fonction distance, et modifient l'algorithme classique de «snake» en incorporant, à chaque itération, la forme a posteriori du contour à segmenter. Une autre solution, proposée dans [LK00], consiste à élaborer un modèle de contour décrit par un maillage, ayant la forme du type d'objet à segmenter, et muni d'un certain nombre de points caractéristiques («landmarks»). L'utilisateur commence par mettre en correspondance ces points avec des points de l'image. Le modèle est ainsi déformé de manière à ce que ses points caractéristiques coïncident avec ceux que l'utilisateur a choisi. Le modèle déformé a dès lors à peu près la forme de l'objet présent dans l'image ; il subit alors une relaxation destinée à le rapprocher au mieux du contour réel, en utilisant l'information donnée par le gradient de l'image. En ce qui concerne les approches statistiques, une segmentation par analyse en composantes principales (ACP) est présentée dans [TCDJ95]. Mais cette approche est fondée sur une description des contours faite à partir de points caractéristiques, et se limite à des images en deux dimensions. Il est important que les points caractéristiques soient déterminés avec précision, ce qui en trois dimensions est particulièrement difficile, voire impossible, lorsque les objets à segmenter sont des organes d'une grande variabilité et ayant une surface lisse, comme c'est le cas pour le foie. Tout en conservant cette idée de modèle déformable contraint par les statistiques, nous avons préféré décrire notre modèle par un maillage simplexe soumis à des forces externes dues à l'image et à des forces internes de régularisation, comme dans [MD98] et [Del94].
Une des bases du travail présenté ici est la notion de forme statistique, liée à celle de recalage. [G.K89] introduit les notions de formes et d'espaces de formes, ainsi que de lois de probabilités sur l'espace des formes associé à k points de $\mathbb{R}^m$, en se limitant à des formes données par un petit nombre de points en dimension 2, et donne quelques applications à l'archéologie, l'astronomie, la géographie et la chimie. [Dry] applique la notion de forme à l'identification de substances biologiques par comparaison des électrophorèses. Le lecteur intéressé par la notion de forme moyenne dans l'espace des formes pourra se reporter à [Pen96a] ou [Zie94]. Enfin, [Pen00] montre comment recaler deux images bruitées de manière optimale.
Une méthode d'analyse statistique plus poussée que l'analyse en composantes indépendantes est très en vogue ces temps-ci. Il s'agit de l'analyse en composantes indépendantes ; celle-ci est exposée de façon exhaustive dans [ICA98], et de manière plus théorique dans [Com94]. Un début d'application à l'analyse statistique des formes d'un organe est présentée dans [Hug00], mais n'est pas appliqué au cas tridimensionnel.

1.6 À propos du foie

Le foie est un organe particulièrement important, qui joue de nombreux rôles essentiels au bon fonctionnement de l'organisme :
Il produit la bile (qui élimine les substances toxiques et facilite la digestion), ainsi que de nombreuses protéines essentielles ; il assainit le sang, régule le taux de glucose dans le sang, le taux de cholestérol, l'apport en vitamines et en minéraux ; il équilibre la circulation de nombreuses hormones...
C'est l'organe le plus volumineux du corps humain. Il mesure typiquement 28 cm dans le sens transversal, 16 cm de haut et 8 cm d'épaisseur. Il est d'une consistance plutôt ferme, fragile mais très malléable : sa variabilité est très importante. Sa forme dépend des ligaments qui le maintiennent et des organes qui se trouvent à côté de lui : les poumons, le cœur, l'estomac, le pancréas, la rate, la vésicule biliaire, les intestins, le diaphragme, les côtes... Par ailleurs, certaines pathologies du foie peuvent influer fortement sur la taille et la forme de celui-ci.
Le foie est doté d'un étonnant pouvoir de régénération, lui permettant de recouvrer la totalité de sa masse quatre mois après une ablation des trois quarts de l'organe. Ce pouvoir de régénération dépasse celui des cancers les plus graves. L'ablation, même massive, d'une partie du foie est donc une solution particulièrement efficace pour traiter certaines pathologies hépatiques graves.
On considère souvent que le foie comprend plusieurs parties, dénommées segments. Ce genre de description est fondé sur la vascularisation du foie. Si une tumeur se développe dans le foie, elle va utiliser la vascularisation porte pour se propager ; il est donc préférable d'ôter tout le segment correspondant, et il n'est pas nécessaire, si l'on s'y prend à temps, de toucher aux segments non concernés. Il existe plusieurs segmentations du foie, mais celle qui est la plus utilisée est celle de Couinaud (voir figures 1.1 et 1.2).

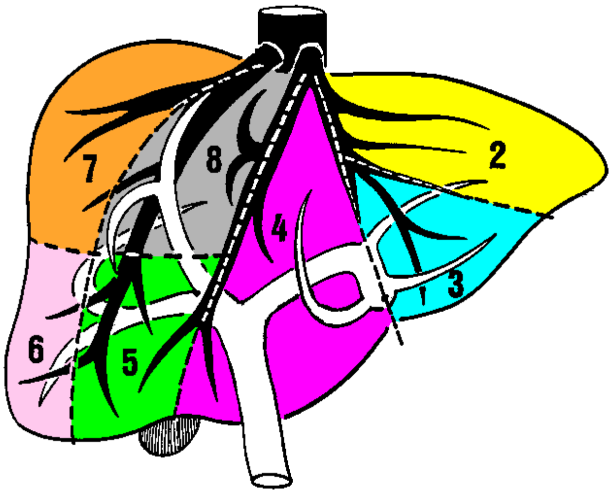


FIG. 1.1 – *schéma de la segmentation de Couinaud*

1.7 Le Visible Human Project de la *National Library of Medicine*

Le *Visible Human Project* de ls *National Library of Medicine* (ou NLM) est une collection d'images scannographiques, IRM et photographiques de deux corps humains des deux sexes (les *Visible Man* et *Visible Woman*), réalisée pour la recherche.
Les images scanner sont des coupes axiales régulièrement espacées de 1 mm. La résolution des

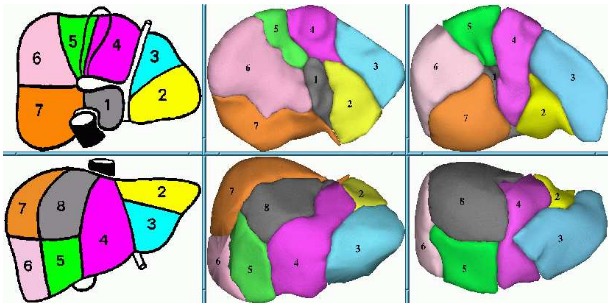


FIG. 1.2 – *Segmentation de Couinaud. De gauche à droite : schéma, puis deux exemples de modèles 3D de foies réels. En haut, vue de dessous ; en bas, vue de face*

images IRM est plus faible, mais celle des images photographiques est identique : les corps ont été cryogénisés, puis découpés en tranches de 1 mm, lesquelles ont été photographiées.
Les images photographiques et scannographiques ont l'avantage de coïncider, ce qui permet de faire d'intéressantes comparaisons.
En ce qui nous concerne, le foie du *Visible Man* fournit un modèle de référence de foie sain.

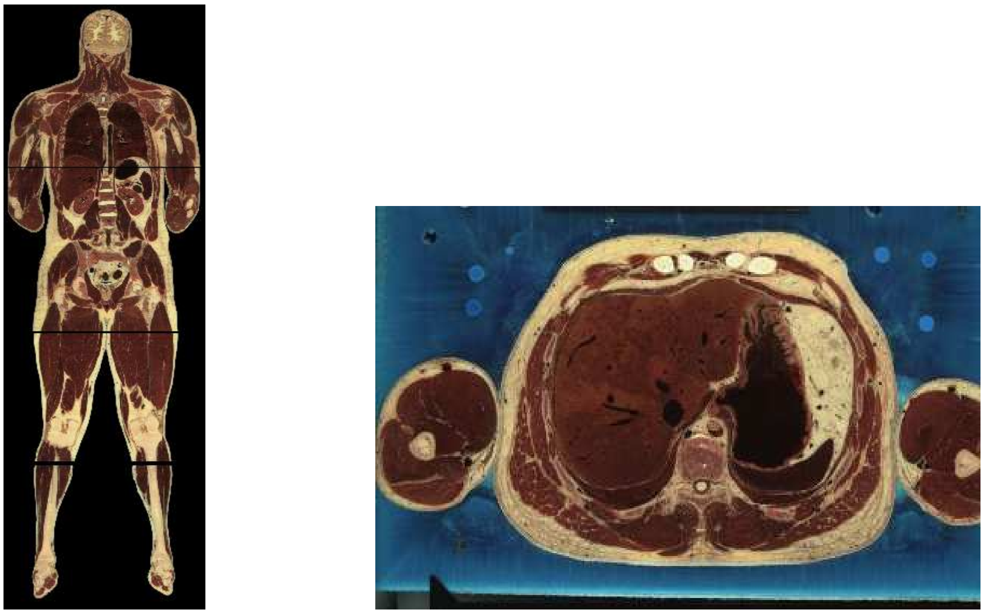


FIG. 1.3 – *le Visible Man et son foie*

Pour plus d'informations sur la *National Library of Medicine*, voir leur site : www.nlm.nih.gov.
Sur le *Visible Human Project* : www.nlm.nih.gov/research/visible.
A noter que le site polytechnique fédérale de Lausanne donne accès à un logiciel en ligne qui permet de visualiser une coupe d'orientation quelconque du Visible Man.

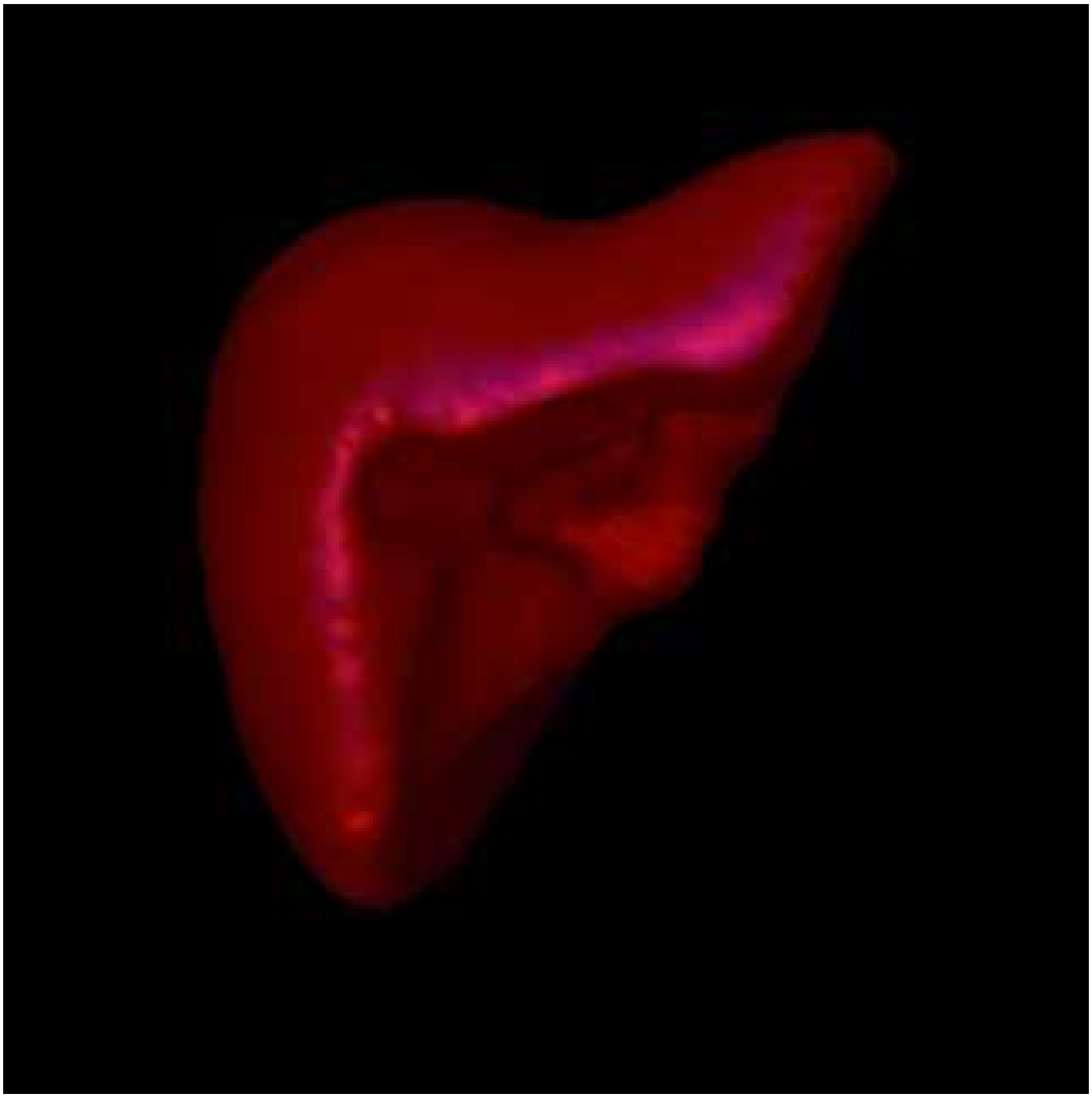


Fig. 1.4 – *Modèle du foie du Visible Man*

Chapitre 2

Maillages simplexes

Afin de rendre utilisables les données des images tridimensionnelles dont nous disposons, nous allons représenter la surface d'un foie par un vecteur. Pour cela, diverses solutions sont envisageables : les surfaces explicites (analogues des snakes en trois dimensions), les surfaces B-splines, les super-quadriques, les triangulations...
Nous avons choisi de représenter nos modèles de foies par des maillages simplexes, cette méthode étant déjà utilisée au sein du projet ÉPIDAURE ; cela nous permet, entre autres choses, de disposer d'outils de visualisation évolués, développés par des membres d'ÉPIDAURE.
Les maillages simplexes sont des maillages discrets topologiquement duaux des triangulations. Ils ont été introduits par Hervé Delingette en 1994. Nous allons voir que c'est une représentation qui permet une grande précision dans la description des organes étudiés.
Avant tout, définissons ce qu'est un maillage simplexe :

2.1 Définitions

Définition 1 *Un k-maillage simplexe de $\mathbb{R}^n$ est un maillage dont chaque sommet $P(i)$ possède exactement $k+1$ voisins : $(PP_1(i), \ldots, PP_{k+1}(i))$; et tel que les propriétés suivantes soient vérifiées :*

- *Un sommet ne peut être son propre voisin.*
- *Les $k+1$ voisins d'un sommet sont tous distincts deux à deux.*
- *La relation de voisinage est réflexive : si $P(i)$ est un voisin de $P(j)$, l'inverse est vrai aussi.*

Définissons aussi les arêtes et les faces d'un maillage simplexe :

Définition 2 *On appelle arête du maillage simplexe M un ensemble de deux sommets qui sont voisins l'un de l'autre.*

Une seule arête est susceptible de relier deux sommets.

Définition 3 *On appelle face d'un 2-maillage simplexe M un ensemble de sommets $PF_i(0), \ldots, PF_i(m-1)$ tel que :*

- *Fermeture : $\{PF_i(0), PF_i(1)\}$, $\{PF_i(1), PF_i(2)\}$, $\ldots$, $\{PF_i(m-1), PF_i(0)\}$ sont des arêtes.*
- *Unicité : Aucune arête ne partage une face en deux :*
  si $PF_i(j)$ et $PF_i(l)$ sont voisins, alors $l = j \pm 1 \ mod[m]$.

Attention : une face n'est a priori pas contenue dans un plan.

Définition 4 *Deux faces qui ont une arête en commun sont dites adjacentes.*

2.2 Orientation

La numérotation des voisins d'un sommet induit une orientation de ce sommet. On impose à tous les sommets d'une face d'être orientés de manière cohérente ; ceci permet de donner une orientation à la face en question. Les faces du maillage sont alors toutes orientées de manière cohérente. Voir figure 2.1.

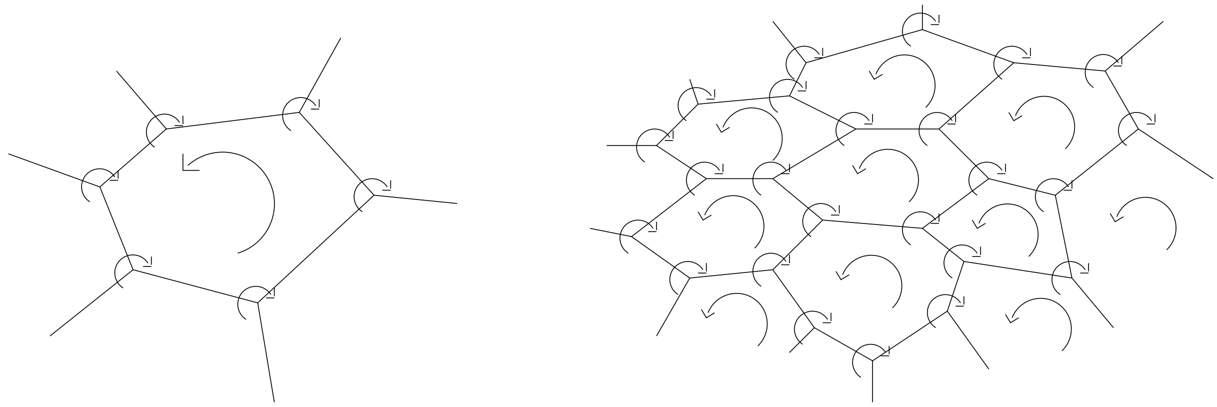


FIG. 2.1 – *Orientation d'une face et de ses sommets ; vue plus générale avec plusieurs faces.*

2.3 Dualité avec les triangulations

Définition 5 *On dira qu' un 2-maillage simplexe est complet si :*
- *il est orienté,*
- *chaque arête fait partie de deux faces exactement,*
- *deux faces adjacentes n'ont qu'une arête en commun.*

Le 2-maillage simplexe de la figure 2.2 n'est pas complet, puisque deux faces ont deux arêtes en commun.
Hervé Delingette montre qu'un 2-maillage simplexe complet est le dual topologique d'une triangulation.
Pour obtenir un rendu à partir d'un 2-maillage simplexe, ou pour calculer son aire, il faut le trianguler. H.Delingette introduit deux méthodes intéressantes pour ce faire : la première consiste à considérer que les sommets de la triangulation duale sont les centres de gravité des faces du maillage. La deuxième conserve les sommets du maillage, mais relie les sommets d'une face avec le centre de gravité de celle-ci. Voir figure 2.3
L'avantage de la première méthode est de fournir une triangulation avec moins de sommets et de triangles, ce qui permet d'accélérer le rendu graphique. En revanche, elle sous-estime la courbure : le rendu n'est pas très bon. La seconde méthode, plus coûteuse, est aussi plus précise. C'est celle que l'on utilisera pour calculer l'aire d'une face ou le volume du maillage.

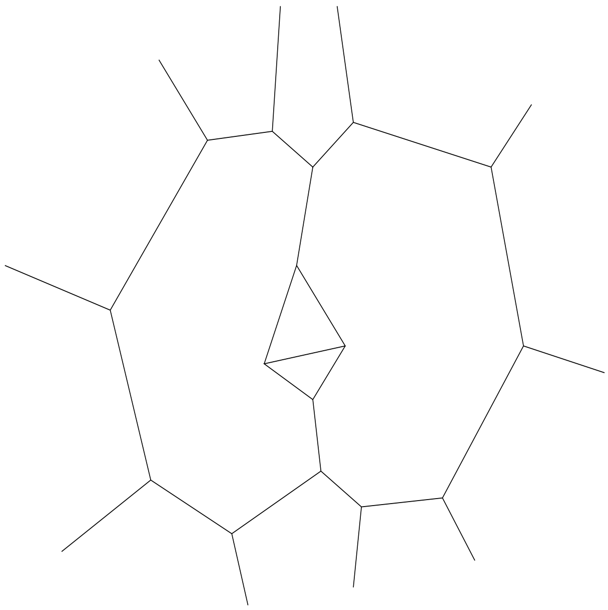


FIG. 2.2 – *2-maillage simplexe non complet*

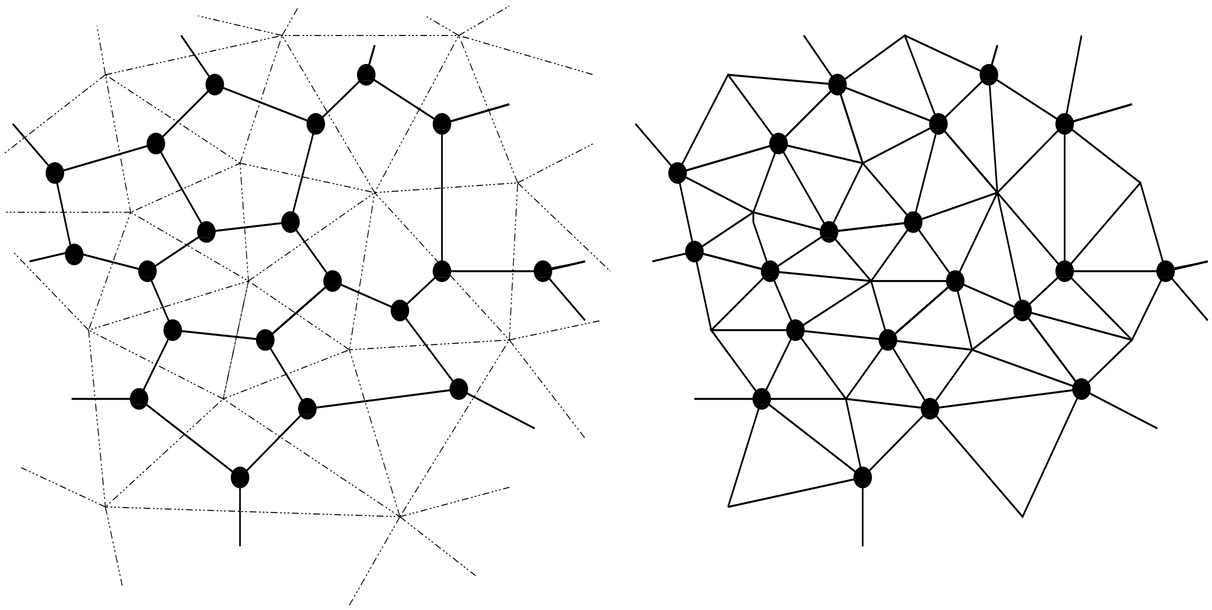


FIG. 2.3 – *Deux méthodes de triangulation d'un 2-maillage simplexe*

Pour plus de précisions sur les maillages simplexes et sur les transformations que l'on peut leur faire subir (notamment le raffinement), voir [Del94], ainsi que [Mon99].

2.4 Stabilité des emplacements des sommets

Nous allons, dans tout ce qui suit, chercher à recaler des maillages de foies sur d'autres maillages de foies ; c'est-à-dire que l'on va chercher la transformation rigide, similitude au affine, qui, appliquée à un maillage M_1, le rapproche le plus possible d'un maillage M_2. Pour cela, nous allons apparier chaque sommet i du maillage M_1 avec le sommet i du maillage M_2, et minimiser la distance entre M_1 et M_2, considérés comme des vecteurs. Cette approche suppose que les sommets de même indice de deux maillages différents représentent les mêmes points sur le foie. En d'autres termes, on fait l'hypothèse que chaque sommet conserve à peu près le même emplacement d'un maillage sur l'autre.
C'est une hypothèse très forte ; et pour que celle-ci soit valide, nous avons utilisé des maillages provenant tous du maillage du foie du Visible Man de la NLM : lors de la segmentation d'une image de foie, nous avons pris comme initialisation le maillage du foie de la NLM, correctement recalé sur l'image. Ainsi les coordonnées initiales d'un sommet étaient proches des coordonnées finales du même sommet.
Cela nous a permis de garantir le fait que deux sommets qui sont éloignés dans un maillage le soient aussi dans un autre, et qu'inversement, deux sommets proches dans un des maillages restent proches dans un autre maillage.
L'appariement n'est donc pas exact, mais il est tout de même cohérent ; cette approche est validée, dans une certaine mesure, par les résultats obtenus par la suite. Elle a tout de même ses limites, comme on pourra le voir par la suite.

2.5 Segmentation semi-automatique corrigée manuellement

L'intérêt d'une segmentation automatique est évident, mais sa mise en œuvre est rarement satisfaisante, et il est fréquent qu'il soit nécessaire de la modifier manuellement. Le résultat est alors beaucoup plus convainquant. Cette technique est bien moins fastidieuse qu'une segmentation entièrement manuelle, qui est très coûteuse en temps.
Partie automatique : la figure 2.4 donne un exemple vu en coupe. Un maillage (celui de la NLM), initialement recalé sur l'image, est ensuite déformé localement pour coller aux contours du foie dans l'image.

Partie semi-automatique :
Exemple d'une correction semi-manuelle : dans la figure 2.5, la zone jaune est le résidu de l'attraction du maillage par la veine sus-hépatique, que l'on ne désire pas garder dans notre modèle. Il faut par conséquent aplatir cette zone de manière à n'avoir plus que le contour du foie. Pour cela, on délimite la zone à modifier (en jaune sur la figure 2.5), puis on annule les forces d'attraction vers les contours, ainsi que la rigidité, et l'on applique une force tendant à minimiser l'aire de la zone en question. La figure 2.5 montre les états initial et final, ainsi que deux états intermédiaires :

Un des premiers objectifs du stage consistait à se procurer un ensemble de foie correctement segmentés, destiné à constituer l'ensemble d'apprentissage des statistiques. Nous disposions de 45 acquisitions tridimensionnelles de foies, fournies par l'IRCAD (Institut de Recherche contre

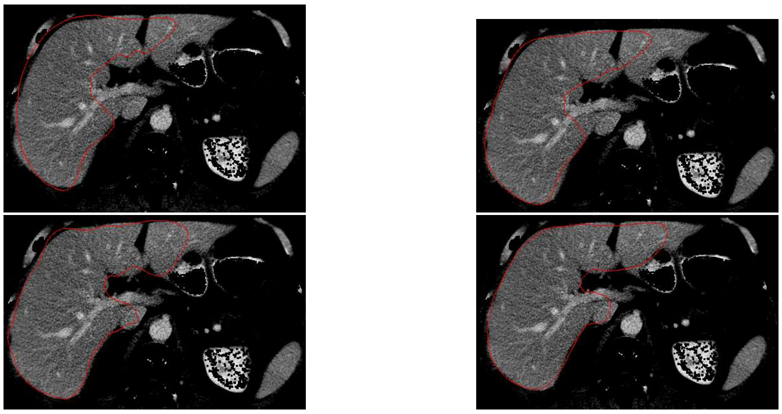


FIG. 2.4 – *Coupe d'une segmentation automatique initialisée avec le foie de la NLM*

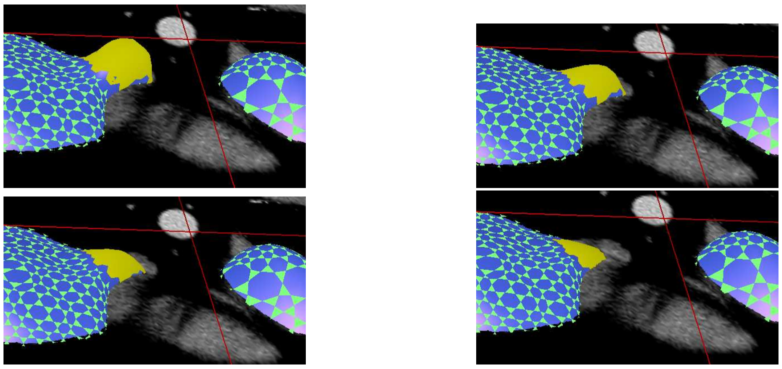


FIG. 2.5 – *Aplatissement d'une zone correspondant à l'attraction du contour par la veine sus-hépatique*

les Cancers de l'Appareil Digestif), et segmentées automatiquement de la manière suivante : le maillage est soumis à chaque itération à des forces externes (réglées par un paramètre β, dues à l'attraction des points de fort gradient de l'image, ainsi qu'à des forces internes de régularisation (réglées par un paramètre α). Le maillage a par ailleurs une certaine rigidité, qui limite les déformations trop importantes. Pour plus d'informations sur les procédés de segmentation automatique des maillages simplexes, le lecteur pourra se référer à [Mon99] et [MD98].
Une première difficulté est que la majorité des images de foie fournies provenaient de patients malades. Nombre d'entre eux avaient subi une ablation au niveau du foie, ce qui rendait l'organe non-représentatif d'un foie sain (cf figure 2.8). Il fallait alors faire un choix : prenait-on en compte, dans notre modèle, les foies ayant subi une ablation ? Le problème est que la variabilité de ceux-ci est excessivement grande, puisque l'on peut enlever les trois-quarts du foie d'un malade sans que cela pose de problème de régénération. Le statistiques risquaient alors de permettre à peu près n'importe quelle déformation, ce qui est justement ce que l'on voulait éviter. Pour valider notre méthode de segmentation contrainte par les statistiques, il nous a semblé plus cohérent de se limiter à des foies dont la forme soit représentative. Par ailleurs, ce choix peut permettre de repérer facilement une ablation, les statistiques empêchant une segmentation adéquate au niveau de celle-ci.
Par ailleurs, certaines acquisitions ne contenaient pas le foie dans son intégralité, interdisant une bonne segmentation (cf figure 2.7).
La discrimination entre foies représentatifs et foies non-représentatifs a été validée par Yves Chau, radiologue en stage au sein du projet ÉPIDAURE.
Ainsi notre ensemble d'apprentissage se trouvait considérablement réduit.
Par ailleurs, les segmentations automatiques n'étaient pas satisfaisantes : le maillage avait souvent été attiré par la veine sus-hépatique ainsi que par le cœur, ce qui lui donnait une forme incorrecte. Il a donc été indispensable de modifier ces segmentations semi-manuellement, ce qui prit un temps non négligeable.
Il restait finalement 13 maillages que l'on estimait représenter des foies sains correctement segmentés.

2.6 Segmentation automatique contrainte par les statistiques

Pour une segmentation précise du foie, il faut un nombre conséquent de sommets par maillage. Mais un trop grand nombre de sommets pourrait ralentir excessivement la segmentation. La segmentation semi-automatique par déformation d'un modèle statistique suivant des modes devra donc se faire avec un maillage assez rudimentaire ne prenant pas en compte les petites variations de la surface de l'organe. Les maillages que l'on utilise ici comprennent 3716 sommets.
La première étape consistera à calculer, à partir de nos 13 maillages d'apprentissage, une maillage qui représente une forme moyenne de foie. On calculera ensuite les déformations les plus probables grâce à une analyse statistique du type Analyse en Composantes Principales (abrégé en ACP ou PCA) ou Analyse en Composantes Indépendantes (ICA).
En ce qui concerne la segmentation d'une image de foie par notre modèle statistique, la première étape consistera à recaler le maillage moyen sur l'image. Ce recalage se fera par transformation rigide (rotation et translation), par similitude (on rajoute un facteur d'échelle), ou par transformation affine (transformation linéaire inversible et translation). Dans un premier temps, avant de faire un recalage automatique, on procédera à un recalage rigide manuel, à l'aide du logiciel «sm» développé par Hervé Delingette et Johan Montagnat, ce qui se fait très rapidement. Ceci permettra au recalage automatique d'être plus efficace.
Dans [Mon99], Johan Montagnat propose d'utiliser l'algorithme d'Iterative Closest Point (ICP) pour recaler au mieux un maillage simplexe sur une image tridimensionnelle. Le principe en est

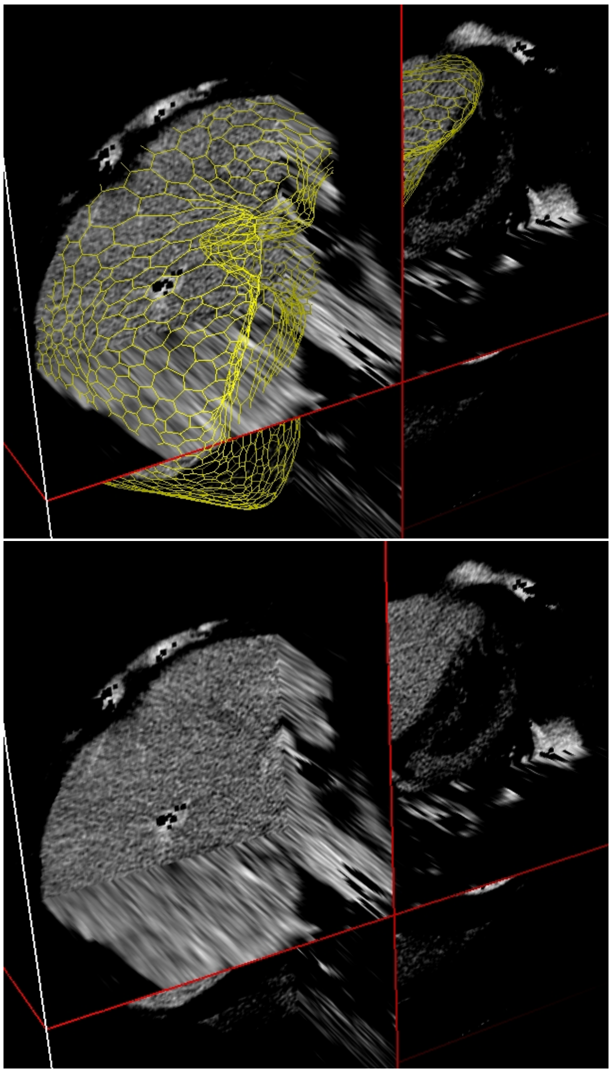


FIG. 2.6 – *Image tridimensionnelle de foie, avec et sans maillage*

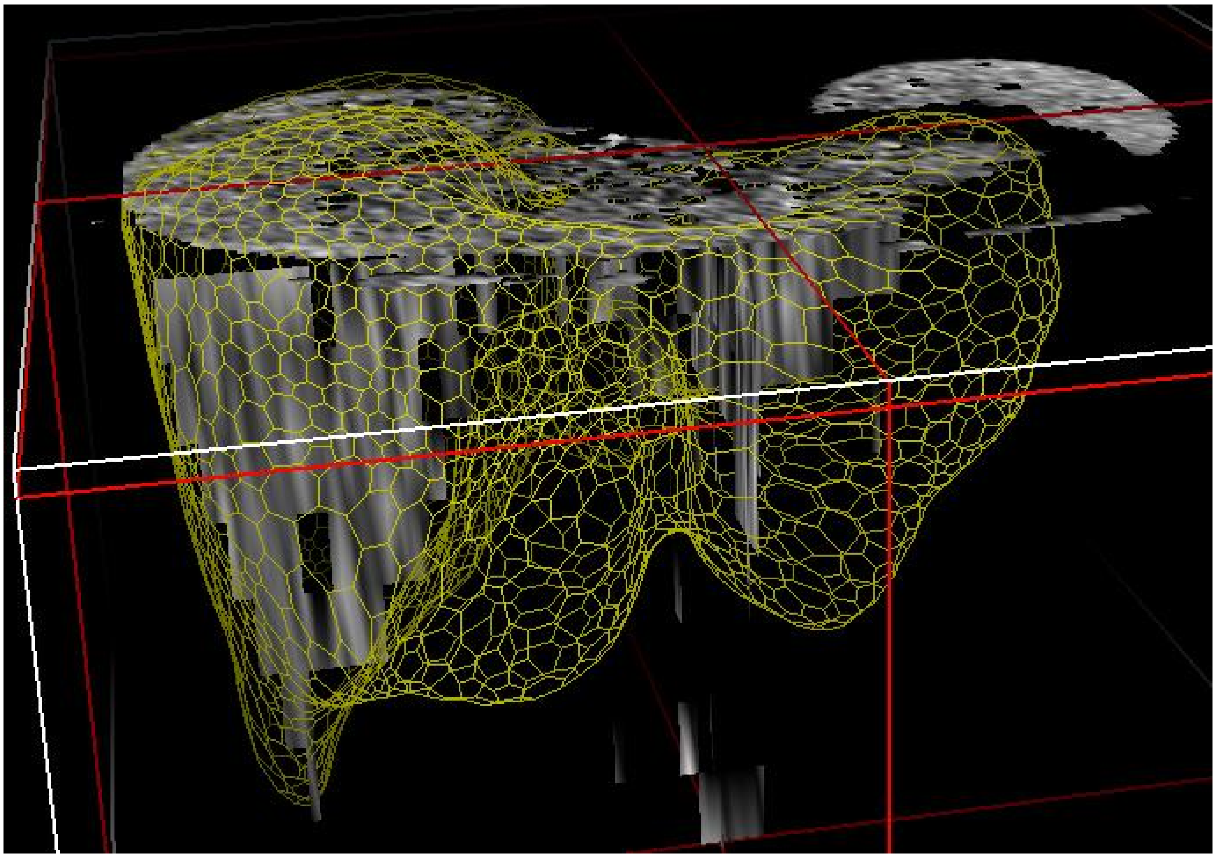


FIG. 2.7 – *Sur cette image, il manque des coupes, d'où un aplatissement anormal du maillage (en haut de l'image)*

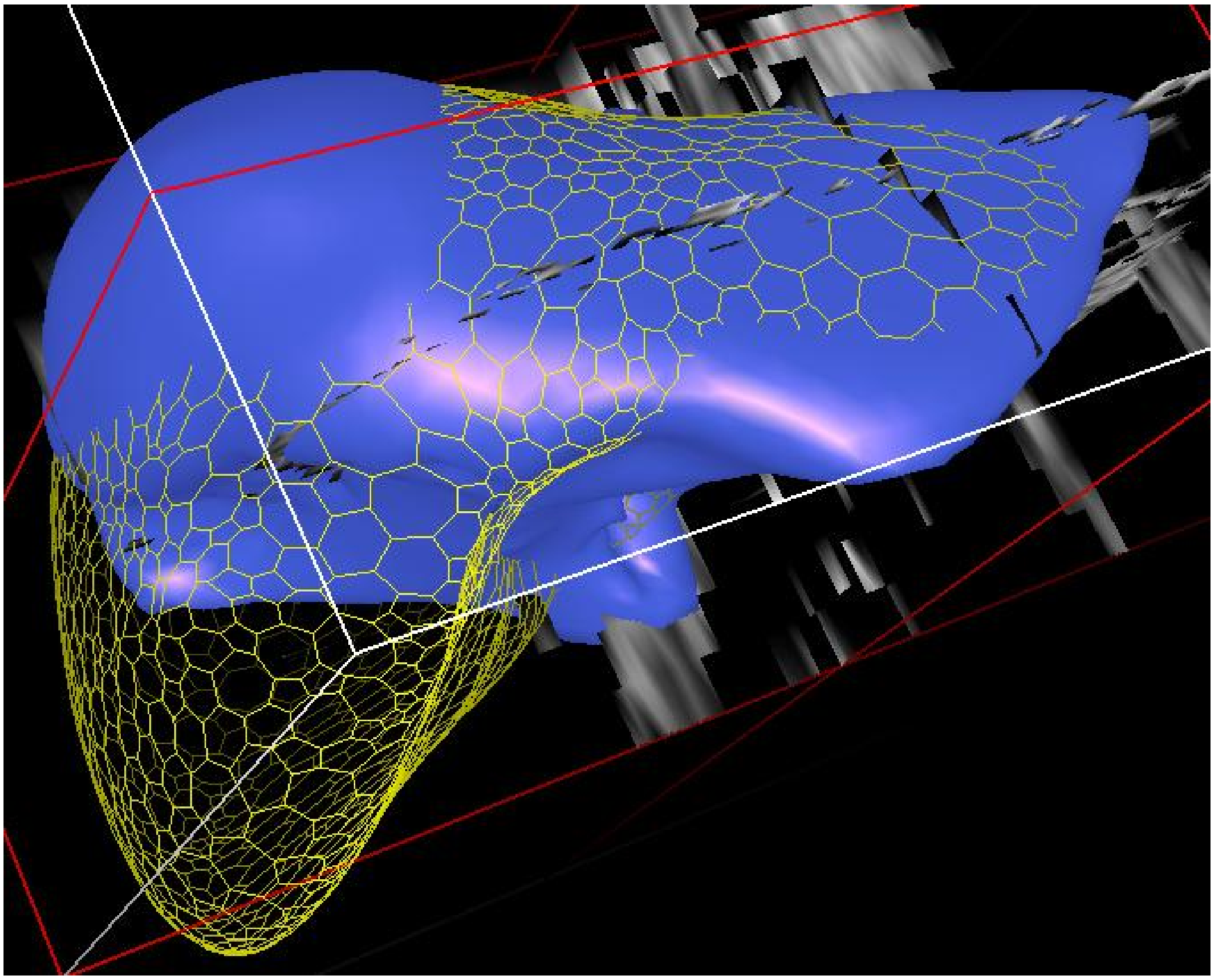


Fig. 2.8 – *Rendu du maillage d'un foie qui a subi une ablation et maillage moyen. La comparaison montre très clairement la partie manquante*

le suivant : on apparie chaque sommet du maillage avec le point de plus fort gradient dans la direction de la normale, qui soit à une distance inférieure à une longueur fixée l (à régler en fonction de l'image) ; ensuite, on applique le recalage qui rapproche le plus le maillage de l'ensemble des poins appariés (il existe plusieurs méthodes pour faire cela. Xavier Pennec, dans [Pen96b], propose une méthode utilisant les quaternions, ainsi qu'une approche fondée sur la décomposition en valeurs singulières de la matrice de corrélations entre les sommets du maillage moyen et leurs appariements).
Ayant recalé le «foie moyen» sur l'image tridimensionnelle, on va imposer au maillage de ne subir des déformations que si elles sont statistiquement probables. Cela permettra d'éviter les erreurs grossières, et restreindra considérablement l'étendue des déformations possibles.
On pourra ensuite, une fois que notre maillage est suffisamment proche du foie réel, continuer la segmentation par une technique plus classique de type «snakes», pour que le maillage épouse parfaitement la forme du foie.
Ainsi notre modèle statistique devrait essentiellement servir à une pré-segmentation, préalable à une segmentation plus précise.
L'utilité d'un tel procédé est évidente : un simple recalage du modèle moyen sur l'image réelle tridimensionnelle ne permet pas, étant donnée la variabilité importante du foie, de s'approcher suffisamment des contours à segmenter. Les nombreux organes qui jouxtent le foie sont autant de formes dont les contours sont susceptibles d'attirer les contours actifs : il est vraiment essentiel de parvenir à une initialisation très proche de la segmentation idéale.
La figure 2.9 montre l'exemple d'une segmentation où le modèle a été attiré par le cœur. On le voit bien sur la coupe, en haut à droite. Sur le modèle tridimensionnel (vue de dessus), la petite excroissance sur la droite représente ce défaut :

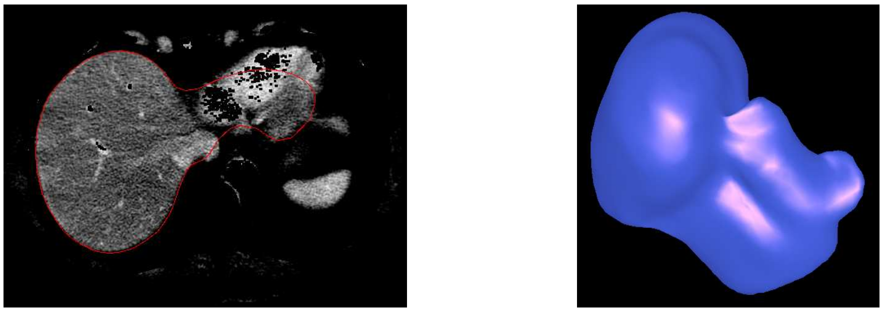


FIG. 2.9 – *Une coupe avec la trace du modèle, et le modèle vu de dessus*

2.7 Modèles d'apprentissage

Nous disposons de treize maillages correspondant à la segmentation de treize images de foies sains. Voir figure 2.10.

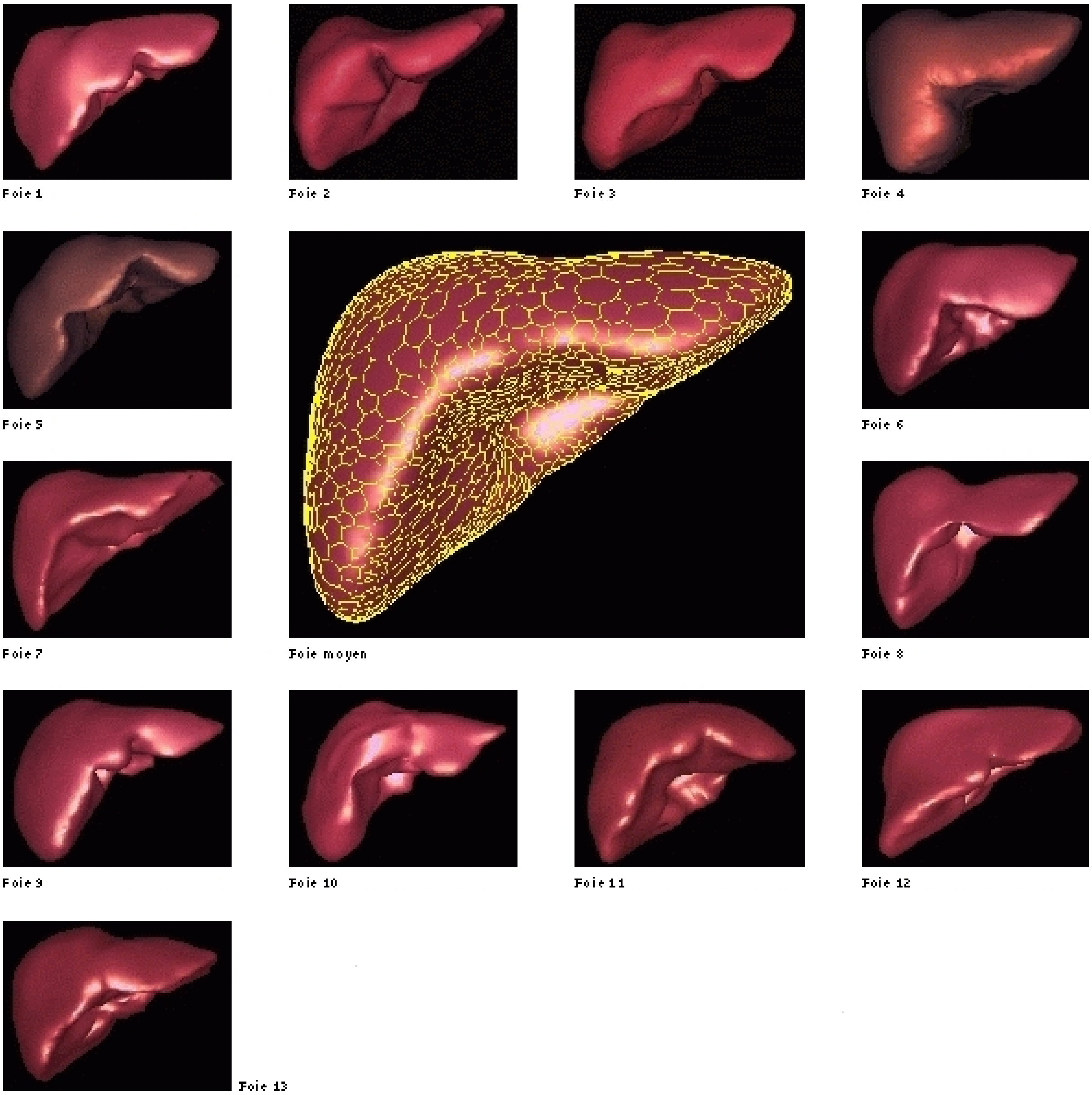


FIG. 2.10 – *Les 13 foies ayant servi à réaliser une étude statistique, ainsi que le foie moyen.*

Chapitre 3

Calcul d'un maillage moyen

Les maillages de foie dont on dispose peuvent être éloignés les uns des autres, ou avoir des orientations différentes. Si l'on faisait directement la moyenne de nos maillages, on obtiendrait très certainement quelque chose qui ne ressemblerait que fort peu à un foie, voire pas du tout. Or on veut à tout prix éviter cela. Notre «foie moyen» doit avoir une forme de foie, car il est destiné à servir d'initialisation pour des segmentations d'images tridimensionnelles du foie.
Il est important de définir ici ce que l'on entend par «forme». Nous utiliserons la définition donnée par Xavier Pennec dans [Pen96a] ou par Herbert Ziezold dans [Zie94] :

Définition 6 *Soient X et Y deux vecteurs de $\mathbb{R}^k$, et $\mathfrak{G}$ un groupe de transformations agissant sur $\mathbb{R}^k$. On dira que X et Y ont la même forme modulo $\mathfrak{G}$, s'il existe un élément g de $\mathfrak{G}$ tel que $Y = g.X$.*

Typiquement, $\mathfrak{G}$ sera le groupe des transformations rigides (rotations et translation), des similitudes (rotations, translation et homothéties) ou des transformations affines (applications linéaires inversibles et translations) :
Soit M un maillage à n sommets notés $M_1, \ldots, M_n$. Une transformation rigide de M sera la déformation qui transforme M en un maillage M' tel que ses sommets M'_i vérifient $M'_i = R.M_i + t$, où R est une matrice de rotation de taille 3×3, et où t un vecteur de $\mathbb{R}^3$. Une similitude sera la déformation qui transforme M en un maillage M' tel que ses sommets M'_i vérifient $M'_i = sR.M_i + t$, où s est un réel strictement positif, R une matrice de rotation de taille 3×3, et où t un vecteur de $\mathbb{R}^3$. Enfin, une transformation affine de M sera la déformation qui transforme M en un maillage M' tel que ses sommets M'_i vérifient $M'_i = A.M_i + t$, où A est une matrice inversible de taille 3×3, et où t un vecteur de $\mathbb{R}^3$.

Approfondissons le cas où $\mathfrak{G}$ est l'ensemble des transformations rigides.

3.1 Cas des recalages rigides

Pour obtenir un maillage moyen qui ressemble à un foie, on va appliquer à chacun de nos maillages une transformation de $\mathfrak{G}$, de manière à les amener dans une configuration où ils soient le plus proche possible les uns des autres. On dira alors qu'ils sont dans une position optimale les uns par rapport aux autres. Définissons ce terme de façon rigoureuse :

Définition 7 *Soient $X_1, \ldots, X_N$ des vecteurs de $\mathbb{R}^k$. On dira que les X_i sont dans une position optimale les uns par rapport aux autres si :*

$$\sum_{1\leq i<j\leq N} ||X_i - X_j||_2^2 \;=\; \inf_{g_1,\ldots,g_N \;\in\mathfrak{G}} \sum_{1\leq i<j\leq N} ||g_i.X_i - g_j.X_j||_2^2$$

(Nos maillages sont des vecteurs de $\mathbb{R}^{3n}$)

Le fait que les foies soient dans une position optimale les uns par rapport aux autres est d'une importance majeure pour le calcul d'une forme moyenne de foie. Dès lors que nos maillages sont dans une telle configuration, le calcul de leur moyenne arithmétique donne un maillage qui ressemble effectivement à un foie.
D'après [Pen96b], le maillage moyen obtenu est une forme moyenne au sens suivant :

Définition 8 *Le vecteur M est une forme moyenne des vecteurs X_i si :*

$$M \;\in\; \inf_{Y\in\;\mathbb{R}^k}\; \inf_{g_1,\ldots,\;g_N\;\in\;\mathfrak{G}} \sum_{1\leq i\leq N} ||Y \;-\; g_i.X_i||^2$$

Remarquons qu'il n'y a pas unicité de la forme moyenne, celle-ci étant définie modulo $\mathfrak{G}$.
Montrons le théorème suivant, qui affirme que si les maillages sont dans une position optimale, alors leur moyenne est une forme moyenne :

Théorème 1 *Le vecteur M est une forme moyenne des X_i si et seulement s'il existe $g_1,\ldots,g_N \in \mathfrak{G}$ tels que les $g_i.X_i$ soient dans une configuration optimale et tels que $M \;=\; \frac{1}{N}\sum_i g_i.X_i$.*

Démonstration 1 *Pour $g_1,\ldots,\,g_N \;\in\; \mathfrak{G}$ et $M \;\in\; \mathbb{R}^k$:*

$\sum_{1\leq i<j\leq N} ||g_i.X_i \;-\; g_j.X_j||^2 \;=\; \frac{1}{2} \sum_{1\leq i\leq N}\sum_{1\leq j\leq N} ||g_i.X_i \;-\; g_j.X_j||^2$

$\;=\; \frac{1}{2} \sum_{1\leq i\leq N}\sum_{1\leq j\leq N} (||g_i.X_i - M||^2 \;+\; ||g_j.X_j - M||^2 \;-\; 2\,\langle g_i.X_i - M,\, g_j.X_j - M\rangle)$

$\;=\; \frac{1}{2} \sum_{1\leq i\leq N} \left(N||g_i.X_i - M||^2 \;+\; \sum_{1\leq j\leq N} ||g_j.X_j - M||^2 \;-\; 2N\,\langle g_i.X_i - M,\, \frac{1}{N}\sum_{1\leq j\leq N} g_j.X_j - M\rangle \right)$

$\;=\; N \sum_{1\leq i\leq N} ||g_i.X_i - M||^2 \;-\; N^2\; ||\frac{1}{N}\sum_{1\leq i\leq N} g_i.X_i \;-\; M||^2$

Ce qui se récrit :

$$\sum_{1\leq i\leq N} ||g_i.X_i - M||^2 \;=\; N\;||\frac{1}{N}\sum_{1\leq i\leq N} g_i.X_i \;-\; M||^2 \;+\; \frac{1}{N}\sum_{1\leq i<j\leq N} ||g_i.X_i \;-\; g_j.X_j||^2$$

Par définition, M est une forme moyenne des X_i si et seulement si il existe des g_i tels que le terme de gauche soit minimal (en minimisant sur les g_i et sur$M$).
Soient $f_1,\ldots,\,f_N \;\in\; \mathfrak{G}$ tels que les $f_i.X_i$ soient dans une position optimale. On a :

$$\frac{1}{N}\sum_{1\leq i<j\leq N} ||f_i.X_i - f_j.X_j||^2 \;\leq\; \frac{1}{N}\sum_{1\leq i<j\leq N} ||g_i.X_i - g_j.X_j||^2 \;\leq\; N\;||\frac{1}{N}\sum_{1\leq i\leq N} g_i.X_i - M||^2 + \frac{1}{N}\sum_{1\leq i<j\leq N} ||g_i.X_i \;-\; g_j.X_j||^2$$

avec égalité entre les deux extrêmes si et seulement si $M \;=\; \frac{1}{N}\sum_{1\leq i\leq N} g_i.X_i$ et si les $g_i.X_i$ sont dans une position optimale.
Donc $\sum_{1\leq i\leq N} ||g_i.X_i - M||^2$ est minimal si et seulement si les $g_i.X_i$ sont dans une position optimale et si $M \;=\; \frac{1}{N}\sum_{1\leq i\leq N} g_i.X_i$.

L'algorithme que nous avons utilisé est décrit sur la figure 3.2 : On commence par recaler tous les maillages sur le foie du Visible Man de la NLM, afin qu'ils soient correctement orientés, ce qui nous permet d'effectuer une première moyenne qui ressemble à un foie. Puis nous itérons la procédure suivante : nous recalons encore une fois les nouveaux maillages sur leur ancienne moyenne ; nous calculons leur nouvelle moyenne, que nous recalons enfin sur le maillage du foie

du Visible Man.
Le fait de recaler systématiquement les maillages sur le maillage moyen pour recalculer la moyenne nous assure qu'après convergence, les nouveaux maillages d'apprentissage transformés soient recalés au mieux sur leur moyenne : il n'existe alors pas de recalage qui puisse rapprocher l'un des maillages d'apprentissage de leur moyenne.
En revanche, cela n'implique pas que ces maillages soient dans une configuration optimale. En effet, imaginons que nous ayons $N = 2p$ maillages de formes identiques, chacun de barycentre nul, tels que pour $1 \leq i \leq p$, $X_i = -X_{p+i}$. Alors leur moyenne est nulle, et ils sont tous recalés au mieux sur leur moyenne. Par ailleurs, il est clair qu'ils ne sont pas dans une position optimale : une telle position serait la superposition exacte de ces maillages. Cela dit, il s'agit d'un exemple où les maillages sont recalés au mieux sur leur moyenne, mais de manière instable, si bien qu'une petite variation rend leur moyenne non-nulle et force les maillages à se recaler les uns sur les autres.
Pour éviter ce genre de phénomène, on commence par recaler tout les maillages sur celui du Visible Man de la NLM.
En outre, le recalage du modèle moyen sur le modèle de la NLM à la fin de chaque itération a aussi pour but de stabiliser la position du modèle moyen dans un repère fixe.
Pratiquement, on stoppe la boucle lorsque le maillage moyen s'est stabilisé. Le critère de stabilisation utilisé est que la distance entre deux maillages moyens successivement calculés soit inférieure à 10^{-10} millimètres par sommet.

3.2 Recalages non rigides

En fait, dans le cas de recalages non rigides, la notion de position optimale se définit de manière différente, et les démonstrations sont plus complexes. En effet, le facteur d'échelle devient problématique : la position optimale est obtenue lorsque celui-ci tend vers zéro, et tous les maillages sont concentrés en un point. Il faudrait donc modifier notre algorithme. Mais les maillages que nous utilisons sont tous à peu près à la même échelle, et le premier recalage sur le maillage du foie du Visible Man de la NLM impose qu'ils aient tous la même échelle. Nous conservons donc ce même algorithme, qui donne un résultat sensiblement identique dans le cas des recalages par similitude.

Pour les expériences, nous avons choisi le recalage par similitude. Il y a deux raisons à cela : l'une d'elles est que les recalages par similitude conservent la «forme» au sens intuitif du terme. En outre, le logiciel de segmentation développé et utilisé au sein du projet ÉPIDAURE utilise beaucoup ces recalages pour les combiner avec d'autres techniques de segmentation. Il était donc nécessaire, pour faire des comparaisons pertinentes, de rester dans le cadre des similitudes.
Le maillage moyen obtenu est tout à fait représentatif d'un foie, et convient parfaitement à l'usage que nous allons en faire.

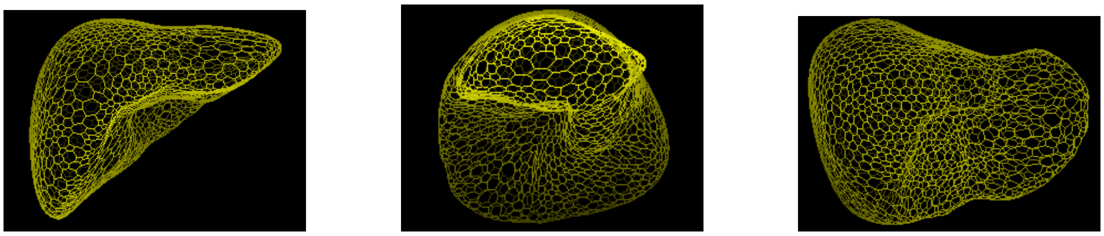


FIG. 3.1 – *Maillage moyen*

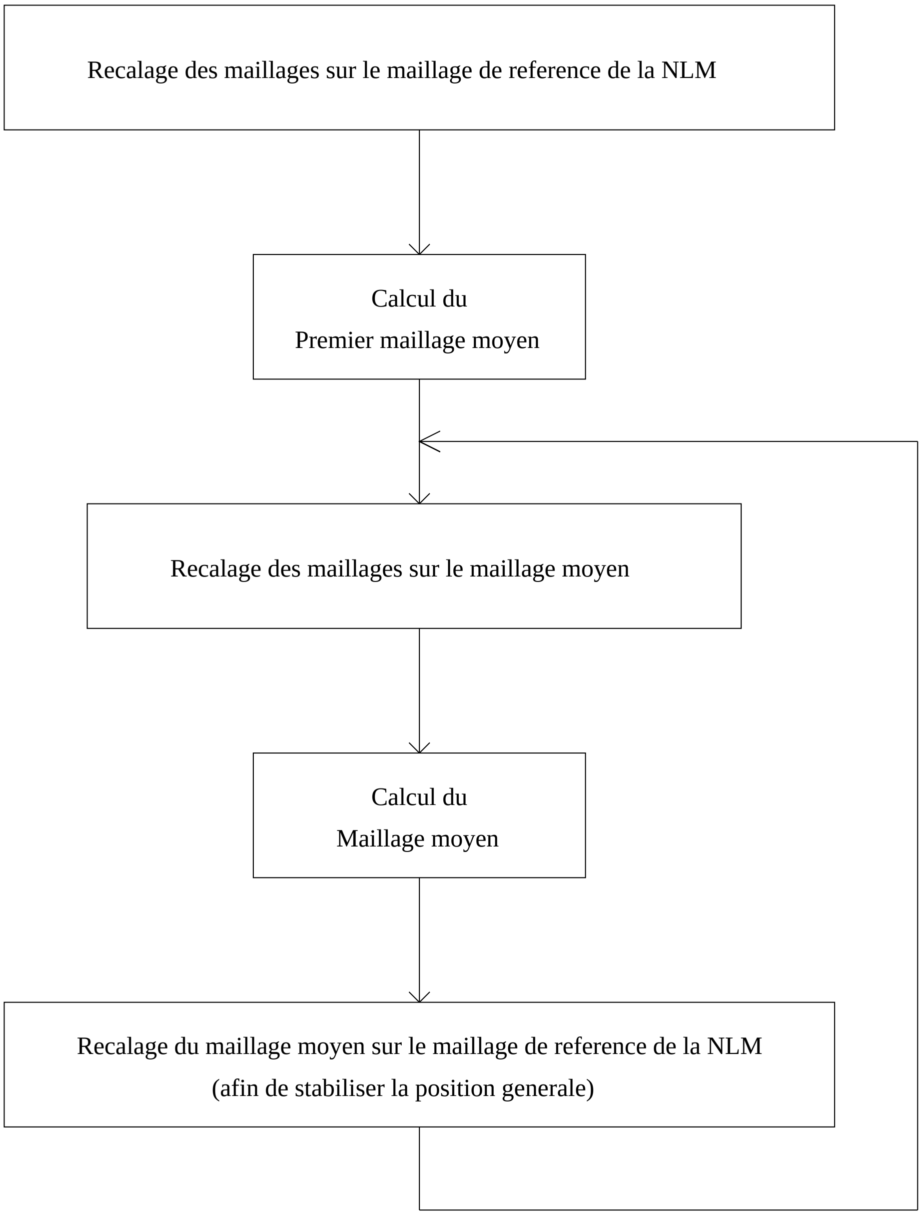


FIG. 3.2 – *Calcul du foie moyen*

Chapitre 4

Analyse en composantes principales (ACP)

L'analyse en composantes principales, ou ACP, est une méthode statistique permettant d'approcher de nombreux échantillons, représentés par des vecteurs, par un petit nombre de vecteurs de même dimension : la moyenne, et quelques déformations statistiquement probables.

4.1 Point de vue théorique

L'approche présentée ici est sensiblement celle de Laurent Younes dans [You].
Soient N vecteurs Z_1, Z_2, ..., Z_N d'un espace de Hilbert H de dimension d. On voudrait décrire au mieux ces N vecteurs par leur moyenne $\bar{Z}$ et par p autres vecteurs V_j, où p est un entier inférieur à N, en écrivant :

$$Z_i = \bar{Z} + \sum_{1 \le j \le p} a_{i,j} V_j + \epsilon_i$$

où les $a_{i,j}$ sont des scalaires, et où ϵ_i est l'erreur d'approximation commise, que l'on veut minimiser.
C'est la combinaison linéaire $\sum_{1 \le j \le p} a_{i,j} V_j$ elle-même qui importe, et non les vecteurs V_j ou les scalaires $a_{i,j}$; ceci nous permet d'imposer à la famille des vecteurs V_j d'être orthonormée. Les $a_{i,j}$ vérifient alors $a_{i,j} = \langle X_i, V_j \rangle$, où $X_i = Z_i - \bar{Z}$.
On veut minimiser la somme des erreurs ϵ_i en norme quadratique, c'est-à-dire la quantité suivante :

$$\sum_{1 \le i \le N} ||X_i - \sum_{1 \le j \le p} \langle X_i, V_j \rangle V_j||^2 = \inf_{W_1, \ldots, W_p \text{ famille orthonormale}} \sum_{1 \le i \le N} ||X_i - \sum_{1 \le j \le p} \langle X_i, W_j \rangle W_j||^2$$

$$= \sum_{1 \le i \le N} ||X_i||^2 - \sum_{1 \le j \le p} \sum_{1 \le i \le N} \langle X_i, V_j \rangle^2$$

On cherche donc à maximiser $\sum_{1 \le j \le p} \sum_{1 \le i \le N} \langle X_i, V_j \rangle^2$.

L'application $f(U,V) = \frac{1}{N} \sum_{1 \le i \le N} \langle X_i, U \rangle \langle X_i, V \rangle$ de H^2 dans $\mathbb{R}$ est une forme linéaire symétrique positive ; elle est donc diagonalisable dans une base orthonormée E_j de H.

Soient λ_1, $\lambda_2,\dots,\lambda_d$ les valeurs propres associées. Celles-ci sont réelles et positives, si bien que l'on peut réordonner les E_j de telle façon que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$.
La quantité à maximiser est $\sum_{1\leq j\leq p} f(V_j,V_j)$.

$$\text{On a : } f(V_j,V_j) = f\left(\sum_{1\leq k\leq d} \langle V_j,E_k\rangle\ E_k,\ \sum_{1\leq l\leq d} \langle V_j,E_l\rangle\ E_l\right) = \sum_{1\leq k,\ l\leq d} \langle V_j,E_k\rangle\ \langle V_j,E_l\rangle\ f(E_k,E_l)$$

$$f(V_j,V_j) = \sum_{1\leq k\leq d} \langle V_j,E_k\rangle^2\ \lambda_k.$$

Ainsi :

$$\sum_{1\leq j\leq p} f(V_j,V_j) = \sum_{1\leq j\leq p}\ \sum_{1\leq k\leq d} \lambda_k\ \langle V_j,E_k\rangle^2 = \sum_{1\leq k\leq d} \lambda_k\ \sum_{1\leq j\leq p} \langle V_j,E_k\rangle^2$$

Comme les V_j forment une famille orthonormale : $\sum_{1\leq j\leq p} \langle V_j,E_k\rangle^2 \leq ||E_k||^2 = 1$.
Or $\sum_{1\leq k\leq d}\ \sum_{1\leq j\leq p} \langle V_j,E_k\rangle^2 = \sum_{1\leq j\leq p} ||V_j||^2 = p$.
Posons $r_k = \sum_{1\leq j\leq p} \langle V_j,E_k\rangle^2$ pour $1 \leq k \leq d$. On veut maximiser $\sum_{1\leq k\leq d} \lambda_k\ r_k$ avec $0 \leq r_k \leq 1$ et $\sum_{1\leq k\leq d} r_k = p$. Il faut donc avoir $r_k = 1$ pour $1 \leq k \leq p$ et $r_k = 0$ sinon.
Donc pour tout $1 \leq k \leq p$, $\sum_{1\leq j\leq p} \langle V_j,E_k\rangle^2 = 1 = ||E_k||^2$, ce qui implique que E_k est dans l'espace engendré par $V_1,\dots,V_p$: $E_k \in Vect(V_1,\dots,V_p)$. Donc $Vect(V_1,\dots,V_p) = Vect(E_1,\dots,E_p)$.

Un choix optimal est donc $V_i = E_i$ pour tout $1 \leq i \leq p$. Ce choix a l'avantage d'être indépendant de l'entier p.
On approxime donc chaque X_i par sa projection orthogonale sur l'espace engendré par E_1, E_2, $\dots$, E_p. Les vecteurs propres E_i sont appelés les modes propres ou modes principaux.

Remarques importantes :

- L'erreur commise en remplaçant les X_i par leurs projections suivant les p premières composantes est facilement calculable à partir des valeurs propres des modes non pris en compte :

$$\sum_{1\leq i\leq N} ||\epsilon_i||^2 = \sum_{1\leq i\leq N} (||X_i||^2\ - \sum_{1\leq j\leq p} \langle X_i,E_j\rangle^2)$$

$$= \sum_{1\leq i\leq N}\ \sum_{p+1\leq j\leq d} \langle X_i,E_j\rangle^2 = N \sum_{p+1\leq j\leq d} \lambda_j$$

  Ainsi, la quantité $\frac{\sum_{1\leq j\leq p} \lambda_j}{\sum_{1\leq j\leq d} \lambda_j}$ permet d'estimer le nombre de modes nécessaires pour représenter un pourcentage donné de la variabilité des X_i.

- Si l'on considère les X_i comme des réalisations d'une variable aléatoire $\mathbf{X}$:
  λ_j n'est autre que la variance empirique des X_i selon la direction du mode j :

$$\lambda_j = \frac{1}{N} \sum_{1 \leq i \leq N} \langle X_i,\, E_j \rangle^2$$

C'est-à-dire que plus λ_j est grand, plus le mode j décrit de façon pertinente la variabilité des X_i.
Les projections de $\mathbf{X}$ suivant deux modes distincts i et j sont non-corrélées :

$$\mathbb{E}(\langle \mathbf{X},\, E_i \rangle \langle \mathbf{X},\, E_j \rangle) = 0$$

En effet, l'espérance empirique précédente s'écrit :

$$\frac{1}{N} \sum_{1 \leq k \leq N} \langle X_k, E_i \rangle \langle X_k, E_j \rangle,$$

expression qui est nulle d'après l'orthogonalité des E_i pour la forme f.

4.2 Mise en oeuvre de l'ACP

Les vecteurs X_i que nous allons considérer représentent des maillages simplexes à n sommets en trois dimensions, et seront donc des vecteurs de dimension $3n$: $X_i \in \mathbb{R}^{3n}$.
N'oublions pas que l'on veut obtenir des modes de déformation qui soient appliqués à un maillage moyen représentant un foie moyen : on ne peut pas se contenter de faire directement la moyenne des maillages des foies, car ceux-ci ont des orientations a priori différentes. Il faut donc les recaler de façon à ce qu'ils se superposent tous au mieux les uns sur les autres avant de faire leur moyenne et de réaliser l'ACP. Cette procédure a été décrite dans la partie «Calcul d'un maillage moyen».
Remarquons que si l'on fait des recalages rigides, au moins l'un des modes intègre le paramètre d'échelle. Si on recale par similitudes ou par transformations affines, alors le paramètre d'échelle n'intervient plus.

4.2.1 Un problème de temps de calcul

Notons X la matrice de taille $3n * N$ dont le i-ème vecteur colonne est X_i.
La forme quadratique f s'écrit alors :

$$\begin{aligned} f(U,V) &= \frac{1}{N} \sum_{1 \leq i \leq N} {}^tU\ X_i\ {}^tX_i\ V \\ &= {}^tU \left(\frac{1}{N} \sum_{1 \leq i \leq N} X_i\ {}^tX_i \right) V \\ &= {}^tU\ \frac{1}{N} X\ {}^tX\ V \end{aligned}$$

Il s'agit donc de diagonaliser la matrice $S = \frac{1}{N} X\ {}^tX$. Ses vecteurs propres sont les modes propres des X_i, et ses valeurs propres en sont les variances selon les modes associés. Mais S étant de taille $3n * 3n$, et n étant de l'ordre de quelques milliers en ce qui nous concerne, la diagonalisation de S est très coûteuse en temps de calcul.
Voici deux manières de contourner cette difficulté :

4.2.2 Première solution

Johannes Hug, Christian Brechbühler et Gábor Szélkely (dans [HBS99]) donnent une méthode intéressante qui consiste à remarquer que l'on peut travailler dans un espace de dimension bien moindre que $3n$:
en effet, les X_i engendrent un espace vectoriel de dimension au plus $N-1$ (on perd une dimension par la relation $\sum X_i = 0$). En procédant à une orthonormalisation de Gram-Schmidt de la famille des X_i, on obtient K vecteurs orthonormaux $M_1, \ldots, M_K$, avec $K \leq N-1$. Soit M la matrice dont les vecteurs colonnes sont les M_i. Nous noterons ceci de la manière suivante : $M = [M_1| \ldots |M_K]$.
Posons $\widetilde{X_i} =^t M \ X_i$: les coordonnées de $\widetilde{X_i}$ sont les produits scalaires de X_i avec les M_j.
Si $\widetilde{X_i} =^t [\widetilde{x}_i^1, \ldots, x_i^K]$, on a : $X_i = \sum\limits_{1 \leq j \leq K} \widetilde{x}_i^j \ M_j = M \ \widetilde{X_i}$
Le vecteur $\widetilde{X_i}$ est une autre représentation du vecteur X_i : il n'y a aucune perte d'information. On peut donc faire une analyse en composantes principales sur les $\widetilde{X_i}$ plutôt que sur les X_i. On obtient ainsi des vecteurs propres $\widetilde{V_i}$. Il suffit ensuite de revenir dans le bon espace : les $V_i = M \ \widetilde{V_i}$, une fois renormalisés, sont les vecteurs propres que l'on recherche.
Or les vecteurs $\widetilde{X_i}$ ont K composantes. On a donc réussi à réduire considérablement le temps de calcul, en représentant nos maillages par des vecteurs de taille inférieure à $N-1$.

4.2.3 Deuxième solution

La méthode présentée ici, utilisée par T.F.Cootes, C.J.Taylor, D.H.Cooper and J.Graham dans [TCDJ95], est moins élégante, mais plus facile à mettre en œuvre.
On peut s'affranchir du calcul de S, en le remplaçant avantageusement par la diagonalisation de la matrice $\Sigma = \frac{1}{3n} \ {}^tX \ X$, qui est de taille $N * N$, ce qui est beaucoup plus raisonnable, l'intérêt de l'ACP étant de travailler avec N petit devant la dimension de l'espace, à savoir $3n$.
De fait, la forme de la matrice S lui impose d'avoir un rang inférieur ou égal à N. Mais la contrainte $\sum X_i = 0$ nous indique que le rang de S est au plus $N-1$. Il s'avère que S et Σ ont les mêmes valeurs propres non nulles à un facteur multiplicatif près, et que les vecteurs propres associés aux valeurs propres non nulles de S (c'est-à-dire précisément ceux qui nous intéressent), se déduisent facilement des vecteurs propres associés aux valeurs propres non nulles de Σ.
Précisons cela :
Il existe $V \in SO_{3n}(\mathbb{R})$ et une matrice diagonale D de diagonale $(\lambda_1, \ \ldots, \lambda_{N-1}, 0, \ \ldots, 0)$ tels que $S = V \ D \ {}^tV$. De même. il existe $U \in SO_N(\mathbb{R})$ et une matrice diagonale Δ de diagonale $(\delta_1, \ \ldots, \delta_{N-1}, 0)$ tels que $\Sigma = U \ \Delta \ {}^tU$.
Soient $V_1, \ \ldots, V_{3n}$ les vecteurs colonne de V, et $U_1, \ \ldots, U_N$ ceux de U. Chaque V_i est un vecteur propre de S associé à la valeur propre λ_i. Il en va de même des U_i et des δ_i vis-à-vis de Σ. On a :

$$S \ V_i = \lambda_i \ V_i, \quad \Sigma \ U_i = \delta_i \ U_i.$$

C'est-à-dire :

$$X \ {}^tX \ V_i = N\lambda_i \ V_i, \quad {}^tX \ X \ U_i = 3n\delta_i \ U_i,$$

d'où :

$${}^tX \ X \ {}^tX \ V_i = N\lambda_i \ {}^tX \ V_i, \quad X \ {}^tX \ X \ U_i = 3n\delta_i \ X \ U_i,$$

soit :

$$\Sigma \ ({}^tX \ V_i) = \frac{N}{3n}\lambda_i \ ({}^tX \ V_i), \quad S \ (X \ U_i) = \frac{3n}{N}\delta_i \ (X \ U_i).$$

Si $\lambda_i \neq 0$ est une valeur propre de S : il existe un vecteur propre normal associé V_i.
On a $\Sigma\ ({}^tX\ V_i) = \frac{N}{3n}\lambda_i\ ({}^tX\ V_i)$, et par ailleurs, ${}^tX\ V_i \neq 0$ (sinon on aurait $\frac{1}{N}X\ {}^tX\ V_i = S\ V_i = 0$, ce qui est absurde). Par conséquent, $\frac{N}{3n}\lambda_i$ est une valeur propre non nulle de Σ, dont un vecteur propre associé est $\frac{{}^tX\ V_i}{||{}^tX\ V_i||}$.
Il existe donc un indice j tel que $\delta_j = \frac{N}{3n}\lambda_i$.

Si $\delta_i \neq 0$ est une valeur propre de Σ : il existe un vecteur propre normal associé U_i.
On a $S\ (X\ U_i) = \frac{3n}{N}\delta_i\ (X\ U_i)$, et $X\ U_i \neq 0$ sinon on aurait ${}^tX\ X\ U_i = 3n\ \Sigma\ U_i = 3n\delta_i\ U_i = 0$, ce qui est impossible. $\frac{3n}{N}\delta_i$ est donc une valeur propre non nulle de S, dont un vecteur propre associé est $\frac{X\ U_i}{||X\ U_i||}$.
Donc il existe un indice j tel que $\lambda_j = \frac{3n}{N}\delta$.

On en déduit donc que S et Σ ont les mêmes valeurs propres non nulles à un facteur $\frac{3n}{N}$ près.
On supposera, quitte à faire un changement d'indexation, que pour tout i tel que $1 \leq i \leq N-1$, on a $\lambda_i = \frac{N}{3n}\delta_i$. Par ailleurs, on supposera que $\lambda_1 \geq \lambda_2 \geq\ \ldots \geq \lambda_{N-1}$ (ce qui entraîne les mêmes inégalités pour les δ_i).

Marche à suivre :
En ce qui nous concerne, comme nous diagonalisons la matrice Σ, qui est de taille $N * N$, on a accès aux valeurs propres δ_i et à leurs vecteurs propres associés U_i.
Mais nous cherchons en fin de compte ceux de S, et il est aisé de les trouver :
Pour tout i tel que $\delta_i \neq 0$, on pose : $\lambda_i = \frac{3n}{N}\delta_i$,
et un vecteur propre associé à cette valeur propre non nulle de S est $\frac{X\ U_i}{||X\ U_i||} = \frac{X\ U_i}{\sqrt{3n\delta_i}}$
Si $i \neq j$, et même si $\lambda_i = \lambda_j$, on a : ${}^t(X\ U_i)(X\ U_j) ={}^t U_i\ 3n\Sigma\ U_j = 3n\delta_j\ {}^tU_i\ U_j = 0$.
Les vecteurs $(\frac{X\ U_i}{||X\ U_i||})_{i\ t.q.\ \lambda_i \neq 0}$ forment donc une famille orthonormale de vecteurs propres de S.

4.3 Résultats de l'analyse

4.3.1 quelques modes de déformation

Sur les figures 4.1, 4.2 et 4.3, on a représenté les trois premiers modes de l'ACP. Les différents maillages sont séparés les uns des autres pour une meilleure visualisation. Le maillage du milieu est toujours le foie moyen ; celui-ci est déformé selon chaque mode par un facteur de $-3\sqrt{\lambda_i}$, $-3/2\sqrt{\lambda_i}$, $3/2\sqrt{\lambda_i}$ ou $3\sqrt{\lambda_i}$.
Rappelons que le nombre de sommets du modèle est de 3716.

Il apparaît que le premier mode représente approximativement l'épaisseur du foie, le deuxième la flexion de celui-ci en son milieu, et le troisième, l'étendue de la partie basse.

4.3.2 Prépondérance des premiers modes ?

Une observation très intéressante en ce qui concerne les modes de déformation fournis par l'analyse en composantes principales est bien le fait que l'importance d'un mode est d'autant plus significative que la valeur propre associée est grande ; ainsi le mode associé à la plus grande valeur propre est celui qui décrit la déformation la plus caractéristique.

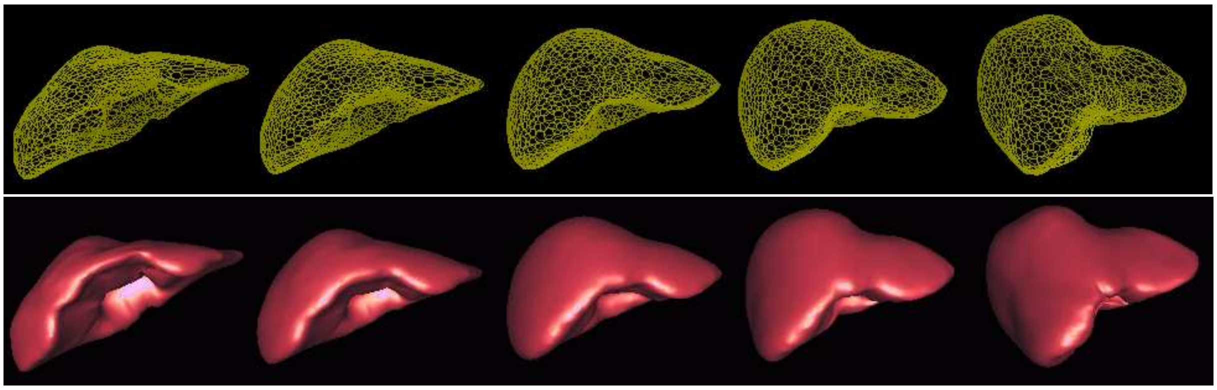


FIG. 4.1 – *Premier mode de l'ACP*

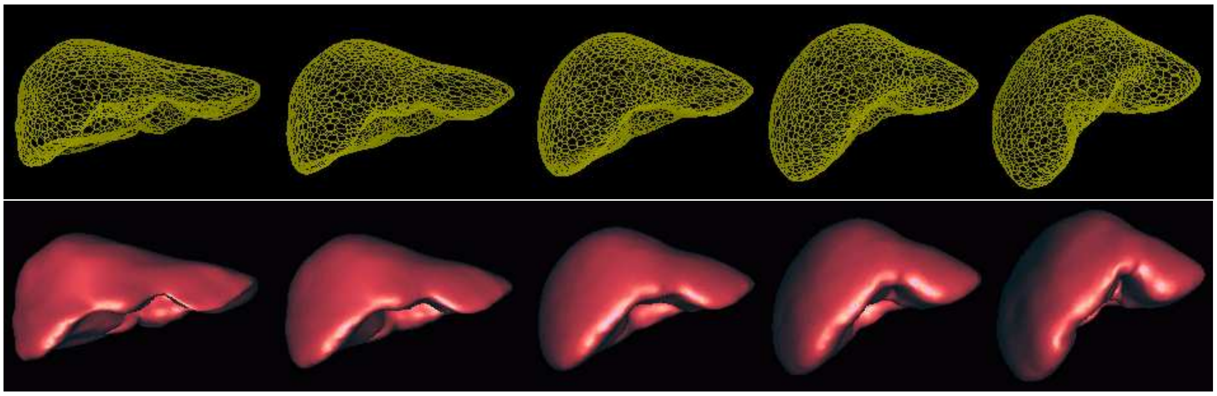


FIG. 4.2 – *Deuxième mode de l'ACP*

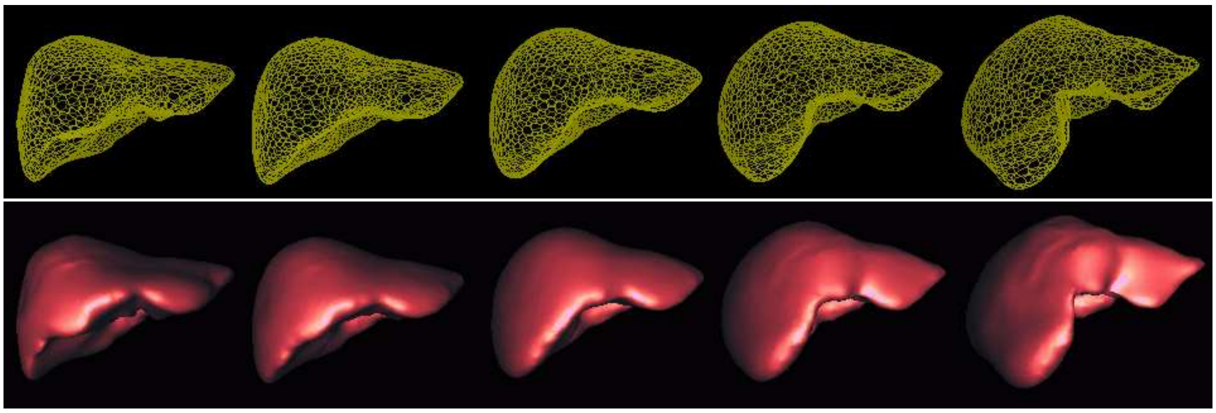


FIG. 4.3 – *Troisième mode de l'ACP*

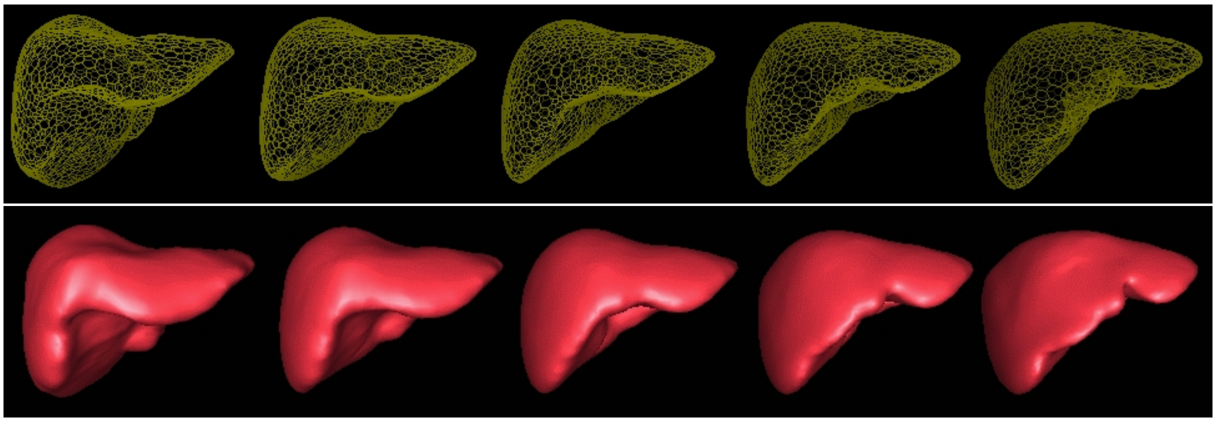


FIG. 4.4 – *Quatrième mode de l'ACP*

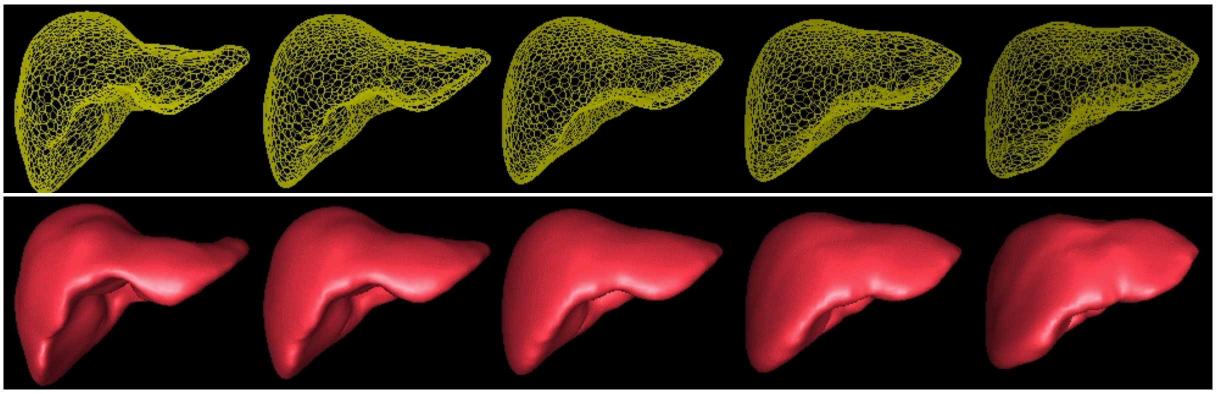


FIG. 4.5 – *Cinquième mode de l'ACP*

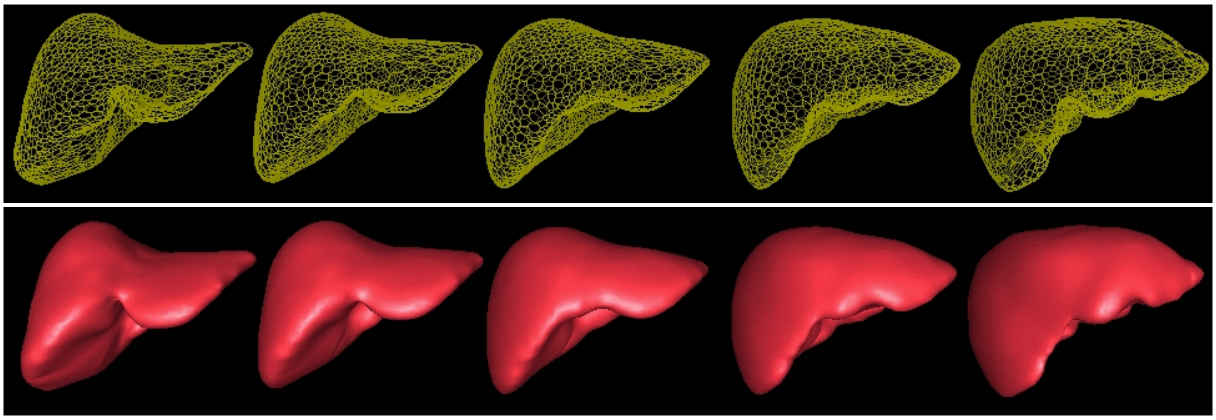


FIG. 4.6 – *Sixième mode de l'ACP*

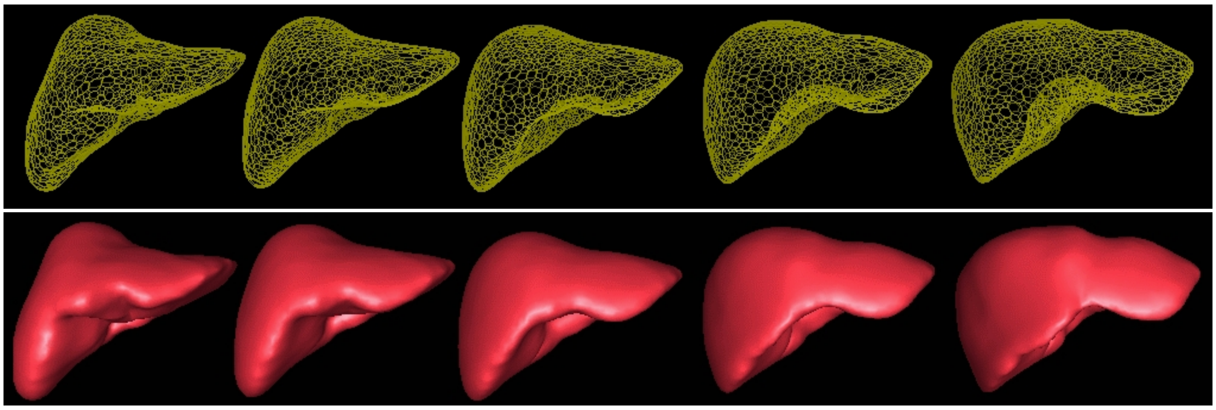


Fig. 4.7 – *Septième mode de l'ACP*

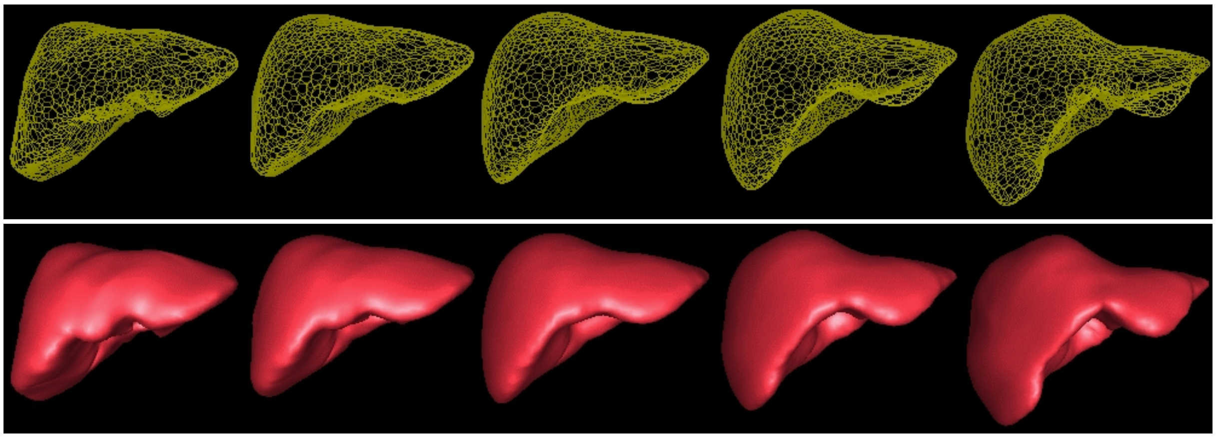


Fig. 4.8 – *Huitième mode de l'ACP*

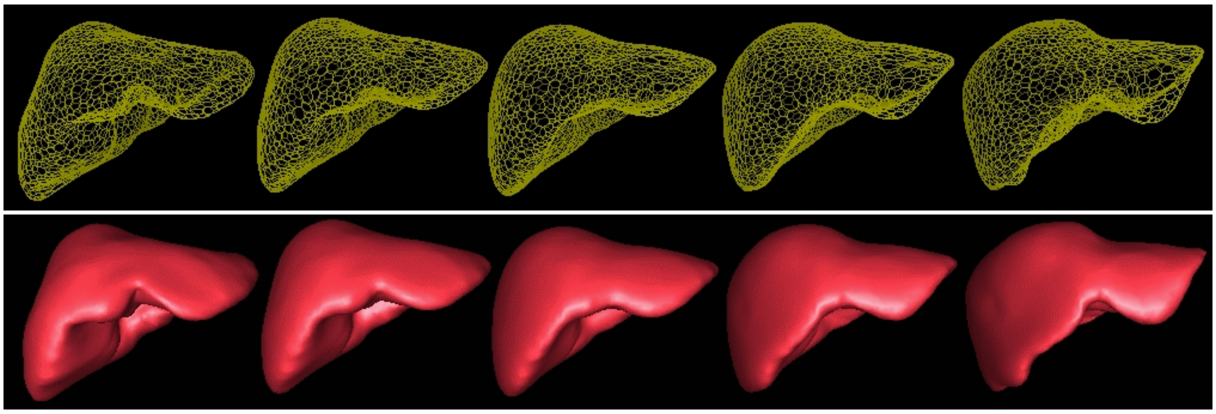


Fig. 4.9 – *Neuvième mode de l'ACP*

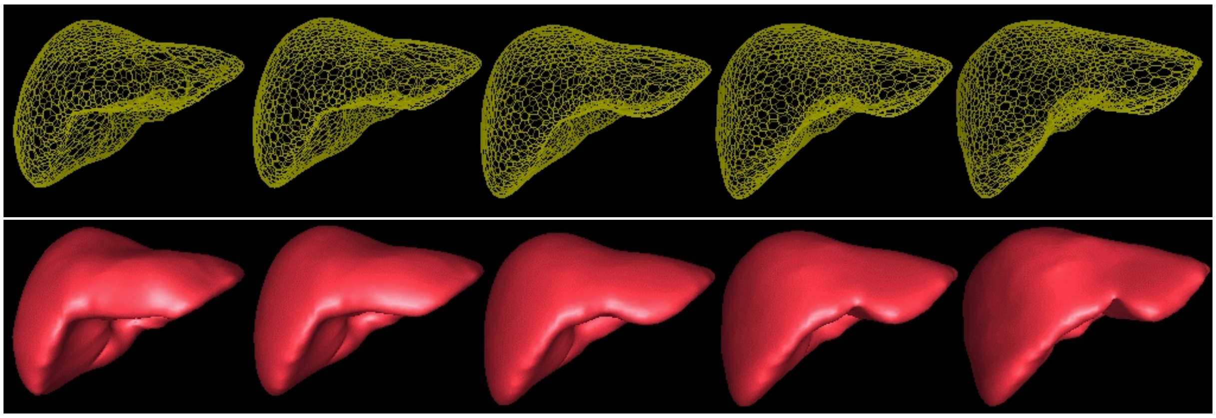


FIG. 4.10 – *Dixième mode de l'ACP*

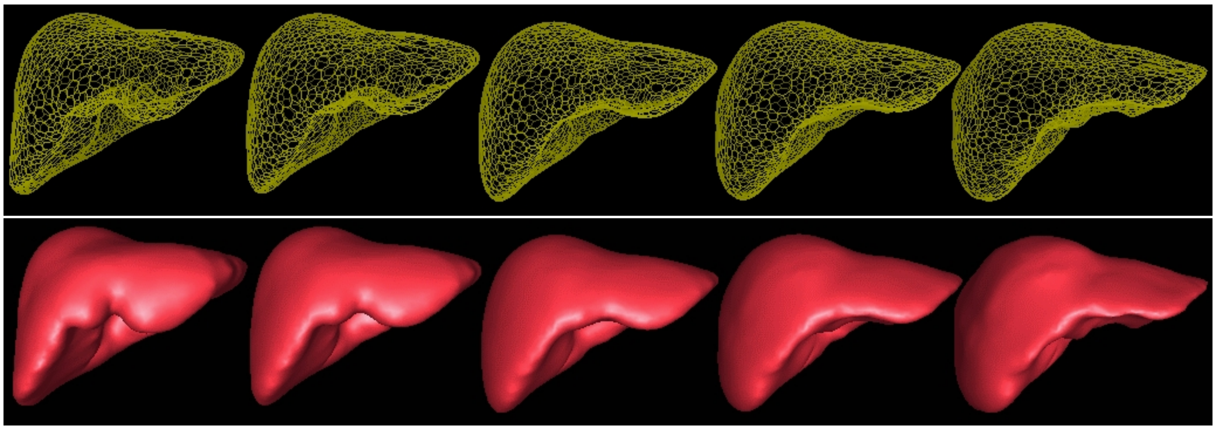


FIG. 4.11 – *Onzième mode de l'ACP*

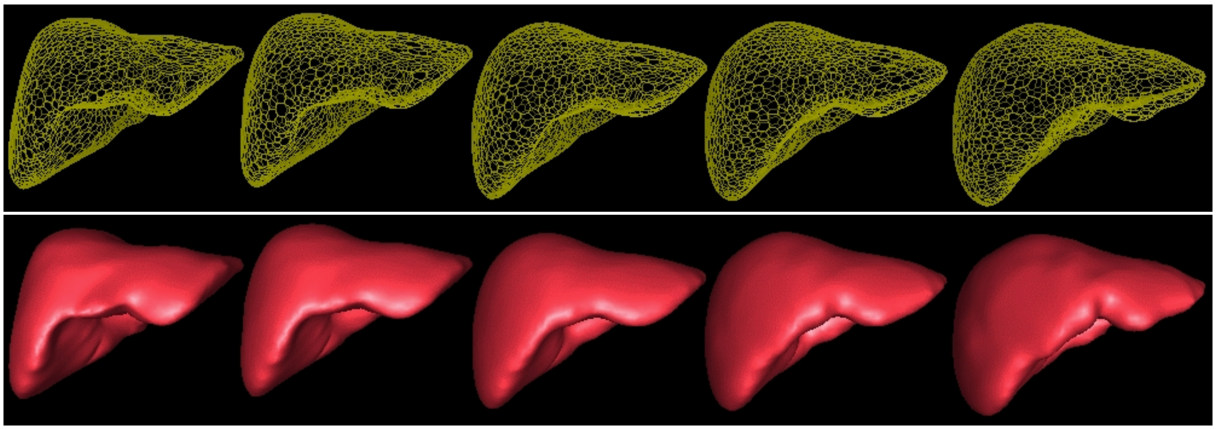


FIG. 4.12 – *Douzième mode de l'ACP*

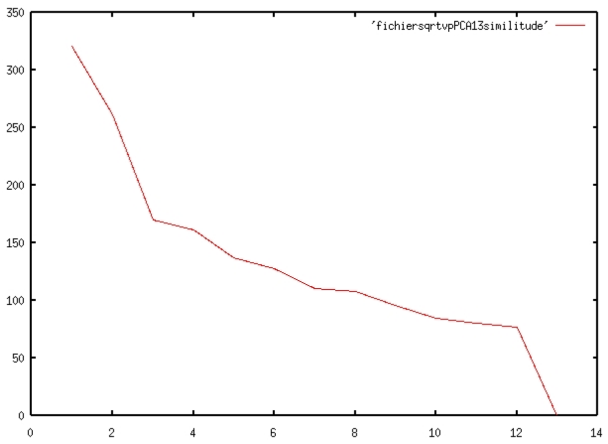


FIG. 4.13 – *Racines carrées des valeurs propres (en mm)*

Si les statistiques sont faites avec un nombre suffisamment grand de modèles (ce qui n'est pas vraiment le cas ici), un faible nombre de modes (comparés au nombre de modes disponibles) suffira à décrire l'essentiel des variations possibles.
Cette affirmation découle de ce qui suit :
Soit $p \leq N$.

$$\begin{aligned}\sum_{1\leq k\leq N} ||X_k - \sum_{1\leq i\leq p} \langle X_k, V_i\rangle\ V_i||^2 &= \sum_{1\leq k\leq N} ||\sum_{i>p} \langle X_k, V_i\rangle\ V_i||^2 \\ &= \sum_{i>p} \sum_{1\leq k\leq N} \langle X_k, V_i\rangle^2 \\ &= N \sum_{i>p} \lambda_i\end{aligned}$$

Ainsi, si l'on remplace les modèles d'apprentissage par leurs projections suivant les p premiers modes, l'erreur commise est donnée par la somme résiduelle des valeurs propres correspondant aux modes non pris en compte. Le graphe de la figure 4.14 montre, pour chaque p, la somme normalisée des p premières valeurs propres.

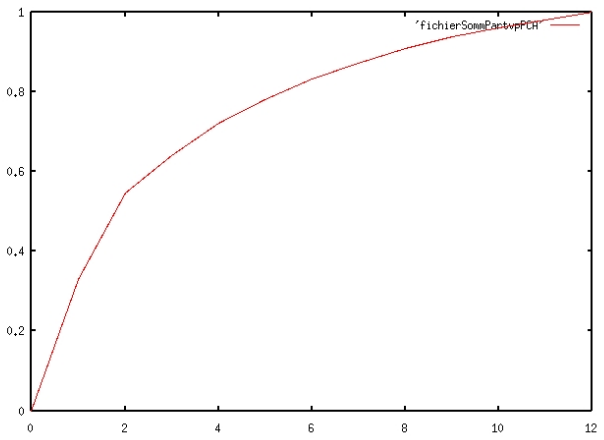


FIG. 4.14 – *Cumul des valeurs propres, normalisé*

On observe que les valeurs propres ne décroissent pas très rapidement, ce qui semble montrer que notre ensemble d'apprentissage n'est pas suffisamment grand pour qu'un petit nombre de modes suffise à représenter l'intégralité des déformations possibles. Pourtant, on verra dans le chapitre sur les résultats de la segmentation que, même avec aussi peu de modèles d'apprentissage, la moitié des modes seulement peut suffire à segmenter correctement une image de foie.

4.4 Les limites de l'ACP

On fait l'hypothèse que la distributions de nos échantillons est gaussienne (l'ACP ne prend en compte que les statistiques d'ordre 2, et c'est l'hypothèse gaussienne qui dans ce cas est la plus «quelconque», dans le sens où elle minimise l'information a priori). Ainsi l'on considérera que les réalisations pertinentes de notre modèle statistique correspondent aux points (de $\mathbb{R}^{3n}$ dans le cas de nos maillages) qui sont dans un ellipsoïde centré autour de la moyenne, dont les axes sont colinéaires aux modes détectés par l'ACP, et dont la longueur des axes est proportionnelle aux racines carrées des valeurs propres des modes (typiquement, la longueur d'un demi axe est de $3\sigma = 3\sqrt{\lambda}$, où λ est la valeur propre associée au mode considéré, soit encore la variance empirique selon cette direction). Ceci signifie que l'ACP donne une bonne représentation de notre variable aléatoire si l'espace des états peut effectivement être approximé par un ellipsoïde, ce qui est loin d'être évident et demande à être vérifié expérimentalement.
Afin de bien comprendre le problème, voici un schéma représentant plusieurs nuages de points (dans $\mathbb{R}^2$ au lieu de $\mathbb{R}^{3n}$, mais cela illustre bien le propos) ayant les mêmes statistiques d'ordre 2, à savoir même moyenne et mêmes variances.

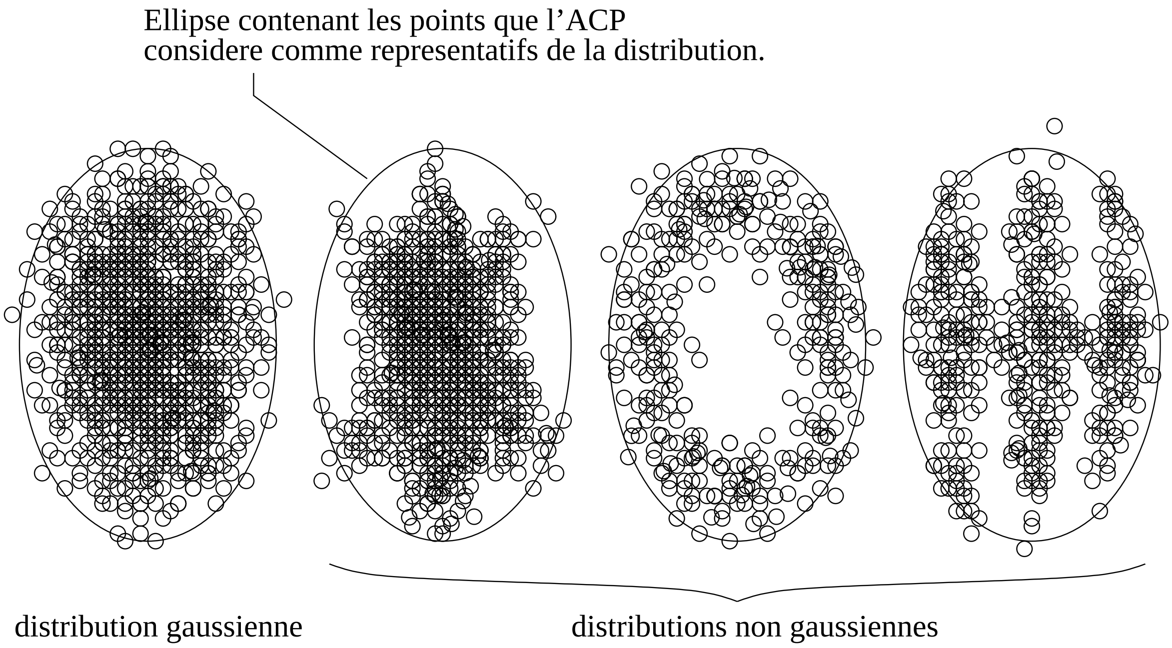


FIG. 4.15 – *distributions aléatoires de même moyenne et mêmes variances*

On voit bien que l'hypothèse gaussienne peut dans certains cas être loin de la réalité, dans le sens où un point pris dans l'ellipsoïde censé regrouper les échantillons probables peut en fait être un évènement quasi-impossible.
On ne peut exclure, par exemple, comme dans le deuxième exemple de distribution non gaussienne de la figure 4.15 que la moyenne arithmétique des réalisations ne soit pas elle-même une réalisation du processus aléatoire étudié.
Ainsi, le fait de supposer que le maillage moyen correspond bien à la forme d'un foie, et le fait de considérer que les formes «acceptables» sont celles qui sont situées dans un ellipsoïde centré

sur le maillage moyen et dont chaque demi-axe a une longueur égale à trois fois l'écart-type dans la direction du mode considéré, sont des approximations qui ne sont valables que parce que l'on constate visuellement que c'est une description cohérente. Des problèmes peuvent apparaître au niveau de la limite que nous avons arbitrairement fixée.

Chapitre 5

Analyse en composantes indépendantes (ICA)

Nous avons vu que l'analyse en composantes principales décorrèle les signaux, ce qui est l'approximation aux statistiques d'ordre deux de l'indépendance.
L'analyse en composantes indépendantes, ou ICA, cherche quant à elle une décomposition en signaux indépendants : on suppose que nos échantillons X_i sont des combinaisons linéaires de composantes indépendantes. C'est-à-dire que l'on a $\mathbf{X} = \Omega.\mathbf{C}$, où Ω est une matrice (inconnue) que l'on supposera inversible, et $\mathbf{C}$ un vecteur aléatoire dont les coordonnées sont des variables aléatoires indépendantes.
On cherche donc une matrice $W = \Omega^{-1}$ telle que si $\mathbf{X}$ est le vecteur aléatoire dont les X_i sont des réalisations, les composantes du vecteur $\mathbf{C} = W.\mathbf{X}$ soient le plus indépendantes possible.
On peut avoir le même point de vue avec l'ACP, pour laquelle les $\langle \mathbf{X}, V_i \rangle$ sont décorrélées. Il s'agit des composantes du vecteur ${}^tV.\mathbf{X}$, où V est la matrice orthonormale dont les vecteurs colonne sont les V_i.
L'ICA a de nombreuses applications, consistant essentiellement à séparer des signaux qui ont été linéairement combinés, en faisant l'hypothèse qu'ils sont indépendants les uns des autres.
Une application intéressante consiste à séparer divers signaux sonores superposés, par exemple une conversation ayant lieu sur fond musical. Une autre application impressionnante, mais plus artificielle, consiste à retrouver n images à partir de p combinaisons linéaires de ces images. Une des contrainte est d'avoir au moins autant de capteur qu'il y a de sources distinctes (dans notre exemple, un micro par source, ou au moins autant de combinaisons linéaires que d'images à retrouver : $n \leq p$).
Mais l'application qui nous concerne ici est encore relativement innovante. Cependant, il n'est pas évident a priori que le fait d'obtenir des composantes indépendantes soit plus efficace qu'une simple analyse en composantes principales. L'objet de cette partie est donc de comparer l'ICA avec l'ACP pour la description des déformations possibles.

5.1 Diversité des approches

Il existe de multiples manières d'appréhender l'ICA, tant du point de vue théorique que du point de vue algorithmique.
Il s'agit d'une méthode relativement récente et dont certains aspects échappent encore à une présentation rigoureuse. L'approche empirique est encore de mise.
Le lecteur intéressé pourra se reporter au livre [ICA98] de Te-Won Lee, ainsi qu'à l'article [HO99] de Aapo Hyvärinen.

Pour illustrer la diversité des approches possibles, nous allons en exposer deux. La première est celle de Laurent Younes dans [You], la seconde est présentée dans [HO99] et dans [Hug00].

5.2 Principe de l'ICA

Nous disposons de N échantillons X_i de vecteurs de $\mathbb{R}^d$. En ce qui nous concerne, il s'agit de nos maillages centrés, et afin de conserver des notations cohérentes avec ce qui précède, on prendra $d = 3n$.
On considère que nos vecteurs X_i sont des réalisations du vecteur aléatoire $\mathbf{X}$. Nous recherchons une application linéaire qui décompose les X_i en composantes le plus indépendantes possible. Appliquer ce principe de façon brute serait une erreur. En effet, cela reviendrait à chercher une matrice W telle que les composantes du vecteur $W.\mathbf{X}$ soient le plus indépendantes possibles. Mais cela donnerait $3n$ composantes, ce qui, dans notre cas, est énorme. Or, contrairement à l'ACP, il n'existe pas de méthode permettant de classer les composantes de l'ICA par ordre d'importance, comme c'était le cas pour l'ACP. C'est donc un réel problème. Par ailleurs, cela mène à des calculs matriciels beaucoup trop lourd : une simple multiplication de deux matrices $3n \times 3n = 11148 \times 11148$ est problématique... Il est donc indispensable de réduire la dimension de l'espace dans lequel on travaille. Pour cela, nous allons utiliser la base donnée par les vecteurs propres de l'ACP.

5.2.1 Réduction de la dimension de l'espace de travail

Afin de réduire le nombre de composantes, et de ne pas être confronté à des calculs matriciels trop lourds, on effectuera une analyse en composantes principales des X_i.
On sait que la matrice d'autocorrélation des X_i a au plus $N - 1$ valeurs propres non nulles (comptées avec ordre de multiplicité). Soit donc m le nombre de valeurs propres non nulles de $S = \frac{1}{N} \sum\limits_i X_i \, {}^tX_i$. Soient $\lambda_1 \geq \lambda_2 \geq \dots \lambda_m > 0$ les valeurs propres de S. Pour chaque λ_i, soit E_i un vecteur propre associé, de norme 1. Soit E la matrice dont les vecteurs colonne sont les E_i, et D la matrice diagonale dont les coefficients de la diagonale sont les λ_i. On a :

$$S = E \, D \, {}^tE$$

Soit encore $\boxed{\mathbf{Z} = {}^tE.\mathbf{X}}$:

$$\forall i \in \{1, \dots, N\},\ Z_i = {}^tE.X_i = \begin{bmatrix} \langle E_1, X_i \rangle \\ \vdots \\ \langle E_m, X_i \rangle \end{bmatrix}.$$

On a vu que $\forall i \in \{1, \dots, 3n\}$, $\lambda_i = \frac{1}{N} \sum\limits_j \langle E_i, X_j \rangle^2$, et $\forall i \geq m + 1$, $\lambda_i = 0$, donc pour tout j, X_j est dans l'espace engendré par $E_1, \dots, E_m$. Par conséquent, Z_i est une représentation exacte de X_i : on ne perd aucune information en remplaçant les X_i par les Z_i, pourvu que l'on garde en mémoire la matrice E, qui permet de passer des uns aux autres (on a aussi $X_i = E.Z_i$).

On peut donc faire une analyse en composantes indépendantes sur les Z_i, ce qui nous donnera m composantes. Par ailleurs, le passage des X_i aux Z_i ou inversement n'est pas excessivement coûteux en calculs.

Comme nous ne travaillons plus avec les vecteurs X_i de taille $3n$ mais avec les vecteurs Z_i de taille m, nous allons aussi changer de notation pour le vecteur de composantes

indépendantes, qui n'a plus que m composantes. Notons-le c. On cherche donc une matrice W, que l'on supposera inversible, telle que les composantes de $\mathbf{c} = W.\mathbf{Z}$ soient le plus indépendantes possibles, en un sens à préciser. On verra plus loin comment faire cela. Supposons que l'on y soit parvenus.
Réunissons toutes les réalisations Z_i de $\mathbf{Z}$ dans une matrice $Z = (Z_1|\ldots|Z_m)$. Cette notation signifie que le $i$ème vecteur colonne de Z est le vecteur Z_i. De même, les réalisations c_i de $\mathbf{c}$ peuvent être réunies dans une matrice $c = W.Z = W.[Z_1|\ldots|Z_m] = [c_1|\ldots|c_m]$.
Soit $\Omega = [\Omega_1|\ldots|\Omega_m] = W^{-1}$ dont chaque vecteur colonne est un c_i.

Rappel des notations :

- Les N vecteurs X_i de taille $3n$ sont nos échantillons, qui sont censés être des réalisations du vecteur aléatoire $\mathbf{X}$.
- On a exprimé les X_i dans la base des vecteurs propres de l'ACP, ce qui nous a donné N vecteurs Z_i de taille m, qui sont des réalisations d'un vecteur aléatoire $\mathbf{Z}$. Nous avons regroupés ces vecteurs Z_i dans une matrice Z de taille $m \times N$ pour avoir des expressions simplifiées.
- De même, les vecteurs $c_i = W.Z_i$ sont censées être des réalisations du vecteur aléatoire $\mathbf{c}$, et les vecteurs c_i ont été regroupés dans une matrice c de taille $m \times N$.
- Ω est l'inverse de la matrice W, et ses vecteurs colonne sont notés $\Omega_1, \ldots, \Omega_m$.
- On notera $\Lambda_i = E.\Omega_i$ le vecteur de taille $3n$ dont Ω_i est le représentant dans la base des modes de l'ACP.

On a : $Z_i = \Omega.c_i = [c_i]_i\,\Omega_1 + \ldots + [c_i]_m\,\Omega_m$, où $[c_i]_j$ est la $j$ème coordonnée du vecteur c_i. Donc nos vecteur de déformation (modes), correspondant aux composantes $\mathbf{c}_i$, sont les vecteurs Ω_i.
On passe de Z_i à X_i par :

$$X_i = \langle E_1, X_i\rangle E_1 + \ldots + \langle E_m, X_i\rangle E_m = [Z_i]_1\,E_1 + \ldots + [Z_i]_m\,E_m = E.Z_i$$

où $[Z_i]_j$ est la $j$ème coordonnée du vecteurZ_i.
Par conséquent, les «modes» de l'ICA dans l'espace d'origine de dimension $3n$ sont les :

$$\Lambda_i = E.\Omega_i$$

On a :

$$X_i = E.Z_i = E.\left([c_1]_i\,\Omega_1 + \ldots + [c_m]_i\,\Omega_m\right) = [c_1]_i\,\Lambda_1 + \ldots + [c_m]_i\,\Lambda_m$$

Soit encore, si on veut le récrire avec des vecteurs normalisés :

$$X_i = [c_1]_i\ ||\Lambda_1||\ \frac{\Lambda_1}{||\Lambda_1||} + \ldots + [c_m]_i\ ||\Lambda_m||\ \frac{\Lambda_m}{||\Lambda_m||}$$

Cette normalisation permet de comparer les variances des «modes» de l'ICA avec celles de l'ACP. Mais cette comparaison est relativement artificielle.

Il reste donc à calculer la matrice W.

5.2.2 Calcul de la matrice W

Position du problème

Les coordonnées du vecteur $\mathbf{c} = W.\mathbf{Z} = \begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_m \end{bmatrix}$ sont à valeurs dans $\mathbb{R}$. Nous allons les ramener dans l'intervalle [0,1] afin d'exploiter le fait que l'entropie d'un vecteur aléatoire à valeurs dans $[0,1]^m$ est maximale lorsque ses composantes sont indépendantes et uniformément distribuées. Il restera ensuite à maximiser l'entropie du vecteur aléatoire obtenu. C'est ce critère qui permettra de dire que les $\mathbf{c}_i$ sont le plus indépendantes possible.

Mais il faudrait ramener les coordonnées du vecteur $\mathbf{c}$ dans [0,1] de manière à ce que les coordonnées du vecteur obtenu aient des densités uniformes. C'est théoriquement possible, mais il faut pour cela connaître les densités f_i des $\mathbf{c}_i$:

Démonstration 2 *Soit* $F_i(x) = \int_{-\infty}^{x} f_i(u)du$, *et* $\mathbf{y}_i = F_i(\mathbf{c}_i)$.

Si α, β *in*[0,1],

$P(\mathbf{y}_i \in [\alpha,\beta]) = P(F_i(\mathbf{c}_i) \in [\alpha,\beta]) = P(\mathbf{c}_i \in [\underset{u}{ArgMin}\ \{F_i(u) = \alpha\},\ \underset{u}{ArgSup}\ \{F_i(u) = \beta\}])$

$= F_i(\underset{u}{ArgSup}\ \{F_i(u) = \beta\}) - F_i(\underset{u}{ArgMin}\ \{F_i(u) = \alpha\}) = \beta - \alpha$

Les $F_i(\mathbf{c}_i)$ *sont donc bien des variables uniformément distribuées sur* [0,1].

Mais nous ne disposons pas des F_i. Nous allons donc remplacer les fonctions F_i par des fonctions du type $\widetilde{q}(x) = \frac{e^{\lambda x}}{e^{\lambda x}+e^{-\lambda x}}$. Il est difficile de justifier le choix de ces fonctions. Certains auteurs considèrent que d'autres sont meilleures, et introduisent des raffinements. Il s'agit en fait que celle-ci soit empiriquement stable. L'idée est que l'on n'estime que les statistiques d'ordre deux ; celles-ci étant données, la densité qui introduit le moins d'information a priori est la densité gaussienne. On cherche donc une fonction qui s'en rapproche, et qui permette des calculs relativement simples. C'est-à-dire que la dérivée de $\widetilde{q}$ doit ressembler à une gaussienne. Dans le cas de la fonction sigmoïdale que nous avons choisi, on a $\widetilde{q}'(x) = 2\widetilde{q}(x)(1 - \widetilde{q}) = \frac{2\lambda}{\left(e^{\lambda x}+e^{-\lambda x}\right)^2}$, qui ne décroît pas aussi vite qu'une gaussienne, mais qui y ressemble quand-même. Remarquons que chaque terme λ peut être incorporé dans la matrice W, ce qui permet d'écrire $\widetilde{q}(x) = \frac{e^x}{e^x+e^{-x}}$.
Nous allons changer de variable encore une fois, et passer de $\mathbf{Z}$ à $\mathbf{Y}$:
soit $\mathbf{Y} = q(\mathbf{c}) = {}^t\,[\widetilde{q}([\mathbf{c}]_1), \ldots, \widetilde{q}([\mathbf{c}]_m)]$.
La notation $[\mathbf{c}]_i$ désigne la $i$ème coordonnée du vecteur $\mathbf{c}$.
Si l'on note par ${}^t w_i$ la $i$ème ligne de la matrice W, on a $\mathbf{Y} = {}^t\left[\widetilde{q}({}^t w_1.\mathbf{Z}), \ldots, \widetilde{q}({}^t w_m.\mathbf{Z})\right]$.
On veut maximiser l'entropie de $\mathbf{Y}$ sous la condition que W soit une matrice inversible :

$$\boxed{W = \underset{W\in\, GL_m(\mathbb{R})}{ArgMax}\ H(\mathbf{Y}) = \underset{W\in\, GL_m(\mathbb{R})}{ArgMax}\ -\int\cdots\int_{[0,1]^m} F(y_1, \ldots, y_m)\ln(F(y_1, \ldots, y_m))\ dy_1 \ldots\ dy_m}$$

où F est la densité associée à $\mathbf{Y}$.
Soit Φ le difféomorphisme de $\mathbb{R}^m$ tel que $\Phi(z) = q(W.z)$. On a $\mathbf{Y} = \Phi(\mathbf{Z}) = q(W.\mathbf{Z})$.
Si F est la densité de $\mathbf{Y}$ et G celle de $\mathbf{Z}$: $F(y) = \frac{G(\Phi^{-1}(y))}{|J_\Phi(\Phi^{-1}(y))|}$, où $|J_\Phi(\ )|$ est le module du déterminant du Jacobien de Φ.

Démonstration 3 *Changement de variable :*

En effet, on a, si f est une fonction de $\mathbb{R}^m$ dans lui-même :
$\mathbb{E}(f(\mathbf{Y})) = \int f(y)\ F(y)\ dy = \int f(\Phi(\xi))\ F(\Phi(\xi))\ |J_\Phi(\xi)|\ d\xi$
Soit $h(y) = F(y)\ |J_\Phi\left(\Phi^{-1}(y)\right)|$.
$\int f(\Phi(\xi))\ h(\Phi(\xi))\ d\xi = \int f(y)\ F(y)\ dy = \mathbb{E}(f(y)) = \mathbb{E}(f(\Phi(\xi))) = \int f(\Phi(\xi))\ G(\xi)\ d\xi$
Ceci étant vrai pour toute fonction f, on a : $G(\xi) = h(\phi(\xi))$, d'où :
$F(y) = \frac{G(\Phi^{-1}(y))}{|J_\Phi(\Phi^{-1}(y))|}$

Donc :

$$\begin{aligned} H(\mathbf{Y}) &= -\int \ln(F(y))\ F(y)\ dy = -\int \ln\left(\frac{G(\Phi^{-1}(y))}{|J_\Phi(\Phi^{-1}(y))|}\right)\ \frac{G(\Phi^{-1}(y))}{|J_\Phi(\Phi^{-1}(y))|}\ dy \\ &= -\int \ln\left(\frac{G(z)}{|J_\Phi(z)|}\right)\ G(z)\ dz = -\int \ln(G(z))\ G(z)\ dz + \int \ln(|J_\Phi(z)|)\ G(z)\ dz \\ &= H(\mathbf{Z}) + \mathbb{E}(\ln(|J_\Phi(\mathbf{Z})|)) \end{aligned}$$

Le terme $H(\mathbf{Z})$ ne dépend pas de W. On doit donc maximiser la quantité $\mathbb{E}(\ln(|J_\Phi(\mathbf{Z})|))$, qui sera estimée empiriquement. La matrice J_Φ dépend de la matrice W, et pour ne pas oublier cela, nous allons la noter J_W.
L'estimation empirique de $\mathbb{E}(\ln(|J_\Phi(\mathbf{Z})|))$ donne :

$$\mathbb{E}(\ln(|J_\Phi(\mathbf{Z})|)) = \mathbb{E}(\ln(|J_W(\mathbf{Z})|)) = \frac{1}{N} \sum_{1 \le i \le m} \ln(|\det J_W(Z_i)|)$$

Le problème est donc : $\boxed{W = \underset{W \in GL_m(\mathbb{R})}{ArgMax} \left(\frac{1}{N} \sum_{1 \le i \le m} \ln(|\det J_W(Z_i)|) \right)}$

Pour le résoudre, nous allons utiliser un algorithme de remontée de gradient.
Soit $\boxed{L(W) = \frac{1}{N} \sum_{1 \le i \le m} \ln(|\det J_W(Z_i)|)}$ la quantité à maximiser.
On a : $L(W+H) = L(W) + \langle \nabla_W L, H \rangle + o(H)$
L'algorithme de gradient s'écrit alors :

$$\boxed{W_{t+1} = W_t + \epsilon \nabla_W.L}$$

ϵ étant un pas à choisir convenablement.
Il faut donc introduire un produit scalaire sur l'espace tangent à celui des matrices carrées de taille $m \times m$, c'est-à-dire sur $\mathcal{M}_{m,m}(\mathbb{R})$. On va choisir un produit scalaire qui dépendra de l'endroit sur $GL_m(\mathbb{R})$ où se trouve la matrice W.

Choix du produit scalaire

Nous allons maximiser une fonctionnelle sur la variété $GL_m(\mathbb{R})$ par un algorithme de gradient. Pour cela, nous allons utiliser une métrique invariante à gauche, qui contraint d'autant plus fortement les variations que l'on est vers le bord de $GL_m(\mathbb{R})$.
Avoir une métrique invariante à gauche signifie que l'on a :

$$\forall W_1, W_2 \in GL_m(\mathbb{R}),\ ||H||_{W_1} = ||W_2 H||_{W_2.W_1}$$

Ceci implique que :

$$\forall W \in GL_m(\mathbb{R}),\ ||H||_W = ||W^{-1}H||_I$$

On pose
$$\langle H_1,\, H_2\rangle_W \;=\; Tr({}^tH_1.{}^tW^{-1}.W^{-1}.H_2)$$
On a donc : $||H||_W^2 \;=\; Tr({}^tH.{}^tW^{-1}.W^{-1}.H)$.

L'intérêt de cette métrique est qu'elle contraint d'autant plus les variations que W est proche d'une matrice non inversible.

Revenons à notre problème

Notons $\frac{dL}{dW} \bullet H$ le résultat de $\frac{dL}{dW}$ calculée en W appliquée à H.
On a : $\frac{dL}{dW} \bullet H \;=\; \sum_{k,l} (\frac{dL}{dW})_{k,l} H_{k,l} = \; Tr({}^t\frac{dL}{dW}.H)$
Mais, par définition du gradient, $\frac{dL}{dW} \bullet H \;=\; \langle \nabla_W L,\, H\rangle_W$, d'où :

$$\forall\, H \;\in\; \mathcal{M}_{m,m}(\mathbb{R}),\;\; Tr({}^t\frac{dL}{dW}.H) \;=\; Tr({}^t\nabla_W L.{}^tW^{-1}.W^{-1}.H)$$

Donc :

$$\boxed{\nabla_W L \;=\; W.{}^tW.\frac{dL}{dW}}$$

Résolution du problème

Exprimons J_W en fonction des données, afin de pouvoir exprimer $\frac{dL}{dW}$ en fonction des différents paramètres.
Rappelons que nous notons ${}^tw_k$ la $k$ème ligne de la matrice W. La $l$ème coordonnée de ${}^tw_k$ sera notée $w_{k,l}$, étant donné que c'est le terme en position k,l de la matrice W. Nous noterons $y_k = \widetilde{q}({}^tw_k.z)$ la $k$ème coordonnée du vecteur $y = q(W.z)$.
Pour tout $z \;\in\; \mathbb{R}^m$: $(J_W)_{k,l}(z) = \;\frac{\partial y_k}{\partial z_l} \;=\; \frac{\partial \widetilde{q}({}^tw_k.z)}{\partial z_l} \;=\; w_{k,l}.\widetilde{q}'({}^tw_k.z)$
Or $\widetilde{q}'(x) \;=\; 2\;\widetilde{q}(x).(1\;-\;\widetilde{q}(x))$, d'où :

$$(J_W)_{k,l}(z) \;=\; 2\; w_{k,l}.\widetilde{q}({}^tw_k.z).(1\;-\;\widetilde{q}(w_k.z)) = \;2\; w_{k,l}.y_k.(1\;-\;y_k)$$

La matrice J_W est donc la suivante :

$$J_W \;=\; 2 \begin{bmatrix} y_1(1-y_1) & & & \\ & y_2(1-y_2) & & O \\ & O & \ddots & \\ & & & y_m(1-y_m) \end{bmatrix} \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,m} \\ w_{2,1} & w_{2,2} & \dots & w_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m,1} & w_{m,2} & \dots & w_{m,m} \end{bmatrix}$$

Soit $K_W(z)$ la matrice diagonale des $y_k(1-y_k)$.
On a :
$$J_W(z) = \;2\; K_W(z).W$$
Rappelons que nous cherchons à calculer $\frac{dL}{dW}$, où $L(W) = \;\frac{1}{N} \sum_{1\leq i\leq m} \ln(|\det J_W(Z_i)|)$. Pour cela, nous allons calculer un développement limité à l'ordre 1 en H de $J_{W+H}(z)$, puis un développement limité à l'ordre 1 en H de $\ln(|\det J_{W+H}(z)|)$, pour avoir l'expression de la différentielle de L en W.

Calcul de $\frac{dL}{dW}$ grâce à des développements limités

Pour tout $H \in \mathcal{M}$, on a :

$$J_{W+H}(z) \; - \; J_W(z) = \; 2.(K_{W+H}(z) - K_W(z)).W \; + \; 2.K_W(z).H \; + \; o(H)$$

Notons ${}^t w_k$ le $k$ème vecteur ligne de W, et ${}^t h_k$ celui de H.

Démonstration 4 *Faisons le calcul pour chaque terme de la matrice $J_{W+H}(z) \; - \; J_W(z)$:*

$$\begin{aligned}
[J_{W+H}]_{k,l}(z) \; - \; [J_W]_{k,l}(z) &= (w_{k,l} + h_{k,l}).\widetilde{q}'({}^t w_k.z + {}^t h_k.z) \; - \; w_{k,l}.\widetilde{q}'({}^t w_k.z) \\
&= w_{k,l}.[\widetilde{q}'({}^t w_k.z + {}^t h_k.z) - \widetilde{q}'({}^t w_k.z)] \; + \; h_{k,l}.\widetilde{q}'({}^t w_k.z + {}^t h_k.z) \\
&= w_{k,l}.2.[[K_{W+H}]_{k,k}(z) - [K_W]_{k,k}(z)] \; + \; h_{k,l}.\widetilde{q}'({}^t w_k.z + {}^t h_k.z) \\
&= w_{k,l}.2.[[K_{W+H}]_{k,k}(z) - [K_W]_{k,k}(z)] \; + \; h_{k,l}.\widetilde{q}'({}^t w_k.z) \; + \; o({}^t h_k)
\end{aligned}$$

Posons : $y'_k = \; y_k(W+H) = \; \widetilde{q}({}^t(w_k + h_k).z)$, $y_k = \; y_k(W)$, $\omega_k = \; {}^t h_k.z$.
Soit $R_W(H)(z)$ la matrice diagonale des $(1 - 2y_k)({}^t h_k.z) = (1 - 2\widetilde{q}({}^t w_k.z))({}^t h_k.z)$.
On a

$$K_{W+H}(z) - K_W(z) = \; 2 \; R_W(H)(z).K_W(z) \; + \; o(H)$$

.

Démonstration 5 *Faisons le calcul pour chaque terme de la matrice $K_{W+H}(z) \; - \; K_W(z)$:*

$$\begin{aligned}
[K_{W+H}]_{k,k}(z) - [K_W]_{k,k}(z) &= y'_k.(1 - y'_k) \; - \; y_k.(1 - y_k) \\
&= [y_k + 2.y_k.(1 - y_k).\omega_k + o({}^t h_k)] \; [1 - y_k - 2.y_k.(1 - y_k).\omega_k + o({}^t h_k)] \; - \; y_k.(1 - y_k) \\
&= y_k.(1 - y_k) \; - \; 2.y_k^2.(1 - y_k).\omega_k \; + \; 2.y_k.(1 - y_k)^2.\omega_k \; - \; y_k.(1 - y_k) \; + \; o({}^t h_k) \\
&= 2.y_k.(1 - y_k).(1 - 2y_k).\omega_k + \; o({}^t h_k) \\
&= 2.[K_W]_{k,k}.(1 - 2y_k).\omega_k + \; o({}^t h_k)
\end{aligned}$$

Donc $J_{W+H}(z) \; = \; 2 \; K_{W+H}(z).W \; = \; 2 \; (K_W(z) + 2 \; R_W(H)(z).K_W(z) + o(H)).(W + H)$.
Donc : $J_{W+H}(z) = \; J_W(z) \; + \; 2 \; K_W(z).H \; + \; 4 \; R_W(H)(z).K_W(z).W \; + \; o(H)$.
Comme $J_W(z) = \; 2 \; K_W(z).W$:

$$J_{W+H}(z) = \; J_W(z) \; + \; J_W(z).W^{-1}.H \; + \; 2.R_W(H)(z).J_W(z) \; + \; o(H)$$

On a maintenant :

$$\begin{aligned}
\ln \; |\det J_{W+H}(z)| \; &= \; \ln \; |\det \left(J_W(z).\left(I \; + \; W^{-1}.H \; + \; 2 \; J_W(z)^{-1} R_W(z).J_W(z) \; + \; o(H)\right)\right)| \\
&= \; \ln \; |\det J_W(z)| \\
&\quad + \; \ln \; |\det(I \; + \; W^{-1}.H \; + \; 2 \; W^{-1}.K_W(z)^{-1}.R_W(z).K_W(z).W \; + \; o(H)|
\end{aligned}$$

Puisque les matrices $R_W(z)$ et $K_W(z)$ sont diagonales, elles commutent, ce qui nous donne :

$$\ln \; |\det J_{W+H}(z)| = \; \ln \; |\det J_W(z)| \; + \; \ln \; |\det(I \; + \; W^{-1}.H \; + \; 2W^{-1}.R_W(z).W \; + \; o(H))|$$

Il nous faut connaître la dérivée du logarithme du module du déterminant en l'identité pour pouvoir achever ce développement limité.
On a : $\frac{d(\ln \; |det(M)|)}{dM}(I) = \; I$.

Démonstration 6 *Pour cela, utilisons la propriété bien connue de la comatrice $\widetilde{M}$ (dite aussi matrice des cofacteurs) de la matrice M : $M.{}^t\widetilde{M} = \; det(M).I$*
Donc pour tout i : $\sum_k m_{i,k}\widetilde{m}_{i,k} = \; det(M)$
Or les termes $\widetilde{m}_{i,k}$ ne dépendent pas de $m_{i,j}$. Il est donc aisé de dériver cette expression : $\frac{\partial det(M)}{\partial m_{i,j}} = \; \widetilde{m}_{i,j}$
Donc : $\frac{d(det(M))}{dM} = \; \widetilde{M}$, soit, si M est inversible : $\frac{d(det(M))}{dM} = \; det(M).{}^t M^{-1}$
D'où : $\frac{d(\ln \; |det(M)|)}{dM} = \; \frac{1}{det(M)}.\frac{d(\det(M))}{dM} = \; {}^t M^{-1}$

$$
\begin{aligned}
&\ln\,|\det(I\ +\ W^{-1}.H\ +\ 2\ W^{-1}.R_W(z).W\ +\ o(H))| \\
&\quad = \ln\,|det(I)|\ +\ \sum_{i,j}\left[\frac{d(\ln\ |det(M)|)}{dM}(I)\right]_{i,j}\ \left[W^{-1}.H + 2\ W^{-1}.R_W(H)(z).W\right]_{i,j}\ +\ o(H) \\
&\quad = Tr\left(W^{-1}.H\ +\ 2\ W^{-1}.R_W(H)(z).W\right)\ +\ o(H) \\
&\quad = Tr\left(W^{-1}.H\right)\ +\ 2\ Tr\left(W^{-1}.R_W(H)(z).W\right)\ +\ o(H) \\
&\quad = Tr\left(W^{-1}.H\right)\ +\ 2\ \sum_{i,j} h_{i,j}.z_j.(1-2y_i)\ +\ o(H) \\
&\quad = Tr\left(W^{-1}.H\right)\ +\ 2\ Tr\left({}^t\Gamma_W(z).H\right)\ +\ o(H)
\end{aligned}
$$

où $\Gamma_W(z) =\ [(1-2y_i).z_j]_{i,j}$.

Donc : $\ln\,|\det J_{W+H}(z)|\ =\ \ln\,|\det J_W(z)|\ +\ Tr\left(W^{-1}.H\right)\ +\ 2\ Tr\left({}^t\Gamma_W(z).H\right)\ +\ o(H)$

$$
\begin{aligned}
\text{Or } \ln\,|\det J_{W+H}(z)| &= \ln\,|\det J_W(z)|\ +\ \sum_{i,j}\frac{d(\ln\ |\det J_W(z)|)}{dw_{i,j}}.h_{i,j}\ +\ o(H) \\
&= \ln\,|\det J_W(z)|\ +\ Tr\left({}^t\left[\frac{d(\ln\ |\det J_W(z)|)}{dW}\right].H\right)\ +\ o(H)
\end{aligned}
$$

Donc : $\frac{d(\ln|det(J_W(z))|)}{dW} =\ {}^tW^{-1}\ +\ 2.\Gamma_W(z)$.

Rappelons que : $L(W) =\ \frac{1}{N}\sum_{1\leq i\leq m}\ln(|\det J_W(Z_i)|)$

D'où :

$$\frac{dL}{dW} = \frac{1}{N}\sum_{1\leq i\leq m}\left(\ {}^tW^{-1}\ +\ 2.\Gamma_W(Z_i)\right)$$

Conclusion des calculs

Comme $\nabla_W L\ =\ W.{}^tW.\frac{dL}{dW}$,

$$\nabla_W L = \frac{m}{N}\ W\ +\ 2.W.{}^tW.\frac{1}{N}\sum_{1\leq i\leq m}\Gamma_W(Z_i)$$

Ce qui donne finalement :

$$\boxed{W_{t+1}\ =\ \frac{m}{N}\ W_t\ +\ \epsilon\ \left(W_t\ +\ 2.W_t.{}^tW_t.\frac{1}{N}\sum_{1\leq i\leq m}\Gamma_{W_t}(Z_i)\right)}$$

Cette méthode n'a malheureusement pas convergé ; nous allons donc en exposer une autre, qui est réputée rapide. Pour cela, il est nécessaire d'introduire la notion de néguentropie.

5.3 Définition de la néguentropie

Soit $\mathbf{v}$ un vecteur aléatoire de densité de probabilité $p_{\mathbf{v}}$.

Définition 9 *Soit $g_{\mathbf{v}}$ la densité de probabilité d'un vecteur aléatoire gaussien de même moyenne et de même matrice de covariance que $\mathbf{v}$. On appelle néguentropie de $\mathbf{v}$ la distance de Kullback entre $p_{\mathbf{v}}$ et $g_{\mathbf{v}}$:*

$$J(\mathbf{v})\ =\ \int\ p_{\mathbf{v}}(\mathbf{v})\ \ln(\frac{p_{\mathbf{v}}(\mathbf{v})}{g_{\mathbf{v}}(\mathbf{v})})\ d\mathbf{v}$$

Proposition 1 *La néguentropie de $\mathbf{v}$ s'exprime en fonction des entropies de $p_{\mathbf{v}}$ et de $g_{\mathbf{v}}$:*

$$J(\mathbf{v})\ =\ H(g_{\mathbf{v}})\ -\ H(p_{\mathbf{v}})$$

Note : on notera indifféremment $H(p_{\mathbf{v}})$ ou $H(\mathbf{v})$ l'entropie de $\mathbf{v}$, ce qui est justifié car l'entropie de $\mathbf{v}$ ne dépend que de sa densité de probabilité.

Démonstration 7 *Rappelons que* $H(\mathbf{v}) = - \int p_{\mathbf{v}} \ln(p_{\mathbf{v}})\, d\mathbf{v}$.
$J(\mathbf{v}) = \int p_{\mathbf{v}}(\mathbf{v}) \ln(\frac{p_{\mathbf{v}}(\mathbf{v})}{g_{\mathbf{v}}(\mathbf{v})})\, d\mathbf{v} = -H(\mathbf{v}) - \int p_{\mathbf{v}}(\mathbf{v}) \ln(g_{\mathbf{v}}(\mathbf{v}))\, d\mathbf{v} = -H(\mathbf{v}) - \int g_{\mathbf{v}}(\mathbf{v}) \ln(g_{\mathbf{v}}(\mathbf{v}))\, d\mathbf{v}$
La dernière inégalité provient du fait que $g_{\mathbf{v}}$ *et* $p_{\mathbf{v}}$ *ont les mêmes statistiques d'ordre 1 et 2, et que* $\ln(g_{\mathbf{v}}(\mathbf{v}))$ *est une forme quadratique réelle.*
Par conséquent : $J(\mathbf{v}) = H(g_{\mathbf{v}}) - H(p_{\mathbf{v}})$

On retrouve le fait qu'à moyenne et à variance données, la distribution gaussienne est celle qui a l'entropie la plus forte.

Proposition 2 *Une propriété intéressante de la néguentropie est que, contrairement à l'entropie, elle est invariante par l'effet d'une transformation linéaire inversible. La néguentropie hérite cette propriété de la distance de Kullback entre deux distributions.*

Démonstration 8 *Montrons-le :*
Soient $\mathbf{x}_1$ *et* $\mathbf{x}_2$ *deux variables aléatoires de densités respectives* p_1 *et* p_2.
Considérons une application linéaire inversible W, *et notons* $\mathbf{y}_1 = W.\mathbf{x}_1$ *et* $\mathbf{y}_2 = W.\mathbf{x}_2$, *de densités respectives* q_1 *et* q_2.
q_i *s'exprime en fonction de* p_i :

$$\begin{aligned} P(\mathbf{y}_i \in A) &= P(W.\mathbf{x}_i \in A) \\ &= P(\mathbf{x}_i \in W^{-1}.A) \\ &= \int_{W^{-1}.A} p_i(x)\, dx \\ &= \int_A p_i(W^{-1}.y)\, \frac{1}{|\det W|}\, dy \end{aligned}$$

Par conséquent, $q_i(y) = \frac{1}{|\det W|}\, p_i(W^{-1}.y)$.
La distance de Kullback entre les lois de $\mathbf{y}_1$ *et* $\mathbf{y}_2$ *est donnée par :*
$D(q_1,q_2) = \int \frac{1}{|\det W|}\, p_1(W^{-1}.y) \ln(\frac{p_1(W^{-1}.y)}{p_2(W^{-1}.y)})\, dy$
Soit ϕ *telle que* $x = \phi(y) = W^{-1}.y$.

$$\begin{aligned} D(q_1,q_2) &= \int p_1(\phi(y)) \ln(\frac{p_1(\phi(y))}{p_2(\phi(y))})\, |J_\phi(y)|\, dy \\ &= \int p_1(x) \ln(\frac{p_1(x)}{p_2(x)}) dx \\ &= D(p_1,p_2) \end{aligned}$$

($|J_\phi(y)|$ *est le module du déterminant de la matrice jacobienne de* ϕ *en* y*)*

5.4 Une méthode plus rapide

Puisqu'il existe de nombreuses manières d'implémenter une ICA, nous allons en présenter ici une plus rapide, donnée par Johannes Hug dans sa thèse [HBS99] ,et par Aapo Hyvarinen et Erkki Oja dans [HO99].
Il s' agit toujours de calculer la matrice W, mais cette fois-ci d'une manière différente.
Nous voulons rendre les composantes du vecteur $\mathbf{c}$ le plus indépendantes possible. Cela revient à dire que la densité de probabilité $p_{\mathbf{c}}(\mathbf{c})$ de $\mathbf{c}$ doit être le plus proche possible (en un sens à déterminer) du produit des densités marginales $p_{\mathbf{c}_i}(\mathbf{c}_i)$. Nous notons ici $\mathbf{c}_i$ la $i$ème coordonnée de $\mathbf{c}$.
Pour cela, nous allons minimiser la distance de Kullback entre ces deux densités :

$$I(\mathbf{c}) = \int p_{\mathbf{c}}(\mathbf{c}) \ln(\frac{p_{\mathbf{c}}(\mathbf{c})}{\prod_i p_{\mathbf{c}_i}(\mathbf{c}_i)})\, d\mathbf{c}$$

qui est positive, et nulle si et seulement si les $\mathbf{c}_i$ sont indépendantes.

5.4.1 Simplification de $I(\mathbf{c})$

L'information mutuelle $I(\mathbf{c})$ s'exprime en fonction des entropies de $\mathbf{c}$ et des $\mathbf{c}_i$:

$$\begin{aligned}
I(\mathbf{c}) &= -H(\mathbf{c}) - \int p_{\mathbf{c}}(\mathbf{c}) \ln(\prod_i p_{\mathbf{c}_i}(\mathbf{c}_i)) \, d\mathbf{c} \\
&= -H(\mathbf{c}) - \sum_i \int p_{\mathbf{c}}(\mathbf{c}) \ln(p_{\mathbf{c}_i}(\mathbf{c}_i)) \, d\mathbf{c} \\
&= -H(\mathbf{c}) - \sum_i \int \ln(p_{\mathbf{c}_i}(\mathbf{c}_i)) \left(\int p_{\mathbf{c}}(\mathbf{c}) \, d\mathbf{c}_1 \ldots d\mathbf{c}_{i-1} d\mathbf{c}_{i+1} \ldots d\mathbf{c}_m\right) \, d\mathbf{c}_i \\
&= -H(\mathbf{c}) - \sum_i \int \ln(p_{\mathbf{c}_i}(\mathbf{c}_i)) \, p_{\mathbf{c}_i}(\mathbf{c}_i) \, d\mathbf{c}_i \\
&= \sum_i H(\mathbf{c}_i) - H(\mathbf{c})
\end{aligned}$$

On veut donc minimiser la quantité :

$$\sum_i H({}^t w_i.\mathbf{Z}) - H(W.\mathbf{Z})$$

où les ${}^t w_i$ sont les lignes de la matrice W.
Nous allons maintenant nous servir de la néguentropie pour simplifier l'expression de $I(\mathbf{c})$. Soit $g_{\mathbf{c}}$ la densité de probabilité d'un vecteur aléatoire gaussien de même moyenne et de même matrice de covariance que $\mathbf{c}$.
Le but est d'exprimer $I(\mathbf{c})$ en fonction des néguentropies de $\mathbf{c}$ et des $\mathbf{c}_i$.
Nous avons :

$$\begin{aligned}
I(\mathbf{c}) &= \sum_i H(\mathbf{c}_i) - H(\mathbf{c}) \\
&= H(g_{\mathbf{c}}) - H(p_{\mathbf{c}}) - H(g_{\mathbf{c}}) + \sum_i H(\mathbf{c}_i) \\
&= J(\mathbf{c}) - H(g_{\mathbf{c}}) + \sum_i H(\mathbf{c}_i)
\end{aligned}$$

Notons $g_{\mathbf{c}_i}$ la densité de probabilité d'une variable aléatoire gaussienne réelle de même moyenne et de même variance que $\mathbf{c}_i$. Pour simplifier $I(\mathbf{c})$, nous allons exprimer l'entropie de $\mathbf{c}$ en fonction de celles des $\mathbf{c}_i$.
Calculons pour cela l'entropie d'une variable aléatoire gaussienne de $\mathbb{R}^d$, dont la matrice de covariance sera notée V. On peut supposer que la moyenne de cette variable est nulle, quitte à faire un changement d'origine qui laissera inchangée la valeur de l'intégrale définissant l'entropie.

$$\begin{aligned}
H(g_{\mathbf{c}}) &= -\int \ln\left(\frac{1}{(2\pi)^{d/2}|\det V|^{1/2}} e^{-\frac{{}^t c.V^{-1}.c}{2}}\right) \frac{1}{(2\pi)^{d/2}|\det V|^{1/2}} e^{-\frac{{}^t c.V^{-1}.c}{2}} \, dc \\
&= \frac{1}{2}\ln((2\pi)^d|\det V|) + \int \frac{{}^t c.V^{-1}.c}{2} \frac{1}{(2\pi)^{d/2}|\det V|^{1/2}} e^{-\frac{{}^t c.V^{-1}.c}{2}} \, dc
\end{aligned}$$

Posons $u = \phi(c) = V^{1/2}.c$. On a $|J\phi(c)| = \frac{1}{|\det V|^{1/2}}$.

$$\begin{aligned}
H(g_{\mathbf{c}}) &= \frac{1}{2}\ln((2\pi)^d|\det V|) + \frac{1}{(2\pi)^{d/2}} \int \frac{{}^t\phi(c).\phi(c)}{2} e^{-\frac{{}^t\phi(c).\phi(c)}{2}} |J\phi(c)| \, dc \\
&= \frac{1}{2}\ln((2\pi)^d|\det V|) + \frac{1}{(2\pi)^{d/2}} \int \frac{||u||^2}{2} e^{-\frac{||u||^2}{2}} \, du \\
&= \frac{1}{2}\ln((2\pi)^d|\det V|) + \frac{d}{(2\pi)^{d/2}} \left(\int \frac{u^2}{2} e^{-\frac{u^2}{2}} \, du\right) \left(\int e^{-\frac{u^2}{2}} \, du\right)^{d-1}
\end{aligned}$$

Une intégration par parties montre que $\int \frac{u^2}{2} e^{-\frac{u^2}{2}} \, du = \frac{1}{\sqrt{2}} \int e^{-v^2} \, dv = \sqrt{\frac{\pi}{2}}$.
Par ailleurs, $\int e^{-\frac{u^2}{2}} \, du = \sqrt{2} \int e^{-v^2} \, dv = \sqrt{2\pi}$.

Ainsi :

$$\begin{aligned}
H(g_{\mathbf{c}}) &= \frac{1}{2}\ln((2\pi)^m|\det V|) + \frac{m}{(2\pi)^{m/2}} \sqrt{\frac{\pi}{2}} (\sqrt{2\pi})^{m-1} \\
&= \frac{m}{2} + \frac{m}{2}\ln(2\pi) + \frac{1}{2}\ln(|\det V|)
\end{aligned}$$

où V est la matrice de covariance de $\mathbf{c}$, et v_i la variance de $\mathbf{c}_i$.
Et de même, $H(g_{\mathbf{c}_i}) = \frac{1}{2} + \frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln(v_i)$, où v_i est la variance de $\mathbf{c}_i$.
Réinjectons les expressions de $H(g_{\mathbf{c}})$ et de $H(g_{\mathbf{c}_i})$ dans celle de $I(\mathbf{c})$:

$$
\begin{aligned}
I(\mathbf{c}) &= J(\mathbf{c}) - \frac{m}{2} - \frac{m}{2}\ln(2\pi) - \frac{1}{2}\ln(|\det\ V|) + \sum_i H(\mathbf{c}_i) \\
&= J(\mathbf{c}) - \sum_i H(g_{\mathbf{c}_i}) - \frac{1}{2}\ \ln(\frac{|\det\ V|}{\prod_i v_i}) + \sum_i H(\mathbf{c}_i) \\
&= J(\mathbf{c}) + \sum_i (H(\mathbf{c}_i) - H(g_{\mathbf{c}_i})) - \frac{1}{2}\ \ln(\frac{|\det\ V|}{\prod_i v_i}) \\
&= J(\mathbf{c}) - \sum_i J(\mathbf{c}_i) - \frac{1}{2}\ \ln(\frac{|\det\ V|}{\prod_i v_i})
\end{aligned}
$$

Comme $\mathbf{c} = W.\mathbf{Z}$, nous pouvons exploiter le fait que la néguentropie est invariante par changement de repère :
$I(\mathbf{c}) = J(\mathbf{Z}) - \sum_i J(\mathbf{c}_i) - \frac{1}{2}\ \ln(\frac{|\det\ V|}{\prod_i v_i})$
Le terme $J(\mathbf{Z})$ ne dépendant pas de $\mathbf{c}$, il ne reste plus que deux termes à minimiser.
Par ailleurs, si V est diagonale, le dernier terme s'annule.
Nous allons faire en sorte que cela soit vrai.

5.4.2 Nouvelle définition de Z

Nous avions défini $\mathbf{Z}$ de la manière suivante :
$S = \frac{1}{N}\sum_i X_i\ {}^tX_i$ s'écrit $S = E\ D\ {}^tE$, où E est la matrice des vecteurs propres associés aux valeurs propres non nulles de l'ACP, et où D est la matrice diagonale des valeurs propres non nulles de l'ACP.
Nous avions ensuite posé :
$\mathbf{Z} = {}^tE.\mathbf{X}$

Nous allons changer ici cette notation et poser : $\boxed{\mathbf{Z} = D^{-1/2}.{}^tE.\mathbf{X}}$
La matrice de covariance de $\mathbf{Z}$ est alors :

$$
\begin{aligned}
\mathbb{E}(\mathbf{Z}.{}^t\mathbf{Z}) &= \frac{1}{N}\sum_i D^{-1/2}.{}^tE.\mathbf{X}.{}^t\mathbf{X}.E.D^{-1/2} \\
&= D^{-1/2}.{}^tE.S.E.D^{-1/2} \\
&= D^{-1/2}.{}^tE.E.D.{}^tE.E.D^{-1/2} \\
&= I_m
\end{aligned}
$$

en effet, les vecteurs colonne de la matrice (non carrée) E formant une famille orthonormale, on a : ${}^tE.E = I_m$.

5.4.3 Nouvelle simplification

La matrice de covariance de $\mathbf{c}$ est alors :

$$V = \mathbb{E}(\mathbf{c}.{}^t\mathbf{c}) = \mathbb{E}(W.\mathbf{Z}.{}^t\mathbf{Z}.{}^tW) = W.{}^tW$$

Nous allons maintenant imposer aux variances v_i d'être égales à 1, comme c'est déjà le cas pour les variances des Z_i. Cela est possible car on a $\mathbf{c}_i ={}^t w_i.\mathbf{Z}$, et donc la variance de $\mathbf{c}_i$ peut être incorporée dans la $i$ème ligne de W, de telle manière que $\mathbf{c}_i$ ait une variance égale à 1. Par ailleurs, comme les composantes de $\mathbf{c}$ doivent être indépendantes, elles doivent être en particulier décorrélées. Ceci signifie que V doit être la matrice identité, et donc que W doit être une matrice orthonormale.
Nous avons ainsi simplifié le problème, et nous n'avons plus que $m \times \frac{m-1}{2}$ degrés de liberté à estimer. Ainsi, selon l'expression de Aapo Hyvärinen et Erkki Oja, nous avons résolu la moitié du problème de l'analyse en composantes indépendantes.
Le problème s'est réduit à trouver une matrice orthonormale W qui minimise $I(\mathbf{c})$, donc qui

maximise $\sum_i J(\mathbf{c}_i) = \sum_i J({}^t w_i.\mathbf{Z})$. Rappelons que les ${}^t w_i$ sont les lignes de la matrice W.

Le problème est devenu :

$$\boxed{W = \underset{M \in \mathcal{O}_m(\mathbb{R})}{ArgMax} \sum_i J({}^t m_i.\mathbf{Z})}$$

où m_i est la $i$ème ligne de M.
Nous avons bien simplifié le problème, mais cela ne suffit pas, car le calcul des $J(\mathbf{c}_i)$ nécessite de connaître les densités associées aux $\mathbf{c}_i$. Nous allons donc devoir utiliser une approximation de la néguentropie.

5.4.4 Approximation de la néguentropie

Pierre Comon (dans [Com94]) propose une approximation fondée sur la kurtosis, définie de la manière suivante pour une variable aléatoire centrée :

$$kurt(x) = \mathbb{E}(x^4) - 3\,(\mathbb{E}(x^2))^2$$

Notons au passage que la fonction kurtosis peut-être utilisée comme critère de non-gaussiannité, et permet encore une autre approche de l'ICA. Une variable aléatoire de kurtosis positive est dite supergaussienne ; elle est dite subgaussienne si sa kurtosis est négative. Une distribution supergaussienne est typiquement plus concentrée et plus pointue en zéro qu'une distribution gaussienne, et plus concentrée aussi en queue de distribution (elle décroît moins rapidement). Une distribution subgaussienne, quand à elle, sera plus plate en zéro qu'une gaussienne, et décroîtra plus vite en l'infini. Il est à noter que la plupart, et non la totalité, des distributions non-gaussiennes ont une kurtosis non-nulle.
Cette séparation entre variables subgaussiennes et supergaussiennes est à l'origine de quelques raffinements de l'ICA. Nous n'entrerons pas dans ces détails, le but n'étant pas ici de traiter l'ICA de manière exhaustive.
Mais revenons à notre problème :
D'après [Com94], on a, pour une variable aléatoire x de moyenne nulle et de variance 1 :

$$J(x) \approx \frac{1}{12}\,\mathbb{E}(x^3)^2 + \frac{1}{48}\,kurt(x)^2$$

Mais cette approximations n'est pas suffisamment robuste.
Hyvärinen propose de la remplacer par l'approximation suivante :

$$J(x) \approx (\mathbb{E}(G(x)) - \mathbb{E}(G(\nu)))^2$$

où ν est une variable gaussienne de moyenne nulle et de variance 1, et G est une fonction non-quadratique. Les exemples donnés sont :
$G_1(x) = \frac{1}{a}\ln(\cosh(ax))$, $G_2(x) = -exp(-x^2/2)$,
où a est une constante comprise entre 1 et 2.
Nous prendrons ici $G = G_2$.

Le problème devient :

$$\boxed{W = \underset{M \in \mathcal{O}_m(\mathbb{R})}{ArgMax} \sum_{1 \leq i \leq m} \left(\mathbb{E}[G({}^t m_i.\mathbf{Z})] - \mathbb{E}[G(\nu)]\right)^2}$$

où les ${}^t m_i$ sont les lignes de M.

5.4.5 Résolution

Puisque $\mathbb{E}[G(\nu)]$ est une constante, le maximum de $\left(\mathbb{E}[G({}^t w_i.\mathbf{Z})] - \mathbb{E}[G(\nu)]\right)^2$ est atteint en un extremum de $\mathbb{E}[G({}^t w_i.\mathbf{Z})]$.
Nous allons estimer les vecteurs w_i indépendamment les uns des autres. Pour cela, nous allons chercher les extrema de $\mathbb{E}[G({}^t w_i.\mathbf{Z})]$ sous la contrainte que w_i soit de norme 1. L'orthonormalité sera acquise par la suite par une orthonormalisation de Gram-Schmidt.
Écrivons la fonction de Lagrange associée au calcul du vecteur w_i :

$$L(w_i, l) = \mathbb{E}[G({}^t w_i.\mathbf{Z})] - l({}^t w_i.w_i - 1)$$

Les dérivées selon w_i et l doivent être nulles, ce qui donne :

$$\begin{aligned} \mathbb{E}[G'({}^t w_i.\mathbf{Z})\mathbf{Z}] - 2lw_i &= 0 \\ ||w_i||^2 &= 1 \end{aligned}$$

Pour satisfaire ces deux équations, on va chercher à résoudre la première itérativement, en renormalisant w_i à chaque itération, de manière à satisfaire la deuxième.

Nous allons appliquer l'algorithme de Newton-Raphson. Pour cela, il faut calculer la matrice jacobienne de $\Phi(w) = \mathbb{E}[G'({}^t w.\mathbf{Z})\mathbf{Z}] - 2lw$:
Si ${}^t w = (w^1, \ldots, w^m)$,

$$J_\Phi(w) = \left(\frac{\partial \Phi_i(w)}{\partial w^j}\right)_{i,j} = \mathbb{E}[G''({}^t w.\mathbf{Z})\ \mathbf{Z}.{}^t\mathbf{Z}] - 2l\ Id$$

L'algorithme de Newton-Raphson donne alors :

$$w_i(t+1) = w_i(t) - J_\Phi(w_i(t))^{-1}\ \Phi(w_i(t))$$

Pour simplifier encore le problème, Hyvärinen suggère de faire l'approximation suivante :

$$\mathbb{E}[G''({}^t w.\mathbf{Z})\ \mathbf{Z}.{}^t\mathbf{Z}] \approx \mathbb{E}[G''({}^t w.\mathbf{Z})]\ \mathbb{E}[\mathbf{Z}.{}^t\mathbf{Z}] = \mathbb{E}[G''({}^t w.\mathbf{Z})]\ Id$$

Il est alors aisé de calculer l'inverse de la matrice Jacobienne de Φ :

$$J_\Phi(w)^{-1} = \frac{1}{\mathbb{E}[G''({}^t w.\mathbf{Z})] - 2l}\ Id$$

Ce qui donne :

$$w_i(t+1) = w_i(t) - \frac{\mathbb{E}[G'({}^t w_i(t).\mathbf{Z})\mathbf{Z}] - 2lw_i(t)}{\mathbb{E}[G''({}^t w.\mathbf{Z})] - 2l}$$

Mais cet algorithme pose parfois des problèmes de convergence ; et il peut être bon de le stabiliser. Une méthode pour ce faire est d'introduire un paramètre μ comme suit :

$$w_i(t+1) = w_i(t) - \mu\ \frac{\mathbb{E}[G'({}^t w_i(t).\mathbf{Z})\mathbf{Z}] - 2lw_i(t)}{\mathbb{E}[G''({}^t w.\mathbf{Z})] - 2l}$$

μ sera pris proche de 1, et inférieur à 1. Cela dit, prendre $\mu < 1$ ralentit quelque peu la convergence.

Il est maintenant nécessaire de connaître l. ou du moins de l'approximer. Sachant qu'à l'équilibre on doit avoir $\mathbb{E}[G'({}^t w_i.\mathbf{Z})\mathbf{Z}] - 2lw_i = 0$, et en multipliant cette égalité à gauche par ${}^t w_i$, on a : $l = \frac{{}^t w_i \ \mathbb{E}[G'({}^t w_i.\mathbf{Z})\mathbf{Z}]}{2}$.

Enfin, pour satisfaire à l'orthogonalité des w_i, on procède à une orthonormalisation de Gram-Schmidt, en ne conservant de $w_i(t+1)$ que sa projection sur l'espace orthogonal aux vecteurs $w_1, \ldots, w_{i-1}$ déjà calculés.
L'algorithme est donc le suivant :

$$\left| \begin{array}{l} \text{Une fois que } w_1, \ldots, w_{i-1} \text{ sont calculés,} \\ w_i(t+1) \ = \ w_i(t) \ - \ \mu \ \frac{\mathbb{E}[G'({}^t w_i(t).\mathbf{Z})\mathbf{Z}] \ - \ {}^t w_i(t) \ \mathbb{E}[G'({}^t w_i(t).\mathbf{Z})\mathbf{Z}] \ w_i(t)}{\mathbb{E}[G''({}^t w_i(t).\mathbf{Z})] \ - \ {}^t w_i(t) \ \mathbb{E}[G'({}^t w_i(t).\mathbf{Z})\mathbf{Z}]} \\ w_i(t+1) \ = \ w_i(t+1) \ - \displaystyle\sum_{1 \leq j \leq i-1} \langle w_i(t+1), w_j(t) \rangle \ w_j(t) \\ w_i(t+1) \ = \ \frac{w_i(t+1)}{||w_i(t+1)||} \end{array} \right|$$

Si $\Omega = W^{-1}$, et si Ω_i est la $i$ème colonne de Ω, on a $\mathbf{Z} = \Omega.\mathbf{c} = \mathbf{c}_1.\Omega_1 + \ldots + \mathbf{c}_m.\Omega_m$. De plus, $\mathbf{X} = E.D^{1/2}.\mathbf{Z}$. On retrouve donc les modes Λ_i de l'ICA (dans $\mathbb{R}^{3n}$) en posant :

$$\Lambda_i \ = \ E.D^{1/2}.\Omega_i$$

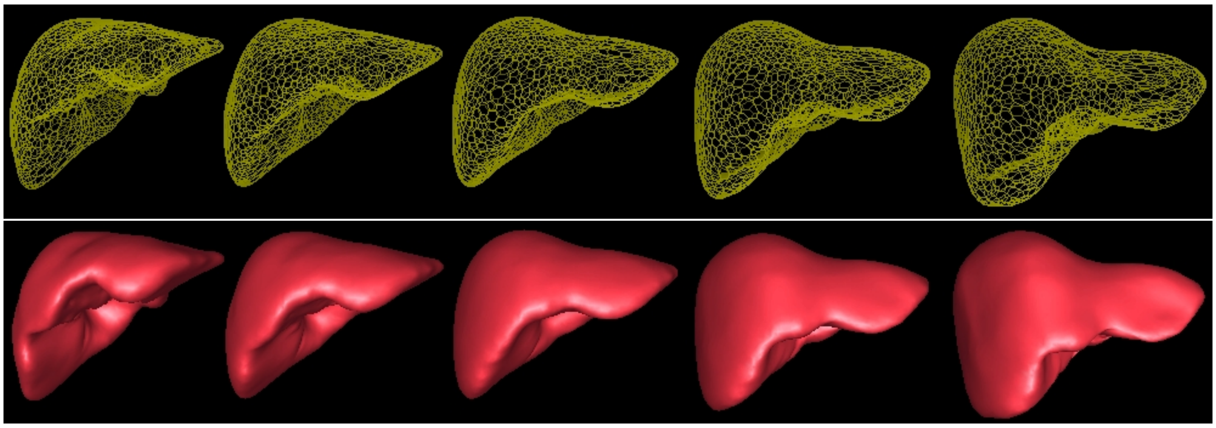


FIG. 5.1 – *Premier mode de l'ICA*

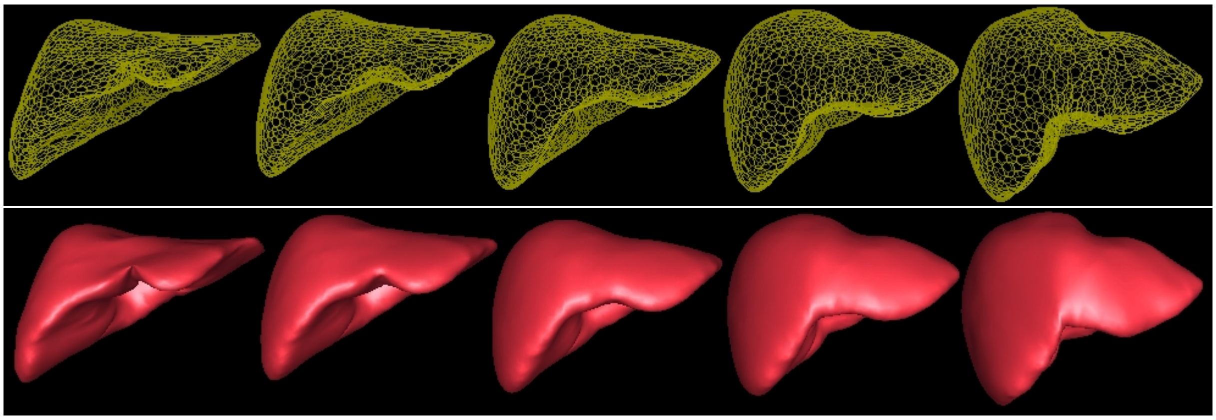


FIG. 5.2 – *Deuxième mode de l'ICA*

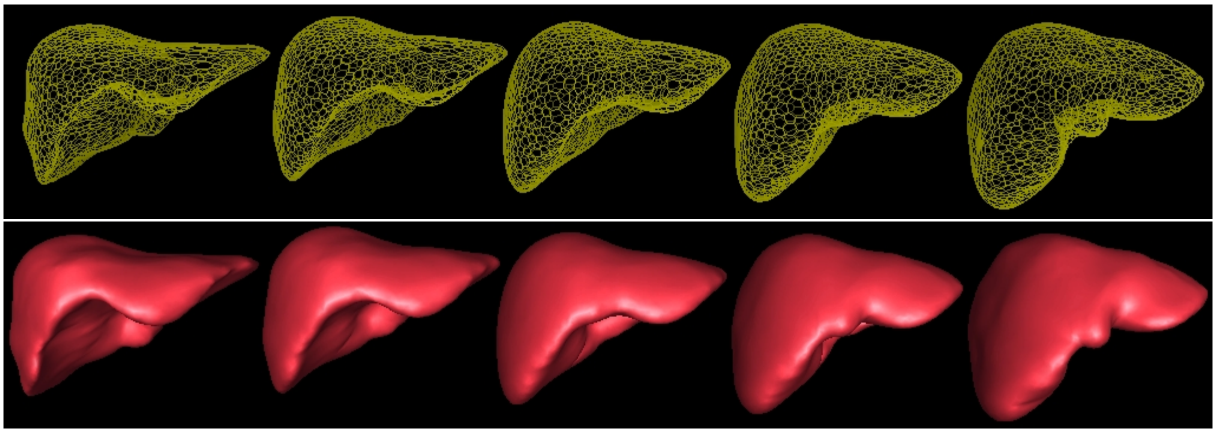


Fig. 5.3 – *Troisième mode de l'ICA*

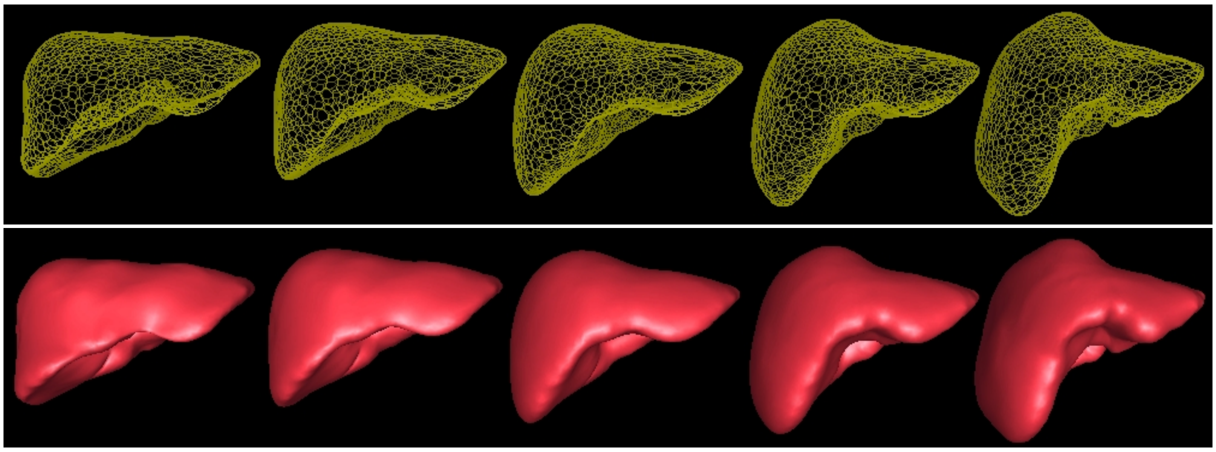


Fig. 5.4 – *Quatrième mode de l'ICA*

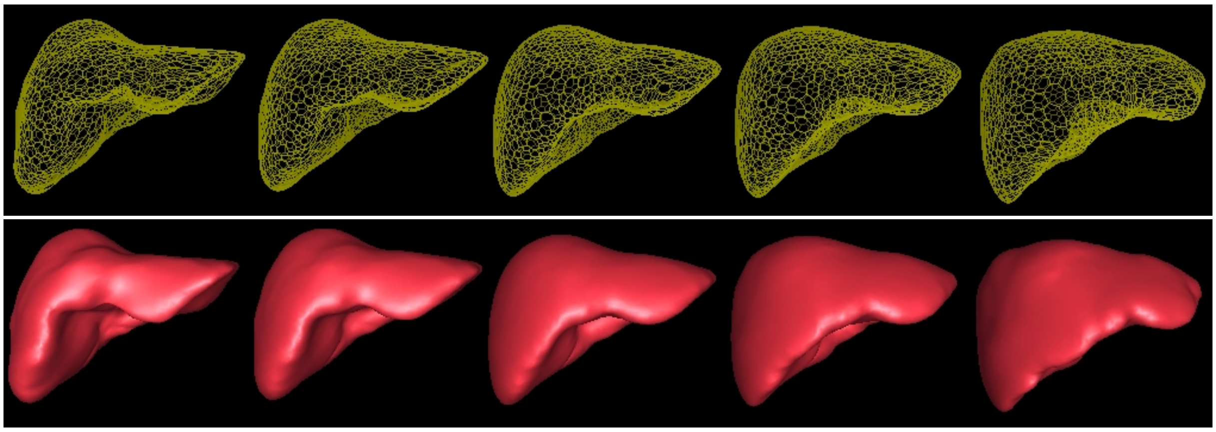


Fig. 5.5 – *Cinquième mode de l'ICA*

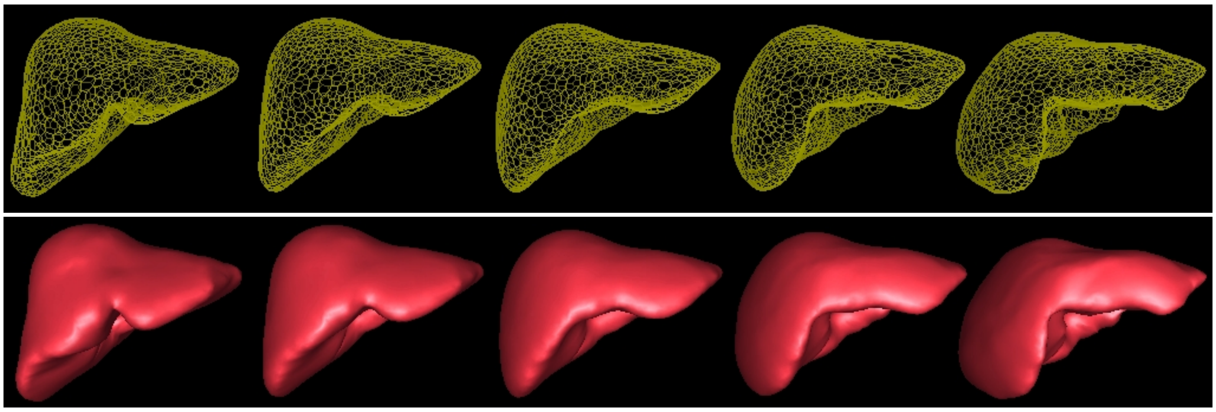


Fig. 5.6 – *Sixième mode de l'ICA*

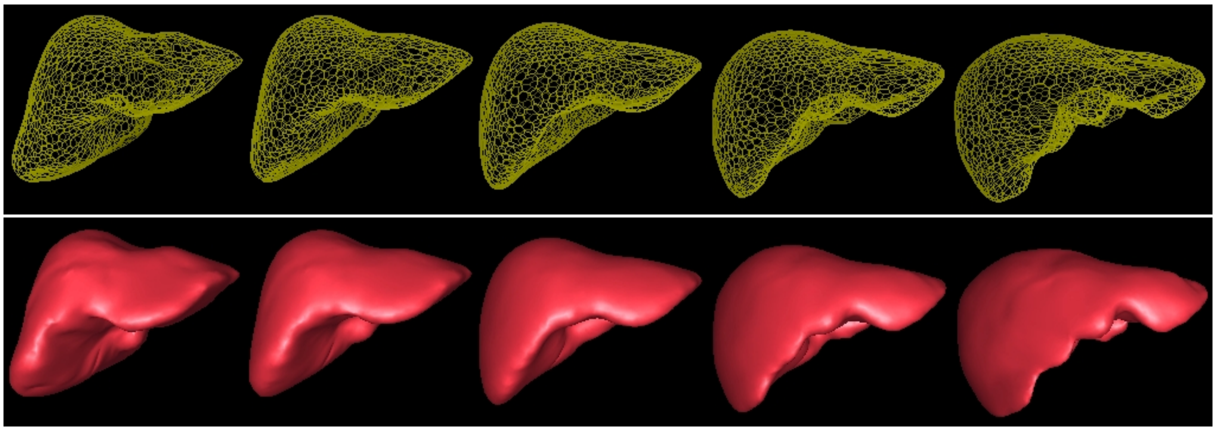


Fig. 5.7 – *Septième mode de l'ICA*

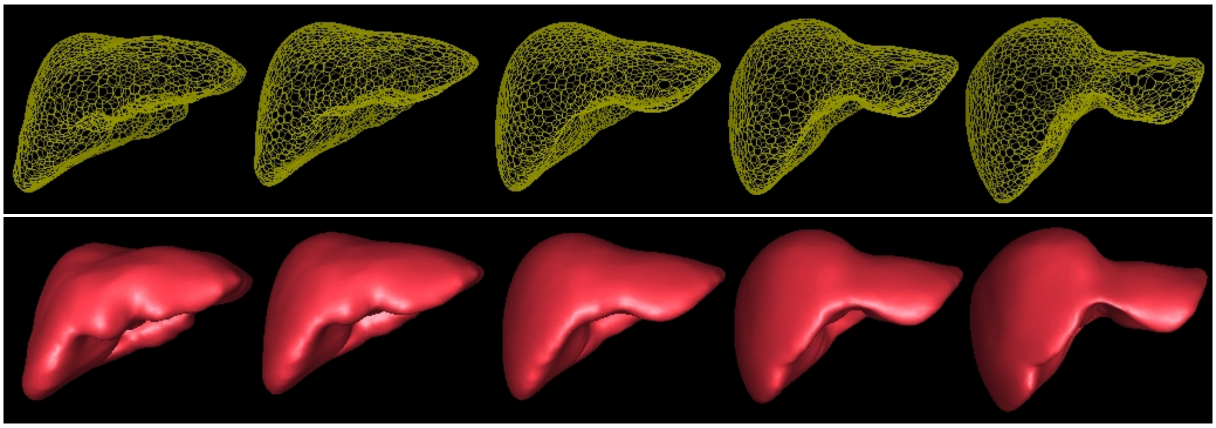


Fig. 5.8 – *Huitième mode de l'ICA*

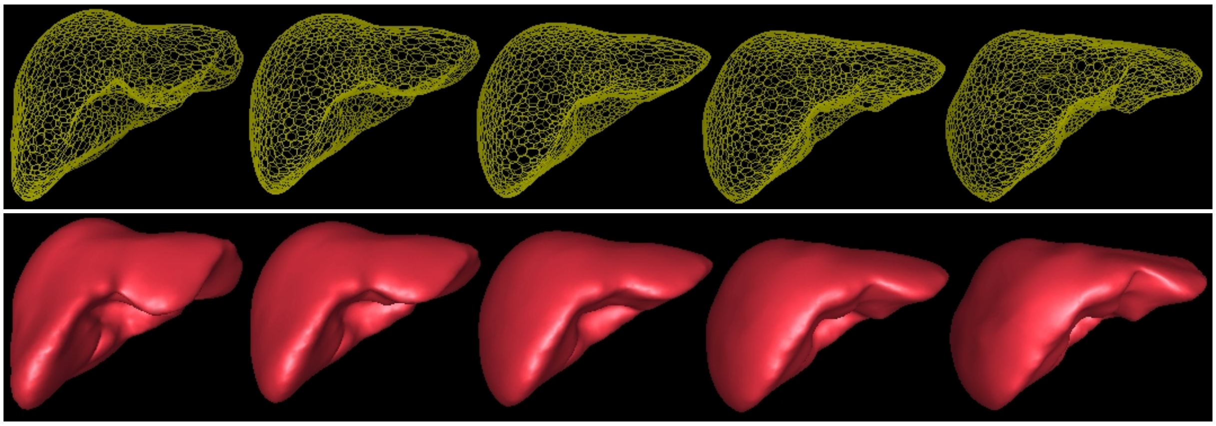


FIG. 5.9 – *Neuvième mode de l'ICA*

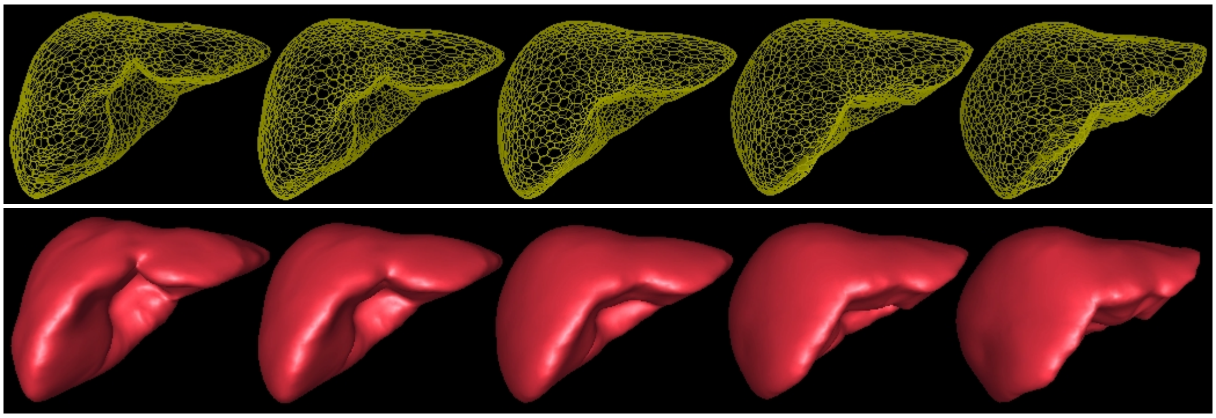


FIG. 5.10 – *Dixième mode de l'ICA*

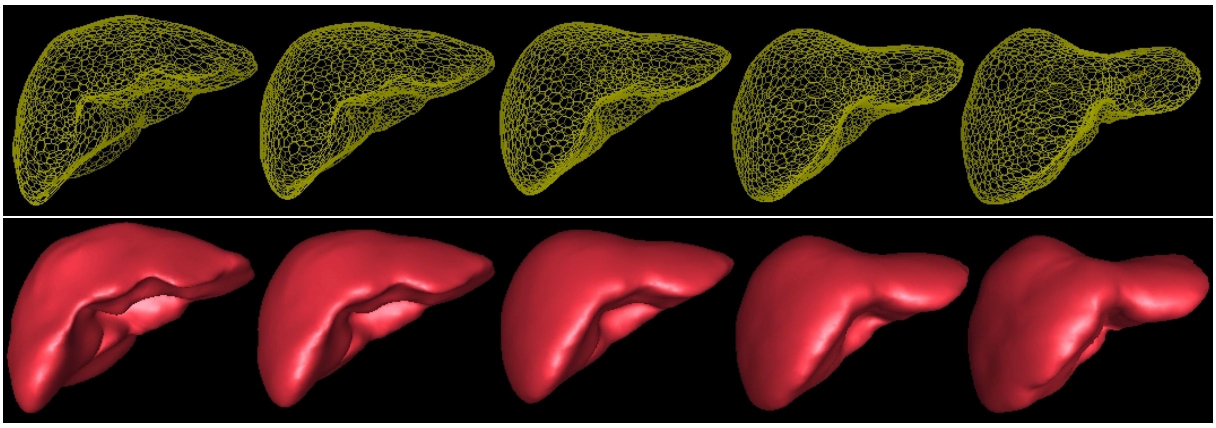


FIG. 5.11 – *Onzième mode de l'ICA*

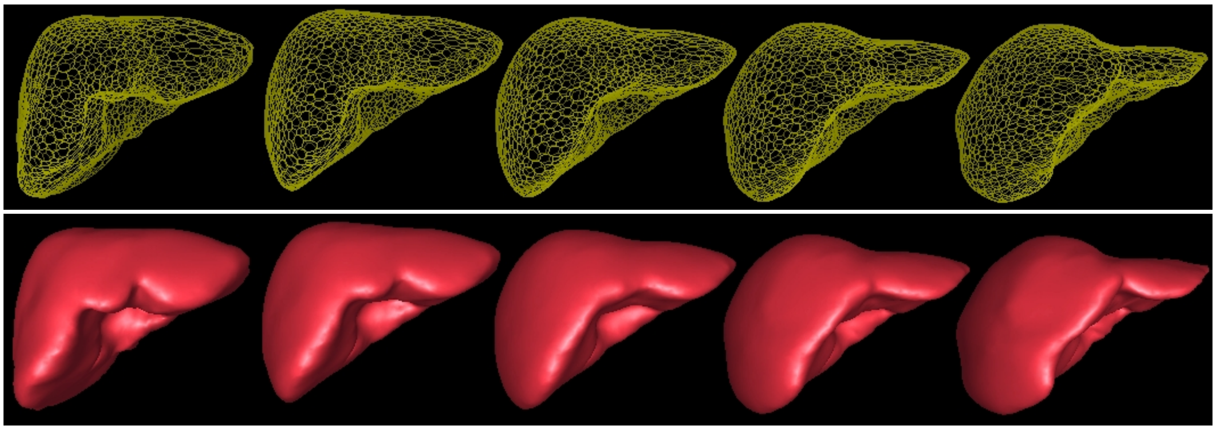


FIG. 5.12 – *Douzième mode de l'ICA*

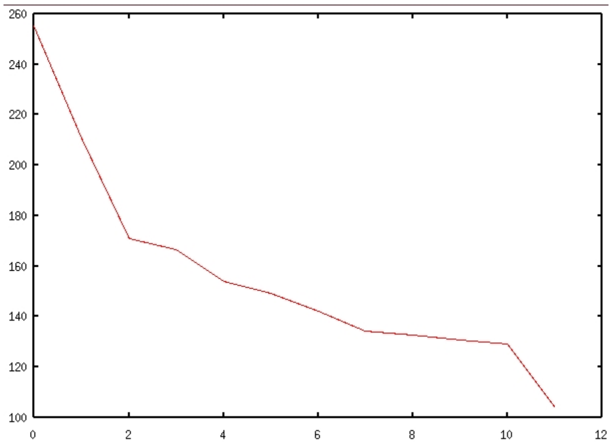


FIG. 5.13 – *Racines carrées des variances associées à chaque mode de l'ICA, en mm*

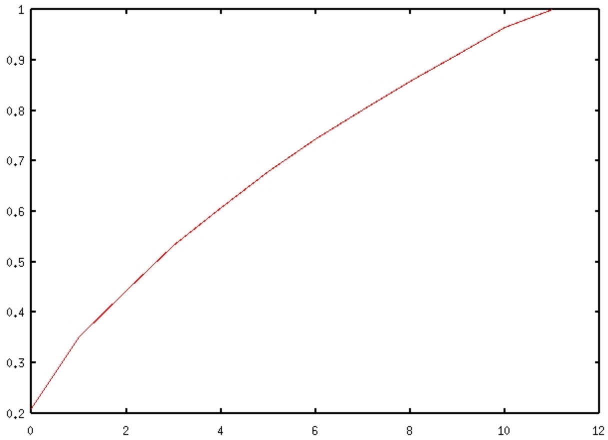


FIG. 5.14 – *Sommes partielles normalisées des variances associées aux modes de l'ICA*

5.5 Avantages et inconvénients de l'ICA dans le cadre de notre application

Le fait que l'ICA décompose des signaux donnés en signaux indépendants nous permet d'espérer que les «modes» de déformation du foie qu'elle génère soient très peu redondants, et donc que chacun d'eux corresponde à un type bien particulier de déformation.
Mais il faut bien se rendre compte que l'ICA ne peut compacter l'information plus que ne le fait l'ACP. C'est-à-dire que les variances associées aux «modes» de l'ICA décroissent nécessairement moins vite que celles des modes de l'ACP. Cela est dû au fait l'ACP combine le plus possible de composantes indépendantes pour obtenir une composante prépondérante. Cela est illustré par la figure 5.15, qui représente deux distribution aléatoire uniformément distribuées sur deux segments sécants en leur milieu, faisant un angle de $\pi/3$. Comme la dimension de l'espace est 2, le nombre de modes est limité à deux. Comme on peut s'y attendre, l'ACP donne un premier mode de forte variance dont la direction est à peu près la bissectrice de l'angle aigu entre les deux segments, et un mode de variance plus faible dont la direction est la bissectrice de l'angle obtus entre les deux segments. En revanche, l'ICA donne les bonnes directions, avec des variances de même ordre. Elle sépare donc mieux les composantes indépendantes, mais n'offre pas une description optimale. La comparaison des figures 4.13 et 5.13 montre que la décroissance des écart-types associés aux modes est plus rapide pour l'ACP que pour l'ICA.
Enfin, au vu des modes donnés par l'ICA, qui ressemblent tout de même à ceux de l'ACP, il n'est pas évident que ceux-ci offrent une description plus fine des transformations possibles.

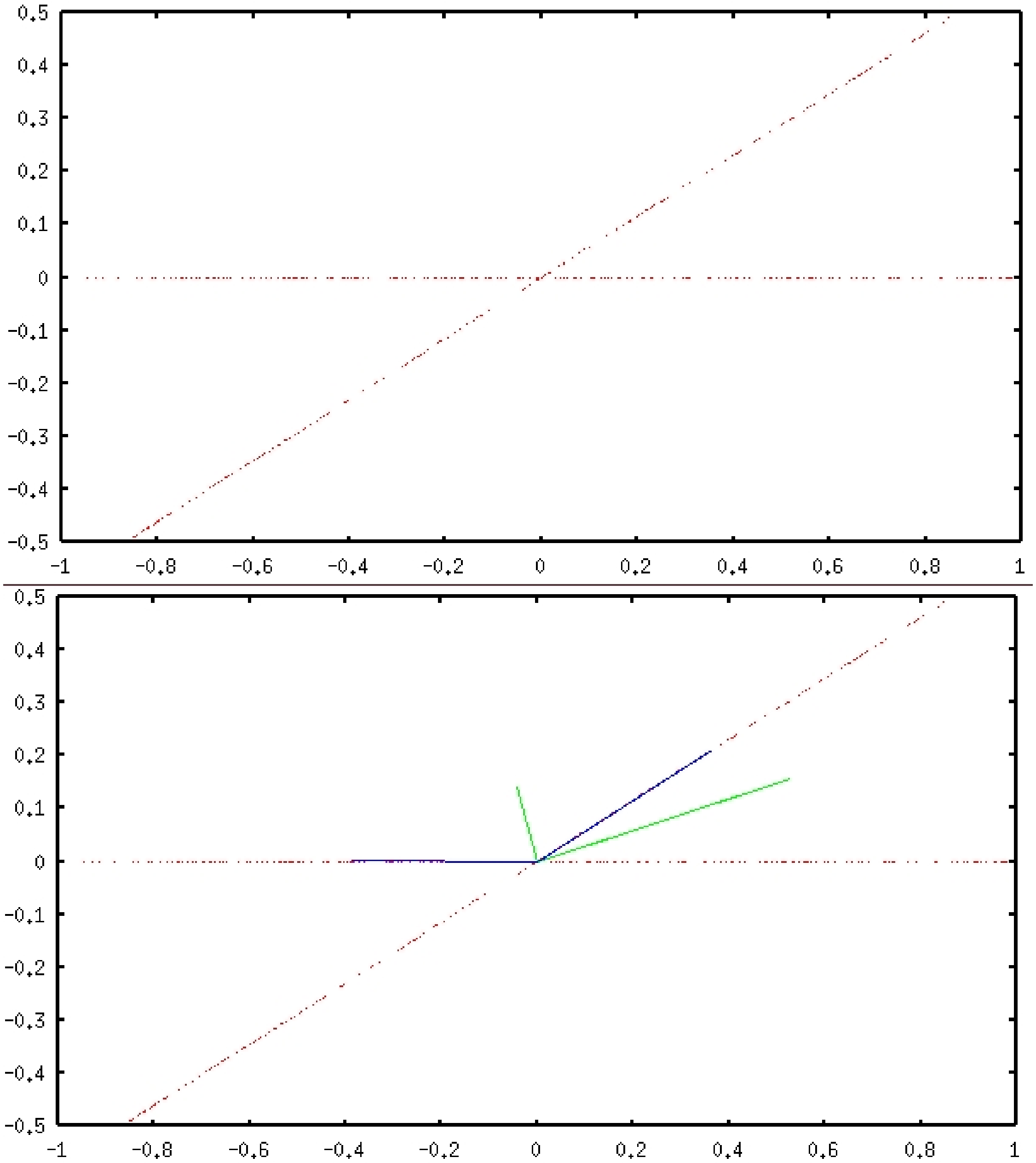


FIG. 5.15 – *ACP et ICA sur une distribution aléatoire en dimension* 2 *(400 points) : 200 points sur le segment* $[-1,1] \times \{0\}$*, et 200 points sur le segment joignant les points* $[-\sqrt{3}/2, \ -1/2]$ *et* $[\sqrt{3}/2, \ 1/2$*. Chaque mode est représenté par un vecteur dont la longueur est l'écart-type associé. En vert, les modes de l'ACP (le vecteur* $[0.958536, \ 0.284971]$*, associé à une variance de 0.304127, et le vecteur* $[-0.284971, \ 0.958536]$ *associé à une variance de 0.0215556) ; en bleu, ceux de l'ICA (le vecteur* $[-1, \ -0.000783683]$*, associé à une variance de 0.14771, et le vecteur* $[0.865995, \ 0.500053]$ *associé à une variance de 0.177973). L'ACP combine les composantes indépendantes pour donner une composante prépondérante et une composante bien moins représentative, alors que l'ICA retrouve les signaux indépendants d'origine, avec des variances de même ordre.*

Chapitre 6

Segmentation à partir d'un modèle statistique déformable

Tout l'enjeu d'une analyse statistique d'un organe réside dans la segmentation d'images tridimensionnelles de cet organe à partir du modèle statistique issu de cette analyse.
La segmentation de l'image consiste à déformer notre modèle de manière à ce que celui-ci prenne la forme de l'organe présent dans l'image.
Ceci a été fait en utilisant l'ACP par [TCDJ95], mais les auteurs se limitaient à des images traditionnelles en deux dimensions. Cristian Lorentz et Nils Krahnstöver ont effectué une ACP sur des maillages tridimensionnels, dans [LK00], sans pour autant l'appliquer à la segmentation d'une image.

6.1 Appariement des sommets du modèle avec des points du contour

La segmentation de l'image nécessite que l'on sache détecter les contours de l'organe. Pour cela, on calcule l'image de gradient : à chaque voxel de l'image de l'organe, on associe le vecteur gradient de l'intensité de niveau de gris correspondant à ce voxel.
Ensuite, pour chacun des sommets du maillage, on calcule, dans la direction de la normale, le maximum du gradient sur une distance préalablement paramétrée. En effet, un contour est caractérisé par une forte variation de l'intensité lumineuse (c'est-à-dire de niveau de gris), donc par un fort gradient. On prend le maximum du gradient dans la direction de la normale, comme suggéré dans [TCDJ95], en espérant que ce point corresponde effectivement au contour et non à du bruit. En général, le bruit engendre des variations d'intensité moins importantes que celles dues à un contour. Si ça n'est pas le cas, l'appariement n'étant faux qu'en un sommet, comme les variations permises par les statistiques imposent une forte cohérence au maillage, l'erreur d'appariement n'aura pas de conséquences désastreuses.
Pourquoi privilègie-t-on la direction de la normale ?
C'est en fait parce que l'on considère que notre modèle est assez proche du contour que l'on veut segmenter (si ça n'est pas le cas, on peut effectuer un recalage automatique au préalable). La direction normale à la surface du modèle est alors intuitivement la meilleure que l'on puisse choisir de manière automatique.
Il est indispensable de régler avec soin la distance maximale sur laquelle on calcule le gradient dans la direction de la normale : si cette distance est trop faible, l'algorithme ne détecte pas tous les points du contour à segmenter, et peut apparier le sommet avec un maximum local de gradient correspondant à du bruit. Par contre, si cette distance était trop importante, le sommet

pourrait être apparié avec un point du contour plus éloigné que le point de contour le plus proche sur la normale, voire même avec un point du contour d'un autre objet ; car n'oublions pas que l'objectif est de segmenter une image médicale, sur laquelle apparaissent plusieurs organes.
Se référer à la figure 6.1 qui est une représentation en coupe de ce que l'on vient d'expliquer.

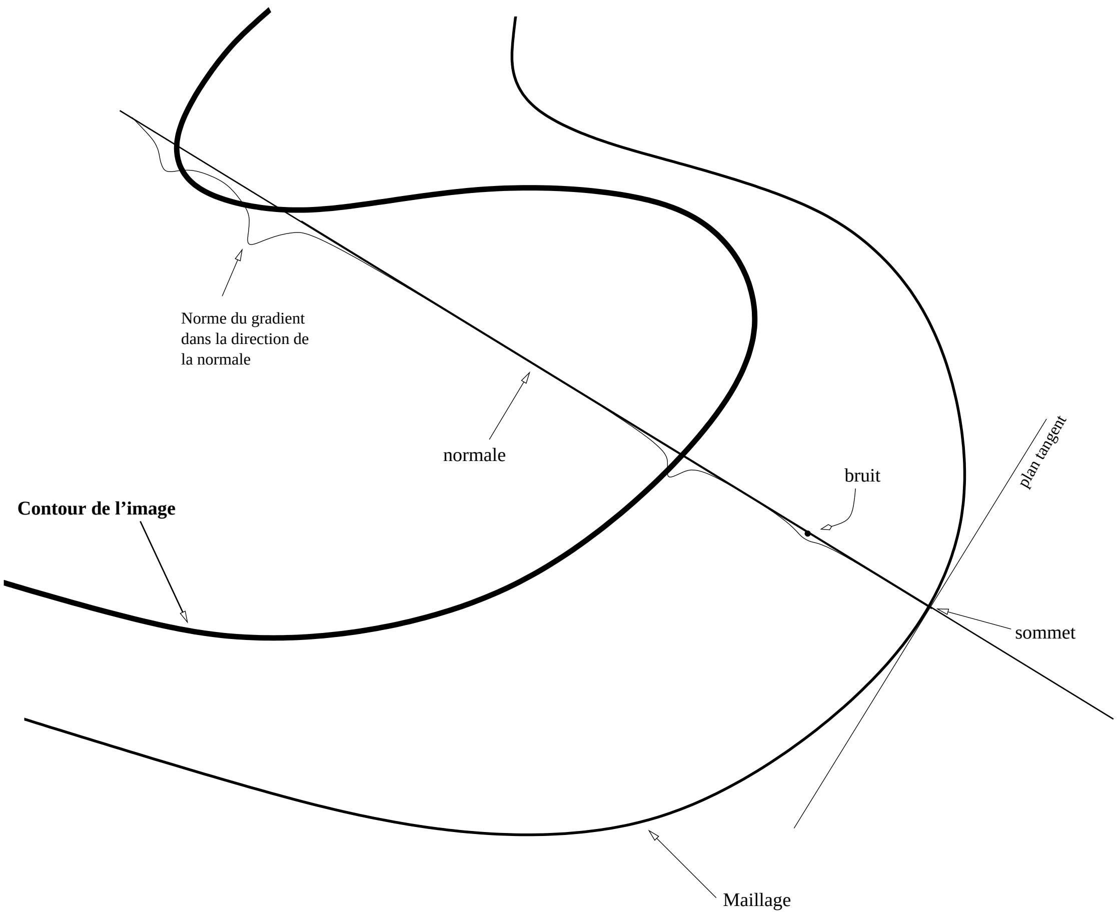


FIG. 6.1 – *Représentation en coupe de l'appariement d'un sommet avec le maximum de gradient suivant la normale*

Remarquons que le calcul du gradient s'accompagne souvent d'une convolution par une gaussienne, dont l'intérêt est ici d'étendre la largeur d'un contour, ce qui peut permettre de détecter un contour bien que le maximum de gradient soit trop éloigné pour être pris en compte. Cela a aussi pour effet de rendre les contours plus flous, mais cela ne change pas la localisation du maximum de gradient suivant la normale, et ce n'est donc pas un problème.
On a ainsi réussi à apparier chaque sommet avec un point de l'image, ce qui nous donne une déformation à appliquer à chaque sommet.

6.2 Déformation suivant les modes de l'ACP

Plaçons-nous dans le cadre d'une segmentation contrainte par les modes de l'ACP. Nous disposons d'un foie moyen et de modes de déformations qui peuvent lui être appliqués. Mais le foie moyen et les modes dont nous disposons ne sont a priori pas orientés de la même manière que le foie de l'image à segmenter, et il est donc nécessaire, avant toute chose, de recaler le maillage de foie moyen sur l'image. Le recalage en question est la composée d'une application linéaire u et d'une translation t, appliqués à chaque sommet du maillage. Les modes de déformation doivent être, eux aussi, «recalés», mais ne doivent pas subir l'effet de la translation : on doit leur appliquer la transformation u.

Ensuite, plusieurs méthodes sont possibles :

Première méthode : Nous avons à notre disposition une déformation dM à appliquer au maillage moyen, sous la forme d'un vecteur à $3n$ composantes. C'est en quelque sorte la déformation idéale à appliquer. Mais nous voulons appliquer une transformation qui soit une combinaison linéaire des modes de déformation que nous avons calculés.

Pour cela, il nous suffit de projeter cette déformation idéale sur l'espace engendré par les modes :

$$dM^* = \sum_{1 \leq i \leq N-1} \langle dM, U_i \rangle \, U_i$$

où les U_i sont les vecteurs des modes de déformation.

Mais on ne peut accepter n'importe quelle combinaison linéaire des modes : si la projection suivant un mode est trop importante, cela ne représente plus une déformation «acceptable». On effectuera donc une troncature telle que, selon chaque mode, la nouvelle position courante soit distante du maillage moyen de moins de trois fois l'écart-type empirique (trois fois la racine carrée de la valeur propre associée pour l'ACP), ce qui permet de décrire 99% des échantillons possibles :

$$dM^{**} = \sum_{1 \leq i \leq N-1} signe(\langle M + dM - \overline{M}, U_i \rangle) \, Min(|\langle M + dM - \overline{M}, U_i \rangle|, 3 * \sqrt{\lambda_i}) \, U_i$$

Pour finir, on met à jour le maillage courant :

$$Mcurr = Mcurr + dM^{**}$$

On itère ce procédé jusqu'à stabilisation.

Deuxième méthode : À chaque itération, préalablement à toute déformation suivant les modes, on recale le maillage courant $Mcurr$ sur la nouvelle position idéale I : cela nous donne un maillage M'. Ceci permet d'être à chaque instant positionné au mieux pour effectuer une déformation suivant les modes :

$$dM^* = \sum_{1 \leq i \leq N-1} \langle I - M', U_i \rangle \, U_i$$

On effectuera à nouveau une troncature, cette fois-ci directement sur dM^*. On tronquera donc de manière plus restrictive, par exemple à une fois la valeur de l' écart-type :

$$dM^{**} = \sum_{1 \leq i \leq N-1} signe(\langle dM^*, U_i \rangle) \, Min(|\langle dM^*, U_i \rangle|, \sqrt{\lambda_i}) \, U_i$$

Pour finir, on met à jour $Mcurr$:

$$Mcurr = Mcurr + dM^{**}$$

On itère ce procédé jusqu'à stabilisation.

On préférera la seconde méthode, qui a l'avantage de mêler recalages et déformations suivant les modes. On espère ainsi être à tout moment recalé au mieux sur l'image pour effectuer une déformation suivant les modes.

6.3 Résultats de segmentation par ACP

Voici les résultats de quelques segmentations contraintes par les modes de déformation de l'analyse en composantes principale.

6.3.1 Validation de la méthode

Trois coupes d'une segmentation d'une image faisant partie de l'ensemble d'apprentissage (cas idéal). En rouge, le foie moyen avant segmentation ; en rose, la segmentation finale par ACP :

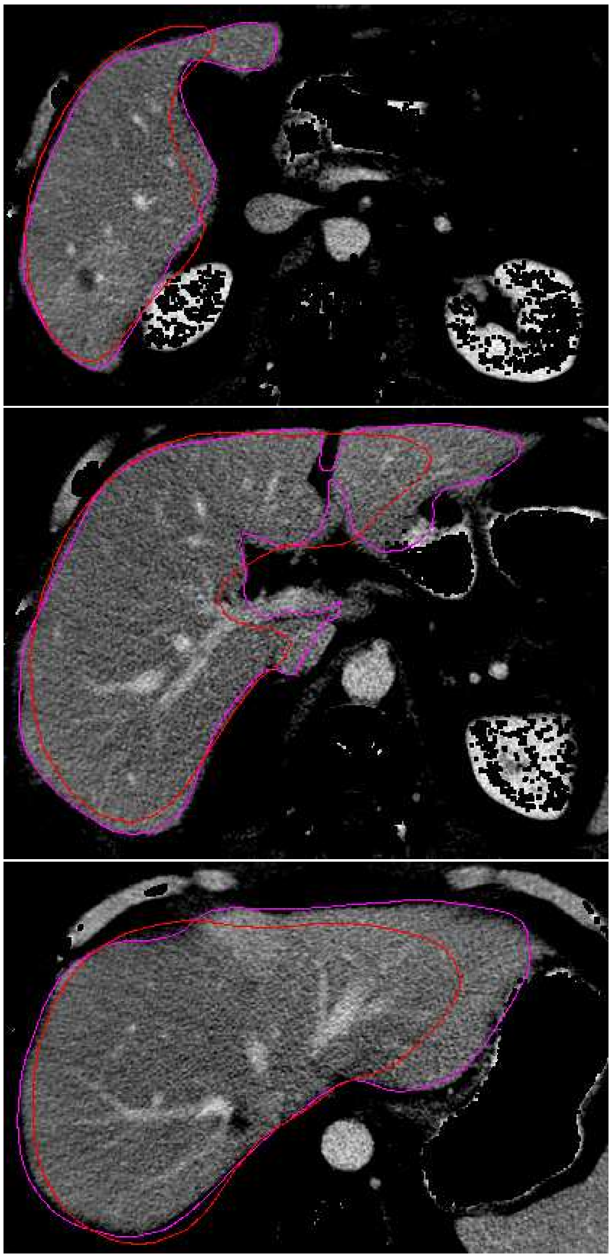


FIG. 6.2 –

La segmentation est quasiment parfaite : on retrouve le maillage de l'ensemble d'apprentissage, à peu de choses près.

Même foie (cas idéal). En rouge, le modèle d'apprentissage ; en rose, le modèle segmenté par ACP :

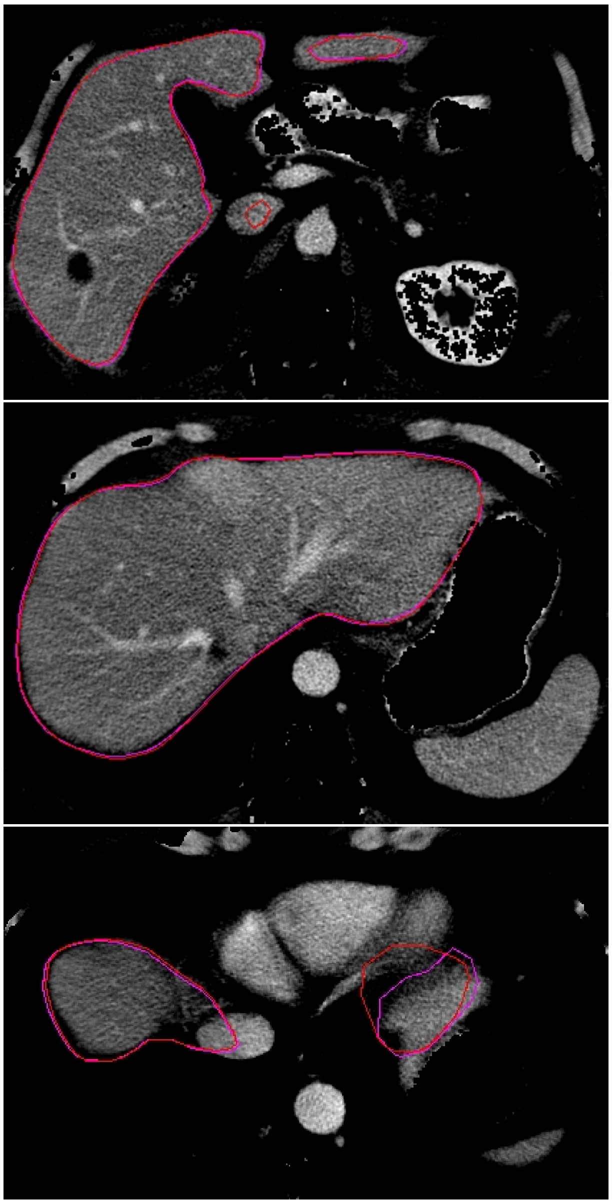


FIG. 6.3 –

Le modèle segmenté par ACP est même un peu meilleur que le modèle d'origine, puisque ce dernier a tendance à segmenter un peu de cœur (dernière coupe, le cœur est la partie un peu plus claire qui se trouve un peu au-dessus, entre les deux bouts de foie).

Trois coupes d'une segmentation d'une image de foie sain ne faisant pas partie de l'ensemble d'apprentissage (cas favorable). En rouge, le foie moyen avant segmentation ; en rose, la segmentation finale par ACP :

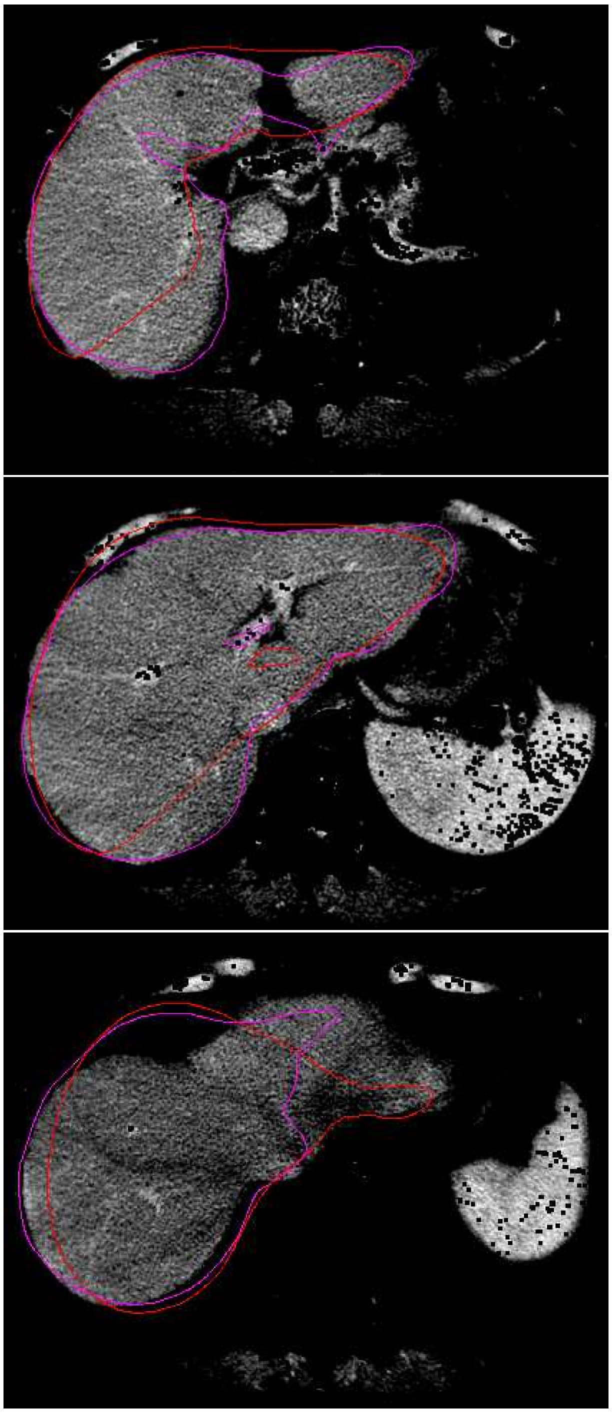


FIG. 6.4 –

Remarquons sur la coupe du milieu que la côte (en haut à gauche de l'image) n'a pas attiré le maillage, ce qui n'aurait pas manqué d'arriver avec une segmentation classique par snakes (voir figure 6.6). Ceci illustre bien le fait que certaines déformations qui ne sont pas caractéristiques d'un foie ne sont pas autorisées. C'est précisément cela que nous espérions.

Même foie (cas favorable). Comparaison avec la segmentation par transformations affines : Sur une image de foie sain n'appartenant pas à l'ensemble d'apprentissage. En rouge, la segmentation affine ; en rose, la segmentation par ACP :

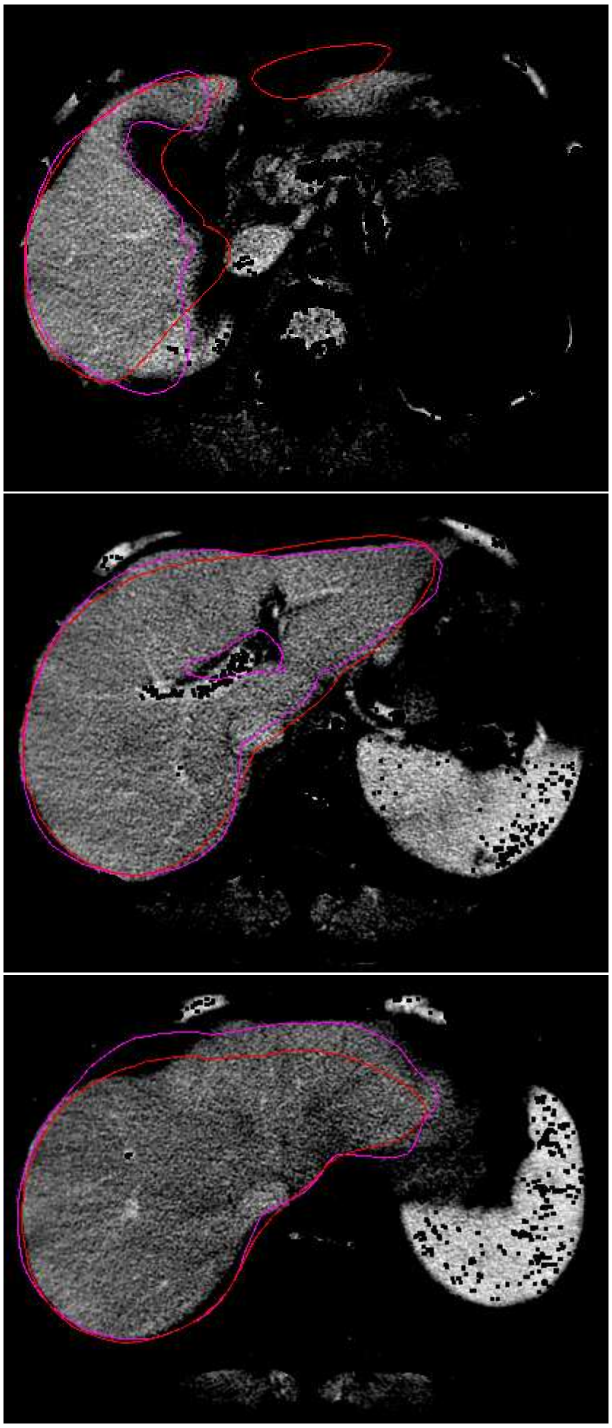


FIG. 6.5 –

Même foie (cas favorable). En rouge, une segmentation automatique classique. En rose, la segmentation par ACP :

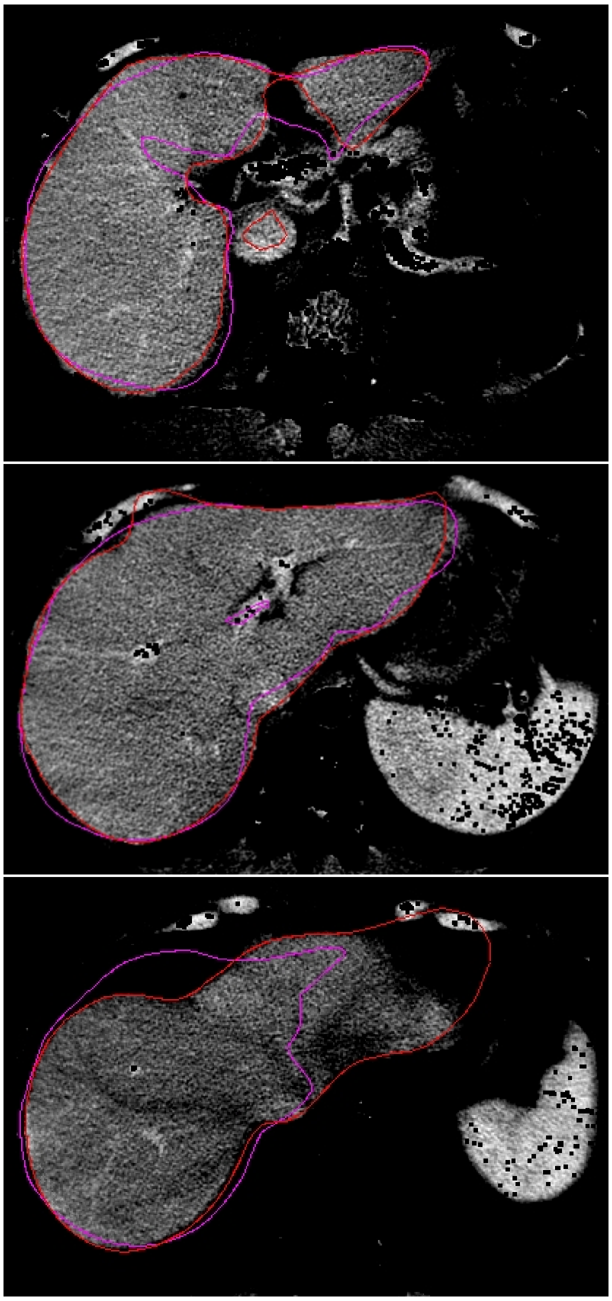


FIG. 6.6 –

Sur la coupe du haut, la segmentation automatique est meilleure. Cela est dû au fait que notre ensemble d'apprentissage n'est pas suffisamment étendu et représentatif.
En revanche, l'avantage d'une segmentation contrainte par les statistiques apparaît clairement sur la coupe du milieu : le maillage déformé par ACP n'a pas segmenté la côte en haut à gauche de l'image, contrairement à l'autre.

Trois coupes d'une segmentation d'une image de foie opéré ne faisant pas partie de l'ensemble d'apprentissage (cas défavorable). En rouge, le foie moyen avant segmentation ; en rose, la segmentation finale par ACP :

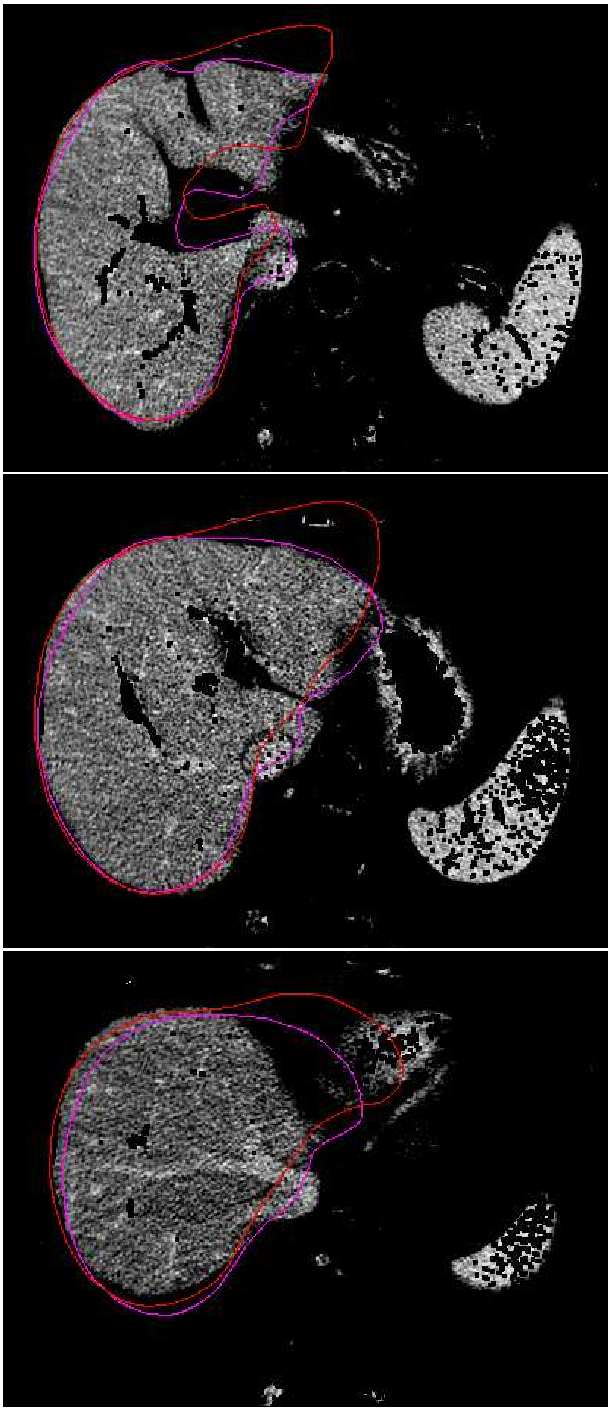


FIG. 6.7 –

Ici, sur la dernière coupe, la plus haute, il apparaît que les deux maillages ont une partie plus proche du cœur (en haut à droite) que du foie. En effet, lors du recalage du foie moyen sur l'image, celui-ci a très certainement été attiré par le cœur. Le maillage segmenté, lui, s'est beaucoup éloigné du cœur, mais reste dans les limites autorisées : si le foie était sain, il devrait y avoir une partie de celui-ci à l'endroit où le maillage s'éloigne du contour du foie.

Cela permet de constater, sans autre information que l'image brute, que le foie n'a pas une forme caractéristique, et qu'il a donc dû subir une ablation.

Même foie (cas défavorable). En rouge, le maillage segmenté automatiquement ; en rose, le maillage segmenté par ACP :

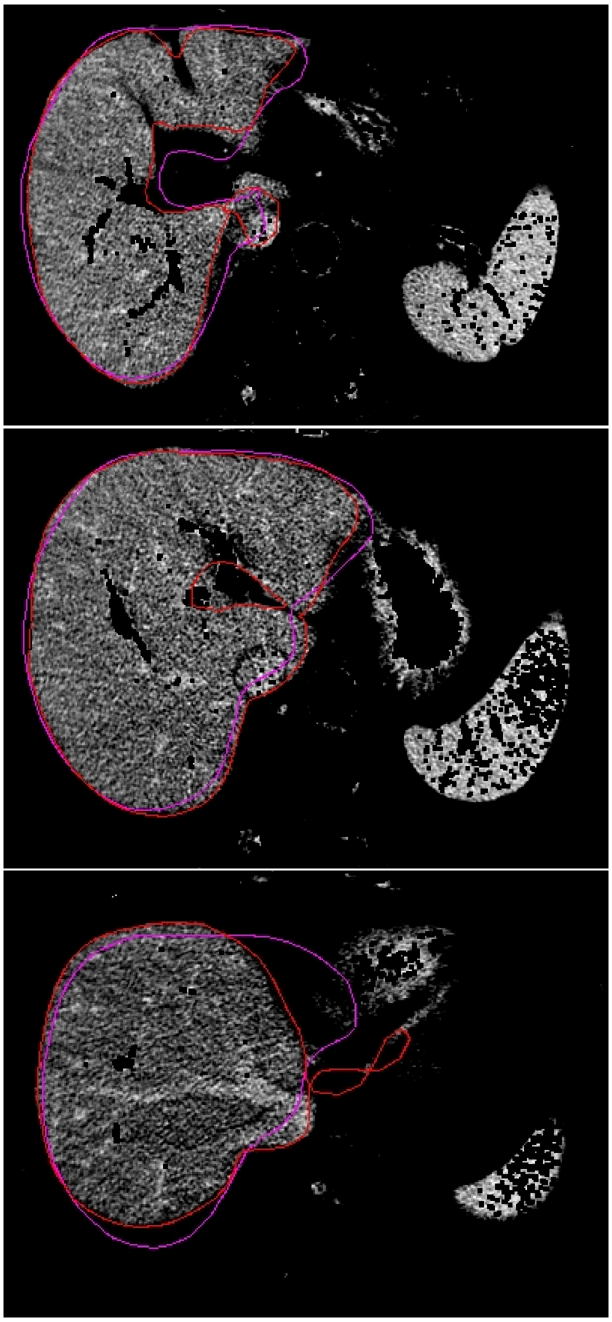


FIG. 6.8 –

6.3.2 Influence du nombre de modes

Le but d'une segmentation basée sur l'ACP étant de décrire l'ensemble des variations possibles avec un nombre minimal de modes, il peut être intéressant de comparer les segmentations obtenues en ne conservant que les premiers modes.
Nous avons donc segmenté une image de foie sain n'appartenant pas à l'ensemble d'apprentissage grâce à notre méthode fondée sur l'ACP, en ne conservant que quelques modes. Nous comparons dans cette section le résultat des segmentations obtenues en ne conservant que trois ou six modes avec la segmentation utilisant l'intégralité des douze modes.

Surface du maillage correspondant à la segmentation avec les trois premiers modes de l'ACP, et maillage correspondant à la segmentation avec tous les modes :

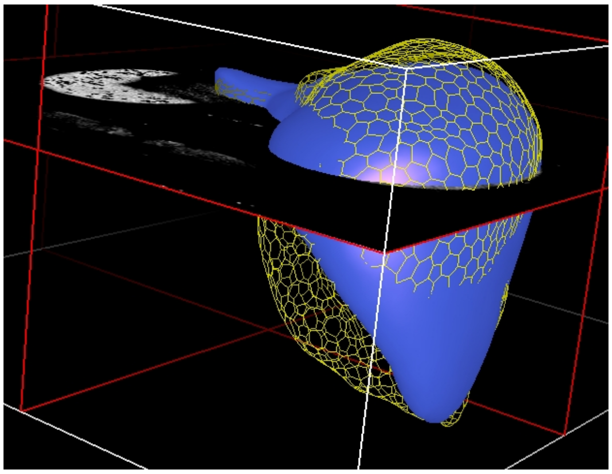


FIG. 6.9 – *En bleu, le résultat de la segmentation avec* 3 *modes. Le maillage représente la segmentation avec les* 12 *modes.*

La figure 6.9 montre l'essentiel des différences entre les segmentations obtenues avec 3 ou 12 modes. Avec seulement 3 modes, certaines régions de ce foie ne peuvent être segmentées : les déformations correspondantes ne figurent pas dans les 3 premiers modes. La figure 6.10 donne quelques vues en coupe.

Même foie (cas favorable). En rouge, le foie segmenté avec 3 modes ; en rose, le foie segmenté avec 12 modes :

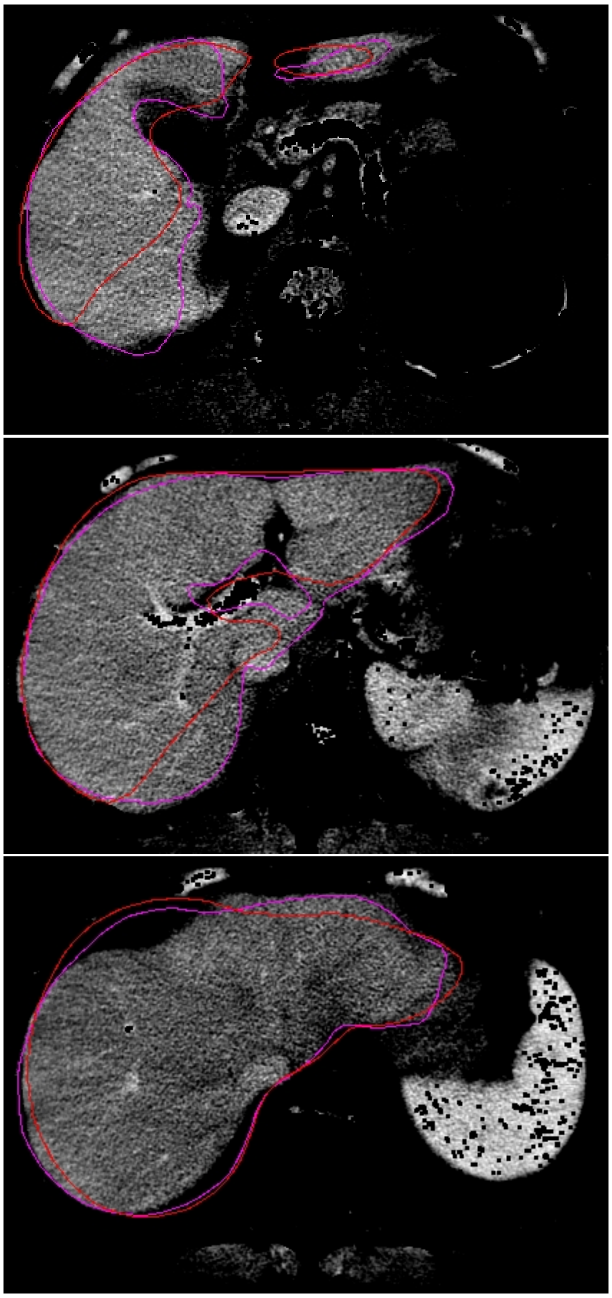


FIG. 6.10 – *Coupes des segmentations avec 12 (rose) et 3 (rouge) modes*

Surface du maillage correspondant à la segmentation avec les trois premiers modes de l'ACP, et maillage correspondant à la segmentation avec tous les modes :

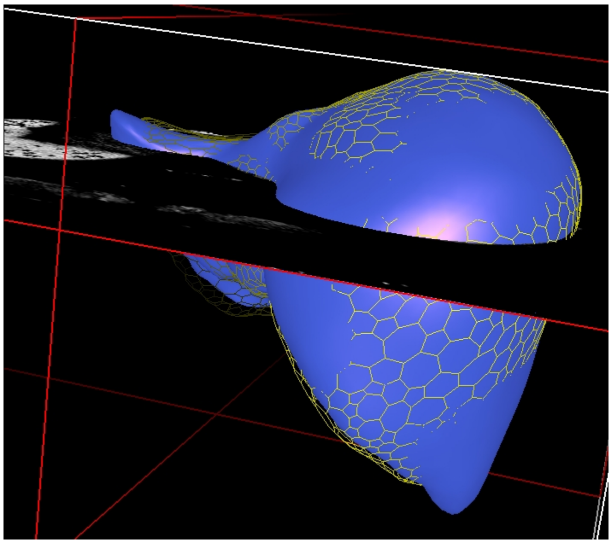


FIG. 6.11 – *En bleu, le résultat de la segmentation avec* 6 *modes. Le maillage représente la segmentation avec les* 12 *modes.*

La figure 6.11 montre l'essentiel des différences entre les segmentations obtenues avec 6 ou 12 modes. Avec la moitié des modes, on arrive à une segmentation relativement proche de celle qui est obtenue avec tous les modes. C'est une bonne illustration du fait que l'ICA compresse l'information : quelques modes peuvent suffire à décrire correctement l'ensemble des variations possibles. La figure 6.12 donne quelques vues en coupe.

Même foie (cas favorable). En rouge, le foie segmenté avec 6 modes ; en rose, le foie segmenté avec 12 modes :

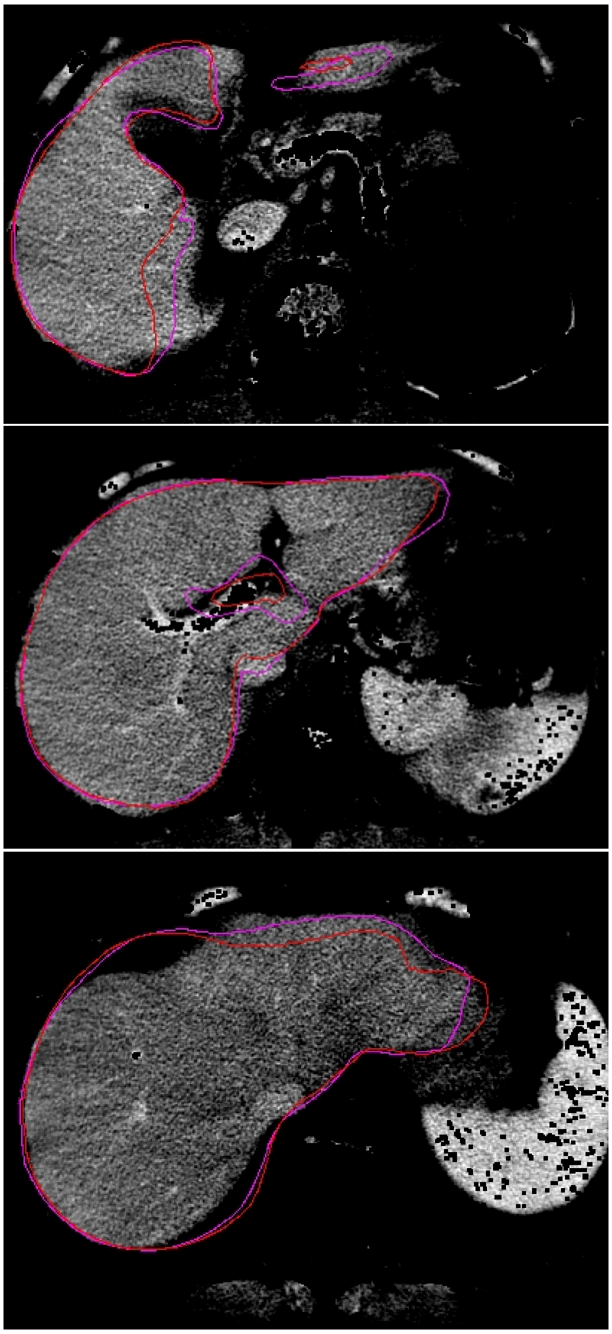


FIG. 6.12 – *Coupes des segmentations avec 12 (rose) et 6 (rouge) modes*

6.4 Couplage avec une segmentation automatique

Notre segmentation contrainte par la PCA nous a permis d'éviter les erreurs grossières consistant à segmenter des bouts d'organes autres que le foie (comme le cœur), mais elle n'est pas parfaite pour autant. Nous sommes cependant parvenus à obtenir des segmentations suffisamment proches de la réalité pour que le résultat serve de point de départ à une segmentation automatique, qui a maintenant toutes les chances de donner de très bons résultats. La figure 6.13 illustre ce fait. Le résultat est particulièrement probant.

Même foie ne faisant pas partie de l'ensemble d'apprentissage. En rouge, le foie segmenté automatiquement ; en rose, le foie segmenté par ACP, puis automatiquement :

6.5 Conclusion

Nous avons segmenté des images de foie grâce à des déformations d'un maillage moyen, contraintes par les modes de l'ACP ; ces segmentations se sont révélées relativement bonnes malgré le faible nombre d'éléments de l'ensemble d'apprentissage ; en effet, elles ont permi d'éviter les erreurs classique des méthodes de segmentation par contours actifs, comme la segmentation de morceaux de cœur, de veine sus-hépatique... On constate que quelques modes suffisent à décrire l'essentiel des variations. Cependant, certains défauts de segmentation nous ont montré que notre ensemble d'apprentissage n'est pas assez fourni, certaines déformations n'étant pas prises en compte par les statistiques, alors qu'elles devraient l'être. La méthode a été mise en défaut dans le cas d'images de foie ayant subi une ablation, ce qui est normal puisque de tels foies ne sont pas représentatifs de notre ensemble d'apprentissage.
La segmentation contrainte par les modes de l'ICA n'a pas été implémentée par manque de temps. Il aurait été nécessaire pour cela de modifier l'algorithme de segmentation, car la famille des modes de l'ICA n'est pas orthonormale, et une simple projection orthogonale de la transformation idéale suivant les modes n'est pas la bonne méthode. Il serait tout de même intéressant de comparer les résultats obtenus avec ceux de l'ACP.
La segmentation d'image tridimensionnelles contrainte par des statistiques s'avère ainsi être une méthode prometteuse, les résultats étant très encourageants.

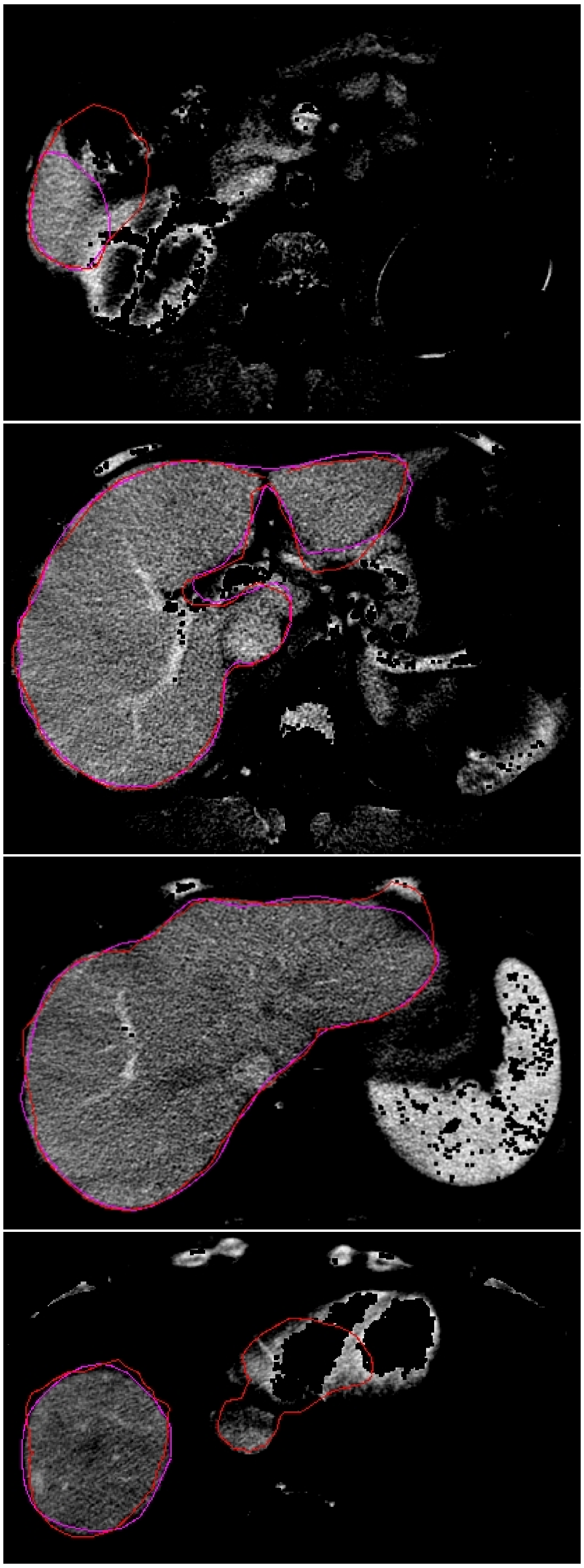


FIG. 6.13 – *Coupes des segmentations automatique en rouge, et ACP+automatique en rose*

Chapitre 7

Fixation de points caractéristiques

Nous cherchons à décrire de manière statistiques les déformations possibles d'un organe, afin de pouvoir segmenter une image tridimensionnelle de cet organe. Mais cette segmentation, comme toute segmentation automatique ou semi-automatique, ne peut être parfaite ; et pour l'améliorer, on peut vouloir imposer à tel sommet du maillage d'avoir telle position dans l'image, sans déroger à la règle de déformation suivant les modes autorisés. Il s'agit alors de calculer la combinaison linéaire des modes qui permet cette translation du sommet en question, tout en modifiant le moins possible le reste du maillage. Ensuite il faudra trouver un moyen de modifier nos modes de déformation de telle sorte qu'ils n'influent plus sur la position du dit sommet.
Tout cecia sera fait dans le cadre de l'ACP. Nous allons exposer ici la méthode proposée par Johannes Hug dans sa thèse ([Hug00]).

7.1 Calcul des bonnes directions à prendre

Nous allons traiter le cas un peu plus général où l'on cherche à fixer les coordonnées de k sommets, dont les indices seront notés : $j_1, \ldots ,j_k$.
Pour commencer, nous allons chercher à calculer $3k$ combinaisons linéaires des modes de déformation, chacune permettant d'incrémenter l'une des coordonnées de l'un de ces k sommets de une unité. Ces combinaisons linéaires s'exprimeront sous la forme de vecteurs $r_{j_1,x}$, $r_{j_1,y}$, $r_{j_1,z}$, $r_{j_2,x}$, $\ldots$, $r_{j_k,z}$ de taille m, tels que si V est la matrice de taille $3n \times m$ dont les colonnes sont les m modes de déformation de l'ACP, alors le vecteur $V.r_{j_{i_0},x}$ (respectivement $V.r_{j_{i_0},y}$, $V.r_{j_{i_0},z}$) est une combinaison linéaire des modes telle que la coordonnée x (resp. y, z) du sommet j_{i_0} soit incrémentée de une unité, et que ses autres coordonnées, ainsi que les coordonnées des autres sommets j_i, soient inchangées.
Mais toutes les combinaisons linéaires ne conviennent pas, et l'on voudrait trouver celles qui déforment le maillage le moins possible.
Pour cela, nous allons minimiser la longueur de Mahalanobis des vecteurs $r_{j_{i_0},x}$, $r_{j_{i_0},y}$ et $r_{j_{i_0},z}$:

$$D(r_{j_{i_0},x}) = \sum_{1 \leq l \leq m} \frac{\left[r_{j_{i_0},x}\right]_l^2}{\lambda_l}$$

$$D(r_{j_{i_0},y}) = \sum_{1 \leq l \leq m} \frac{\left[r_{j_{i_0},y}\right]_l^2}{\lambda_l}$$

$$D(r_{j_{i_0},z}) \;=\; \sum_{1\le l\le m} \frac{\left[r_{j_{i_0},z}\right]_l^2}{\lambda_l}$$

où λ_l est la valeur propre associée au mode l.
D contraint plus fortement les déformations suivant les modes les moins représentatifs. On privilégie ainsi les déformations les plus caractéristiques de l'organe, ce qui a pour objet d'éviter, autant que faire se peut, que l'on déforme le maillage outre mesure.
Notons A la matrice de taille $3k \times m$, extraite de la matrice V, et dont la ligne $3l+1$ est la ligne $3j_l+1$ de V, la ligne $3l+2$ est la ligne $3j_l+2$ de V, et la ligne $3l+3$ est la ligne $3j_l+3$ de V. En d'autres termes, on ne conserve que les lignes de V qui concernent les coordonnées à fixer.
Soit aussi $e_{j_i,x} \;=\; {}^t[0,\ldots,0,1,0,\ldots,0]$ (respectivement $e_{j_i,y}$, $e_{j_i,z}$) le vecteur de taille $3k$ dont toutes les coordonnées sont nulles sauf la coordonnée $3j_i+1$ (resp. y, z), qui vaut 1.
On veut imposer que :

$$\begin{aligned} A.r_{j_i,x} &= e_{j_i,x} \\ A.r_{j_i,y} &= e_{j_i,y} \\ A.r_{j_i,z} &= e_{j_i,z} \end{aligned}$$

On a ainsi $3k$ fonctions de Lagrange associées au problème :

$$\begin{aligned} L_{j_i,x}(r_{j_i,x},\, l_{j_i,x}) &= \sum_{1\le s\le m} \frac{[r_{j_i,x}]_s^2}{\lambda_s} \;-\; {}^t l_{j_i,x}(A.r_{j_i,x} - e_{j_i,x}) \\ L_{j_i,y}(r_{j_i,y},\, l_{j_i,y}) &= \sum_{1\le s\le m} \frac{[r_{j_i,y}]_s^2}{\lambda_s} \;-\; {}^t l_{j_i,y}(A.r_{j_i,y} - e_{j_i,y}) \\ L_{j_i,z}(r_{j_i,z},\, l_{j_i,z}) &= \sum_{1\le s\le m} \frac{[r_{j_i,z}]_s^2}{\lambda_s} \;-\; {}^t l_{j_i,z}(A.r_{j_i,z} - e_{j_i,z}) \end{aligned}$$

Les solutions sont données par la résolution de :

$$\begin{aligned} \frac{\partial L_{j_i,x}}{\partial r_{j_i,x}}(r_{j_i,x},\, l_{j_i,x}) &= \left[\frac{\partial L_{j_i,x}}{\partial\left[r_{j_i,x}\right]_1},\ldots,\frac{\partial L_{j_i,x}}{\partial\left[r_{j_i,x}\right]_m}\right] \\ &= \left[\frac{2r_{j_i,x,1}}{\lambda_1} - ({}^t l_{i_j,x}.A)_1,\ldots,\frac{2r_{j_i,x,m}}{\lambda_m} - ({}^t l_{i_j,x}.A)_m\right] \\ &= 2\;{}^t r_{j_i,x}.\begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & O \\ O & \vdots & \\ & & \lambda_m \end{bmatrix}^{-1} \;-\; {}^t l_{j_i,x}.A \\ &= 0 \end{aligned}$$

et de même, $\frac{\partial L_{j_i,y}}{\partial r_{j_i,y}}(r_{j_i,y},\, l_{j_i,y}) = 0$ et $\frac{\partial L_{j_i,z}}{\partial r_{j_i,z}}(r_{j_i,z},\, l_{j_i,z}) = 0$.
Ceci nous donne les systèmes suivants à résoudre :

$$\left\{\begin{aligned} 2.\Lambda^{-1}.r_{j_i,x} \;-\; {}^t A.l_{j_i,x} &= 0 \\ A.r_{j_i,x} &= e_{j_i,x} \end{aligned}\right. ,$$

$$\left\{\begin{aligned} 2.\Lambda^{-1}.r_{j_i,y} \;-\; {}^t A.l_{j_i,y} &= 0 \\ A.r_{j_i,y} &= e_{j_i,y} \end{aligned}\right. ,$$

$$\left\{\begin{aligned} 2.\Lambda^{-1}.r_{j_i,z} \;-\; {}^t A.l_{j_i,z} &= 0 \\ A.r_{j_i,z} &= e_{j_i,z} \end{aligned}\right. .$$

Soit encore :

$$\left[\begin{array}{c|c} 2\Lambda^{-1} & -{}^tA \\ \hline A & O \end{array}\right] \left[\begin{array}{c} r_{j_i,x} \\ \hline l_{j_i,x} \end{array}\right] = \left[\begin{array}{c} O \\ \hline e_{j_i,x} \end{array}\right]$$

et de même,

$$\left[\begin{array}{c|c} 2\Lambda^{-1} & -{}^tA \\ \hline A & O \end{array}\right] \left[\begin{array}{c} r_{j_i,y} \\ \hline l_{j_i,y} \end{array}\right] = \left[\begin{array}{c} O \\ \hline e_{j_i,y} \end{array}\right], \left[\begin{array}{c|c} 2\Lambda^{-1} & -{}^tA \\ \hline A & O \end{array}\right] \left[\begin{array}{c} r_{j_i,z} \\ \hline l_{j_i,z} \end{array}\right] = \left[\begin{array}{c} O \\ \hline e_{j_i,z} \end{array}\right]$$

On peut récrire ceci de manière plus condensée, en notant :

$$R = \left[\begin{array}{c|c|c|c|c|c|c} r_{j_1,x} & r_{j_1,y} & r_{j_1,z} & \ldots & r_{j_m,x} & r_{j_m,y} & r_{j_m,z} \end{array}\right]$$

,

$$Q = \left[\begin{array}{c|c|c|c|c|c|c} l_{j_1,x} & l_{j_1,y} & l_{j_1,z} & \ldots & l_{j_m,x} & l_{j_m,y} & l_{j_m,z} \end{array}\right]$$

On a alors :

$$\left[\begin{array}{c|c} 2\Lambda^{-1} & -{}^tA \\ \hline A & O \end{array}\right] \left[\begin{array}{c} R \\ \hline Q \end{array}\right] = \left[\begin{array}{c} O \\ \hline I_{3k} \end{array}\right]$$

ou encore :

$$\begin{cases} 2.\Lambda^{-1}.R \; - \; {}^tA.Q \; = \; 0 \\ A.R \; = \; I_{3k} \end{cases}$$

On a donc $R = \frac{1}{2}\Lambda.{}^tA.Q$, donc $\frac{1}{2}A.\Lambda.{}^tA.Q = A.R = I_{3k}$. Les matrices $A.\Lambda.{}^tA$ et Q sont de taille $3k \times 3k$, et leur produit donne l'identité, donc elles sont inversibles, et on a :

$$Q = 2(A.\Lambda.{}^tA)^{-1}$$

Comme $R = \frac{1}{2}\Lambda.{}^tA.Q$, cela donne finalement :

$$\boxed{R = \Lambda.{}^tA.(A.\Lambda.{}^tA)^{-1}}$$

7.2 Mise en œuvre

On peut maintenant déformer le maillage courant $Mcurr$ selon les modes de manière à imposer les coordonnées des sommets $j_1, \ldots, j_m$:

$$Mcurr^* = Mcurr + V.R. \begin{bmatrix} \Delta x_{j_1} \\ \Delta y_{j_1} \\ \Delta z_{j_1} \\ \vdots \\ \Delta x_{j_k} \\ \Delta y_{j_k} \\ \Delta z_{j_k} \end{bmatrix} = Mcurr + \sum_{1 \leq l \leq k} (\Delta x_{j_l}.V.r_{j_l,x} + \Delta y_{j_l}.V.r_{j_l,y} + \Delta z_{j_l}.V.r_{j_l,z})$$

Mais si l'on effectue ensuite une segmentation selon les modes à partir du nouveau maillage obtenu, les coordonnées que nous avons fixé vont immanquablement changer. Il faut donc modifier aussi les modes de variation de telle sorte qu'ils ne permettent pas à ces coordonnées de changer. Chaque maillage d'apprentissage s'écrit :

$$\overline{X} + X_i = \overline{X} + V.b_i$$

où b_i est le vecteur des coefficients représentant X_i dans la base des modes.
On déforme chacun de ceux-ci de manière à ce que les coordonnées des sommets $j_1, \ldots, j_m$ soient les mêmes que pour le maillage $Mcurr^*$:

$$X_i^* = X_i + V.R. \begin{bmatrix} \Delta x_{i,j_1} \\ \Delta y_{i,j_1} \\ \Delta z_{i,j_1} \\ \vdots \\ \Delta x_{i,j_k} \\ \Delta y_{i,j_k} \\ \Delta z_{i,j_k} \end{bmatrix}$$

On effectue alors une nouvelle analyse statistique de ces nouveaux modèles d'apprentissage, ce qui nous donne des modes de déformation laissant invariantes les coordonnées des sommets que l'on désire fixer.
Sur la figure 7.1 apparaît le premier mode de déformation obtenu après fixation du sommet encadré dans l'image.

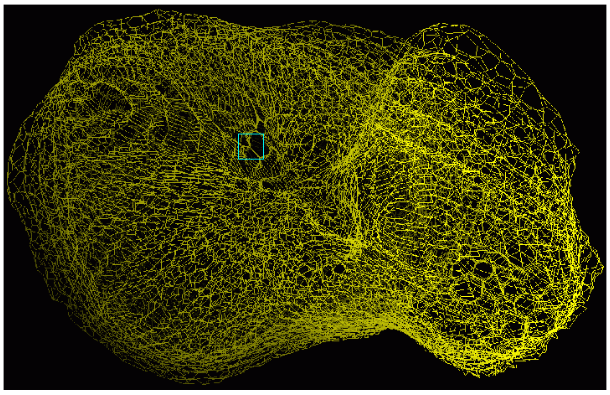


FIG. 7.1 – *Vue de dessous du premier mode de déformation avec un point fixe*

Notons que l'on perd trois degrés de liberté par sommet fixé, et donc que le nombre de modes a été réduit d'autant.

7.3 Choix des sommets à fixer : potentiel de réduction

Il peut être intéressant de déterminer les sommets dont la fixation est optimale, dans le sens où les modes de déformation résultant de la fixation de ces sommets induisent le moins de variations possible.
Ceci s'exprime mathématiquement de la manière suivante : la somme des valeurs propres (des variances) correspondant aux modes de déformation doit être minimale.
Pour déterminer cela, on remplace chaque $X_i = V.b_i$ par :

$$X_i^* \;=\; V.(b_i - R.\begin{bmatrix} x_{j_1}(X_i) \\ y_{j_1}(X_i) \\ z_{j_1}(X_i) \\ \vdots \\ x_{j_k}(X_i) \\ y_{j_k}(X_i) \\ z_{j_k}(X_i) \end{bmatrix}) \;=\; V.(b_i - R.A.b_i) \;=\; V.(I \;-\; R.A).b_i$$

c'est-à-dire que l'on fixe les coordonnées des sommets $j_1, \ldots, j_k$ des X_i à zéro, de manière à ce que les maillages $\overline{X} + X_i^*$ aient des sommets $j_1, \ldots, j_k$ aux mêmes positions que ceux du maillage moyen $\overline{X}$.
Si $\Lambda_{j_1,\ldots,\, j_k}$ est la matrice diagonale des variances associées aux modes de déformation obtenus par fixation des sommets $j_1, \ldots, j_k$, on note $P(\Lambda_{j_1,\ldots,\, j_k}) = -Tr(\Lambda_{j_1,\ldots,\, j_k})$ le potentiel de réduction associé.
On choisira les sommets qui maximisent le potentiel de réduction. C'est-à-dire que l'on veut fixer les sommets qui enlèvent le plus de variabilité.
On peut aussi procéder itérativement, en commençant par fixer un sommet, celui qui a le plus fort potentiel de réduction, puis par fixer celui qui a le plus fort potentiel de réduction après fixation du premier, etc.
Résultats (cas de la fixation d'un sommet) : La figure 7.2 montre les potentiels de réduction associés à la fixation de chacun des sommets.

Le potentiel maximal est atteint pour le sommet 1210, qui correspond à un sommet situé sur la partie supérieure du foie ; il s'agit d'une zone qui varie beaucoup d'un foie à l'autre. Imposer les coordonnées de ce sommet est donc très contraignant. Le plus faible potentiel est atteint pour le sommet 2372, qui se situe, lui, sous le foie ; il s'agit d'un point situé dans une zone de faible variabilité ; il est donc logique que sa fixation soit beaucoup moins contraignante. Les résultats observés sont donc conformes à ceux auxquels on pouvait s'attendre. Il faut toutefois remarquer que les potentiels de réduction se situent pour l'essentiel dans une même gamme de valeurs. Beaucoup de sommets ont un fort potentiel de réduction. Une visualisation des valeurs sur la surface du maillage moyen serait certainement plus parlante : c'est ce que permet la figure 7.3 : chaque sommet porte une couleur d'autant plus chaude (rouge) que son potentiel de réduction est grand. Il apparaît que les régions de faible potentiel de réduction sont situées sous le foie (la plus importante), ce qui est logique, mais aussi sur certaines parties plus variables, ce qui est étonnant. On trouve par ailleurs des sommets isolés de très faible potentiel de réduction dans des zones de fort potentiel. Ce genre d'incohérence peut provenir du fait que la correspondance entre les sommets de même indice de deux maillages différents ne correspond pas nécessairement à une réalité physique. C'est une description qui est peut-être trop restrictive.

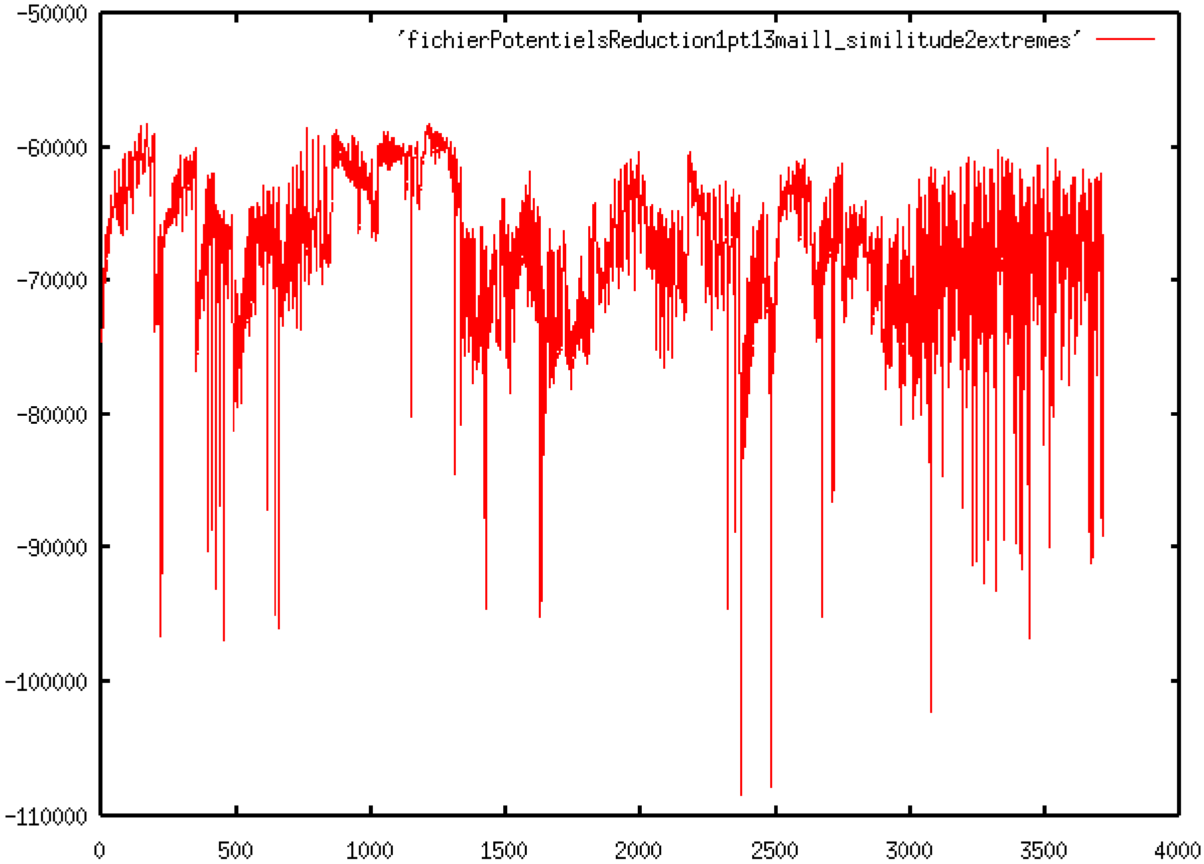


FIG. 7.2 – *Potentiel de réduction associé à la fixation d'un sommet*

Il est toutefois intéressant de comparer le premier mode de déformation associé à la fixation du sommet de plus fort potentiel avec le premier mode correspondant à la fixation du sommet de plus faible potentiel. En effet, le premier mode est celui qui exprime la plus grande variabilité, et donc il devrait témoigner d'une variabilité d'autant plus faible que le potentiel de réduction du sommet fixé était fort. On le vérifie visuellement sur les figures 7.5 et 7.6.

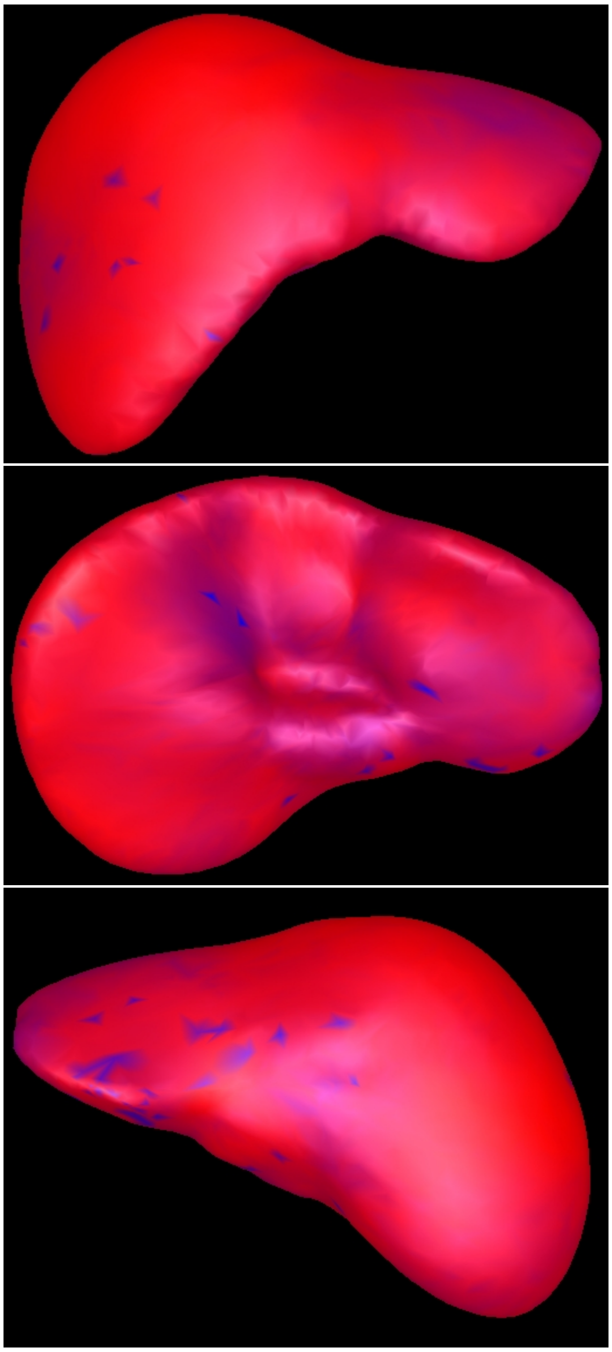


FIG. 7.3 – *Carte des potentiels de réduction associés à la fixation d'un sommet : les potentiels les plus forts sont en rouge, les plus faible en bleu.*

La figure 7.4 montre les potentiels de réduction associés à la fixation de chacun des sommets une fois que le sommet 1210 a été fixé. On peut observer que le potentiel est minimal pour le sommet 1210, puisque dans ce cas on ne rajoute aucune contrainte.

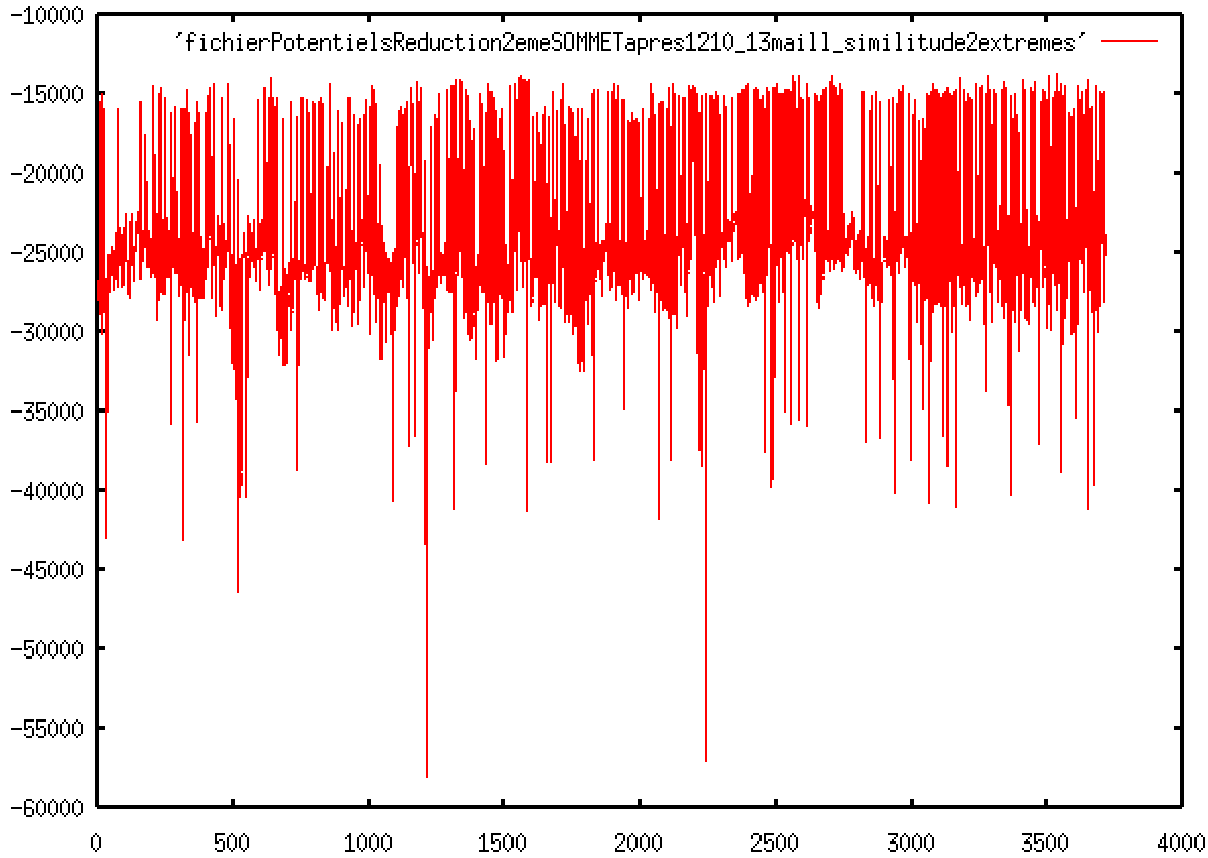
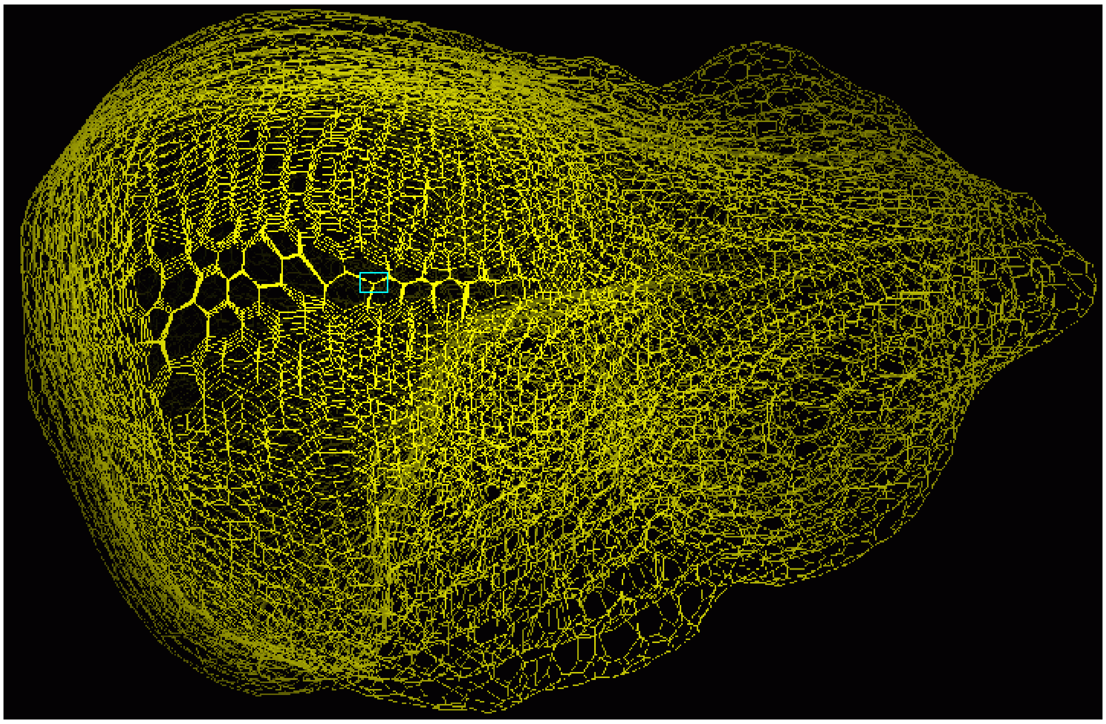


FIG. 7.4 – *Potentiel de réduction associé à la fixation d'un sommet après fixation du premier sommet de plus fort potentiel.*

7.4 Résultats de la segmentation

Nous allons exposer ici trois expériences riches d'informations. Dans chacune d'entre elles, le maillage du foie moyen ainsi que les modes de déformation sont correctement recalés sur l'image.
Expérience 1 :
On cherche à segmenter une image de l'ensemble d'apprentissage, en fixant les coordonnées du sommet de plus fort potentiel de réduction (le sommet 1210). Nous allons faire ceci de manière totalement artificielle, en donnant au sommet en question les coordonnées qu'il a dans le maillage de l'ensemble d'apprentissage qui correspond à l'image. C'est-à-dire que l'on introduit notre connaissance a priori des coordonnées que le sommet 1210 DOIT avoir.
Dans ce cas, la segmentation est deux fois plus rapide que la segmentation classique par ACP (5 itérations au lieu de 10), ce qui montre que la méthode fonctionne bien si l'on connaît exactement les coordonnées que le sommet à fixer doit avoir.
Examinons maintenant un cas plus réaliste :
Expérience 2 :
Nous cherchons à segmenter une image de foie sain qui ne fait pas partie de l'ensemble d'apprentissage. Nous allons donc choisir des nouvelles coordonnées pour le sommet 1210, qui seront



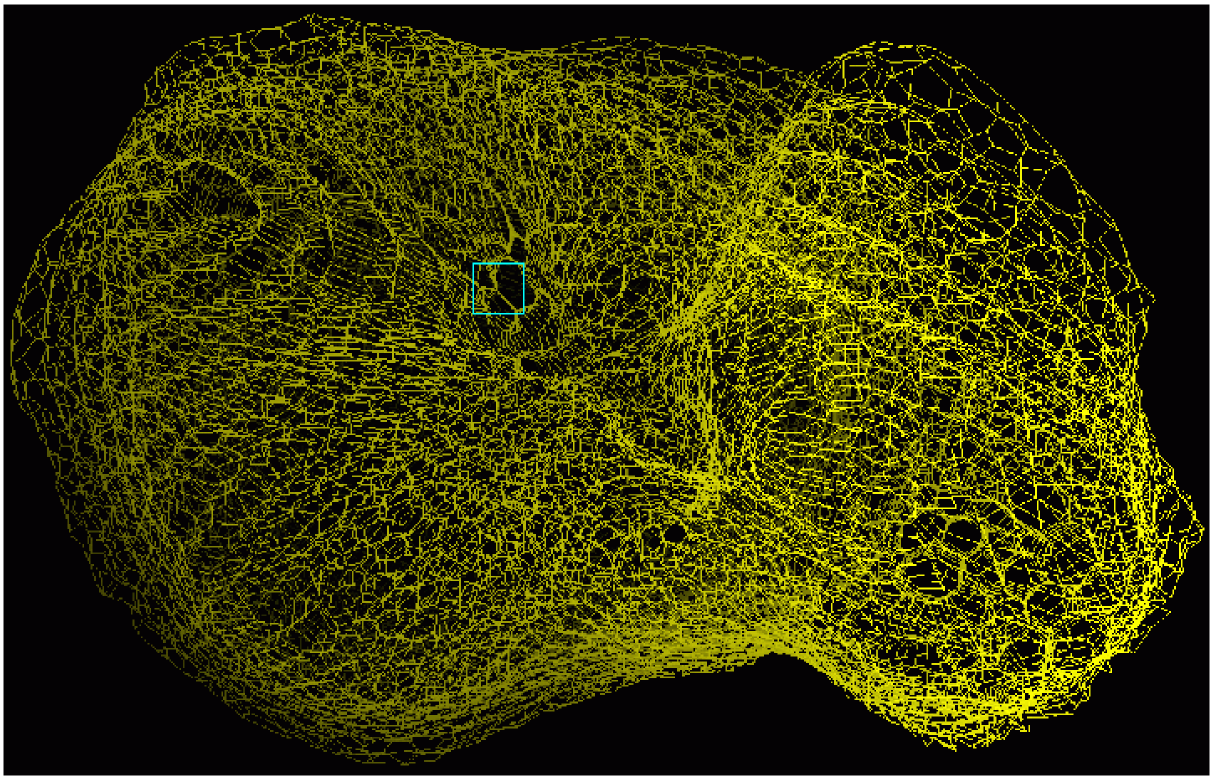


Fig. 7.5 – *Premier mode de déformation avec fixation du point de plus fort potentiel de réduction en haut, de plus faible potentiel de réduction en bas*

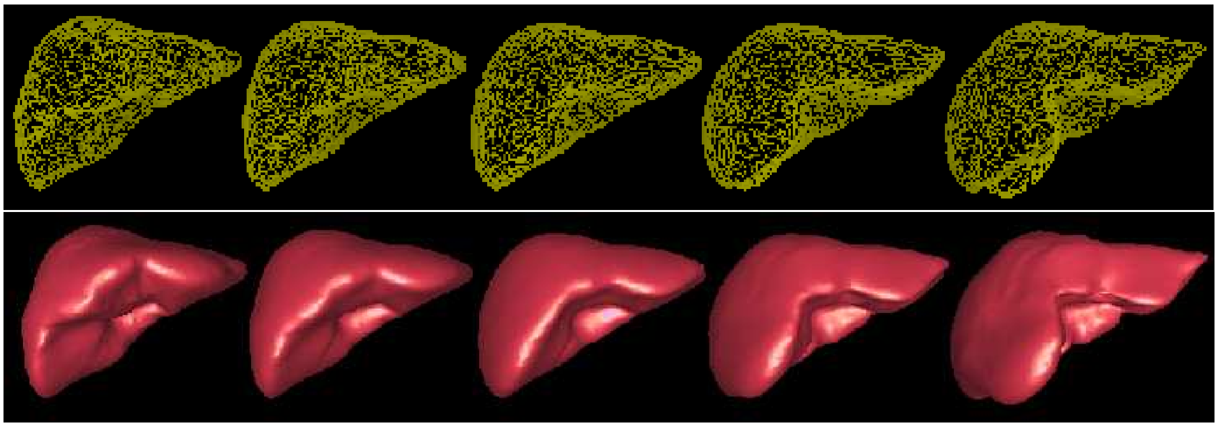


FIG. 7.6 – *Premier mode de déformation avec fixation du point de* plus fort *potentiel de réduction*

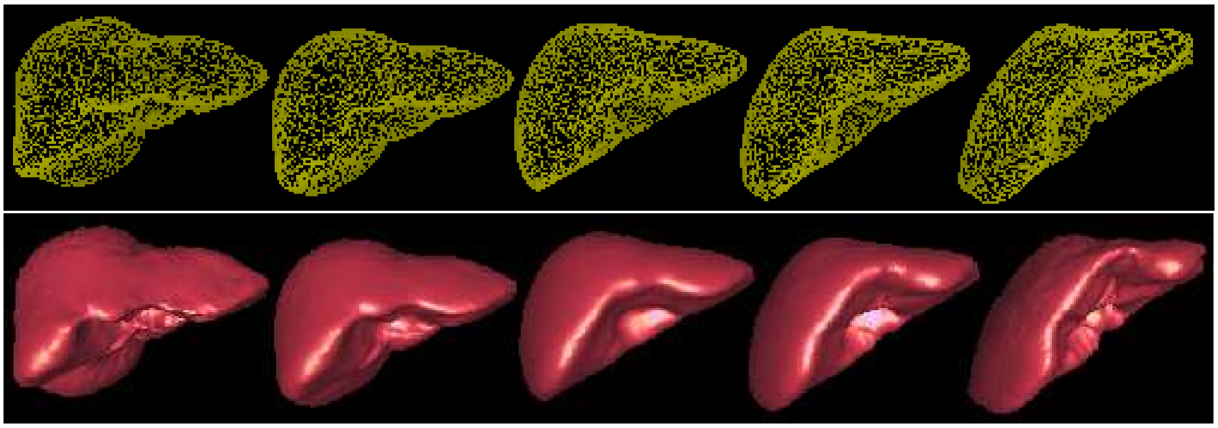


FIG. 7.7 – *Premier mode de déformation avec fixation du point de* plus faible *potentiel de réduction : les variations sont plus importantes.*

celles d'un point de l'image qui se trouve sur la surface du foie et proche des coordonnées initiales du sommet 1210 (anciennes coordonnées : [76.729668, 102.028091, 102.204269], nouvelles coordonnées : [75.880028, 118.857934, 95.732876]). Le nouveau maillage moyen obtenu (figures 7.10 et 7.12) est bien trop déformé pour pouvoir donner lieu à une segmentation correcte : si le sommet 1210 a bien les coordonnées voulues, la majorité des autres sommets se sont beaucoup trop éloignés de leur position d'origine. Le maillage a encore une forme proche de celle d'un foie, mais celle-ci ne correspond plus du tout à ce qu'elle devrait être.
Il semble donc que le choix de la nouvelle position du sommet 1210 n'ait pas été judicieux. Nous allons donc nous placer à nouveau dans un cas irréaliste (et sans intérêt pratique), pour tenter de trouver une meilleure position pour le sommet 1210 :
Expérience 3 :
Nous conservons la même image d'un foie qui ne fait pas partie de l'ensemble d'apprentissage. Cette fois, nous allons utiliser un maillage correspondant à une segmentation automatique (qui n'est pas partout satisfaisante, mais qui l'est autour de la position du sommet 1210) de l'image, faite à partir du maillage de référence du foie de la NLM (donc sur le même modèle que notre maillage moyen). Nous avons fait l'approximation, depuis le départ, que les sommets d'indice donné de deux maillages représentent le même point physique sur le foie, c'est-à-dire que l'appariement entre sommets de même indices sur des maillages différents correspond à une réalité. Si cette approximation est suffisamment précise, la position du sommet 1210 du maillage correctement segmenté (coordonnées [62.7352, 128.437, 106.064]) devrait être une bonne position à imposer à notre maillage moyen. Malheureusement, il se trouve que le résultat est similaire à celui de l'expérience 2 (figures 7.11 et 7.13) ; on n'a rien gagné.

Interprétation :
L'expérience 1 nous montre que la méthode de fixation d'un sommet fonctionne bien s'il existe une combinaison linéaire des modes à faibles coefficients qui permette de réaliser la fixation du sommet en question à la bonne position. On obtient alors un nouveau maillage moyen plus proche de l'image que l'original, et la segmentation est plus rapide. Mais nous avons introduit artificiellement un appariement exact entre le sommet 1210 du maillage moyen et une position idéale.
En revanche, l'expérience 2 montre que dans le cas réaliste où l'on ne connaît qu'à peu près la position à donner au sommet à fixer, notre méthode ne marche plus : rien ne nous assure qu'il existe une combinaison linéaire des modes de déformations qui ne bouleverse pas la forme de notre maillage moyen et qui en même temps réalise l'appariement du sommet à fixer avec sa nouvelle position (déterminée arbitrairement). Il semble donc que la fixation exacte des coordonnées d'un sommet de manière arbitraire soit une contrainte excessive, qui rend invalide l'hypothèse d'appariement exact entre sommets de mêmes indices.
Nous atteignons dans l'expérience 3 les limites de cette hypothèse : la position du sommet 1210 du maillage segmenté de manière automatique ne fait pas partie de celles qui donnent une bonne segmentation avec cette nouvelle méthode.
Il est probable aussi que l'on manque de modes de déformation, c'est-à-dire de maillages d'apprentissage. On peut en effet espérer qu'avec un échantillon plus vaste, nous obtiendrions des modes qui permettent d'imposer arbitrairement (mais raisonnablement) les coordonnées d'un sommet tout en conservant une forme proche du maillage moyen d'origine.
Sur la figure 7.9, on peut voir, en coupe, le premier mode de variation une fois que le sommet 1210 (en haut à gauche) a été fixé.

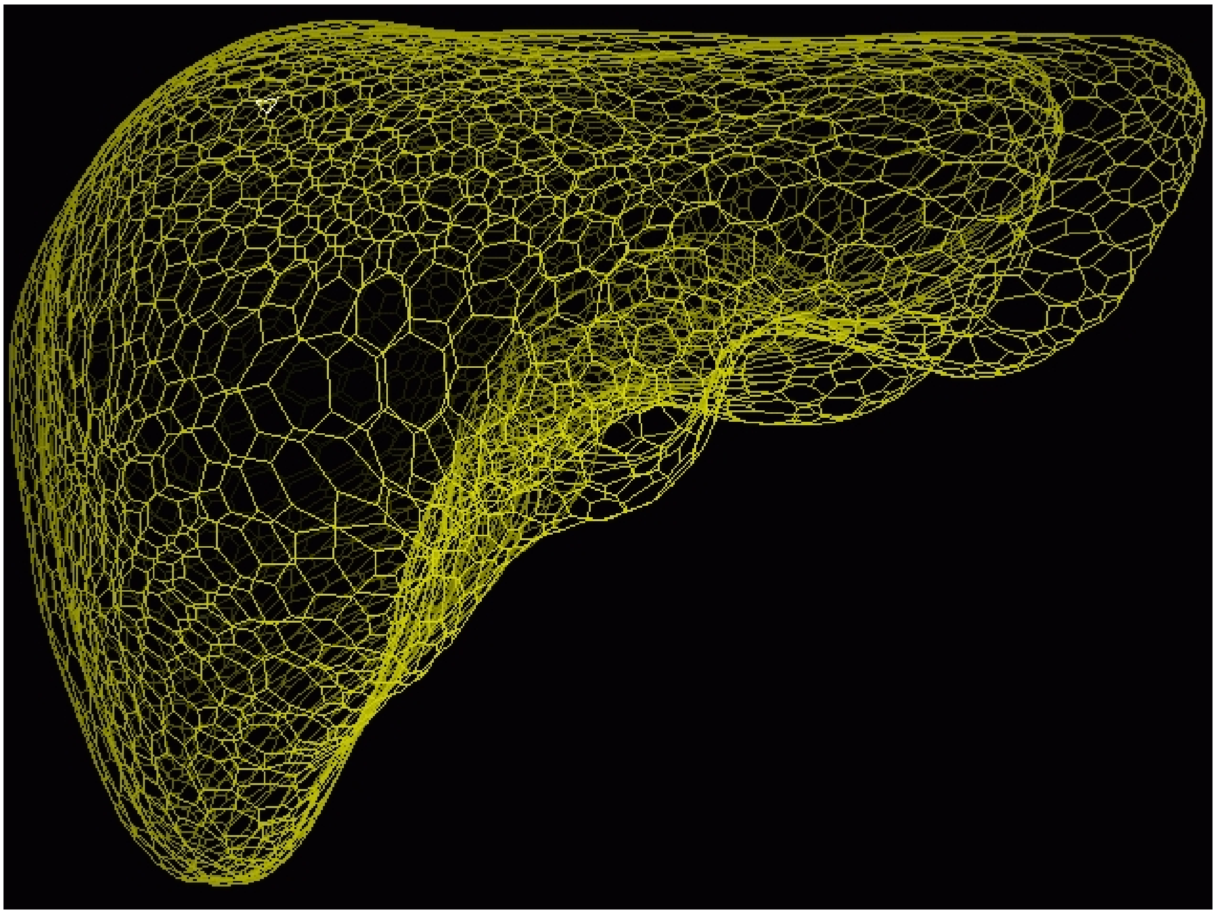


FIG. 7.8 – *Maillage allongé : le nouveau foie moyen de l'expérience 1 ; l'autre maillage est celui du maillage moyen d'origine. Le petit triangle blanc permet de repérer le sommet* 1210 *du maillage de la segmentation*

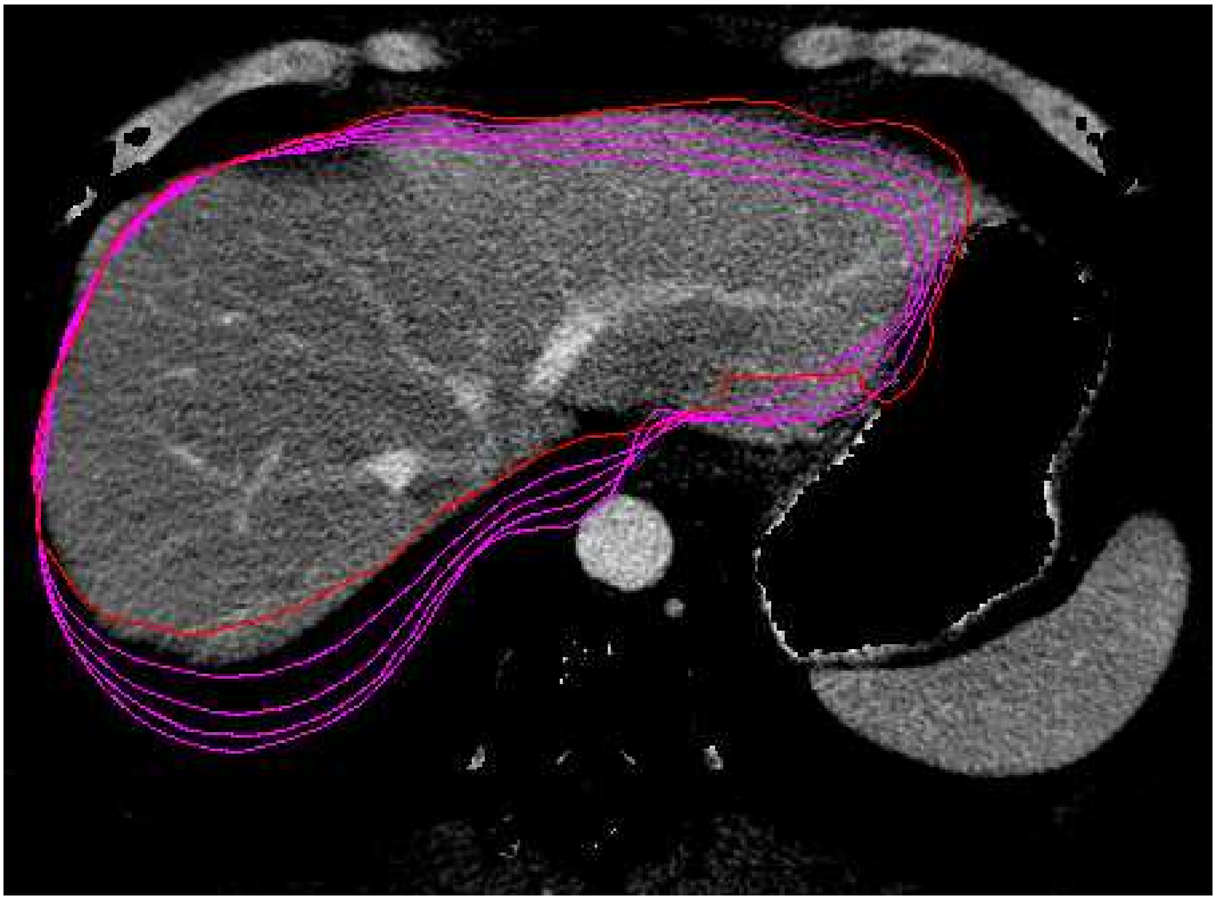


Fig. 7.9 – *Premier mode de déformation avec fixation du point de plus fort potentiel de réduction (en coupe) sur une image de l'ensemble d'apprentissage (expérience 1)*

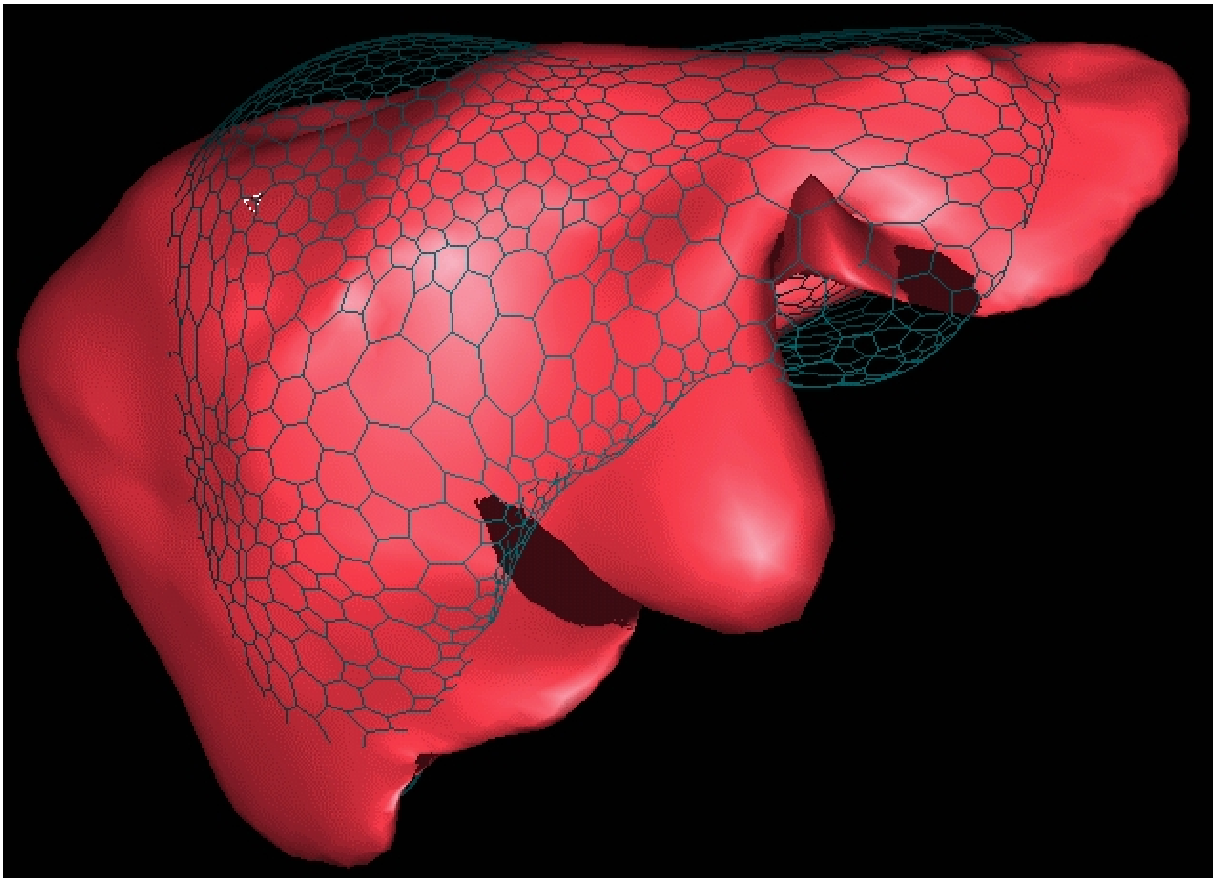


FIG. 7.10 – *En rouge, le nouveau foie moyen de l'expérience 2 ; le maillage non texturé est celui du maillage moyen d'origine. Le petit triangle blanc permet de repérer le sommet* 1210 *du maillage texturé*

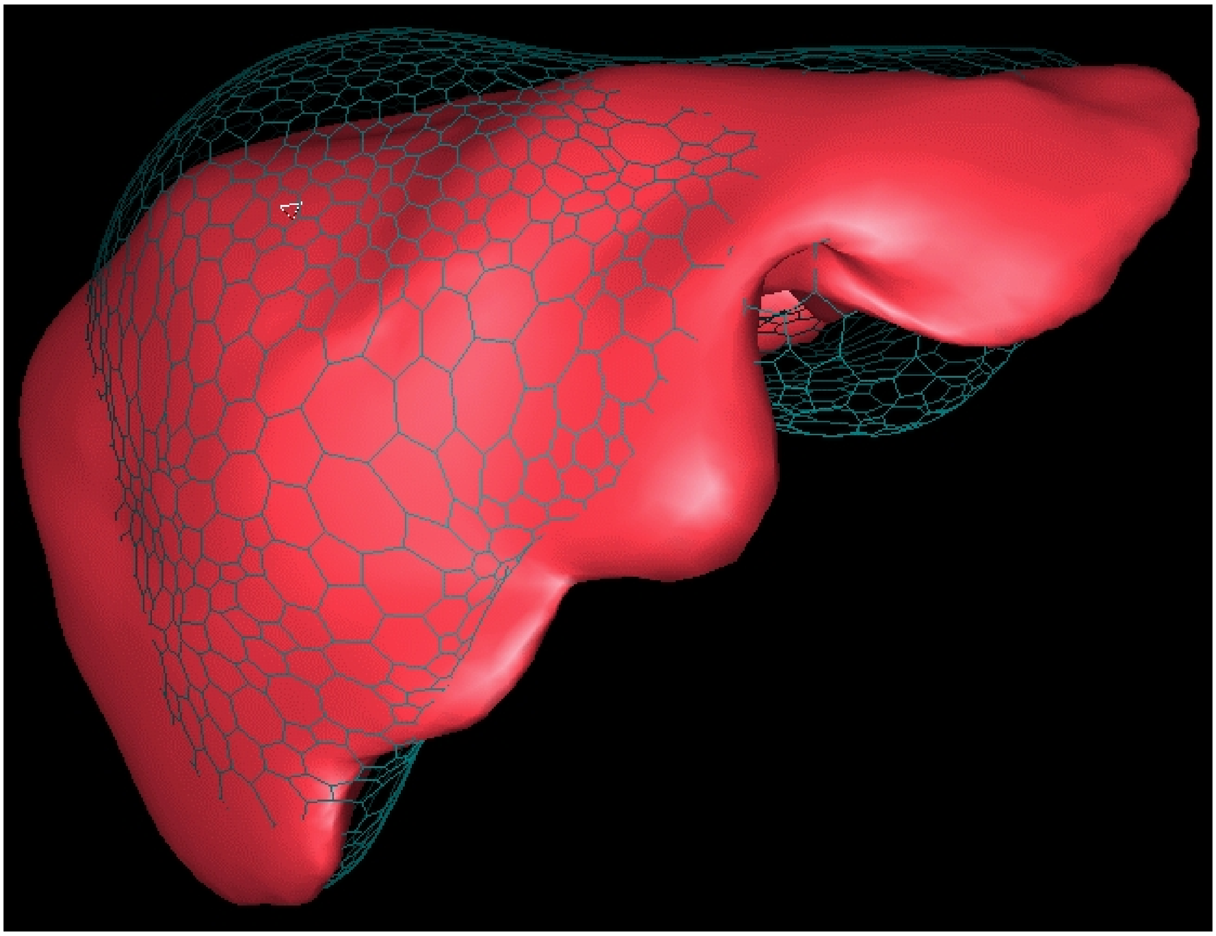


FIG. 7.11 – *En rouge, le nouveau foie moyen de l'expérience 3 ; le maillage non texturé est celui du maillage moyen d'origine. Le petit triangle blanc permet de repérer le sommet* 1210 *du maillage texturé*

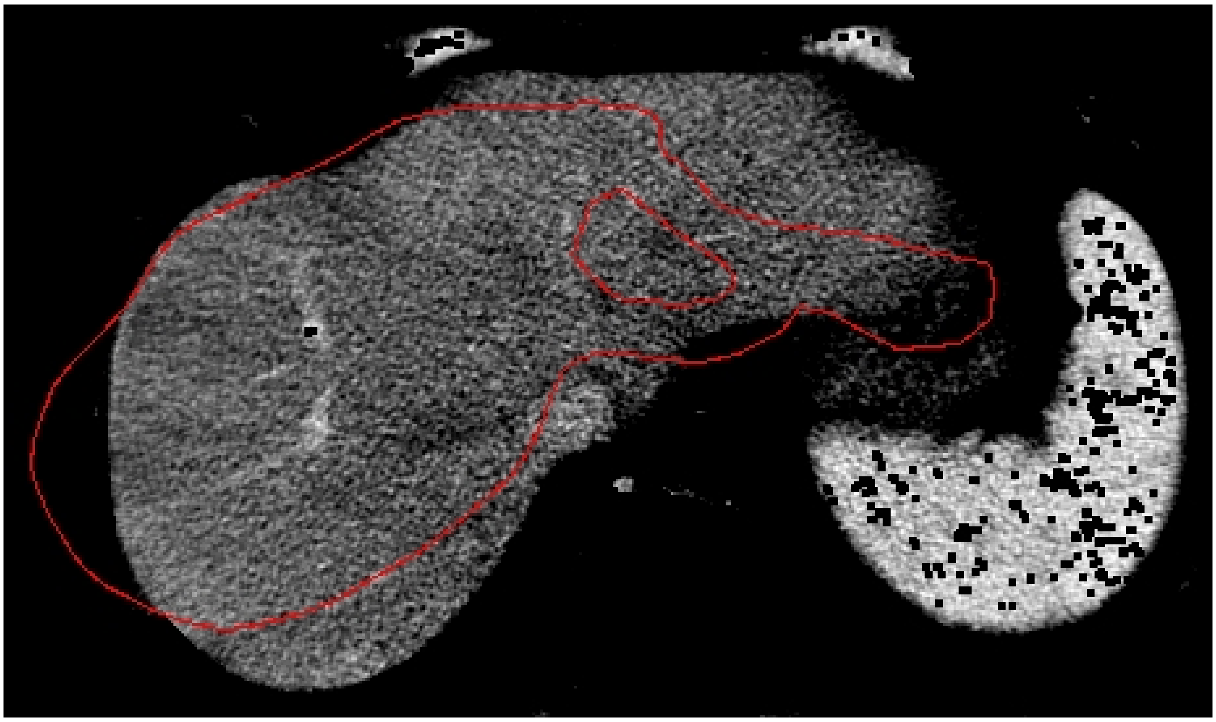


FIG. 7.12 – *Coupe du foie avec trace du maillage segmenté de l'expérience 2. Le sommet* 1210 *se trouve en haut à gauche, légèrement en-dessous d'une côte (tache claire)*

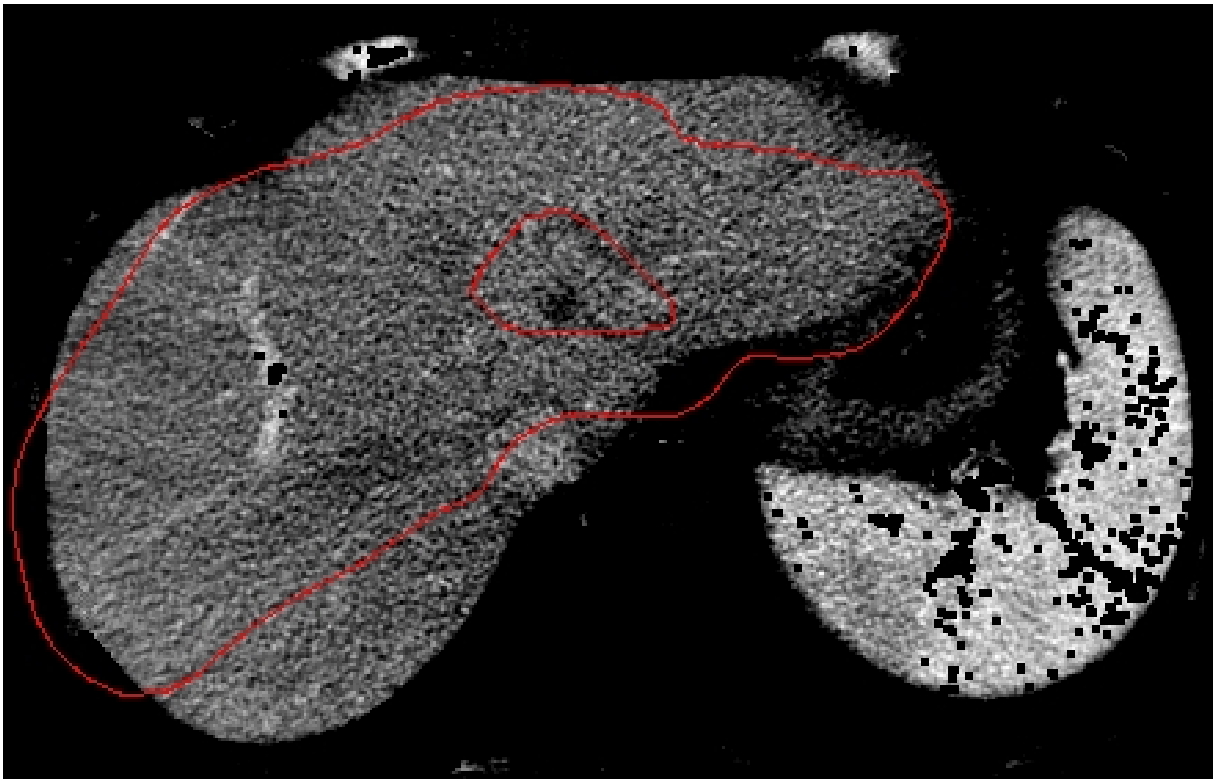


FIG. 7.13 – *Coupe du foie avec trace du maillage segmenté de l'expérience 3. Le sommet* 1210 *se trouve en haut à gauche, légèrement en-dessous et à gauche d'une côte (tache claire)*

Chapitre 8

Conclusion et perspectives

Nous avons exposé une technique de segmentation contrainte par les statistiques, en proposant une approche par ACP et une approche par ICA. Nous avons pu implémenter la première, et nous avons obtenu des résultats très satisfaisant, compte tenu du faible nombre de foies de notre ensemble d'apprentissage. Il n'est pas clair que l'ICA aurait donné des résultats sensiblement meilleurs, mais il aurait fallu un eu plus de temps pour être fixés sur ce point. En ce qui concerne la fixation de sommets pour améliorer la segmentation, celle-ci s'est révélée inappliquable telle qu'elle a été décrite et implémentée. Il est toutefois possible qu'un ensemble d'apprentissage plus conséquent eut rendu cette méthode plus efficace.

8.1 Perspectives : relâchement des contraintes d'appariement

Nous avons vu que notre appariement des sommets de même indices est une procédure qui peut se révéler trop simpliste. Elle pourrait peut-être être raffinée de la manière suivante :
Soient M1 et M2 deux maillages ayant le même nombre de sommets, et que l'on veut apparier. Typiquement, le sommet $S_{i,1}$ numéro i de M1 et le sommet $S_{i,2}$ numéro i de M2 représentent des points physiques du foie assez proches, mais ce n'est pas pour autant le meilleur appariement possible (voir figure 8.1).
Une idée est d'apparier le sommet $S_{i,1}$ de M1 avec un point du plan tangent au maillage M2 en son sommet $S_{i,2}$. On pourrait simplement projeter $S_{i,1}$ sur le plan en question, mais rien n'indique que ce soit la bonne méthode : cette projection, notée H_i peut être trop éloignée de $S_{i,2}$. Il serait peut-être préférable de choisir un compromis entre H_i et $S_{i,2}$. Si la courbure en $S_{i,2}$ dans la direction donnée par $S_{i,2}H_i$ forte, il faut contraindre le point apparié à $S_{i,1}$ à être proche de $S_{i,2}$, car une forte courbure indique une meilleure localisation.
Un meilleur appariemment permettrait d'effectuer des recalages plus précis d'un maillage sur l'autre. Cela pourrait même permettre d'améliorer la segmentation contrainte par les modes, en appariant non plus un sommet M_i du maillage courant avec le point P_i de plus fort gradient selon la normale, mais avec un point de la surface du maillage formé par les points P_i, ce qui serait moins contraignant et donc probablement plus efficace.

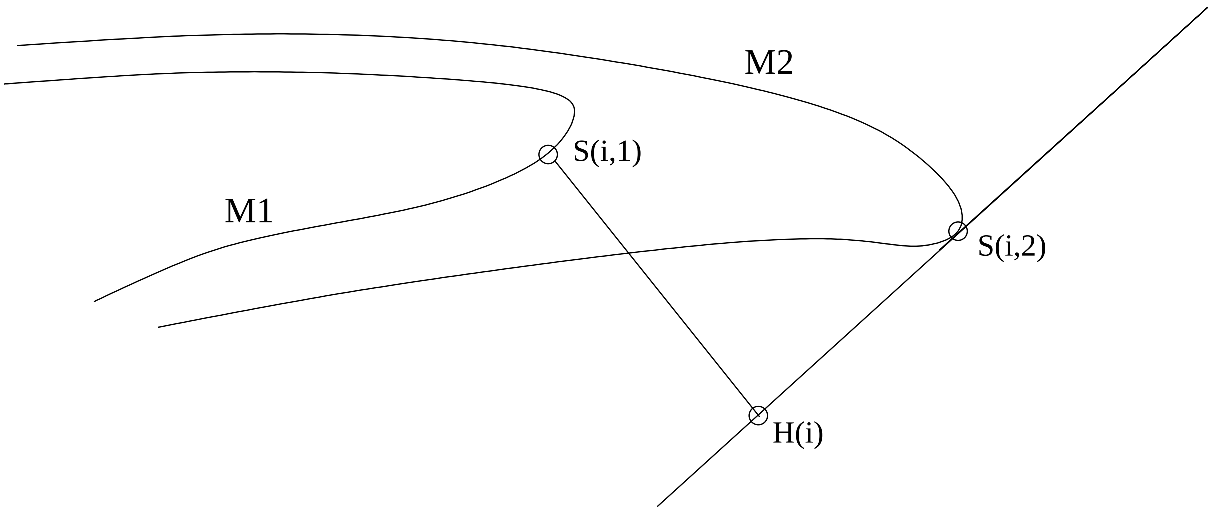


FIG. 8.1 –

Bibliographie

[Com94] Pierre Comon. Independant component analysis, a new concept? In *Signal Processing*, volume 36, pages 287–314. April 1994.

[Del94] Hervé Delingette. *Thèse de Doctorat : Modélisation, déformation et reconnaissance d'objets tridimensionnels à l'aide de maillages simplexes.* PhD thesis, École Centrale Paris, juillet 1994.

[Dry] Ian Dryden. General shape and registration analysis. Technical report, Department of Statistics, University of Leeds.

[ELF00] Michael E.Leventon, W. Eric L.Grimson, and Olivier Faugeras. Statistical shape influence in geodesic active contours. *IEEE*, 2000.

[G.K89] David G.Kendall. A survey of the statistical theory of shape. In *Statistical Science*, volume 4, pages 87–120. 1989.

[HBS99] Johannes Hug, Christian Brechbühler, and Gábor Szélkely. Tamed snake: A particle system for robust semi-automatic segmentation. Technical report, Swiss Federal Institute of Technology, March 1999.

[HO99] Aapo Hyvärinen and Erkki Oja. Independant component analysis: A tutorial. Technical report, Helsinki University of Technology, Laboratory of Computer and Information Science, April 1999.

[Hug00] Johannes Hug. *Semi-Automatic Segmentation of Medical Imagery.* PhD thesis, Swiss Federal Institute of Technology Zurich, 2000.

[ICA98] *Independant Component Analysis, Theory and Applications.* Kluwer Academic Publishers, 1998.

[LK00] Cristian Lorentz and Nils Krahnstöver. Generation of point-based 3d statistical shape models for anatomical objects. *Computer Vision and Image Understanding*, 77:175–191, 2000.

[MD98] Johan Montagnat and Hervé Delingette. Globally constrained deformable models for 3d objects reconstruction. *Signal Processing*, 71:173–186, 1998.

[Mon99] Johan Montagnat. *Thèse de Doctorat : Modèles déformables pour la segmentation et la modélisation d'images médicales 3D et 4D.* PhD thesis, Université de Nice Sophia-Antipolis, décembre 1999.

[Pen96a] Xavier Pennec. Multiple registration and mean rigid shapes - application to the 3d case. Technical report, INRIA, 1996.

[Pen96b] Xavier Pennec. *Thèse de Doctorat : L'incertitude dans les problèmes de reconnaissance et de recalage - Application en imagerie médicale et biologie moléculaire.* PhD thesis, École Polytechnique, 1996.

[Pen00] Xavier Pennec. From noise hypotheses to registration criterions. Technical report, INRIA, 2000.

[TCDJ95] T.F.Cootes, C.J.Taylor, D.H.Cooper, and J.Graham. Active shape models - their training and application. In *Computer Vision and Image Understanding*, volume 61, pages 38–59. Janunary 1995.

[You] Laurent Younes. Modèles statistiques d'images, cours du dea mathématiques, vision, apprentissage.

[Zie94] Herbert Ziezold. Mean figures and mean shapes applied to biological figure and shape distributions in the plane. In *Biom. J.*, volume 36, pages 491–510. 1994.