

Modélisation et analyse automatisée de la radiographie de la main de l'enfant

P. Chassagnet

29 juillet 2010

Résumé : Le but de cette étude est une analyse entièrement automatisée d'une radiographie des os de la main d'un enfant. La forme et les dimensions des régions ossifiées de la main sont des éléments de diagnostic pour un certain nombre de pathologies de la croissance. La mise au point d'un tel outil permettrait une nouvelle approche de l'évaluation de la maturation osseuse, de la mesure de la minéralisation osseuse et de divers paramètres biométriques. Ce document fait le point sur nos travaux sur l'analyse de radiographies de la main de l'enfant, réalisés en collaboration avec le Service de Radiologie Pédiatrique de l'Hôpital Robert Debré [HSN⁺97]. Il vient largement compléter un rapport [Cha96] qui faisait la synthèse des premiers travaux réalisés à ce sujet [El 93, CGE94, Fuc94, Nit96, Sai96, BC98, Tud99, CNHS99, Ste00].

Notre problème était de pouvoir localiser et identifier les régions osseuses qui apparaissent sur une radiographie pour ensuite procéder aux mesures de distance, de surface ou de densité requises pour un diagnostic. Nous ne cherchons pas une reconnaissance explicite des formes. Pour les applications envisagées, quelques mesures élémentaires suffisent pour caractériser les formes d'une manière suffisamment précise pour un diagnostic. Dans cette approche, l'identification et la localisation précise de certains éléments de contours restent essentielles. Une détection des contours par le gradient, sans autre information que celles issues localement de l'image, nous semble la méthode la plus efficace et la plus neutre, en première approche.

La difficulté provient essentiellement des cas de juxtapositions et de superpositions des régions osseuses qui créent diverses ambiguïtés de détection ou d'identification. Dans de tels cas, les méthodes classiques de traitement d'images ne donnent pas de bons résultats. Pour résoudre les ambiguïtés, nous devons introduire des connaissances a priori sur la morphologie du squelette de la main. Cet apport d'information doit être pourtant minimal et suffisamment souple pour assurer une interprétation fidèle de certaines morphologies qui s'écartent sensiblement de la normalité.

Le problème est maintenant résolu de manière satisfaisante en ce qui concerne les os longs (métacarpiens et phalanges) [CNHS99]. La forme allongée des doigts est assez caractéristique pour permettre leur détection en injectant très peu d'information a priori. Une résolution hiérarchisée en plusieurs étapes s'impose alors. Une étape essentielle de ce processus est la construction des axes des doigts. Nous procédons par l'appariement de contours "gauches" et "droits", dans un premier temps relativement grossiers mais sûrs, ce qui permet de mettre en évidence des alignements de points médians qui formeront ces axes. Après une seconde phase, à échelle plus fine pour une meilleure localisation, on peut déterminer tous les contours et ainsi distinguer les régions osseuses successives le long d'un même axe.

À titre de validation, nous avons procédé sur une base d'environ 800 images. La fiabilité des résultats permet d'envisager un certain nombre d'applications diagnostiques, en particulier, la détermination de l'âge osseux et l'analyse d'anomalies morphométriques.

Cela nous permet également de disposer d'une bonne référence géométrique qui est requise pour détecter et identifier les os du carpe. Comme il s'agit de formes plutôt rondes, un modèle basé sur des symétries centrales s'avère bien approprié. On procède encore en deux phases. Certaines constructions géométriques permettent le placement a priori des centres de symétrie d'une manière relativement robuste vis-à-vis de la variabilité biométrique et des éventuelles erreurs sur la localisation des métacarpiens. Les contours ainsi détectés et identifiés permettent de recalculer les centres de symétrie sur l'image pour déterminer une bonne localisation des contours.

Nos résultats peuvent donc servir de base pour des applications médicales directes ou pour d'autres développements dans l'analyse des radiographies qui demanderaient un recalage précis sur certaines parties de l'image ou une adaptation selon le stade du développement osseux.

Ces travaux, notamment la constitution de la base expérimentale, ont été financés en grande partie par le Ministère de la Santé dans le cadre d'un Programme Hospitalier de Recherche Clinique.

Table des matières

Introduction	7
Un peu d'anatomie	7
Notre problème	11
Remarques sur la qualité des images	12
Historique	14
Plan	15
1 Analyse du problème	17
1.1 Difficultés	17
1.1.1 Des images complexes	18
1.1.2 Limites des techniques usuelles de segmentation des images	20
1.1.3 Limites des techniques de reconnaissance des formes	23
1.2 Les grandes lignes de notre méthode	27
1.2.1 Principes d'une approche quantitative	27
1.2.2 Notre méthode	28
1.3 Contexte d'une validation expérimentale	30
1.3.1 Critères de qualité	30
1.3.2 Base d'images de test	31
2 Détection, identification et localisation des os longs	33
2.1 Orientation préliminaire	34
2.1.1 Heuristiques d'orientation	35
2.1.2 Autres techniques d'orientation	35
2.2 Détection des contours	37
2.2.1 Caractérisation des contours	37
2.2.2 Calcul du gradient	37
2.2.3 Validation des pixels de contour	38
2.3 Construction des points médians	39
2.3.1 Détermination de l'échelle initiale	40
2.3.2 Seuillage des contours	41
2.3.3 Appariement des points de contours	43
2.3.4 Validation et adaptation de l'échelle de détection	46
2.4 Construction des axes des doigts	48
2.4.1 Chainage par connexité	48
2.4.2 Extension de la proximité et filtrage	50
2.4.3 Continuité de deux chaînes	52
2.4.4 Regroupement des chaînes	54
2.4.5 Algorithme de construction d'axes	55
2.4.6 Cas des axes coupés	57
2.4.7 Détection des axes des doigts	59

2.4.8	Exemple d'un cas plus difficile	61
2.5	Détermination de l'échelle de localisation	64
2.5.1	Objectif de précision	64
2.5.2	Influence du bruit	65
2.5.3	Erreur de caractérisation	68
2.5.4	Choix de l'échelle de localisation	69
2.6	Localisation et identification de chaque os	70
2.6.1	Construction des contours longitudinaux	71
2.6.2	Localisation et identification des axes	73
2.6.3	Problème de l'échelle pour la segmentation	74
2.6.4	Segmentation longitudinale d'un axe	75
2.7	Bilan, mise en œuvre et généralisation	79
2.7.1	Réglage des paramètres	79
2.7.2	Association avec d'autres techniques de traitements des images	82
2.7.3	Applications à d'autres contextes	82
3	Validation expérimentale	85
3.1	Validation de la chaîne de traitement	85
3.1.1	Robustesse de l'algorithme	85
3.1.2	Influence du nombre d'itérations	86
3.1.3	Erreurs de segmentation	87
3.1.4	Scores de segmentation	89
3.1.5	Premières conclusions	92
3.2	Influence des paramètres de normalité	93
3.2.1	Variations des paramètres	94
3.2.2	Analyse des résultats	96
3.2.3	Corrélations entre les erreurs de segmentation	99
3.2.4	Nouvelles conclusions	103
3.3	Validation des mesures sur un échantillon	103
3.4	Applications	108
3.4.1	Relation à l'âge osseux	108
3.4.2	Profil métacarpo-phalangien	110
3.4.3	Recalage et étude statistique de la forme des axes	111
3.4.4	Difficultés pour un modèle actif de forme des axes	115
3.5	Conclusion	117
4	Définition d'un modèle géométrique pour le carpe	119
4.1	Détection et identification des contours	120
4.1.1	Utilisation d'un point de référence pour la caractérisation des contours	120
4.1.2	Utilisation conjointe de plusieurs points de référence	122
4.1.3	Détails techniques	123
4.1.4	Remarques théoriques	124
4.1.5	Recalage d'un point de référence	125
4.2	Placement a priori des points de référence	128
4.2.1	Définition d'un modèle linéaire	129
4.2.2	Défauts d'un modèle linéaire	130
4.2.3	Solution non linéaire	131
4.3	Résultats et applications	133
4.4	Problèmes non résolus	135

Conclusions	137
A Détails techniques	141
A.1 Rappels sur la détection des contours	141
A.1.1 Caractérisation des contours	141
A.1.2 Calcul du gradient	142
A.1.3 Scale-space et analyse multi-échelle	144
A.2 Précisions sur l'algorithme d'appariement des contours	145
A.2.1 Caractérisation des contours “gauches” et “droits”	145
A.2.2 Contraintes d'appariement	145
A.2.3 Principe du balayage	146
A.2.4 Analyse des configurations d'appariement	147
A.2.5 Traitement des redondances d'un contour	148
A.2.6 Distance minimale entre pixels médians	148
A.2.7 Influence du bruit	149
A.3 Mise en évidence des symétries	149
A.3.1 Principe	150
A.3.2 Résultats	152
A.4 Précisions sur l'algorithme de construction des axes	157
A.4.1 Calcul de la droite de régression linéaire d'une chaîne	157
A.4.2 Définition plus rigoureuse de la “distance” entre deux chaînes	157
Bibliographie	159
Liste des figures	171
Liste des tableaux	175
B Illustrations en couleurs	177

Introduction

Un peu d'anatomie

Durant la croissance, la calcification des os se fait progressivement en se substituant au cartilage. Ceci s'observe facilement sur une radiographie de la main (voir figures 3 et 4) qui met en évidence les parties calcifiées alors que les cartilages, peu contrastés par rapport aux chairs, restent invisibles. D'une manière générale, notons aussi que l'ossification de la couche périphérique d'un os, sa zone corticale, est plus dense qu'à l'intérieur. L'absorption des rayons est donc plus importante lors d'une trajectoire tangentielle à travers cette zone corticale et c'est ce qui explique les "contours" blancs de l'os.

Du point de vue anatomique, la distinction est très nette entre les os longs et les os courts. En ce qui concerne la main, les os des doigts (les phalanges) et les métacarpiens sont des os longs tandis que le carpe est constitué d'os courts. Des différences se révèlent notamment dans leur mode de croissance, comme nous allons le voir ci-dessous.

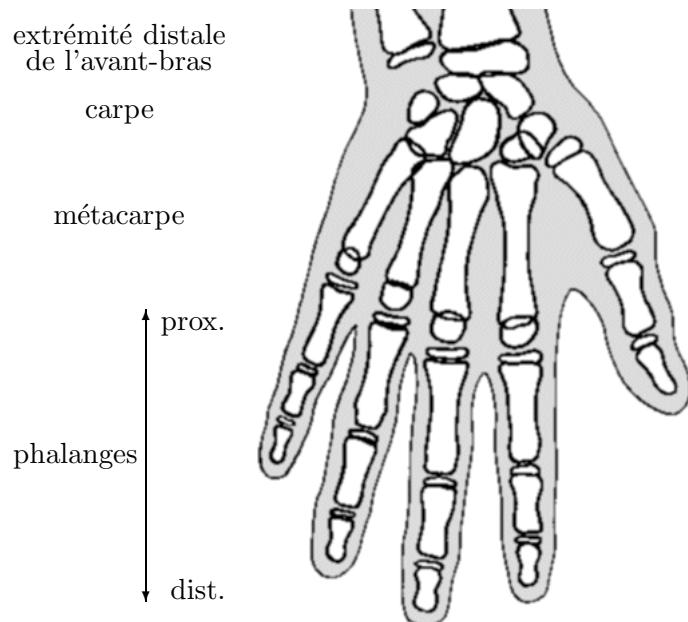


FIG. 1 – Les principales parties de la main et du poignet

Au passage, nous allons introduire un certain nombre de termes anatomiques que nous serons amenés à utiliser chaque fois qu'il s'agira de préciser une localisation.

Notamment, les deux qualificatifs *proximal* et *distal* sont employés pour situer un élément anatomique, relativement à un autre, le long d'un membre. Ils désignent alors, respectivement, le plus *proche* et le plus *distant* du tronc. On les utilise en particulier pour désigner les phalanges proximales et distales (voir figure 1).

Le cas du pouce est très particulier et, d'un point de vue anatomique, il serait plus juste de dire qu'il n'a pas de métacarpien. Nous suivrons pourtant l'usage qui veut que l'on désigne sa phalange proximale comme un métacarpien, et la manière de considérer les deux autres phalanges sera adaptée aux besoins du moment.

Métacarpiens et phalanges

La partie centrale des os longs, appelée la diaphyse, est déjà nettement visible à la naissance. Cette partie ossifiée s'allonge progressivement et on peut alors distinguer un évidemment tubulaire axial, le canal médullaire (voir figures 2 et 3). Les extrémités de la diaphyse qui ferment ce canal, sont appelées les métaphyses.

D'autre part, de nouveaux points d'ossification apparaissent aux extrémités de l'os. Ce sont les épiphyses qui se consolident rapidement pour former les articulations. On remarque qu'il y a une seule épiphysè par os, mais que deux épiphyses peuvent se trouver en vis-à-vis.

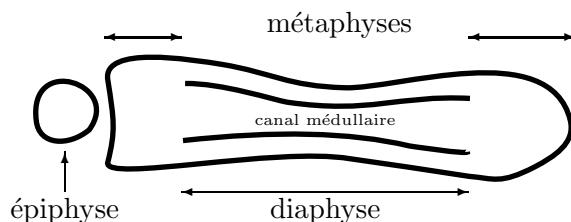


FIG. 2 – Désignation des parties d'un os long

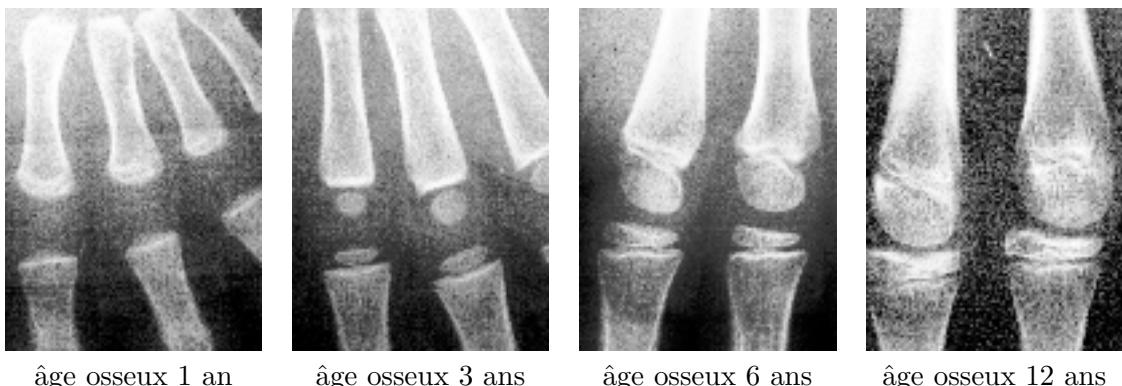


FIG. 3 – Exemples d'ossification des métacarpiens et phalanges proximales en fonction de l'âge

os	apparition	soudure
métacarpes	entre 2 et 3 ans	entre 16 et 18 ans
phalanges	entre 2 et 4 ans	entre 18 et 20 ans

TAB. 1 – Âges normaux d'ossification des épiphyses, selon [Del78]

Jusqu'à la maturité, les épiphyses restent séparées des métaphyses par un mince cartilage qui assure la croissance en longueur (voir figure 3, image de droite). Le tableau 1, tiré de [Del78], donne les âges normaux d'apparition et de soudure des épiphyses, selon cet auteur.

On voit que l'information apportée par la seule détection des épiphyses est relativement pauvre. Pour établir une mesure de la croissance, il est nécessaire de s'intéresser plus finement aux formes (et aux dimensions) des régions ossifiées. Ainsi considérés, les os longs de la main

constituent une source d'informations relativement progressives et fiables pour toutes les tranches d'âge. Notamment, les épiphyses des phalanges médianes et distales ont une évolution assez tardive et elles ne doivent pas être négligées [Kir84].

Os du carpe

À la naissance, les os du carpe sont encore invisibles mais leur ossification se développe assez rapidement ensuite (voir figure 4). Le tableau 2, tiré de [Del78] et [DL52], donne les âges normaux de début d'ossification des os du carpe selon ces auteurs.

Il est d'usage de traiter les épiphyses distales des os longs de l'avant-bras (le radius et le cubitus) comme des os du carpe. Au tout début de leur ossification, ces deux épiphyses sont appelées respectivement point radial et point cubital, par exemple le point radial est visible sur l'image de gauche de la figure 4.

On voit que l'information apportée par le carpe est progressive et potentiellement très riche en bas âge, à condition de s'intéresser à la forme et à l'extension de chacun. Mais, à partir de 7 ou 8 ans, il devient déjà délicat d'identifier clairement les formes de chaque os et l'ossification du carpe devient difficile à interpréter au-delà de 11-13 ans.

Dans le tableau 2, on constate facilement des discordances entre les deux sources. Il est établi qu'il n'existe pas d'ordre chronologique total, ce que l'on voit par exemple avec les quelques images de la figure 6. On en conclut notamment que le nombre d'os visibles dans le carpe n'est pas une mesure précise de l'âge osseux et que l'identification de chaque point d'ossification nécessite aussi sa localisation géométrique.



âge osseux 1 an âge osseux 5 ans âge osseux 14 ans

FIG. 4 – Exemples d'ossification du carpe en fonction de l'âge

os	Delmas	Debré et Lelong		
		min.	moy.	max.
grand os	entre 0 et 1 an	1 mois	6 mois	1 an
os crochu	entre 0 et 1 an	1 mois	6 mois	1 an
point radial	entre 1 et 2 ans	6 mois	1,5 an	2,5 ans
pyramidal	entre 1,5 et 4,5 ans	1 an	2,5 ans	4,5 ans
semi-lunaire	entre 2 et 6 ans	2 ans	4 ans	5,5 ans
trapèze	entre 3 et 6 ans	4 ans	5,5 ans	7,5 ans
trapézoïde	entre 3 et 6 ans	4,5 ans	6 ans	8 ans
scaphoïde	entre 5 et 6 ans	4 ans	5,5 ans	7,5 ans
point cubital	entre 6 et 9 ans	5 ans	7 ans	8,5 ans
pisiforme	entre 10 et 14 ans	9 ans	10,5 ans	12,5 ans

TAB. 2 – Âges normaux de début d'ossification du carpe, selon [Del78] et [DL52]

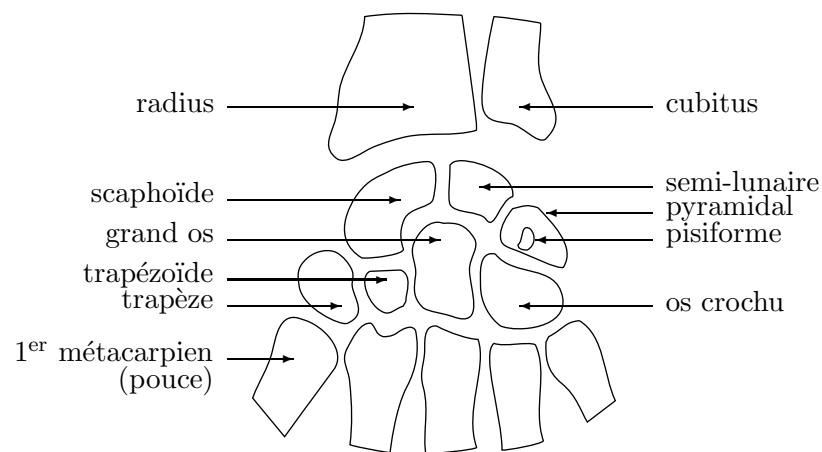


FIG. 5 – Désignation des os du carpe (vue schématique)

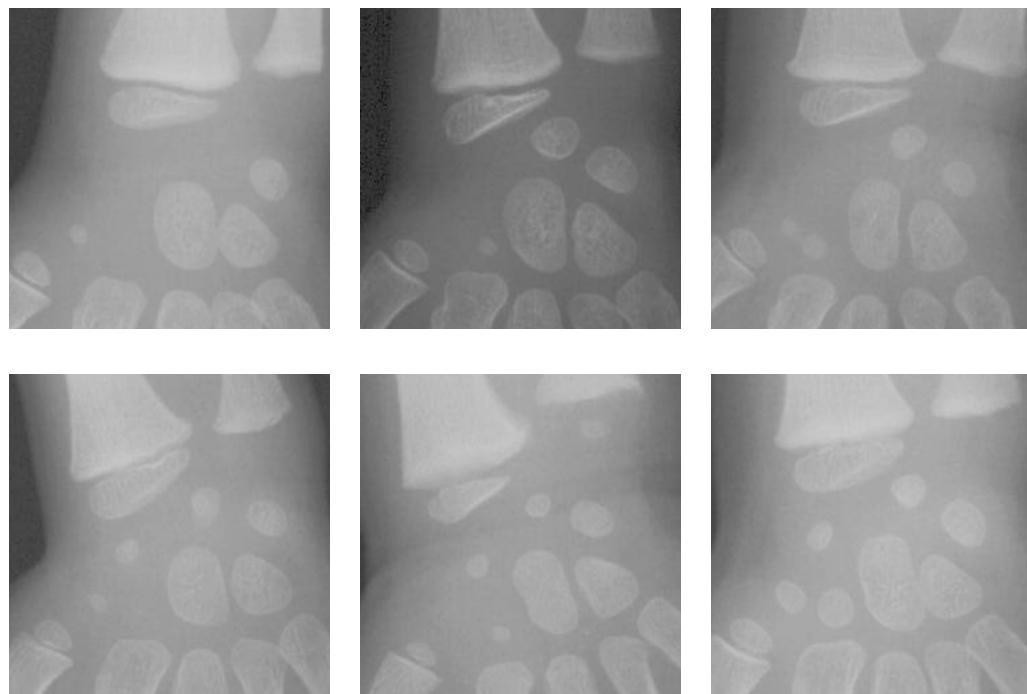


FIG. 6 – Exemples de variation chronologique, pour des âges osseux de l'ordre de 3 à 4 ans

Notre problème

Faire une radiographie de la main est un acte relativement facile et peu traumatisant, même chez l'enfant, et c'est un examen qui est souvent pratiqué. Comme nous venons de le voir, le squelette de la main est une structure complexe, potentiellement très riche en informations. En dehors des traumatismes, les principales indications cliniques incluent l'évaluation de l'âge osseux dans les dysfonctionnements de la croissance, l'analyse morphologique et biométrique pour aider au diagnostic des maladies osseuses constitutionnelles et, enfin, l'estimation de la minéralisation.

La connaissance de l'âge osseux est un élément de diagnostic important en pédiatrie et la détermination de l'âge osseux à partir d'une radiographie de la main est effectivement une technique médicale courante. La pratique la plus classique de cet examen consiste à comparer visuellement une radiographie de la main du jeune patient avec les images d'un catalogue de référence, comme, par exemple, l'atlas de Greulich et Pyle [GP83] et les méthodes FELS [RCT88] ou TW2 [TWC⁺83], plus récemment TW3 [THGC01]. On détecte ainsi des caractéristiques (présence, forme, dimensions relatives, ...) qui permettent de proposer une certaine tranche d'âge avec une certaine fiabilité. C'est un travail qui nécessite une grande expertise et du temps, surtout par la méthode TW2/TW3 qui est la plus précise mais aussi la plus complexe. Cette méthode consiste à examiner une vingtaine de régions d'intérêt dans la main et à affecter à chacune un stade de développement, par référence aux images du catalogue et avec l'aide de commentaires du style :

Discret aplatissement de la surface proximale de l'épiphyse du deuxième métacarpien.

À chaque stade de développement correspond un score et la combinaison des scores donne l'âge osseux par le biais de tables empiriques. L'interprétation reste donc assez subjective et diffère de manière parfois très sensible d'un radiologue à l'autre. On notera aussi que les différentes méthodes aboutissent à des estimations différentes par un même radiologue. Une méthode comme celle de Eklof et Ringertz [ER67] semble plus objective car elle repose sur des mesures, mais la prise de ces mesures est particulièrement fastidieuse et elle peut être la source d'erreurs. Enfin, l'adéquation statistique du catalogue et des tables de référence est toujours déterminante.

Nos travaux visent à développer une procédure d'analyse d'images entièrement automatique. Ils ont été réalisés en collaboration avec le Service de Radiologie Pédiatrique de l'Hôpital Robert Debré. Il s'agit de détecter, d'identifier et de localiser correctement les différentes régions osseuses de la main pour procéder de manière robuste aux mesures voulues de forme, de longueur, de surface, de densité, etc. Ces mesures pourront ensuite faire l'objet d'analyses statistiques et être utilisées dans un certain nombre d'applications diagnostiques [ER67, Poz74]. On peut citer en particulier la technique du profil métacarpo-phalangien [PGNG72, DV91, DV92] qui permet d'analyser les déviations par rapport à la normale des longueurs des 19 os longs et qui est un élément de diagnostic pour un grand nombre de pathologies [PG97].

Un objectif nettement plus ambitieux [HSN⁺97] est de déterminer automatiquement l'âge osseux avec une précision comparable à la pratique des radiologues, soit une marge d'erreur de l'ordre de quelques mois. Dans ces deux exemples, l'âge osseux et le profil métacarpo-phalangien, notre contribution concerne essentiellement le traitement numérique de l'image, en y incluant la prise des mesures requises. Cette partie se veut donc entièrement automatique. L'analyse diagnostique des mesures peut être ensuite plus ou moins automatisée. En pratique clinique, une intervention du praticien reste possible et souhaitable, pour valider les mesures, et la suite du processus diagnostique peut être aussi conçue comme interactive.

D'autre part, des cas pathologiques sont nécessaires pour valider la méthode ou pour cibler des pathologies particulières, mais ils sont parfois rares et il faudra utiliser des images d'archive, souvent de mauvaise qualité.

Un autre objectif (ou du moins une contrainte) est de pouvoir traiter un grand nombre d'images afin de constituer de nouvelles bases de référence plus opportunes, plus précises et plus facilement réactualisables que ne le sont les atlas existants. Par exemple, pour la détermination de l'âge osseux, si l'on retient les chiffres de 100 cas par tranche d'âge et 15 tranches d'âge pour chaque sexe, on a besoin de traiter 3000 images¹ pouvant être considérées comme "normales". Pour cela, des radiographies prescrites suite à un traumatisme accidentel constituent une source intéressante. Elles sont a priori décorrélées de toute pathologie mais les paramètres techniques, notamment le contraste et le cadrage, ne sont pas toujours bien adaptés aux exigences d'un traitement informatique.

Les méthodes développées doivent donc être à la fois rapides, de l'ordre de quelques secondes par image, et robustes vis-à-vis des artefacts, du bruit et des paramètres techniques qui affectent la lisibilité de la radiographie. Enfin, l'analyse doit être entièrement automatique, car il est exclu d'intervenir image par image en cours de traitement pour désigner certaines régions d'intérêt ou pour adapter des paramètres.

Remarques sur la qualité des images

La qualité d'une image radiographique est une notion assez subjective qui varie selon le besoin diagnostic et qui dépend d'un grand nombre de facteurs physiques [Spr]. Nous considérons ici quelques aspects quantifiables qui déterminent le bon fonctionnement de nos algorithmes.

Résolution

Intrinsèquement le film radiographique a une très bonne résolution qui serait largement supérieure aux besoins requis pour un diagnostic. L'écran scintillateur que l'on plaque sur le film et qui convertit les photons X en photons lumineux, eux-mêmes captés par le film, est beaucoup moins fin et il se produit une certaine dispersion, autrement dit du flou. Ce flou affecte la détection et la localisation des détails fins. La résolution effective est de l'ordre de 8 paires de lignes par mm, soit l'équivalent de 16 pixels/mm, soit environ 400 dpi (dots per inch). Les systèmes de radiographie numérique ont une résolution moindre, de l'ordre de 3 à 3.5 paires de lignes par mm, soit 150 à 200 dpi.

Contraste

Le contraste est un sujet techniquement très complexe, d'autant plus qu'il est souvent difficile de concilier contraste global et contraste local. La discréétisation des niveaux de gris introduit une perte irréversible de contraste. Le critère le plus important pour nous sera la différence de niveau de gris moyen entre deux régions adjacentes que l'on veut distinguer. C'est ce qu'on appelle communément le "signal" qui intervient surtout en valeur relative par le rapport signal sur bruit.

Bruit

En radiographie médicale, une des principales sources de bruit est la fluctuation quantique du rayonnement capté. Une explication simplifiée revient à considérer que la formation d'une image radiographique est équivalente au comptage du nombre de photons X incidents sur des petits éléments de surface pendant la durée d'exposition. Si le rayonnement incident moyen correspond à n photons par unité de surface, on peut considérer que le comptage suit une loi gaussienne (approximation normale d'une loi de Poisson) de moyenne n et d'écart-type \sqrt{n} .

¹C'est aussi l'ordre de grandeur de la base de référence pour TW2.

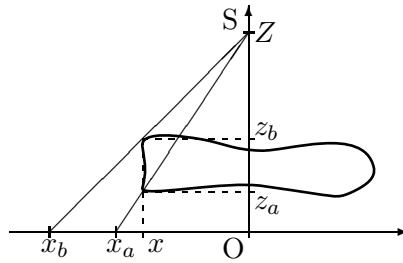


FIG. 7 – Principe projectif de la radiographie

Cela donne un niveau de bruit relatif de $\frac{1}{\sqrt{n}}$. Pour une incidence typique de l'ordre de 20 000 photons/mm² et une résolution équivalente à 72 dpi, soit 8 pixels (compteurs) par mm², cela donne 2 % de bruit.

Dans les zones blanches de l'image qui correspondent à une forte absorption, notamment à l'emplacement des os, n sera plus faible et le bruit relatif sera plus important. Un doublement de la résolution qui divise par 4 la surface d'un capteur élémentaire, revient à doubler l'erreur relative. Un quadruplement de la dose de rayons revient à diviser par deux l'erreur relative. Finalement, les images radiologiques sont toujours assez nettement bruitées car on cherche à limiter la dose. Notons aussi qu'à résolution et dose égales, les images de radiographie numérique directe sont moins sujettes à ce type de bruit, du fait d'un meilleur rendement (le comptage effectif ne porte pas sur n mais sur rn photons X, avec un rendement r compris entre 0 et 1, dépendant de la technologie).

Le bruit quantique est usuellement assimilé à un bruit blanc gaussien. En réalité, il n'est pas additif et il semble présenter une légère corrélation spatiale (texture de "moutonnement").

Les détails de structure interne des os se traduisent par des textures généralement bien visibles dans l'image. Un traitement adapté est difficile dans la mesure où la structure change très rapidement à l'intérieur d'un os, ce qui interdit une estimation statistique des caractéristiques de la texture résultante. À cela s'ajoute des bruits techniques, liés par exemple à la technologie des capteurs CCD qui sont utilisés pour la numérisation. Enfin, dans notre cas, on a un bruit inhérent à la compression JPEG. Ce bruit présente une certaine périodicité, mais il est totalement dépendant du contenu de l'image, donc inconnu a priori.

Le manque d'information sur les modalités d'acquisition interdit de corriger ces défauts a posteriori, notamment par soustraction. Tous ces phénomènes parasites ont en commun d'être d'amplitude modérée, avec une distribution relativement "compacte", et de constituer l'essentiel des hautes fréquences spatiales qui sont présentes dans l'image. Ces deux caractéristiques nous permettent de procéder à leur filtrage, sans affiner leur modélisation ni chercher à les distinguer. On utilise alors le modèle du bruit blanc gaussien additif pour lequel on dispose de résultats analytiques.

Déformations projectives et flou géométrique

En toute généralité, l'image radiographique est une projection perspective qui ne préserve pas les rapports de longueurs, de surfaces, etc. La figure 7 illustre le problème de la mesure des longueurs, avec un effet de perspective très exagéré.

Soit Z la distance de la source S par rapport au plan de formation de l'image (plan $z = 0$). Un point à la distance x de l'axe optique OS et à la distance z du plan de l'image, se projette sur l'image à une distance $\frac{Zx}{Z-z}$. Pour un objet épais, les points à la distance x et tels que $z_a \leq z \leq z_b$, vont se projeter entre $x_a = \frac{Zx}{Z-z_a}$ et $x_b = \frac{Zx}{Z-z_b}$.

Dans les conditions particulières de la main, d'épaisseur faible et posée à plat, presque au contact du plan de l'image, les valeurs extrêmes z_a et z_b sont faibles en regard de Z et $\frac{z}{Z}$ est de l'ordre de 1 à 2 %. La source S n'est pas ponctuelle, avec des dimensions de l'ordre du millimètre, mais le flou (analogique à un effet de pénombre) est négligeable en raison du rapport $\frac{z}{Z}$.

On peut considérer le développement limité au premier ordre : $\frac{Zx}{Z-z} = \frac{x}{1-\frac{z}{Z}} \approx (1 + \frac{z}{Z})x$. Si l'on cherche à estimer la position réelle x à partir de l'image, on peut considérer que l'on fait une surestimation systématique de $\frac{x_a+x_b}{2} - x \approx \frac{(z_a+z_b)x}{2Z}$, avec un flou de demi-largeur $\frac{|x_a-x_b|}{2} \approx \frac{(z_a-z_b)|x|}{2Z}$ qui s'ajoute aux autres flous de formation de l'image.

Ces valeurs sont relativement faibles mais on notera qu'elles sont proportionnelles à x , la distance à l'axe optique. Pour une radiographie de la main bien cadrée, on peut considérer que l'axe optique passe aux environs de l'extrémité distale du troisième métacarpien. Les mesures de longueurs des métacarpiens et des phalanges proximales sont alors entachées d'erreurs de l'ordre de 1 à 2 % en valeurs relatives. Pour les phalanges distales et le carpe, l'erreur absolue de position est multipliée par 2 environ, du fait de l'éloignement par rapport à l'axe optique, mais l'erreur relative sur des longueurs plus faibles peut approcher les 10 %.

On ne connaît pas les paramètres de la projection, ni des mensurations réelles de la main qui permettraient d'inverser le système, et une estimation par la mesure du flou $|x_a - x_b|$ serait trop imprécise pour être utile². En fait, si toutes les radiographies sont prises avec des cadrages assez similaires, on peut faire abstraction des mensurations réelles et considérer que les mesures faites sur plusieurs images sont comparables, à condition d'utiliser des rapports qui permettent d'éliminer les facteurs d'agrandissement.

Historique

Ce projet a démarré en 1993, en collaboration avec le Pr. Max Hassan, chef du Service de Radiologie Pédiatrique de l'Hôpital Robert Debré.

On disposait alors d'un jeu de 9 images, un peu représentatives de différents stades de croissance. Nos premiers travaux ont révélé les difficultés pour parvenir, même simplement sur ces quelques images, à une extraction automatique des contours par des caractérisations purement locales. Par ailleurs, des essais de mise en œuvre de contours actifs ont montré la nécessité d'une bonne localisation préalable pour assurer les initialisations et la paramétrisation qui sont déterminantes dans ce type de techniques. Toutes ces difficultés seront vues plus en détail au cours du chapitre 1.

Nous nous sommes donc rapidement orientés vers une méthode spécifique basée sur la construction des axes des doigts et le découpage os par os le long des axes. Ces premiers résultats sont résumés dans un rapport de 1994 [CGE94]. Cette méthode utilise du gradient dans l'image et un minimum de connaissances a priori sur la structure osseuse d'une main. De nombreuses améliorations seront ensuite apportées, mais cela donnait déjà les grandes lignes du principe que nous avons retenu, comme il est décrit au chapitre 2.

Sur ces résultats, nous avons constitué un dossier de PHRC (Programme Hospitalier de Recherche Clinique) du Ministère de la Santé, intitulé "Modélisation et lecture automatisée de la radiographie de la main de l'enfant". Cela nous a donné notamment les moyens de procéder, durant l'année 1996, à la saisie d'un grand nombre d'images et de constituer ainsi une base d'expérimentation de plusieurs centaines d'images, plus consistante d'un point de vue statistique mais aussi plus riche en cas particuliers. La première version robuste du code de traitement des métacarpiens et des phalanges et sa validation sur cette base d'images datent de 1998. La

²L'idée serait de mesurer l'aplatissement de courbes considérées comme des ellipses qui sont les projections de cercles perpendiculaires au plan de l'image. Par exemple, les extrémités des métaphyses seraient de tels cercles.

méthode et les résultats sont résumés dans [CNHS99]. La mesure de ces os permettaient d'envisager des applications diagnostiques et cela a fait aussi l'objet d'une communication purement médicale [HSN⁺97].

Dans le même temps, nous avons commencé à nous intéresser à un problème très différent, celui des os du carpe. Nous avons ainsi exploré diverses stratégies de coopération entre une technique croissance de régions et le calcul des contours [Nit96, Sai96] et, en particulier, en les intégrant dans un schéma multi-résolution. Il en ressortait que l'apport d'une approche par régions n'est pas avéré et que les résultats obtenus avec les contours seuls sont facilement aussi bons. La voie d'un modèle géométrique qui serait déterminé par les os longs [Fuc94], s'est montrée plus fructueuse. Ces résultats sont résumés dans [Cha96] et [BC98]. Ce modèle géométrique a été par la suite profondément remanié, pour aboutir à ce qui est exposé au chapitre 4, mais nous avons conservé le principe initial qui consiste à placer un point de référence par os.

Plan

Le chapitre 1 recense les difficultés de ce problème, avec référence aux travaux existants dans ce domaine, ce qui permet de situer les grandes lignes de notre méthode. Ce chapitre présente également la base d'images que nous avons constituée pour assurer une validation expérimentale.

Le chapitre 2 expose ensuite la méthode que nous avons mise au point pour la détection, l'identification et la localisation des contours des métacarpiens et des phalanges. Ce chapitre traite notamment de la détection des contours, de leur appariement pour la construction des points médians, de la mise en évidence des axes et, pour finir, de la segmentation le long de chaque axe. Un certain nombre de détails assez techniques qui concernent la mise en œuvre, sont regroupés dans l'annexe A. Pour des raisons pratiques, quelques illustrations relatives au chapitre 2 et qui nécessitent de la couleur sont placées à la fin du document, en annexe B.

Le chapitre 3 est consacré à une validation expérimentale de cette méthode. Sur une base de plusieurs centaines d'images, nous analysons l'efficacité de la méthode et sa sensibilité vis-à-vis des paramètres de normalité. Des mesures effectives, sur un échantillon d'une trentaine d'images, permettent de consolider cette analyse. Cette validation est convaincante pour les phalanges proximales et médianes, en fait pour la localisation de toutes les articulations des doigts, mais elle confirme une insuffisance du modèle pour ce qui concerne les extrémités proximales des métacarpiens et les difficultés inhérentes au manque de contraste pour les phalanges distales.

Enfin, le chapitre 4 explique comment nous pouvons ensuite exploiter la connaissance de la position des os longs pour calibrer un modèle géométrique du carpe. Ce modèle permet d'assurer une détection et une identification fiables dans une région de l'image dont le contenu est très variable. En retour, il permet d'améliorer la localisation des extrémités proximales des métacarpiens.

Chapitre 1

Analyse du problème

Le traitement plus ou moins automatisé de radiographies de la main a fait l'objet d'un nombre relativement important de travaux durant ces quinze dernières années. Certains de ces travaux [SVV90, Eff94, RCN⁺95, GBB95, PMGKH91, PMGHH92, PKKH93, Pie95, PH95, PGP⁺01, PPKGC03, ZGL07, MSC⁺00, MDDS01, dLGAAFAL02, AFdLGMFAL04, NvM⁺03, HJT07, HJCT07, KK07] visent explicitement la détermination de l'âge osseux, parfois sans même aborder les problèmes de traitement de l'image. D'autres [LH89, MN89, Rad93, MCRS93, MCRS94, dLGMAFAL03, MFMFAL03, MFMMMFAL05, LYZ04, HLP07] ne traitent que de la détection et de la localisation des régions osseuses. Ces travaux sont loin d'avoir réglé tous les problèmes. Les méthodes décrites ne sont généralement applicables que sur des images qui sont techniquement bien choisies et elles ne concernent le plus souvent que des mains jeunes. Sauf quelques exceptions relativement récentes [PGP⁺01, HJT07, KK07], on ne trouve pas de validation sur un nombre d'images qui soit probant au regard de la variabilité du problème.

Comme nous allons le voir en 1.1, la variabilité du squelette de la main rend très complexe la définition et l'utilisation d'un modèle détaillé. Nous nous sommes donc orientés vers la définition d'un modèle géométrique simplifié, permettant la réalisation des mesures appropriées à diverses applications et suffisamment général pour convenir à tous les âges. Ce modèle est conçu pour une résolution progressive, avec un apport minimal d'information a priori afin de ne pas forcer l'interprétation de la normalité. Nous introduisons les grandes lignes de cette méthode en 1.2 et elle sera détaillée dans les chapitres suivants.

Notre méthode est validée sur une base de plus de 800 images, de qualité très variée et concernant tous les âges. On trouvera en 1.3 une présentation de cette base, ainsi qu'une première discussion sur des critères pratiques de qualité des images.

1.1 Difficultés

Il est reconnu que les radiographies et notamment celles de la main posent de nombreux problèmes à toutes les techniques usuelles de traitement automatique d'images. Toute méthode de segmentation explicite, sans autre information que celles issues localement de l'image, se heurte à diverses difficultés d'identification des structures et demande finalement un apport d'information a priori, donc requiert un modèle. Une automatisation de la méthode "classique" TW2 [TWC⁺83] par des techniques de reconnaissance de formes conduit à un modèle d'une grande complexité, dans sa construction et dans son utilisation. Il faut tenir compte de la grande variabilité anatomique de la main, entre les individus, en fonction de l'âge et du fait que chaque point d'ossification possède une certaine indépendance de développement. La définition d'un modèle qui s'inspirerait de la croissance physiologique du squelette de la main est donc délicate.

Une approche basée sur la localisation des structures puis leur mesure, apparaît alors comme plus simple de mise en œuvre. Elle est aussi plus générale, puisqu'elle s'ouvre aux applications

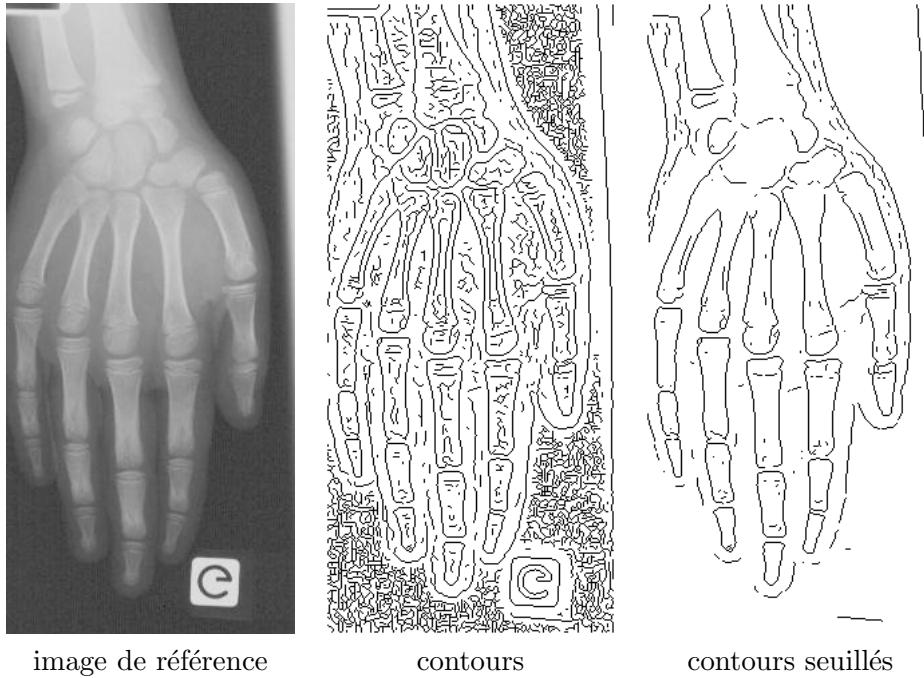


FIG. 1.1 – Les “vrais” contours ne peuvent pas être caractérisés seulement par un seuillage

autres que l’âge osseux [Poz74]. Cela revient à dire que l’on renonce ici à capturer une part de l’expertise du radiologue qui est mise en œuvre dans une méthode comme TW2 [AFdLGMFAL04, Zad99]. En fait, il s’agit de reporter cet apport d’expertise aux étapes dédiées au diagnostic.

1.1.1 Des images complexes

Un problème délicat, assez général en traitement d’images, est celui de la détermination des divers paramètres et valeurs de seuils. Cela devient encore plus délicat dans un contexte d’analyse automatique qui exclut toute intervention adaptée au cas par cas. Pour ce qui nous concerne, le problème du seuillage est notamment compliqué par une forte décroissance de la densité osseuse, en allant du carpe vers les phalanges distales. De plus, un certain nombre de particularités anatomiques se traduisent par de fortes variations d’intensité “parasites”, notamment :

- certaines régions osseuses ont une forme anguleuse, en particulier dans leur troisième dimension (l’épaisseur) avec localement des courbures importantes, ce qui conduit à des variations rapides de l’opacité aux rayons,
- dans le même ordre d’idées, les os longs et principalement les métacarpiens comportent un canal médullaire bien contrasté,
- enfin, l’ossification est plus importante au voisinage des articulations et, plus généralement, dans les zones soumises à de fortes contraintes mécaniques qui induisent cette ossification.

Ces “détails” internes sont souvent mieux contrastés que les “vraies” frontières des régions osseuses que l’on cherche à identifier. En fait, tous les problèmes évoqués ci-dessus vont en augmentant avec l’âge qui accentue et diversifie les variations de densité osseuse. Dans une classification usuelle entre signal et bruit, ces détails sont du signal et ils résistent aux techniques de filtrage et de seuillage qui sont par ailleurs relativement efficaces sur les bruits d’origine technique et sur les textures d’ossification, dans ce cas, assimilables à du bruit (voir figure 1.1). Il faut aussi noter que, lorsqu’ils sont correctement identifiés, ces détails internes peuvent devenir des éléments utiles pour la suite de l’analyse. C’est le cas par exemple pour le canal médullaire.

Une autre difficulté est due à la grande variabilité des paramètres techniques qui interviennent dans le processus radiographique et qui affectent l'image. Certains des travaux cités supposent une très bonne qualité d'images et surtout une grande homogénéité (images issues d'un même appareillage de radiographie et de numérisation, avec des réglages et un cadrage usuel).

C'est le cas notamment pour les premiers travaux de Pietka et al. [PMGKH91, PMGHH92, PKKH93]. Ceux-ci sont à l'origine d'une contribution relativement complète sur l'automatisation de l'âge osseux, mais le besoin du traitement d'images hétérogènes et les problèmes inhérents ne sont considérés que tardivement [PGP⁺01]. De manière encore plus restrictive, Michael et Nelson [MN89] supposent la séparation nette des points de l'image en trois classes d'intensités (fond, chair et cartilage, os). Il faut pour cela que l'image soit obtenue dans des conditions de contraste et de bruit qui limitent le "mélange" des classes. En particulier, l'acquisition des mesures statistiques (histogrammes) [SS04] qui déterminent localement les trois classes doit tenir compte des particularités de distribution de la densité d'ossification, d'une décroissance nette du carpe aux phalanges distales et de l'alternance périodique dans la direction transversale. Le principe même d'une telle classification ne s'applique pas à des mains jeunes qui sont trop faiblement contrastées. D'autres travaux [HJT07, dLGMFAAL03] utilisent aussi plus ou moins d'heuristiques faciles pour le seuillage et la pré-localisation, ce qui lie implicitement ces méthodes à des images de bonne qualité ou à des hypothèses de cadrage très particulières.

Enfin, il faut souligner la présence d'objets étrangers, à des emplacements très variables et avec des caractéristiques trop diverses pour être facilement détectés a priori. Dans le cas des radiographies de la main, on trouve notamment divers appareillages, perfusion, bijou, étiquette, marquage en plomb, bord du champ radiographique, etc. Ces objets produisent généralement des signaux très forts (en intensité, contraste, cohérence géométrique, ...) et leur présence à proximité de la zone d'intérêt va perturber gravement les opérations de mise en évidence. Sur des mains très jeunes, la cohérence géométrique de bords relativement rectilignes peut être un problème, même pour des objets non métalliques, a priori plus anodins, comme les bords des vêtements ou des bandages. Un remède possible est une cascade de prétraitements dédiés aux cas les plus typiques [MGPH92, PGP⁺01] qu'il est difficile de rendre exhaustive. Après avoir aussi expérimenté dans cette voie [Nit96, Sai96], nous avons constaté que l'on parvient aisément aux mêmes résultats par des contraintes structurelles, notamment l'appariement des contours.

Un aspect anatomique, également très problématique, est que certains os se touchent ou se recouvrent. Le phénomène est déjà très net pour les os du carpe vers 7 ou 8 ans. C'est aussi rapidement le cas entre les phalanges. Lorsque ces effets se combinent avec les autres problèmes mentionnés ci-dessus, même l'interprétation visuelle devient très délicate et elle requiert toute la connaissance anatomique d'un radiologue. Les techniques qui procèdent par une segmentation explicite, sans utiliser d'autres informations que celles qui sont issues localement de l'image, produisent alors des résultats incomplets, erronés, voire instables.

Une alternative, pour contourner les problèmes inhérents à une segmentation, est la méthode employée par Rucci et al. [RCN⁺95]. Les différentes étapes de l'analyse d'image sont réalisées par des réseaux de neurones dont la première couche est rétinienne. Deux réseaux de neurones séparés sont dédiés à chaque région d'intérêt TW2. Le premier réseau travaille sur toute l'image à échelle grossière et sert à localiser et orienter la région d'intérêt en question. Le deuxième réseau travaille sur cette région d'intérêt à échelle fine et produit un score qui correspond au stade de maturation TW2 pour cette région. Finalement, l'ensemble des scores est repris par un dernier réseau qui réalise la classification TW2 globale.

Les principaux inconvénients de cette dernière approche sont, d'une part, la lourdeur de la phase d'apprentissage qui est de plus dédiée à TW2, et, d'autre part, le manque de détection et de contrôle dans le traitement des cas anormaux. Mais on peut remarquer, le principe qui consiste à séparer le traitement de chaque région d'intérêt. Au-delà du simple fait qu'il soit ici dicté par

la méthodologie TW2, il y a aussi des considérations techniques. La relative indépendance de croissance des régions osseuses et les autres aspects de variabilité naturelle font qu'un traitement global prendrait une forte dimension combinatoire. D'autant plus que, pour une approche qui est basée sur un apprentissage, cette complexité se retrouve dans taille qui est requise pour la base d'apprentissage, et elle s'avère très pénalisante.

Nous pensons que le recours à une segmentation explicite qui permet d'identifier des régions d'intérêt et de localiser certains réglages, reste préférable en première approche. Il s'agit alors de déterminer une technique de segmentation, à la fois intrinsèquement robuste et aussi adaptée à coopérer avec un apport minimal d'information pour résoudre des ambiguïtés.

1.1.2 Limites des techniques usuelles de segmentation des images

Manos et al. [MCRS93, MCRS94] ont développé une approche par division et fusion de régions. Leur méthode combine des critères de cohérence de région avec des informations de contour. Les techniques de filtrages qu'ils utilisent, sont relativement sophistiquées (filtres de Canny [Can83, Can86] pour les contours et régularisation de l'image par une diffusion non linéaire [Wei97]) et un certain nombre d'heuristiques de fusion sont utilisées en complément. Néanmoins, les défauts principaux de l'approche par région subsistent. En particulier, lorsque plusieurs os se touchent ou se recouvrent, ils sont fusionnés en une même région et les règles qui sont alors utilisées pour les séparer sont très arbitraires. En fait, si le résultat obtenu est visuellement acceptable, il ne peut pas servir de base à des traitements ultérieurs et il conduirait à des mesures approximatives.

Nous avons procédé à une comparaison [BC98] de techniques basées, d'une part, sur les contours et, d'autre part, sur les régions. En règle générale, l'approche par régions [HP74, CL94, WH03] ne donne des résultats conformes que lorsque les contours sont faciles à détecter, c'est-à-dire lorsque les régions osseuses sont clairement séparées et relativement homogènes. Nous avons aussi vérifié [Fuc94] que des informations sur les contours sont utiles pour guider efficacement la construction des régions [PL90], mais sans trouver de réciproque effective. Enfin, des techniques de croissance de régions dirigées par une information a priori [AB94] ne sont pas utilisables dans un contexte où le nombre, la géométrie et les caractéristiques d'intensité des régions sont très variables.

Les techniques de segmentations orientées vers la détection des contours semblent donc les plus pertinentes. Serrat et al. [SVV90] ainsi que Radeva [Rad93] utilisent une caractérisation des contours par le laplacien, selon la méthode de Haralick [Har84]. Dans le cas d'images bruitées ou texturées, il est pourtant préférable d'utiliser des caractérisations par le gradient, par exemple, par la méthode de Canny [Can83, Can86]. Ces dernières sont beaucoup plus robustes, à condition d'utiliser de bons opérateurs de filtre, notamment à réponse infinie et à échelle ajustable (voir annexe A.1). Notons aussi que l'information donnée par le gradient comporte une intensité et une orientation, ce qui en fait une source relativement riche pour exprimer des règles de cohérence.

Pietka et al. [PMGKH91, PMGHH92, PKKH93] utilisent les filtres de Sobel, qui sont très rudimentaires, et les opérateurs de morphologie mathématique qui sont employés ensuite pour régulariser les contours, induisent des déformations notables. Des prétraitements incorporés ultérieurement [PGP⁺01], comme un filtre médian et un filtre de Lee basique [Lee80], vont contribuer à améliorer l'aspect visuel de l'image mais ne peuvent pas améliorer la localisation des contours calculés. Dans le même ordre d'idées, des filtres statistiques, même relativement élaborés [Lee81, KHEK76], relèvent toujours d'une approche très empirique de la caractérisation d'un contour. L'approche par diffusion non linéaire [Wei97, PM90, WB97] qui s'appuie sur la définition différentielle d'un contour, donne de meilleurs résultats [ZGL07] mais son principe itératif conduit à une résolution relativement lente et le paramétrage (contraste et régularisation) reste un problème délicat, sauf si on peut déjà localiser à des zones de caractéristiques connues.

D'autre part, les contours caractérisés par le gradient sont généralement incomplets et si, par construction, la caractérisation alternative par le laplacien fournit bien des contours fermés, ce ne sont pas nécessairement les "bons". On constate essentiellement deux types de défauts qui sont inhérents à la définition et au calcul d'un contour :

- Certaines régions sont trop proches pour que l'on puisse séparer leurs contours. Dans ce cas, la frontière naturelle entre deux régions n'est pas caractérisée comme une paire de contours mais plutôt comme une ligne de maxima de courbure, une "vallée". Le problème peut être partiellement corrigé en utilisant des règles de cohérence locale qui combinent les contours et les extrema de courbure, un sujet qui a été largement développé par Radeva [Rad93]. Mais, on doit adapter des règles à chaque région osseuse pour traduire ses particularités, ce qui suppose déjà leur identification et une bonne localisation.
- À la jonction de plusieurs contours, les grandeurs différentielles caractéristiques manquent de régularité, il en résulte une certaine indétermination dans le suivi d'un contour "principal". On n'obtient alors que des fragments de contours et il est difficile de distinguer correctement les fragments qui correspondent au contour d'un même os.

C'est fondamentalement un problème d'échelle [Can83, Can86, MH80, Wit83, Koe84, MZ92, Lin94], mais le choix de l'échelle est toujours un compromis entre la finesse de localisation et l'atténuation du bruit. Nous en avons expérimenté les limites sur nos images. On est souvent proche de la limite de résolution et il faudrait utiliser des images de meilleure qualité (résolution supérieure ou meilleur rapport signal/bruit). À partir d'un certain stade de maturation osseuse, ces défauts se manifestent couramment dans le carpe et au voisinage des épiphyses.

C'est pourquoi, il est relativement facile d'obtenir de bons contours sur des mains jeunes alors que la difficulté et le manque de pertinence d'une telle approche augmentent ensuite. En particulier, pour évaluer l'avancement de la soudure des épiphyses aux diaphyses qui caractérise les derniers stades de l'âge osseux, il est vain d'utiliser des contours. Pour cela, Pietka et al. ont recours à une analyse dédiée qui est basée sur une décomposition en ondelettes [PH95, PPKGC03]. Une telle approche serait encore plus pertinente si elle était orientée plutôt vers la caractérisation de singularités de divers ordres [MH92], au lieu d'une détection des maxima du gradient [MZ92]. Autrement, des techniques plus purement spectrales (DCT) donnent aussi des résultats exploitables [HJT07, KK07]. De nouveau, on remarquera que cela requiert au préalable une relativement bonne localisation de la zone d'intérêt, en position et en orientation, et aussi un apprentissage des "signatures spectrales" attendues. De ce point de vue, il y a finalement peu de différence entre de telles méthodes plus ou moins spectrales et une méthode purement "pixelique" qui procède à un recalage dense [NvM⁺03, MFMFAL03]. La différence se trouve alors dans la pertinence de tel ou tel critère de distance entre le motif analysé et celui qui est pris dans la base de référence.

Pour assurer une certaine régularité des contours et une cohérence dans leur fermeture, on peut envisager des techniques de contours actifs ou "snakes" [KWT87, MT96]. Ces techniques donnent de bons résultats lorsqu'elles sont utilisées de manière interactive ou alors dans des applications de suivi. L'initialisation est en effet déterminante, car il s'agit de se placer assez près du contour recherché et, d'autant plus près qu'il y a des contours concurrents dans le voisinage. Comme nous l'avons constaté [CGE94], une mise en œuvre automatique pose des problèmes, tant que l'on n'a pas localisé les formes avec précision. De plus, le réglage des différents paramètres qui assurent le contrôle de la convergence est délicat [Ste00, Eff94], en raison notamment de la grande variabilité d'une image à l'autre. Enfin, il faut souligner que ces techniques mettent en œuvre des méthodes itératives, relativement lentes et contrariées par la présence des contours parasites.

Ainsi, Efford [Eff94] procède à une pré-localisation de la main puis il initialise chaque contour actif par le recalage d'un modèle rigide simplifié. Cette méthode par localisation progressive ne

fonctionne correctement que lorsque les os sont bien séparés. Sur des mains plus âgées, on ne peut plus considérer chaque os de manière isolée et il semble nécessaire de traiter la main dans son ensemble dans un processus excessivement plus complexe.

Comme autre exemple, on peut également considérer les travaux réalisés par de Luis-García et al. [dLGMFAAL03] qui parviennent à placer un point à l'intérieur de chaque os long, après des constructions relativement compliquées. Ils peuvent ainsi initialiser un contour actif de type “ballon” [Coh91, CC93], bien adapté à des contours qui présentent des concavités. Pourtant, ces auteurs [dLGMFAAL03] rapportent beaucoup de difficultés à contrôler l'évolution, du fait des détails internes à l'os qui peuvent bloquer et aussi des variations de contraste le long des vrais contours qui peuvent “fuir”. Ces problèmes sont contournés par le recours à des essais successifs qui sont paramétrés en fonction d'un contrôle a posteriori de la cohérence du résultat.

En général, l'introduction d'un champ de gradient artificiel (Gradient Vector Flow) [XP98b, XP98a], obtenu par la diffusion d'un gradient “brut”, constitue une amélioration très significative du principe de contour actif. Cette technique permet en effet de régulariser l'intensité et la direction du vecteur gradient et d'étendre ainsi relativement loin la “zone d'attraction”, ce qui améliore nettement la convergence, notamment dans le cas de concavités et de discontinuités du contour. Indirectement, l'utilisation d'un GVF peut aussi simplifier l'initialisation car celle-ci devient moins critique. Par exemple, dans un contexte relativement propre, sur les os du carpe pour des mains jeunes, c'est-à-dire avec des os bien séparés et sans contour intérieur, Lin et al. [LXYZ04] obtiennent une bonne convergence à partir d'une initialisation assez grossière. On notera qu'ils utilisent en préliminaire un lissage par diffusion de type Perona et Malik [PM90], pour atténuer les textures internes à l'os.

L'initialisation reste pourtant délicate et déterminante en présence de parasites forts. On peut d'ailleurs formaliser des critères pour construire une initialisation qui pourra conduire à la convergence sur le “bon” contour. Il faut placer correctement [TBA05] le contour initial par rapport aux points de divergence du gradient utilisé, ce qui est effectivement d'autant plus simple que la carte de gradient est régulière, de part ou d'autre du contour cible. Quand ce n'est pas le cas, le placement du contour initial par rapport aux détails internes ou aux os voisins, reste un problème difficile. C'est ce que montrent, par exemple, les résultats obtenus sur des épiphyses par Han et al. [HLP07].

On peut noter que certaines des améliorations [CC93, XP98b] apportées au concept initial de contour actif [KWT87] s'éloignent d'une alternative à la détection des maxima du gradient (qui devient un préalable) et ne sont surtout qu'un substitut pour ce qui serait la phase de chaînage dans une détection de contours classique [Can83, Can86]. Ces techniques sont aussi relativement coûteuses car elles demandent de nombreuses itérations. Enfin, les difficultés de paramétrage restent plus délicates que le problème du seuillage. L'apport des contours actifs se limite ainsi à des contraintes de régularité qui permettent de contourner des ambiguïtés de chaînage, mais qui sont locales par définition.

Certaines variantes dans les schémas d'évolution (de type Level Set [CCCD93, MSV95, CKS97]) permettent d'éviter toute forme de description explicite du contour et de contrainte globale, notamment des aspects topologiques. Ceci ne semble pas souhaitable dans notre cas. Au contraire, le recours à une paramétrisation explicite particulière du contour actif semble préférable car elle permet de spécifier un peu la forme, ensuite de l'identifier facilement et même d'apporter directement une certaine mesure du stade de développement. De plus, ce type de modélisation paramétrique permet une résolution relativement rapide. Néanmoins, de telles descriptions de la forme restent trop générales pour permettre de diriger la convergence ou même de valider a posteriori. Nous reviendrons ci-dessous sur diverses approches plus spécifiquement dirigées par la forme.

En résumé, les techniques de contours actifs restent peu adaptées pour assurer la détection et la validation des contours, sur des images relativement complexes et variées, et d'autant

plus comme une première étape dans un système entièrement automatique. Une détection de contours “brute”, même imparfaite, nous semble ainsi préférable. Celle-ci a l'avantage d'être très rapide et de ne pas forcer de mauvaises interprétations. Ensuite seulement, on peut envisager une approche plus globale pour gérer des relations entre les différents fragments de contours et ainsi valider leur identification. Notre expérience [Cha96, CGE94, BC98] est qu'une détection des contours que l'on aura caractérisés par le gradient, pour la robustesse, et calculés par des filtres linéaires bien adaptés, pour l'efficacité et la précision, reste la méthode la plus efficace pour aboutir à une première segmentation. Il est ensuite relativement plus facile d'envisager la mise en œuvre de techniques spécifiques à certaines zones d'intérêt qui seront plus sophistiquées mais qui demandent une bonne localisation en préalable. Enfin, cette localisation par les contours constitue une base commune, à partir de laquelle on pourra dériver des applications dédiées à divers diagnostics.

1.1.3 Limites des techniques de reconnaissance des formes

En ce qui concerne le problème particulier de l'âge osseux, l'approche qui est généralement retenue est une automatisation de la méthode TW2 “classique” [TWC⁺83] par des techniques de classification de formes plus ou moins supervisées. Pour mettre en œuvre une telle méthode, il faut procéder dans un premier temps à l'apprentissage des formes de référence de TW2. Dans le cas de techniques supervisées, l'identification et l'annotation des formes caractéristiques pour l'apprentissage est une tâche complexe et, même pratiquée par un expert, elle est sujette à des ambiguïtés. L'absence de référence absolue et les variabilités d'interprétation d'un expert à l'autre amplifient encore cela. L'apprentissage est un processus qui est aussi très complexe du fait qu'il doit couvrir un grand nombre de paramètres de variabilité anatomique et technique d'une radiographie. De plus, une telle approche est souvent trop hermétique et complètement dédiée à l'âge osseux et elle ne peut pas être réutilisée. C'est le cas, notamment, pour une méthode basée sur des réseaux de neurones [RCN⁺95]. Une alternative plus ouverte consiste dans le recalage [DM98, MV98] avec des images d'un atlas de formes de référence, voir par exemple [NvM⁺03, MFMFAL03].

Dans le cas d'une modélisation globale du squelette de la main, ces techniques s'avèrent très complexes. La position relative et l'apparence des os varient en fonction des individus, de l'âge et aussi de la “pose” anatomique de la main. La variabilité du processus de maturation osseuse interdit aussi d'utiliser des hypothèses fortes pour lier géométriquement ou chronologiquement la reconnaissance des différents os (voir tableau 2 et figures 4 et 6 de l'introduction). Pour limiter le caractère combinatoire, il convient de séparer le traitement de chaque région d'intérêt. Toutefois, Serrat et al. [SVV90] introduisent le principe d'une résolution géométrique progressive, où la localisation de quelques os, dont la détection est la plus fiable, permettrait de prédire la localisation et une forme pour les autres os. Un contrôle *a posteriori* est plus facile à mettre en œuvre et ces mêmes auteurs utilisent plutôt les règles entrevues comme un système de cohérence pour valider les stades de maturation.

Comme autre exemple, Martín-Fernández et al. [MFMFAL03, MFMMMFAL05] procèdent à l'estimation d'un recalage global sur un petit nombre de points clés qui peuvent être considérés comme des invariants anatomiques. Cette approche est plus fiable que l'utilisation de primitives de l'image, obtenues par une caractérisation locale, même s'il existe des techniques robustes pour cela [SMB00]. En l'occurrence, ces points sont placés manuellement, ce qui contourne le besoin de traitements préliminaires complexes. Ce premier recalage permet d'initialiser un recalage dense (basé sur les valeurs des pixels), os par os et sur chaque image de l'atlas. Le stade de maturation effectif est ainsi déterminé par la reconnaissance de la forme mémorisée la plus approchante. Cette technique est donc coûteuse, d'autant plus qu'elle demande un atlas relativement exhaustif de toutes les variations possibles de la forme et de l'apparence de chaque os, selon des critères

à la fois techniques et anatomiques.

Dans une approche qui serait basée sur une quantification des formes, la différence entre deux formes peut se mesurer comme la distance entre deux vecteurs numériques, caractéristique chacun de l'une des formes considérées. Au-delà de l'intérêt algorithmique qui est procuré par des possibilités d'indexation, une telle approche semble aussi plus facile à concevoir [ER67, MSC⁺00] et à évaluer et aussi plus générale dans ses applications [PG97]. Dans une telle approche, le traitement de l'image peut être envisagé comme non supervisé. C'est-à-dire que, même si un apprentissage est requis, la quantification des formes se fait alors sur des critères purement géométriques ou statistiques. Le traitement de l'image peut être ainsi conçu comme une étape d'acquisition préliminaire et indépendante de l'application de diagnostic qui fera seule usage de l'expertise médicale, voir par exemple [GBB95, AFdLGMFAL04].

Une première voie qui reste basée sur les contours, est la description paramétrique des formes. On réalise ainsi un codage de la forme à analyser par un vecteur de paramètres qui la décrive géométriquement, et on aura fait de même pour des formes de référence. Par exemple, Serrat et al. [SVV90] utilisent des cubic-splines et Efford [Eff94] considère plutôt des descriptions de Fourier [SD92, SKBG96].

Dans cette catégorie, on peut également inventorier les nombreuses techniques de contours actifs paramétriques [MT96] et aussi toutes les techniques d'estimation des paramètres par une régression explicite sur les pixels d'un contour déjà calculé. On notera que ces techniques comportent des étapes de régularisation qui permettent de gommer les aléas locaux de traitement de l'image. La description reste très générale et permet de représenter toute forme qui présente une certaine régularité, mais c'est aussi une faiblesse. Il est possible d'apporter des contraintes de forme a priori par un modèle statistique [SD92, SKBG96]. Des contraintes a posteriori sont également faciles à vérifier. La taille du vecteur de paramètres ne dépend pas du nombre de pixels du contour, il est possible d'exprimer analytiquement diverses mesures de la forme à partir des paramètres et, de même, il est possible de définir une distance relativement pertinente entre deux formes.

Les formes de référence pour TW2 varient beaucoup en fonction de l'âge (voir figures 3 et 4, pages 8 et suivantes de l'introduction). La description d'une forme demande donc un nombre important de paramètres, pour supporter la variabilité requise et pour assurer une certaine régularité d'une forme à la suivante. Efford envisage ainsi d'utiliser jusqu'à 80 paramètres qui sont les coefficients requis pour coder 20 harmoniques [SD92], afin de décrire correctement un métacarpien d'âge avancé et il souligne donc la difficulté à mettre ainsi en œuvre la méthodologie TW2. La forme d'un os présente pourtant un nombre limité de singularités et on peut penser qu'une autre représentation, par ondelettes par exemple, permettrait de réduire le nombre de paramètres. Nous allons voir maintenant une technique beaucoup plus générale de décorrélation entre des paramètres de forme.

Les formes des os varient aussi beaucoup selon les populations et les individus et c'est ce qui contribue à la difficulté de TW2. Un apprentissage statistique de ces variations est possible par une régression sur une base d'exemples suffisamment large.

C'est ce qui est réalisé par la construction d'un modèle de distribution de points (Point Distribution Model) [CTCG95] où chaque forme est échantillonnée par la position d'un ensemble de points clés (les "landmarks") qui sont situés, usuellement, le long d'un contour, mais d'autres primitives [SMB00] peuvent aussi convenir. Ces points doivent être mis en correspondance, d'une image à l'autre, ce qui revient à leur numérotation. La phase d'apprentissage est alors fastidieuse, car elle demande de traiter un nombre important d'images et de faire sur chacune le marquage qui assure la mise en correspondance. Pour une application réelle, un traitement manuel est fastidieux et une automatisation est délicate. On notera que cela demande au préalable une détection et une identification fiables des primitives de l'image.

Les phases suivantes de la construction d'un modèle de points sont relativement bien décrites dans la littérature [CTCG95, DM98] et d'application relativement universelle. Il s'agit, d'une part, du recalage itératif des formes de référence sur le modèle en cours de construction (méthode dite de Procrustes [Goo91]), le recalage permet de séparer la "pose" (usuellement une similitude) et la forme, et, d'autre part, de l'analyse en composantes principales qui consiste à diagonaliser la matrice de covariance des positions des points clés après ce recalage. Une forme peut alors s'approximer par une combinaison linéaire des principaux vecteurs propres (ou modes dans la terminologie des PDM), avec des coefficients qui sont les écarts normalisés par rapport à la forme moyenne.

Efford [Eff94] a expérimenté un modèle de points pour un métacarpien et il retient de l'ordre de 9 à 16 paramètres (modes principaux) pour capturer correctement de 96 % à 99 % d'une forme¹. La base d'apprentissage, comporte 50 exemples avec 40 points et semble inclure des cas avec épiphys mais celle-ci est très mal restituée. D'autre part, pour une phalange proximale, Mahmoodi et al. [MSC⁺00] retiennent 35 modes pour capturer correctement 99 % de la forme, sur une base de 57 exemples avec 38 points. Dans les deux cas, les points clés sont placés en partie manuellement et en partie par interpolation, ce qui contourne une des principales difficultés. Ils sont ainsi assez régulièrement espacés, ce qui est assez mal adapté et pourrait expliquer le besoin d'un nombre important de modes pour aboutir à une modélisation correcte.

Il semble naturellement préférable d'avoir une plus grande densité de points clés au voisinage des métaphyses et des épiphyses, alors que ce n'est pas utile le long de la diaphyse. Pour un modèle dédié à une analyse TW2, Niemeijer et al. [NvM⁺03] ont expérimenté un modèle de points pour l'articulation entre les phalanges proximale et médiane. Ce modèle est défini par 11 points clés, déterminés manuellement et complétés par 70 points interpolés et il concentre justement environ un tiers des points sur l'épiphys. L'apprentissage a été réalisé sur 119 exemples qui ne couvraient que les 5 derniers stades TW2 et ce modèle semble insuffisant pour décrire assez finement les variations de formes significatives de ces différents stades.

La sélection automatique des points clés peut être vue comme la mise en correspondance partielle entre des ensembles de points plus vastes, en déterminant à la fois des appariements de points [HZ83] et une transformation géométrique qui les lie [Sto87, LW88]. C'est un problème qui est relativement classique [BM92, AK96] lorsqu'il s'agit de faire un recalage entre seulement deux ensembles de points, mais déjà plus compliqué lorsqu'il s'agit de le coupler avec la construction d'un modèle statistique [RCB97, CRZL04]. Le placement approprié d'un nombre minimal de points clés, est un problème également difficile et coûteux [BT00, DTC⁺02] qui peut être posé comme l'optimisation d'une mesure de la complexité globale du système (notion de Minimum Description Length [Ris78]). On perd alors de vue la réalité anatomique de la constitution d'un os et le résultat est très dépendant de la représentativité statistique de la base d'apprentissage.

Dans notre cas, la difficulté principale réside dans la construction d'un ensemble de points clés qui soit reproductible sur toutes les images. En effet, ces points doivent être identifiables et persistants à tous les stades de maturation. L'identification est délicate car il y a peu de singularités sur le contour d'un os, surtout en bas âge où les formes sont très "rondes", plus particulièrement pour le carpe. On peut augmenter le nombre de points clés par interpolation, mais cela demande tout de même un certain nombre de points de référence qui doivent être liés à des invariants anatomiques ou géométriques.

Le problème de l'invariance à travers les âges se pose notamment avec les épiphyses qui sont d'abord absentes, puis apparaissent isolées avant de se souder progressivement (voir figure 3 de l'introduction). Si on peut potentiellement traiter des parties manquantes, le problème est que celles-ci concerneraient systématiquement un élément essentiel de variabilité de la forme. Il semble ainsi nécessaire de distinguer au moins 3 modèles différents d'un os long pour prendre en

¹Plus exactement, ce pourcentage de la variance totale est porté par les valeurs propres associées à ces modes.

compte les topologies correspondantes. Pour une description plus fine, Niemeijer et al. [NvM⁺03] concluent que le nombre de modèles devrait approcher le nombre de stades TW2. On pourrait aussi expliciter que la croissance d'un os n'apparaît pas comme un allongement élastique mais comme une extension aux extrémités. De même, les figures 4 et 6 de l'introduction montrent qu'il faudrait distinguer un grand nombre de modèles si l'on veut traiter le carpe comme un ensemble.

Le recours à une analyse en composantes principales, plus exactement le fait de formuler la sélection des modes principaux comme un problème de moindres carrés, pose la question de la validité statistique et requiert que la distribution des formes soit gaussienne ou, en pratique une bonne approximation (unimodale et assez symétrique). Pour considérer les évolutions des os, il faut a priori une densité uniforme des exemples d'apprentissage selon les différents âges. Vu os par os, cela se traduit par des distributions de formes très déséquilibrées avec, en particulier, une accumulation sur la forme définitive lorsque celle-ci est atteinte rapidement. On est alors tenté de traiter le stade de maturation comme une dimension supplémentaire pour analyser les relations avec l'âge comme des corrélations. Mais la maturation de chaque os est une inconnue a priori et la multiplicité et la relative indépendance de développement de chaque région font que le problème est nettement plus complexe. Notons enfin qu'une analyse en composantes principales ne permet pas de modéliser correctement un ensemble de cas rares, propres à certaines pathologies. Au mieux, on pourra détecter la présence d'une anomalie si son écart par rapport au modèle gaussien est suffisamment significatif.

La résolution de l'apprentissage demande alors un partitionnement préalable, guidé par exemple par une maximisation de l'espérance [DLR77], en fonction d'un critère similaire à un stade de maturation. Compte tenu de la taille du problème, ceci n'est praticable que si l'on procède à une annotation des exemples, en apportant explicitement l'information essentielle de la maturation. On revient ainsi à une technique d'apprentissage supervisé, du style TW2, avec ses inconvénients pour la mise en œuvre. Pour rester dans le domaine linéaire qui est intéressant pour sa rapidité et son efficacité en grande dimension, une technique de type machine à vecteurs de support (Support Vector Machine [Vap95]) semble déjà mieux appropriée que l'analyse en composantes principales, dans la mesure où elle s'intéresse plus à établir la séparation entre des classes qu'à déterminer leurs distributions respectives. Si l'on veut rétablir la quantification statistique des formes, il convient de déterminer ensuite un modèle de distribution de points sur chaque classe, en supposant que celle-ci soit bien représentée dans la base d'apprentissage.

En résumé, les techniques de reconnaissance de forme sont mal adaptées à notre problème, principalement du fait qu'elles requièrent un apprentissage très complexe. Les lois apparentes qui expriment la distribution des formes en fonction de la maturation sont aussi très éloignées des hypothèses statistiques qui fondent les techniques usuelles d'apprentissage. Notamment, une classification explicite sur des critères de croissance apparaît comme un préliminaire nécessaire pour aiguiller sur plusieurs variantes du modèle. Compte tenu de tous les autres aspects de variabilité à y intégrer, cet apprentissage demande de traiter une base conséquente et il ne peut pas être raisonnablement déterminé manuellement. Un apprentissage automatique s'impose donc, mais il pose le problème de l'identification et de la localisation préalable des formes ainsi que l'évaluation d'un stade de maturation. Cela suppose un certain nombre de prétraitements qui reviennent finalement à une analyse déjà détaillée de l'image.

Quand ces problèmes sont correctement résolus, on peut envisager un apprentissage selon un protocole automatique. En aval de la phase d'apprentissage, c'est-à-dire pour le traitement de nouvelles images, les problèmes sont moins critiques, car le modèle peut guider correctement la reconnaissance. En particulier, l'initialisation devient moins déterminante.

On peut citer à ce propos les techniques de modèle de forme actif (Active Shape Model) de Cootes, Taylor et al. [CTCG95], avec en particulier des exemples d'application aux os de

la main [MSC⁺⁰⁰, NvM⁺⁰³]. Cette approche utilise un schéma d'optimisation itérative qui est assez similaire à celui d'un contour actif paramétrique, mais avec une différence essentielle qui est que les variations de forme explorées sont contraintes, autour de la forme moyenne et selon les modes principaux appris lors de la construction d'un modèle de distribution de points (PDM).

Le principe d'un modèle de forme actif est d'alterner deux étapes jusqu'à atteindre le critère de convergence. La première étape consiste à déplacer chaque point du modèle PDM dans l'image, autour de l'hypothèse de forme courante, de manière à optimiser le recalage local sur un profil d'intensité, de gradient, . . . , appris statistiquement lui aussi. Ce principe autorise plus de souplesse de caractérisation que par les champs de gradient qui sont utilisés pour des contours actifs, et cela permet généralement des évolutions plus rapides. La seconde étape consiste à appliquer les contraintes en projetant la forme résultante sur le sous-espace déterminé par les seules composantes principales du PDM qui sont considérées comme significatives. Ceci donne une nouvelle hypothèse de forme pour l'itération suivante. Dans notre cas, la difficulté est que la localisation de l'hypothèse de forme initiale doit être assez précise pour échapper aux nombreuses ambiguïtés. Il faut aussi pouvoir prendre en compte la grande variabilité des profils selon la maturation et certains facteurs techniques de la radiographie.

Dans ce cas, on peut envisager d'utiliser des modèles d'apparence [TP91] qui utilisent les niveaux de gris, ou plutôt des modèles qui combinent [Cov97, CET98] la forme (sur des portions de contours identifiables) et l'apparence (sur des zones bien localisées). Notons alors que l'apparence devrait être plutôt gérée comme un atlas [SVV90, NvM⁺⁰³] ou un mélange gaussien, car une loi unimodale ne peut pas modéliser correctement les différents stades de répartition spatiale de l'ossification.

Plus généralement, une question essentielle est celle de la pertinence d'utiliser un a priori de normalité pour guider la localisation des éléments dans l'image. Dans un contexte qui reste très complexe tant que l'on ne connaît pas le stade de croissance, une approche dirigée par la détection des primitives de l'image, avec des contraintes de structure assez légères, semble préférable. Néanmoins, une approche dirigée par un modèle d'apparence reste une technique complémentaire intéressante pour traiter les derniers stades de la croissance, quand l'interprétation des contours devient particulièrement délicate.

1.2 Les grandes lignes de notre méthode

1.2.1 Principes d'une approche quantitative

Ce que nous appelons une approche quantitative consiste à localiser les différentes régions osseuses puis à procéder à des mesures utiles pour l'application diagnostique envisagée. Nous voulons séparer la prise des mesures qui sera systématique, de la phase de diagnostic qui pourra interagir avec le praticien. Les mesures considérées ici seront relativement élémentaires, à support continu (du moins en théorie) ou faiblement discrétilisé. Ainsi, Gross et Boone [GBB95] ont utilisé de manière probante un réseau de neurones pour calculer un âge osseux à partir de mesures quantitatives (manuelles). Si on limite l'usage des techniques de classification automatique à une analyse de mesures synthétiques, on aboutit à un protocole d'apprentissage réaliste. Ce principe s'applique à diverses techniques alternatives [ZGL07, AFdLGMFAL04, HJT07, KK07].

Comme on l'a vu ci-dessus, certaines méthodes basées sur TW2 peuvent entrer dans ce cadre, en considérant qu'elles reposent sur la mesure des formes [SVV90, Eff94, MSC⁺⁰⁰, AFdLGMFAL04]. Par contre, un stade de maturation TW2 n'est pas une mesure car elle serait à la fois subjective et trop grossièrement discrétilisée. En fait, la reconnaissance automatique d'un stade TW2 peut être vue comme étant déjà le résultat d'une classification sur un vecteur de mesures élémentaires.

Notons qu'une telle approche purement quantitative serait fastidieuse si les mesures devaient être effectivement réalisées par le praticien. De plus, certains types de mesure sont de nature beaucoup plus calculatoire que le cas d'une longueur. Citons par exemple, une mesure d'aire ou, de manière encore plus sophistiquée, une signature spectrale (ou d'ondelettes) qui peut quantifier l'avancement d'une soudure épiphysaire. Un autre exemple est la mesure de la concavité des bords longitudinaux d'un os long qui est relativement bien corrélée avec la maturation. C'est une mesure apparemment facile mais qui manque de précision si elle est faite ponctuellement. Des expériences [MSC⁺00, Eff94] laissent voir que cette concavité est bien décrite comme un mode principal dans un modèle de distribution de points ou par les premiers harmoniques d'une description de Fourier. L'automatisation des mesures est alors bien indispensable et elle est un objectif raisonnable, d'autant plus que l'on peut revendiquer une certaine fiabilité.

Les travaux de Pietka et al. constituent l'une des contributions les plus développées dans le sens d'une approche quantitative. Ces auteurs traitent des mesures des doigts [PMGKH91, PMGHH92] et du carpe [PKKH93], de la validation des résultats et de leur exploitation par de la logique floue pour en déduire un âge osseux [Pie95, PPKGC03]. Malheureusement, on l'a déjà remarqué, le traitement en amont est discutable et l'analyse ne traite que quelques mesures, faites sur les os dont la détection est la plus robuste. Malgré cela, le nombre de mauvaises mesures est relativement important.

Au-delà des critères de précision intrinsèque des techniques d'extraction et de localisation des primitives, il est utile de procéder à des mesures redondantes, en considérant qu'il existe certaines corrélations naturelles entre les grandeurs mesurées alors que les erreurs de mesure sont indépendantes. Il faut alors pouvoir effectivement accéder à toutes les mesures envisageables et la méthode doit aussi être robuste vis-à-vis des erreurs d'identification des structures qui conduisent à des mesures fausses. Cela s'appliquera à échelle fine, par exemple pour la prise d'une mesure sur tous les pixels d'un contour, ou sur tous les os, dans une régression particulière à l'application.

Dans une approche beaucoup plus globale, Levitt et Hedgcock [LH89] mettent en œuvre un modèle très développé où les mesures élémentaires sont validées par un système d'inférence bayésienne établi à partir de statistiques biométriques. Ces travaux ne concernent que des mains d'adultes et ils sont difficilement généralisable à un contexte pédiatrique. En effet, le modèle statistique devrait prendre en compte l'indépendance d'évolution de chaque noyau d'ossification et cela augmenterait considérablement la complexité d'un tel modèle et donc sa construction et son utilisation. Enfin, il est à craindre qu'une telle méthode qui est trop nettement dirigée par le modèle biométrique ne force l'interprétation de la normalité.

Nous aurons l'occasion d'illustrer ces problèmes au chapitre 3 et de voir la difficulté à décider automatiquement entre la mesure correcte d'une structure anormale et la fausse mesure due à une erreur d'identification. Nous verrons aussi cela en 4.2 à propos de la définition d'un modèle géométrique du carpe. Cette ambiguïté d'interprétation de la normalité impose une certaine prudence envers l'idée d'un diagnostic entièrement automatique et le principe d'une intervention du praticien reste donc souhaitable. On peut l'envisager comme une dernière étape de validation ou sous une forme plus interactive permettant de pondérer la pertinence des mesures.

1.2.2 Notre méthode

Nous nous sommes orientés vers la définition d'un modèle simple, dans lequel une définition précise des formes n'est pas requise, car nous voulons réaliser des mesures quantitatives, plus élémentaires. Du point de vue anatomique, les os longs, métacarpiens et phalanges, constituent une source d'information relativement pertinente pour toutes les tranches d'âge et leurs mesures suffisent déjà à de nombreuses applications diagnostiques. Notre méthode permet d'effectuer ces mesures de manière robuste, car elles seront peu affectées par la fragmentation des contours, par les erreurs d'interprétation, ou encore par les contraintes du modèle.

Dans cette approche, l'identification et la localisation précise de certains éléments de contour restent essentielles. Nous avons vu que ce sont des préliminaires indispensables, avec souvent aussi une estimation du stade de maturation, pour la mise en œuvre de diverses techniques d'analyse plus sophistiquées qui demandent une bonne initialisation et parfois un apprentissage automatique. Une détection des contours par le gradient, sans autre information que celles issues localement de l'image, nous semble alors la méthode la plus efficace et la plus neutre, en première approche.

Il faut ensuite introduire des informations supplémentaires pour résoudre les ambiguïtés. Notre modèle ne cherche donc pas à décrire les formes et il n'introduit que l'information qui s'avère strictement nécessaire pour valider l'identification des contours. Une résolution hiérarchisée en plusieurs étapes s'impose naturellement. La localisation préalable d'une entité englobante permet d'adapter des critères pour des zones restreintes de l'image, ce qui autorise des caractérisations simples et malgré tout sans ambiguïté. Par exemple, les contours de l'os d'une phalange ne sont détectés et localisés avec précision qu'après avoir localisé le doigt, défini en première approximation par son axe et sa largeur.

La forme allongée des doigts est suffisamment caractéristique pour permettre leur détection en injectant très peu d'information a priori. Une étape essentielle de ce processus est la construction des axes des doigts. Une caractérisation directe, par exemple par l'analyse des niveaux de gris, ne serait pas aisée. On remarquera en particulier que, le long d'un axe, les intensités sont élevées (une "bosse") quand l'os est "plein" mais relativement plus faibles (un "creux") si le canal médullaire est perceptible.

Nous procédons par l'appariement de contours caractérisés comme "gauches" et "droits". Pour commencer, il s'agit de contours relativement grossiers, mais sûrs, ce qui permet de mettre en évidence des alignements de points médians qui seront des points d'axes potentiels. Après une seconde phase, à échelle plus fine pour une meilleure localisation, on peut déterminer tous les contours et ainsi distinguer les régions osseuses successives d'un même doigt.

En ce qui concerne le carpe, dès l'âge de 7 ou 8 ans, il faut des informations particulières pour résoudre les ambiguïtés dans l'affectation des fragments de contours aux différentes régions osseuses. La mise en œuvre de ces informations demande une localisation précise et nous utilisons pour cela un certain nombre de constructions géométriques. Un modèle robuste est ainsi recalé sur l'image grâce à la localisation préalable des os des doigts.

La méthode que nous utilisons pour les doigts pourrait s'appliquer aussi au radius et au cubitus. Mais, pour une bonne détection des axes de ces os, il faudrait un cadrage relativement large de la main, ce qui est préjudiciable par ailleurs. Notre méthode ne requiert donc pas la détection de ces deux axes. D'autre part, les épiphyses radiales et cubitales qui se sont pas suffisamment déterminées par la position des doigts ne peuvent pas non plus être traitées comme des composantes du carpe.

En suivant cette démarche, la suite des opérations requises pour la détection, l'identification et la localisation des métacarpiens et des phalanges, est décrite plus en détail dans le chapitre 2. En complément, à l'annexe A, sont regroupées un certain nombre de précisions techniques sur les justifications mathématiques et sur la mise en œuvre. Le chapitre 3 analyse les résultats expérimentaux que nous avons obtenus sur une base de plusieurs centaines d'images. Celle-ci sera présentée à la fin du présent chapitre. La validation expérimentale nous permettra de vérifier l'efficacité de la méthode, d'en noter les limites et aussi d'esquisser une méthode d'estimation de l'âge osseux. Le chapitre 4 traite de la définition et de l'utilisation d'un modèle géométrique pour le carpe qui est construit à partir des résultats acquis sur les os longs. Ce modèle permet la détection, l'identification et la localisation des os du carpe et cela permet ensuite de corriger rétroactivement certaines erreurs constatées sur les métacarpiens.

1.3 Contexte d'une validation expérimentale

1.3.1 Critères de qualité

Nous faisons ici un inventaire des défauts de réalisation des images qui sont susceptibles d'empêcher un traitement correct. Il s'agit de définir les critères de qualité, purement techniques, qui seront requis pour assurer une interprétation automatique complète et fiable. Ces critères peuvent être facilement satisfaits, par quelques précautions lors de l'acquisition des images, et leur respect doit éviter de réaliser une radiographie inutilisable. De telles recommandations sont tout à fait comparables à celles énoncées pour une méthode classique car elles concernent aussi la lecture "visuelle" de l'image.

1. Les paramètres techniques de radiographie et de numérisation doivent déterminer un bon contraste de l'image. En particulier, la présence d'éléments absorbants (le plus souvent métalliques, par exemple, bijou ou aiguille de perfusion) peut conduire à un tassement exagéré de l'échelle de gris disponible sur les détails anatomiques de la main qui devraient utiliser l'essentiel de l'échelle des gris. À l'inverse, une saturation sur les os de la main conduit à la déformation ou à la perte de certains contours.
2. Même si leur intensité n'altère pas le contraste, les objets parasites créent des contours forts qui peuvent masquer, "dévier" ou même "annuler" les contours des os. Leur effet n'est pas prévisible et il n'est généralement pas possible de retrouver la forme osseuse qui est affectée. Pour des mains très jeunes, très faiblement calcifiées, cette remarque concerne aussi les produits tissés, bords des manches, bandages, ..., et matières synthétiques qui créent des signaux d'intensité comparable aux structures osseuses.
- En règle générale, les objets situés sur la main même, ou à proximité, ne sont donc pas acceptables. Les objets périphériques au champ de l'image posent beaucoup moins de problèmes, à condition qu'ils ne débordent pas trop vers l'intérieur.
3. Dans les toutes premières étapes du traitement, les dimensions de l'image sont considérées *a priori* comme représentatives de celles de la main. Elles servent à déterminer un certain nombre de paramètres et aussi l'orientation générale. Notre méthode suppose que les doigts, y compris le pouce, sont sensiblement orientés selon la plus grande dimension de l'image, au pire, à $\pm \frac{\pi}{4}$. Les doigts ne doivent donc pas être trop écartés et l'image doit être relativement bien cadree.

On a vu dans l'introduction que la normalisation du cadrage était également le moyen de faire abstraction des déformations perspectives. Un relâchement sur cette contrainte nous permet d'apprécier la robustesse de l'algorithme vis-à-vis des variations morphométriques. Dans le cadre d'une application diagnostique, il s'agirait cependant d'être plus vigilant.

Le respect des critères énumérés ci-dessus ressort comme "généralement nécessaire et suffisant *a priori*" pour obtenir une bonne analyse des images, comme nous le verrons aux chapitres 2 et 3. Remarquons qu'il ne s'agit que de formulations qualitatives qui permettent d'apprécier visuellement la qualité des images. Les défauts éventuels et leur influence respective dans la dégradation des résultats sont difficilement quantifiables et ils co-agissent de manière complexe. Il n'est donc pas possible de quantifier un seuil limite pour ces critères. En général, des petits défauts vont conduire à des petites erreurs locales et donner des résultats moins précis. Des défauts plus importants ou plus nombreux peuvent mettre en échec l'identification des structures.

On notera que le bruit ne fait pas partie des critères de qualité. Comme nous l'avons dit en introduction, toutes nos images sont relativement bruitées et un certain niveau de bruit doit être considéré comme inévitable. De plus, un niveau de bruit est très difficile à apprécier visuellement². En fait, l'influence sur la qualité du résultat s'exprime au travers du rapport

²Et même automatiquement, car la mesure du bruit suppose de le séparer du signal.

qualité visuelle	complet	dossier incomplet	inexistant	total
bonne	512 images	70 images	19 images	601 images
intermédiaire	126 images	18 images	2 images	146 images
mauvaise	84 images	21 images	9 images	114 images
total	722 images	109 images	30 images	861 images
	695 dossiers	107 dossiers	0 dossier	802 dossiers

TAB. 1.1 – Répartition par qualité d'image et de dossier

signal/bruit, dont le contrôle repose sur celui du contraste, un élément capital du savoir-faire radiographique.

Toutefois, au fur et à mesure du traitement, il devient possible de réaliser des mesures techniques sur la qualité d'une image, comme un rapport signal/bruit, ou sur la normalité des résultats obtenus. Ces mesures ne sont réalisables qu'après un certain avancement du processus d'identification et de localisation et elles ne peuvent que constituer des critères de confiance a posteriori.

1.3.2 Base d'images de test

Nous utilisons des images de résolution moyenne (de l'ordre de 600×1000), en 256 niveaux de gris, obtenues par numérisation, à 75 ou 150 dpi, de clichés conventionnels d'origines diverses. Il s'agit, soit du "tout venant" résultant de l'activité normale du service de radiologie de l'Hôpital Robert Debré, soit d'archives concernant des cas rares. Les images ont été compressées, d'un facteur de l'ordre de 10, par un algorithme JPEG intégré au système de numérisation. La perte d'information et l'ajout de bruit inhérents à cette compression n'ont lieu qu'une seule fois, à la saisie des images. Le résultat peut être considéré comme "sans perte visuellement perceptible" et il semble effectivement que la compression/décompression n'altère pas le diagnostic fait par un radiologue.

Au total, nous disposons de 861 images utilisables qui correspondent à 802 cas cliniques. Pour certains cas, plusieurs radiographies ont été numérisées ou la même radiographie a été numérisée plusieurs fois avec des paramètres différents. Ces images sont parfois mal contrastées, elles sont souvent relativement bruitées et comportent un certain nombre de marquages et autres artefacts périphériques inhérents à la réalisation du cliché ou à sa numérisation. Ces images sont donc de qualité très variable. Par une analyse visuelle rapide et selon les critères de qualité énoncés ci-dessus, nous les avons réparties arbitrairement en trois catégories :

- 601 images ont été déclarées de "bonne" qualité ; elles satisfont globalement ces critères.
- 146 images sont de qualité "intermédiaire" ; elles ne respectent pas tous les critères requis.
- 114 images sont de "mauvaise" qualité ; elles ne satisfont pas les critères les plus élémentaires. Parmi ces dernières, 8 ne sont d'ailleurs pas des images de main. Nous verrons au chapitre 3 qu'il y a une nette corrélation entre cette qualité d'image estimée a priori et les résultats que nous obtenons, ce qui justifie les exigences techniques énoncées ci-dessus.

En vue d'une application à la détermination de l'âge osseux, quelques données comme le sexe, l'âge effectif au moment de l'examen et l'âge déterminé par le radiologue doivent être associées à l'image. Pour certaines images, le dossier d'information est soit incomplet soit inexistant.

Une répartition par qualité d'image et de dossier est donnée dans le tableau 1.1. Au mieux, seules les images qui correspondent aux dossiers complets et que nous parvenons à traiter d'une manière décidée correcte, pourraient être réellement exploitées comme une base pour le développement d'une application diagnostique. De plus, la répartition par tranches d'âges est très inégale. Pour 50% des classes d'âge, nous ne disposons que de 3 images par classe, au plus.

Cette base est donc insuffisante pour le développement d'une application. Pour constituer une bonne référence statistique, un nombre beaucoup plus important d'images de bonne qualité serait nécessaire. Mais ces 861 images permettent déjà la validation de nos procédures de traitement et il est alors très utile de disposer de cas difficiles en nombre suffisant pour pouvoir étudier les divers problèmes techniques.

Chapitre 2

Détection, identification et localisation des os longs

Ce chapitre donne une description plus détaillée des techniques que nous avons développées pour la détection, l'identification et la localisation des phalanges et des métacarpiens. Ces opérations consistent à prendre en entrée l'image de la radiographie pour donner en sortie une description fine des contours attachés à chacun de ces os.

La forme allongée des doigts est suffisamment caractéristique pour arriver au résultat en injectant très peu d'information a priori. Nous utilisons un modèle structurant ad-hoc, mais il sert surtout à valider la détection, à faire le lien entre détection et localisation et à assurer l'identification finale. Le processus de reconnaissance est donc principalement dirigé par les informations extraites de l'image, ce qui assure un traitement neutre des anomalies anatomiques.

Notre méthode comporte essentiellement six étapes, durant lesquelles l'image est analysée successivement à au moins deux échelles, une relativement grossière pour la détection et une autre plus fine pour la localisation :

- **orientation préliminaire**

Il s'agit simplement de déterminer l'orientation de la main dans l'image, ce qui requiert très peu d'information a priori. Le fait que la main soit replacée approximativement dans une orientation standard simplifie les étapes suivantes.

- **détection des contours**

Il s'agit de techniques classiques menées à une échelle relativement grossière qui privilégie la fiabilité de détection.

- **construction des points médians**

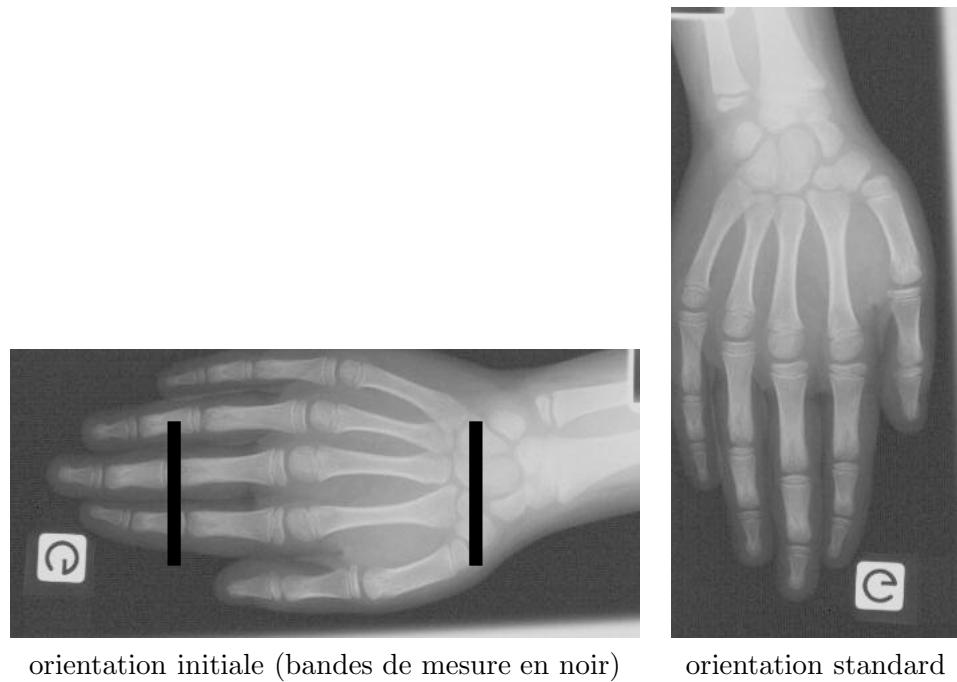
Il s'agit de construire des paires de points de contours en satisfaisant des contraintes qui traduisent la symétrie longitudinale. Chaque médian d'une telle paire est alors un point d'axe potentiel. Si nécessaire, cette étape est reprise avec des contours calculés à une échelle légèrement différente pour améliorer la consistance des points médians. Le sens de variation de l'échelle est déterminé par une analyse de leur distribution.

- **construction des axes des doigts**

Les axes recherchés sont ensuite mis progressivement en évidence comme des alignements d'un nombre suffisant de points médians. Des critères de normalité sur la position et la longueur des axes sont alors utilisés pour permettre de ne conserver finalement que 5 axes, affecté chacun à un doigt.

- **détermination de l'échelle de localisation**

L'analyse se poursuit à une échelle nettement plus fine pour obtenir des contours bien localisés. Le choix de cette échelle doit prendre en compte une erreur de caractérisation systématique et l'influence estimée du bruit.



– localisation et identification de chaque os

La difficulté est alors d'identifier les contours respectifs de chaque os. On commence par déterminer les fragments de contours longitudinaux en se basant sur leur symétrie par rapport aux axes. Il reste alors à déterminer les contours frontières entre les os d'un même doigt. À ce stade, un apport de données a priori sur les longueurs normales des os est nécessaire pour diriger la validation des informations extraites de l'image.

Ce chapitre va maintenant entrer dans le détail de ces étapes, soit respectivement, 2.1 pour l'orientation, 2.2 pour la détection des contours, 2.3 pour le calcul des points médians, 2.4 pour la construction des axes, 2.5 pour le choix de l'échelle de localisation et 2.6 pour la localisation et l'identification des os. Un certain nombre de considérations plus techniques, et relativement classiques pour certaines, sont reportées en annexe A. En 2.7, nous traiterons de la mise en œuvre et les résultats obtenus expérimentalement seront largement analysés au chapitre 3.

2.1 Orientation préliminaire

Pour diverses raisons pratiques, l'orientation “brute” des images disponibles est souvent quelconque [MGPH92]. Les phases ultérieures supposeront que la main est orientée d'une manière standard, en position verticale, le carpe en haut et les doigts vers le bas, comme sur la figure 2.1. Le recours à une orientation canonique a pour premier avantage de faciliter l'exposé de la méthode en permettant un usage non ambigu de termes usuels comme gauche, droite, descendre, etc. Nous n'avons pas à faire la différence entre la radiographie d'une main droite et celle d'une main gauche numérisée à l'envers. Cette distinction qui reposera sur la reconnaissance du marqueur G ou D, ou sur des renseignements extérieurs à l'image, est inutile dans les traitements qui nous concernent.

Un autre avantage, en fait le plus important, est qu'après cette orientation préliminaire, le programme se trouve fortement simplifié et accéléré par la détermination statique des directions et notamment le codage implicite d'un certain nombre de contraintes relationnelles. Notons que le coût d'une rotation de π ou $\pm \frac{\pi}{2}$ pour aboutir à cette orientation canonique est relativement

faible au regard des traitements ultérieurs et notamment de l'extraction des contours. De plus, ces opérations sont réalisables a priori et sans aucune altération des valeurs des pixels, alors que des balayages dans des directions quelconques demanderaient à être particulièrement bien intégrés dans les opérations qui les utilisent.

2.1.1 Heuristiques d'orientation

On suppose que l'orientation verticale de la main correspond aussi à une orientation verticale de l'image, ce qui revient à dire qu'une main est plus longue que large et que l'image est relativement bien cadrée. Notons qu'il pourrait y avoir des problèmes de détermination avec des mains jeunes (qui sont relativement peu allongées) si les doigts étaient très écartés ou la main trop en diagonale dans l'image. Heureusement, de telles configurations ne correspondent pas aux pratiques radiographiques et toutes les images de notre base, même les plus jeunes, sont traitées correctement. Pour commencer, si l'image initiale est plus large que haute, on décide donc qu'elle doit être tournée de $\pm \frac{\pi}{2}$, sans encore connaître le sens de cette rotation.

Ensuite, pour déterminer le côté du carpe, on compare les niveaux de gris moyens que l'on a mesurés sur deux bandes transversales. Ces bandes, marquées en noir sur la figure 2.1, sont placées arbitrairement, en fonction des dimensions de l'image, de manière à ce qu'elles se situent généralement l'une sur la paume et l'autre là où il n'y a plus que trois doigts. Cela reprend, d'une manière qui semble plus robuste, la méthode qui est exposée dans [PMGKH91]. Cela suppose encore une fois que l'image n'est pas trop mal cadrée et que les artefacts restent localisés à la périphérie de l'image. On utilise alors le fait que la densité est nettement plus élevée du côté du carpe.

Le rapport de densité entre les deux bandes reste suffisamment discriminant, quel que soit l'avancement de l'ossification, et cela autorise une grande tolérance sur les dimensions et le placement des bandes. Ainsi, cette méthode d'orientation empirique fonctionne correctement pour les 861 images de notre base de test, même pour les plus mal cadrées et avec des artefacts.

2.1.2 Autres techniques d'orientation

Nous avons également expérimenté d'autres méthodes pour déterminer avec plus de précision la position, les dimensions et l'orientation de la main dans l'image.

On peut notamment envisager d'utiliser les moments des niveaux de gris¹. En effet, les moments d'ordre 1 déterminent un "centre de gravité"², situé a priori du côté du carpe, et les moments d'ordre 2 donnent une "matrice d'inertie"³ dont la diagonalisation fournit des valeurs propres, donnant des dimensions caractéristiques utiles, et les directions principales, non orientées. Mais ces calculs sont très sensibles à l'importance de l'avant-bras et à la présence d'artefacts, d'autant plus que ceux-ci sont de forte intensité et sont situés à la périphérie de l'image. Ces artefacts sont présents dans presque toutes les images. En pratique, on aboutit à une direction principale qui peut s'écarte de plusieurs dizaines de degrés de l'orientation évidente de la main. En dehors de divers traitements très spécifiques pour écarter les artefacts [MGPH92], un masquage approprié est nécessaire pour limiter le domaine de calcul des moments. Les diverses techniques pour le calcul d'un masque de la main [Sai96, Eff94, PMGKH91, PGP⁺01, PPKGC03, ZGL07, MN89] restent relativement fragiles. Un tel masque s'obtient essentiellement par un seuillage adaptatif [SS04] et cela demande une orientation préliminaire approximative, pour

¹Soient $m = \sum_{ij} w_{ij}$, où w_{ij} est le niveau de gris du pixel de coordonnées (i, j) , puis $m_x = \frac{1}{m} \sum_{ij} i.w_{ij}$, $m_y = \frac{1}{m} \sum_{ij} j.w_{ij}$, $m_{xx} = \frac{1}{m} \sum_{ij} i^2.w_{ij}$, $m_{xy} = m_{yx} = \frac{1}{m} \sum_{ij} i.j.w_{ij}$ et $m_{yy} = \frac{1}{m} \sum_{ij} j^2.w_{ij}$.

²Les coordonnées du "centre de gravité" sont alors (m_x, m_y) .

³La "matrice d'inertie" est définie comme $\begin{bmatrix} m_{xx} - m_x^2 & m_{xy} - m_x m_y \\ m_{xy} - m_x m_y & m_{yy} - m_y^2 \end{bmatrix}$.

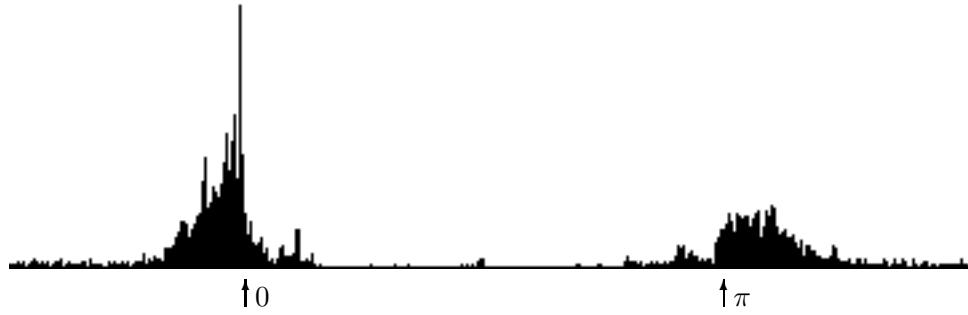


FIG. 2.2 – Mise en évidence d'une direction principale des contours

définir des voisinages non-isotropes et gérer correctement la décroissance des niveaux en allant du carpe vers les phalanges.

Le calcul d'un masque de la main permet, dans certaines conditions, de localiser les doigts et d'amorcer ainsi le recalage d'un modèle plus fin. Par exemple, on peut construire un squelette morphologique de ce masque [dLGMFAAL03] et en extraire les axes des doigts. On peut aussi détecter les extrema de courbure du contour au bout des doigts et entre les doigts [MSC⁺00] qui peuvent servir de points clés pour un recalage. Mais ces techniques sont très sensibles à la qualité des images, déjà pour la construction du masque lui-même. Ensuite les doigts doivent être bien séparés pour que le contour du masque soit l'équivalent d'un contour "peau". Enfin, l'extrapolation qui permettrait la localisation des métacarpiens (voir 2.6.2) et, à plus forte raison, du carpe (voir 4.2) est mal conditionnée.

Si l'on ne s'intéresse qu'à la mise en évidence d'une direction principale, une technique plus robuste consiste à construire un histogramme des orientations du gradient aux points de contour. Par exemple, pour une carte des contours comme celle qui est donnée en figure 2.8, page 44, on obtient l'histogramme de la figure 2.2 où les ordonnées donnent la fréquence correspondant aux orientations codées en abscisse. On retrouve facilement les deux accumulations, ici sur les directions voisines de 0 et π , qui correspondent respectivement aux contours "gauches" et "droits". Le pic très marqué, proche de l'orientation 0 , correspond au contour très rectiligne de la marge blanche sur le bord droit de l'image. Dans ce cas particulier, il pourrait être ainsi facilement identifié comme un artefact et donc écarté, mais ce n'est pas toujours aussi évident.

Une méthode du type accumulateur de Hough [DH72, Ris89] pourrait apporter une meilleure séparation des réponses parasites, en ajoutant à la direction un paramètre de position dans l'image. Mais, en fait, la mise en évidence des tangentes aux contours n'est pas envisageable car la direction varie trop le long du contour d'un os. Des entités rectilignes relativement longues sont nécessaires pour déterminer des droites identifiables. Des modèles plus sophistiqués [Bal81] supposeraient une définition simple et a priori de la forme d'un os. Une recherche des axes de symétrie [RWY95, PEMF98] serait nettement plus pertinente mais plus coûteuse (voir en annexe A.3). Toutes ces techniques supposent un paramétrage préalable, au moins pour l'échelle de calcul des contours, et elles perdent alors de leur intérêt pour un traitement préliminaire. À défaut, une recherche itérative du bon paramétrage, par exemple une analyse multi-échelle, est nécessaire. Notons aussi que la détermination d'une direction principale globale devient moins pertinente lorsque les doigts sont plus écartés et que l'on peut difficilement estimer l'orientation de chaque doigt et de chaque os à partir d'un histogramme des directions.

Tout compte fait, les heuristiques simples que nous avons exposées en 2.1.1, sont plus rapides et au moins aussi efficaces. Elles permettent de déterminer une orientation globale à $\pm \frac{\pi}{4}$, au pire, et généralement mieux du fait d'un cadrage raisonnable. En ce qui concerne l'estimation de dimensions caractéristiques de la main, nous nous baserons également sur les dimensions de l'image. Nous allons voir que cela suffit pour construire de manière robuste les axes des doigts et

ceux-ci permettront des traitements beaucoup plus riches en information que tout ce que nous venons d'envisager.

2.2 Détection des contours

La technique retenue pour la détection des contours est des plus classiques et nous ne faisons ici que motiver nos principaux choix. Pour plus de détails sur les aspects mathématiques et sur la mise en œuvre, on se reportera à l'annexe A.1.

2.2.1 Caractérisation des contours

Un point de contour est caractérisé comme un maximum local du module du gradient dans la direction du gradient, c'est-à-dire un point qui vérifie $\partial_g(g^2) = 0$, où g et ∂_g notent respectivement le gradient de l'image et l'opérateur de dérivée dans la direction de ce gradient. Une caractérisation équivalente est $g^T \mathbf{H} g = 0$, où \mathbf{H} note le hessien de l'image.

D'un point de vue formel, cette caractérisation est donc différente de la caractérisation "naturelle" par les passages par zéro du laplacien [Har84, Mar82]. En pratique, pour tout ce qui concerne la localisation des os, on va surtout s'intéresser à des portions de contour ayant une faible courbure et les écarts entre ces deux caractérisations sont alors négligeables. Le problème se posera effectivement (voir 2.5.3) lorsqu'il sera question de formes nettement plus "pointues", comme par exemple certaines épiphyses.

Un avantage essentiel de la caractérisation $\partial_g(g^2) = 0$ est qu'elle peut se réaliser par une recherche directe des maxima de g , avec l'algorithme de *Suppression des Non-Maxima* (NMS) de Canny [Can83, Can86], un rappel sur cette méthode est donné en annexe A.1. On évite ainsi d'utiliser explicitement des dérivées secondes qui, dans notre contexte, s'avèrent nettement trop perturbées par les bruits et les textures. En effet, le niveau relatif des hautes fréquences spatiales augmente avec le degré de dérivation, ce qui demande un lissage plus important et a pour effet de déformer les contours.

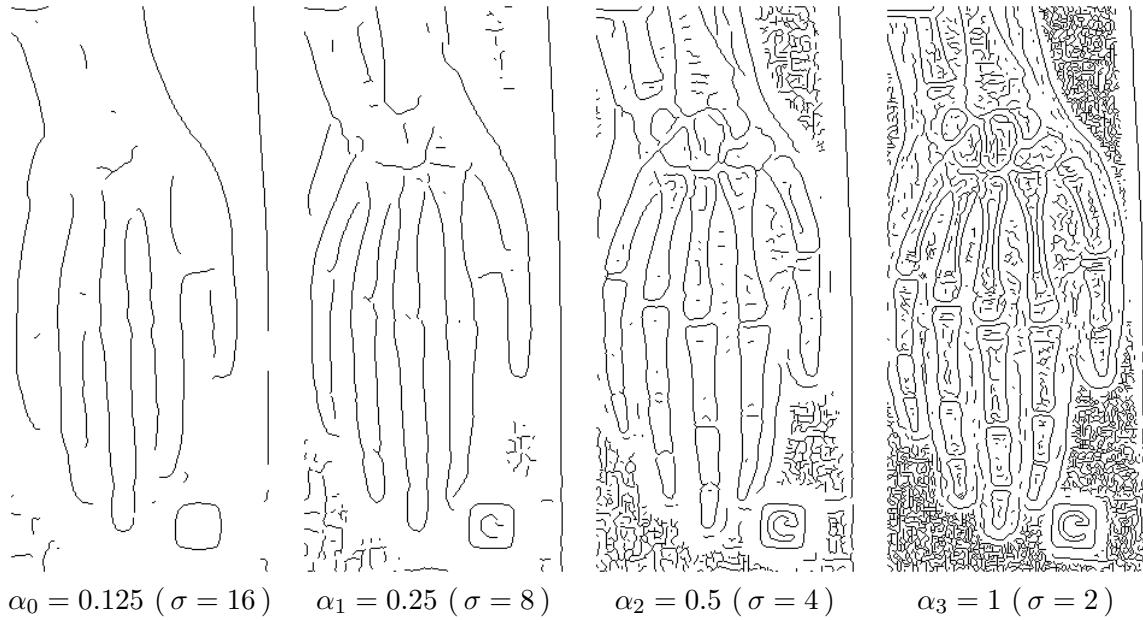
La caractérisation par les maxima du gradient nous semble ainsi la plus appropriée.

2.2.2 Calcul du gradient

Pour calculer les composantes du gradient, nous avons retenu les filtres de Deriche [Der87, Fau93]. Ces filtres sont "optimaux" pour les critères de Canny [Can83] qui expriment une certaine qualité de détection (rapport signal/bruit et unicité) et de localisation (erreur sur la position) des maxima du gradient (voir annexe A.1). Les filtres de Deriche ont l'inconvénient connu d'induire une forte anisotropie dans l'estimation du module du gradient [ZF93] mais une correction [HM95] permet de rétablir une isotropie suffisante pour pratiquer convenablement la suppression des non-maxima (NMS) et des seuillages.

Ces filtres sont contrôlés par un paramètre d'échelle⁴, noté α , qui permet de jouer facilement entre une bonne détection (avec un critère de qualité en $\frac{1}{\sqrt{\alpha}}$) et une bonne localisation (avec un critère de qualité en $\sqrt{\alpha}$). La figure 2.3 illustre l'intérêt de travailler avec différentes valeurs du paramètre α . Avec une valeur α_1 suffisamment faible, c'est-à-dire avec un lissage relativement important, le bruit et autres petites structures sont bien effacés et on détecte essentiellement les contours des doigts. Mais on peut voir que la localisation est mauvaise, avec des déplacements de plusieurs pixels, on constate notamment que les angles sont fortement arrondis. Nous utiliserons un résultat de ce type pour construire les axes des doigts, phase dans laquelle la fiabilité de la

⁴La correspondance avec un filtre gaussien de "largeur" σ est établie par la relation $\sigma = \frac{2}{\alpha}$. Formellement, α est l'inverse d'une dimension et s'exprime donc en pixel^{-1} . Le terme d'échelle est un abus de langage usuel [Wit83, Koe84].

FIG. 2.3 – Maxima du gradient en fonction du paramètre d'échelle α

détection est plus importante que la finesse de la localisation. Il est à noter que pour des valeurs encore plus faibles, comme α_0 , la structure des doigts serait elle aussi effacée.

Une valeur α_3 plus élevée, c'est-à-dire très peu de lissage, permet de localiser précisément tous les détails de l'image. Mais le bruit est important, la carte des maxima du gradient est donc très complexe et souvent ambiguë. Il faut alors apporter de l'information pour l'exploiter correctement et efficacement. Cette information sera issue de la localisation préalable des axes. Il est à noter que pour des valeurs de α encore plus élevées (échelles trop fines), le bruit peut contrarier la localisation.

Les filtres de Deriche ont l'avantage d'être séparables et récursifs, ce qui permet une mise en œuvre relativement efficace, notamment en comparaison avec les filtres gaussiens. Malgré cela, le calcul du gradient reste de loin la phase la plus coûteuse de tout le processus d'analyse d'une image. Pour calculer une carte des contours, principalement pour le calcul du vecteur gradient, il faut compter une petite centaine d'opérations en arithmétique flottante, par pixel et pour chaque échelle. Il convient donc de limiter le nombre d'échelles examinées.

Le choix d'une échelle dépend à la fois de la qualité de l'image (caractéristiques du bruit, contraste) et de son contenu (dimensions des formes pertinentes). Comme nous n'avons pas de connaissance préalable sur ces paramètres, l'estimation d'une échelle appropriée doit se faire de manière empirique. Des techniques d'exploration relativement dense du scale-space [Wit83, Ber87, Lin93, EZ96, EZ98, Lin98] (voir annexe A.1) permettent de déterminer, en chaque point, une échelle localement “optimale”. Les explorations dyadiques [BA83, MZ92, CM97, WL98], où les échelles suivent une progression géométrique de raison 2, sont beaucoup plus rapides, mais la figure 2.3 montre que l'on peut avoir besoin de quelques échelles intermédiaires. Pour le cas particulier de la main, nous allons utiliser une approche plus spécifique, guidée par la construction des axes, pour limiter le nombre d'échelles requises.

2.2.3 Validation des pixels de contour

L'étape qui termine usuellement un calcul des contours consiste à valider les solutions de la caractérisation différentielle pour ne conserver que de “vrais” pixels de contour. Sauf à échelles très grossières, un simple seuillage sur le module du gradient ne suffit pas, il est en fait impossible

de séparer correctement les contours du bruit. La caractérisation différentielle est par définition très locale et on doit utiliser des contraintes de cohérence plus globales. Une méthode, en général efficace, est la mise en correspondance avec un modèle de forme défini sur les contours [SD92, CTCG95]. Cela supposerait que l'on puisse effectivement construire (ou apprendre) et mettre en œuvre ce modèle, mais les formes des contours à décrire sont trop variables, comme nous l'avons vu au chapitre précédent.

Dans une approche purement guidée par l'image, on ajoute principalement de la cohérence topologique, relativement locale, en construisant des chaînes de pixels connexes [ED85, Gir87]. Cette construction va de pair avec un “amincissement” des contours, chaque pixel à l'intérieur d'une chaîne n'ayant plus que deux voisins (au sens de la 8-connexité) dans une direction qui est approximativement perpendiculaire au gradient. Une chaîne est alors validée globalement si elle contient suffisamment de pixels de fort gradient. On arrive ainsi à écarter les bruits relativement isolés et c'est généralement suffisant pour les larges échelles, mais le problème du seuillage reste encore délicat pour les échelles fines, quand le bruit peut s'organiser en grandes structures.

Un traitement multi-résolution “coarse to fine” [Wit83] permet d'exploiter la cohérence des contours entre échelles voisines dans une approximation discrète du scale-space. Cela consiste donc à considérer plusieurs cartes de maxima du gradient, pour une suite d'échelles de plus en plus fines. Un pixel de contour est alors validé s'il est assez “proche” d'un pixel déjà validé à une échelle plus grossière et on parvient ainsi à propager une sélection des contours depuis l'échelle de détection jusqu'à l'échelle de localisation [MZ92, CD95].

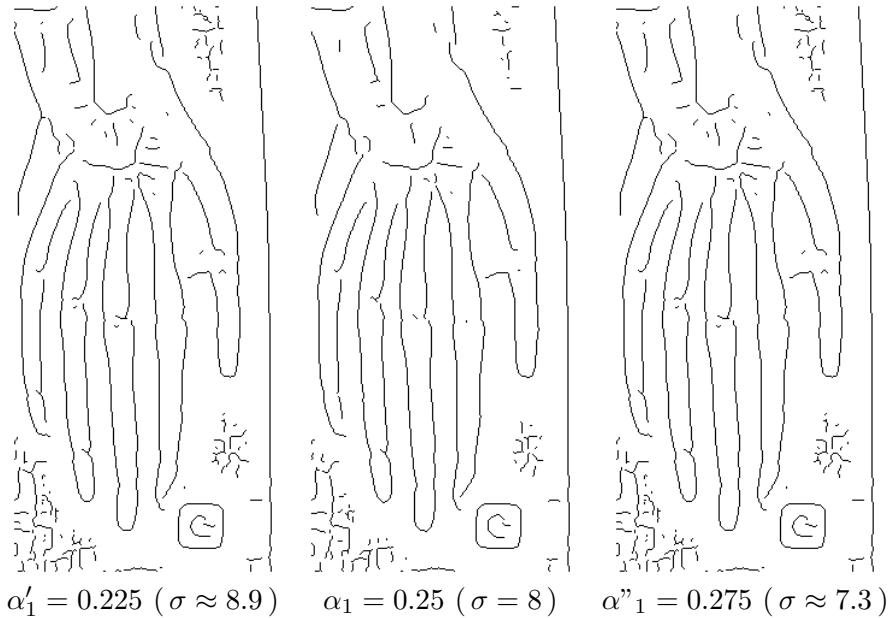
À ce propos, nous avons noté [BC98] l'importance d'une expression de la cohérence qui utilise à la fois la localisation des pixels et l'orientation du vecteur gradient. Ce dernier critère permet d'assurer une forte contrainte de régularité entre les échelles. Mais cela suppose toujours une analyse relativement dense dans les échelles (voir annexe A.1) et la méthode s'avère finalement trop coûteuse en regard du gain de qualité obtenu. Le problème essentiel que l'on rencontre en suivant une telle approche, est que les seules mesures de qualité disponibles sont locales, par définition du scale-space [Wit83, Koe84, MH92]. Notamment, cela ne permet pas de quantifier la pertinence effective d'un contour par un critère bien choisi de cohérence globale, par exemple, un critère de forme.

Dans notre cas, la validation des contours grossiers peut être considérablement renforcée par la construction des axes des doigts, ce qui permet d'exprimer des contraintes particulières de symétrie. Alors qu'un contour peut se déplacer notablement quand on traverse les échelles, l'axe entre deux contours relativement symétriques est beaucoup plus stable. La connaissance de ces axes est ainsi utilisée pour sélectionner des pixels de contour à une échelle plus fine. Cette utilisation combinée des données issues de l'image et d'un modèle de symétrie, persistant à toutes les échelles usuelles, permet d'assurer directement la cohérence entre deux échelles extrêmes. On fait ainsi “le grand saut”, au lieu d'un suivi “à petits pas” par l'approche multi-résolution.

Enfin, l'échelle de détection et l'échelle de localisation seront fixées a priori pour toute l'image. Cette simplification est permise par une relativement bonne homogénéité dans les dimensions des structures. Nous déterminons l'échelle de détection par quelques approximations qui reposent sur un cadrage convenable de la main dans l'image. Dans le cas contraire, le besoin d'une correction éventuelle est apprécié a posteriori, en fonction des points médians obtenus ; ce mécanisme sera détaillé en 2.3.4. Nous verrons ensuite, en 2.5, comment déterminer au mieux l'échelle de localisation.

2.3 Construction des points médians

Le but est d'obtenir les axes de 5 structures très allongées et orientées dans une direction générale proche de la verticale de l'image. Cela se fait à partir d'une carte des contours calculée

FIG. 2.4 – Petites variations du paramètre d'échelle α

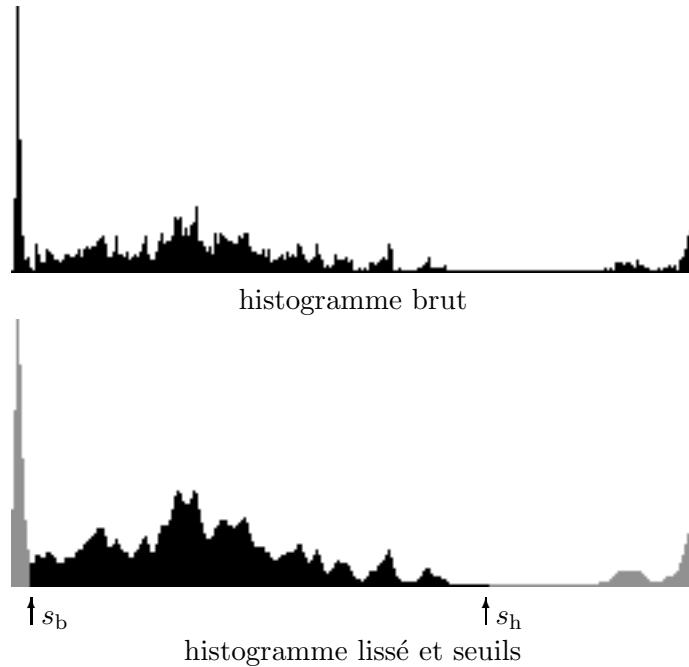
à une échelle qui est déterminée pour privilégier la qualité de détection. Une échelle initiale est choisie de manière empirique et un processus adaptatif permet d'améliorer la détection si besoin. On peut ainsi procéder à un appariement qui détermine des pixels médians relativement sûrs. Ces pixels seront ensuite chaînés en se basant sur leur connexité (ou forte proximité) pour construire effectivement les axes.

2.3.1 Détermination de l'échelle initiale

Le choix de l'échelle de détection est contraint par des considérations anatomiques (on veut distinguer les contours des 5 doigts) et par l'importance du bruit. Il est difficile d'obtenir a priori une mesure réaliste du bruit qui permettrait de fixer une échelle minimale (α maximal), on essaie donc de se situer relativement près de l'échelle maximale.

Cette échelle maximale est celle qui assure l'existence des contours recherchés et elle doit être compatible avec la détection des objets dont la dimension caractéristique est de l'ordre de la largeur d'un doigt. Pour des échelles trop grandes, les gradients opposés des deux contours vont se déplacer et s'affaiblir en se combinant entre eux. Notre problème présente des difficultés particulières car la largeur d'un doigt est comparable à l'écart entre deux doigts et des détails internes comme les contours du canal médullaire donnent eux-aussi des gradients antagonistes. Ainsi, l'estimation de l'échelle maximale est relativement critique.

À défaut d'autre information initiale, nous avons choisi de déterminer l'échelle de détection en fonction des dimensions de l'image. Sous l'hypothèse que l'image est relativement bien cadrée, ces dimensions peuvent être considérées comme représentatives de celles de la main et, par là, de celles de doigts, des os, etc. L'idéal serait bien sûr de disposer d'une information plus pertinente, comme la largeur des doigts, mais celle-ci n'est pas encore connue à ce stade du traitement. Notre choix s'est porté sur la hauteur de l'image (H) car le cadrage en largeur est nettement plus affecté par l'écartement des doigts et les variations d'orientation autour de la verticale. L'échelle de détection initiale est donc calculée par une formule du type $\alpha = \frac{K}{H}$, où K est une constante déterminée empiriquement et où la loi en $1/H$ traduit directement que α est l'inverse d'une dimension.



L’axe des abscisses porte les valeurs des maxima du module du gradient et l’axe des ordonnées donne la fréquence correspondante.

FIG. 2.5 – Histogramme des maxima du gradient et seuils de détection, pour $\alpha = 0.25$

Pour notre exemple, l’échelle de détection initiale sera ainsi choisie de l’ordre de $\alpha = 0.25$, soit $\sigma = 8$, alors que la largeur d’appariement qui sera estimée plus tard, est de l’ordre de 20 pixels. On verra en 2.3.4 (et figure 2.10) que l’échelle maximale serait de l’ordre de $\sigma = 10$ ($\alpha = 0.2$). On verra aussi qu’il serait possible d’utiliser des échelles nettement plus fines, sans que le bruit ne prenne trop d’importance. En fait, la carte des maxima du gradient ne change que très lentement quand l’échelle varie autour d’une valeur adaptée à la détection. On le constate par exemple sur la figure 2.4, avec des variations de 10 % de l’échelle.

C’est une remarque très générale, valable pour presque toutes les images que nous avons considérées. Cela nous a permis de valider expérimentalement la formule $\alpha = \frac{K}{H}$, avec peu de contraintes sur le cadrage et sur le choix de la constante K . En pratique, une surestimation d’un facteur 2, voire 4, est souvent acceptable dans la détermination de K et donc α .

Nous verrons au chapitre 3 que cette estimation initiale est très généralement satisfaisante, dans la mesure où elle permet d’aboutir directement à la construction des axes des doigts pour la plupart des images de notre base de test. Il arrive pourtant qu’un niveau de bruit relativement élevé demande d’utiliser une valeur d’échelle plus grande (α plus faible) ou, au contraire, qu’un mauvais cadrage conduise à une sous-estimation de la valeur pour α . Ainsi, pour notre exemple, si la main été cadrée dans une image 2 fois plus haute, on utiliserait une échelle de l’ordre de 0.125, totalement inadaptée comme le montre la figure 2.3. Dans de tels cas, le besoin et le sens d’une correction seront décidés a posteriori en analysant la qualité des axes obtenus (voir 2.3.4).

2.3.2 Seuillage des contours

Le bruit qui crée des maxima du gradient est imputable pour l’essentiel aux diverses étapes de production de l’image, processus radiographique, numérisation, codage JPEG, etc. La texture osseuse intervient pour une part moindre. À une échelle grossière, ces contours parasites sont d’intensité relativement faible et ils peuvent être éliminés par seuillage sans écarter trop de

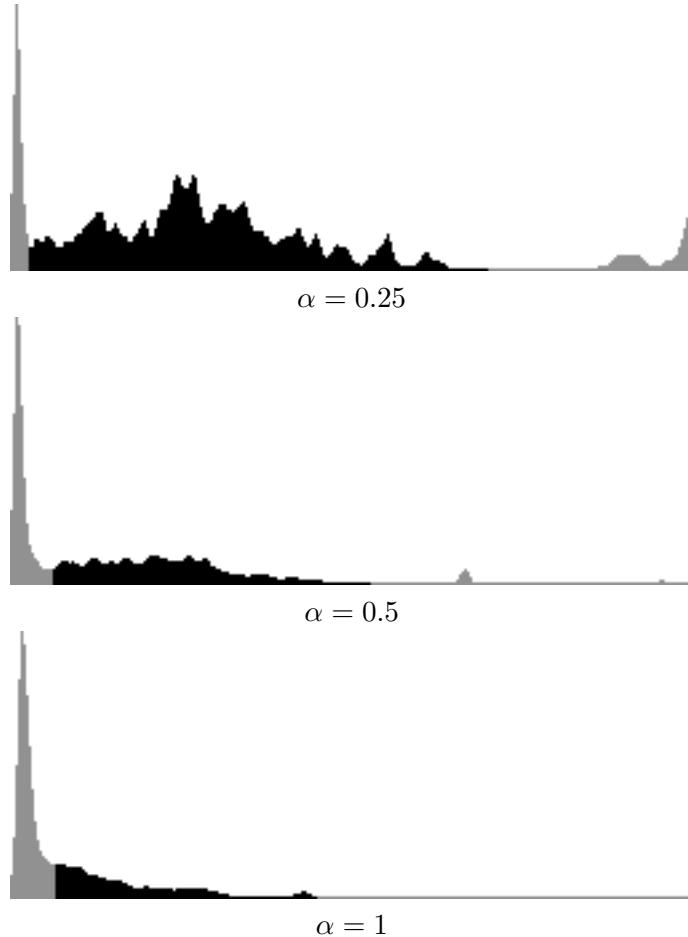


FIG. 2.6 – Histogramme des maxima du gradient en fonction du paramètre d'échelle α

“vrais” contours. Mais les seuils doivent être adaptés à une distribution des maxima du gradient qui est particulière à chaque image et qui dépend aussi bien des paramètres techniques que de l’importance de la minéralisation osseuse. Nous avons mis en œuvre une adaptation automatique des seuils [Niç96], en utilisant un histogramme des maxima du gradient (voir figure 2.5). On peut noter la distribution relativement large et complexe de l’intensité des contours, avec de nombreux pics, ce qui traduit la grande diversité des structures de la main. Cela met en échec les principales techniques de seuillage [SS04] qui supposent une distribution nettement plus simple.

On observe facilement le premier pic très marqué. Il y a en effet de nombreux maxima de gradient de faible intensité qui correspondent en grande part au bruit et aux textures. Après une légère régularisation de l’histogramme, nous pouvons déterminer, par balayage de gauche à droite, un seuil bas s_b qui est l’abscisse du premier minimum local qui suit ce premier pic. Cela reste assez classique [Sez90] et revient à considérer que les niveaux de maxima du gradient dûs au bruit et aux textures suivent une loi unimodale relativement “pointue” et sont d’occurrence très rare au-delà du seuil s_b .

La figure 2.5 correspond à un champ de gradient calculé à l’échelle de détection estimée pour cet exemple. Quand α augmente, c’est-à-dire avec un lissage plus faible, les structures parasites donnent des valeurs de gradient relativement plus importantes qui peuvent se confondre avec celles de “vrais” contours. Comme on le constate sur la figure 2.6, la limite à droite du premier pic devient donc plus floue. Sur cet exemple, pour $\alpha = 1$, la construction du seuil bas s_b n’a alors même plus de sens.

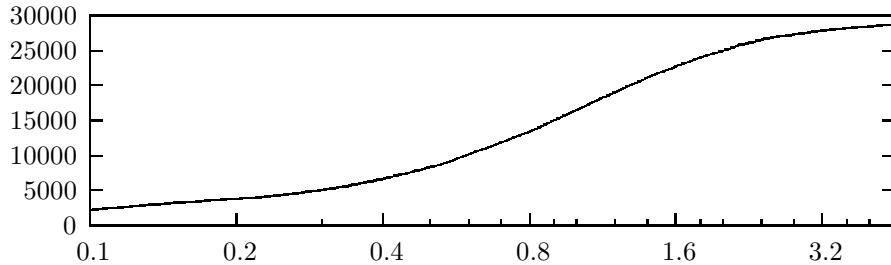


FIG. 2.7 – Nombre de maxima du gradient en fonction du paramètre d'échelle α

D'autre part, tout à droite de l'histogramme pour $\alpha = 0.25$, on peut remarquer les deux accumulations de maxima qui sont dues à des artefacts (marque G et bande blanche à droite de la radiographie). Ce sont des valeurs très proches du maximum absolu du gradient pour cette image et à cette échelle. Pourtant, quand α augmente, certains parasites du gradient, alors moins filtrés, peuvent atteindre des valeurs relatives nettement plus élevées. Ainsi, pour $\alpha = 0.5$, ces mêmes artefacts ne se situent qu'aux $2/3$, environ, de la dynamique des valeurs de gradient. Dans le même temps, la “fréquence” de ces maxima est diminuée car leur nombre varie peu alors que le nombre total de maxima comptabilisés est nettement plus important (voir figure 2.7).

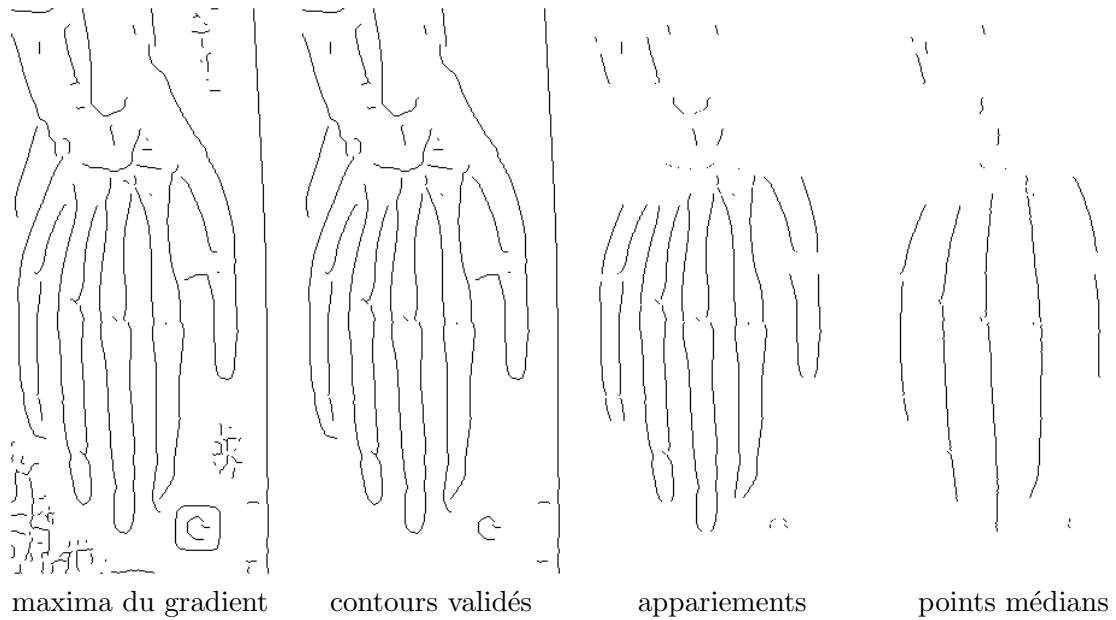
On peut essayer d'écartier de tels artefacts par seuillage. Sans hypothèse particulière sur la distribution des maxima du gradient, un seuil haut est déterminé comme $s_h = \mu + k\sigma$, où μ et σ^2 notent respectivement la moyenne et la variance des maxima du module de valeur $\geq s_b$ (donc hors bruit). Cette formule permet une bonne adaptation à la dynamique propre à chaque image et la valeur de la constante k est fixée empiriquement à $\frac{3}{2}$. On préserve ainsi l'essentiel des “vrais” contours, pour toutes les “bonnes” images et pour une plage d'échelles assez large. Lorsque la distribution des maxima du gradient est trop étalée, cette formule produit une valeur de seuil sans effet car très haute.

Par ailleurs, un seuillage similaire, mais appliqué aux niveaux de gris de l'image, au lieu de le faire sur le module du gradient, permet de masquer certains artefacts très contrastés. Pour une moitié des images environ, la combinaison des deux techniques arrive à supprimer l'essentiel des artefacts et, dans quelques cas, cela joue même un rôle important pour la bonne qualité des axes. Pour les autres images, un seuil satisfaisant ne peut pas être déterminé automatiquement, même localement et même quand un tel seuil existe effectivement, car les artefacts ont des caractéristiques photométriques trop proches de celles de certains os. Une localisation précise permettrait de mieux guider un tel processus, mais elle n'est pas encore acquise à ce stade.

En fait, la présence de ces artefacts résiduels n'est généralement pas critique pour la suite du traitement. Le plus souvent, leurs contours ne vont pas satisfaire les contraintes structurelles, dont la contrainte d'appariement pour commencer. Ensuite, les points médians éventuels se trouvent trop isolés et sont exclus lors du chaînage des axes. La figure 2.8 montre le résultat pour notre image d'exemple. Le contour intérieur de la marque G et celui de la bande blanche à droite résistent au seuillage mais ne donnent pas d'appariement conséquent.

2.3.3 Appariement des points de contours

Il s'agit maintenant de mettre en évidence la symétrie des os longs et, plus généralement, des doigts en associant les pixels pour constituer des paires dont les médians feront des points d'axes potentiels. On constate rapidement que la symétrie longitudinale recherchée n'est que très approximative, contrairement aux apparences. Cette perception est nettement favorisée par l'orientation verticale. C'est un phénomène visuel très général qui se vérifie même sur les formes

FIG. 2.8 – Seuillage et appariement des contours (échelle $\alpha = 0.25$)

les plus abstraites [Mar82]. Dans notre cas, s'ajoutent des mécanismes cognitifs et subjectifs qui prédisposent à la reconnaissance des axes des doigts, alors qu'il existe de nombreuses autres symétries intra ou inter-digitales.

Les techniques usuelles de mise en évidence des symétries par l'accumulation de réponses locales [Low85] demandent des calculs denses et coûteux⁵, sur les pixels [BG87, GP93a] ou le gradient [MPL94, PBC⁺94, RWY95] et, dans notre cas, ils ne sont pas particulièrement efficaces (voir annexe A.3). En raison d'un grand nombre de symétries potentielles et de leurs chevauchements, on ne peut pas aboutir facilement à une mise en évidence qui soit à la fois quantitativement suffisante et précise des seuls axes des doigts. Les techniques par accumulation sont efficaces pour filtrer du bruit mais, ici, le problème provient de structures parasites très cohérentes. La distinction repose sur des caractérisations particulières avec un *a priori* sur la distance d'appariement. Le réglage de ce paramètre et le seuillage des réponses est délicat. Sans *a priori* particulier, la résolution passe par l'essai de différents paramétrages.

Une caractérisation par les lignes d'extrema de courbure [GP93b, EGM⁺94, MDDS01] serait plus rapide mais elle n'est pas adaptée. La définition de l'axe de symétrie d'un doigt en terme de courbures est très variable suivant le contexte et, aux échelles requises, les interactions entre structures voisines entraînent des déviations importantes des lignes d'extrema.

Un intérêt particulier du gradient est qu'il fournit une direction qui peut être utilisée pour formuler les contraintes de symétrie. On peut aussi simplifier le problème de l'appariement du gradient en ne s'intéressant qu'aux pixels des contours [PGP⁺01, dLGMFAAL03, RWY95, Low85, MN92] et, comme les doigts ont été orientés plutôt verticalement, on peut facilement distinguer des contours "gauches" et des contours "droits", selon le signe de la composante horizontale du gradient [CGE94, CNHS99]. Notons que les contours détectés de chaque côté d'un doigt ne sont pas toujours symétriques en intensité, voire en nature. Il peut s'agir des contours des os et/ou de ceux de la peau ou d'un seul contour plus grossier et cela peut changer le long et de part et d'autre d'un doigt en fonction des contrastes ou pour des raisons structurelles. Notamment, pour les métacarpiens, il n'y a pas de contour de la peau, sauf le contour extérieur de la main, et on détecte assez fréquemment les contours des canaux médullaires.

⁵Il s'agit d'explorer, pour chaque pixel, un voisinage de dimensions comparables à la largeur d'un doigt.

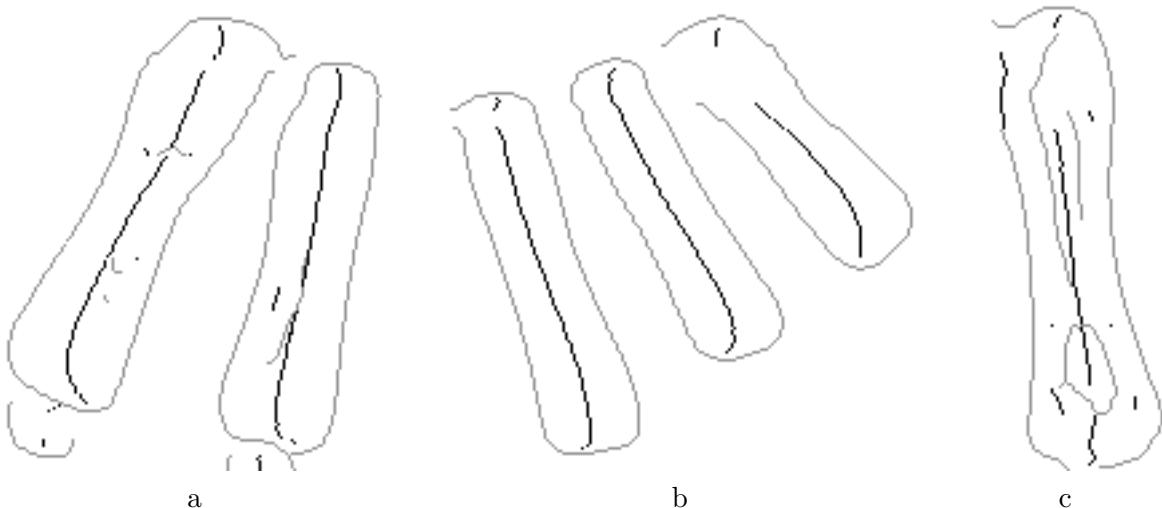


FIG. 2.9 – Exemples de construction des points médians

Une approche rigoureuse serait de privilégier les appariements dans la direction localement perpendiculaire à l'axe attendu. On n'a aucun moyen de déterminer exactement la direction d'appariement sans itérer sur la mise en évidence des axes. Une direction globale de la main peut être mise assez facilement en évidence (voir 2.1.2). Le problème est plus difficile pour un doigt seul, d'autant plus que les doigts qui posent généralement le plus de problèmes du fait de leur inclinaison exagérée, le pouce et l'auriculaire, sont aussi les plus courts et, à plus forte raison, dans le cas du pouce qui peut être plié. Notons aussi que les doigts disposent d'une grande liberté d'orientation individuelle et qu'il n'est pas envisageable d'estimer la direction de l'un à partir de celles des autres, à l'aide d'un modèle angulaire "en éventail". Construire un appariement correctement dirigé sans *a priori* est donc un problème qui reste relativement lourd.

Les techniques générales [GP93a, MPL94, RWY95] sont isotropes mais posent les problèmes que l'on a vus. Pour acquérir cette information de direction, on va commencer par construire des paires horizontales dont les points médians formeront des axes approximatifs.

En pratique, un tel point médian est assez proche de l'axe théorique, tant que l'angle β entre cet axe et la perpendiculaire à la direction d'appariement (ici, la verticale) reste dans les limites $\pm \frac{\pi}{4}$ et surtout que la largeur d'appariement ne varie pas trop rapidement le long de l'axe. Pour un écart δ_l entre les largeurs réelles (mesurées perpendiculairement à l'axe) aux deux points d'une paire, la distance du médian à l'axe est $\frac{\delta_l}{4\cos^2\beta}$. L'écart δ_l est fonction de la distance (approximativement $l \tan \beta$) entre les projections sur l'axe des deux points d'une paire. Les seuls écarts notables se constatent aux extrémités des os inclinés (voir par exemple figure 2.9). Ces défauts bien caractérisés seront écartés lors de la construction des axes (voir 2.4) et les axes ainsi obtenus permettront ensuite de diriger correctement de nouveaux appariements (voir 2.6).

Les appariements horizontaux vont s'obtenir au cours d'un balayage ligne par ligne, qui est l'opération minimale pour relire la carte des contours. Ils sont réalisés "au vol" et c'est une opération relativement peu coûteuse. Au cours de cette construction, il serait d'ailleurs prématûr de procéder à une validation verticale, en utilisant la continuité des appariements d'une ligne de balayage à la suivante, car il y a un risque non négligeable de propager ainsi de mauvaises hypothèses. Une faible corrélation des caractérisations élémentaires est fondamentale dans une mise en évidence par accumulation [Low85]. On n'applique donc, à ce stade, que des règles qui assurent une cohérence horizontale, restreinte à la ligne analysée.

Notre méthode est relativement robuste pour des inclinaisons jusqu'à $\pm \frac{\pi}{4}$. Si on accepte de traiter 3 directions d'appariement (l'horizontale et les deux diagonales), on limite les déviations

d'axe avec une limite d'inclinaison relative de $\pm \frac{\pi}{8}$. Avec en plus un essai d'appariement vertical, on peut même déterminer ainsi l'orientation de l'image. Cela reste nettement plus rapide que les techniques générales car le seul calcul lourd, celui du gradient, n'est fait qu'une fois. Le principe de base de la construction des appariements est alors simple.

Durant le balayage d'une ligne, si on rencontre un pixel de contour "gauche", que l'on notera ' $'$, suivi d'un pixel de contour "droit", noté ' $)$ ', on valide la paire ' $()$ ' et on produit le pixel médian. Avec une bonne carte des contours, ce type d'appariement est de loin le plus courant et cette règle suffit pour obtenir la quasi totalité des points médians des figures 2.8, 2.9-a et 2.9-b. Mais on va aussi rencontrer des cas moins simples qui ne sont pas toujours négligeables car ils peuvent représenter jusqu'à plus de 15 % des paires valides dans certaines images.

Pour une bonne mise en évidence des axes, il est donc important de traiter convenablement tous ces cas. Par exemple, si les contours du canal médullaire sont détectés (voir figure 2.9-c), il faut traiter globalement une séquence ' $()()$ ' pour lui assigner un seul point médian central. Mais ce cas ne doit pas être confondu avec celui de deux doigts côté à côté. L'espacement sur la ligne de balayage joue donc un rôle important pour décider d'un choix structurel ou pour simplement écarter du bruit.

Cela nous a conduit à utiliser deux paramètres qui fixent des espacements minimal et maximal entre les pixels d'une paire, ce qui permet de distinguer convenablement les contours des canaux médullaires de ceux de doigts voisins. On ne connaît pas encore la "largeur" d'un doigt qui serait déterminante mais, à partir de la largeur de l'image, on peut estimer a priori des valeurs qui s'avèrent convenables, même quand les doigts sont inclinés, très écartés ou, au contraire, très serrés. Nous avons fixé, de manière assez naturelle, la distance maximale d'appariement à un cinquième de la largeur de l'image. En ce qui concerne la distance minimale, un trentième de la largeur de l'image apparaît comme un bon compromis.

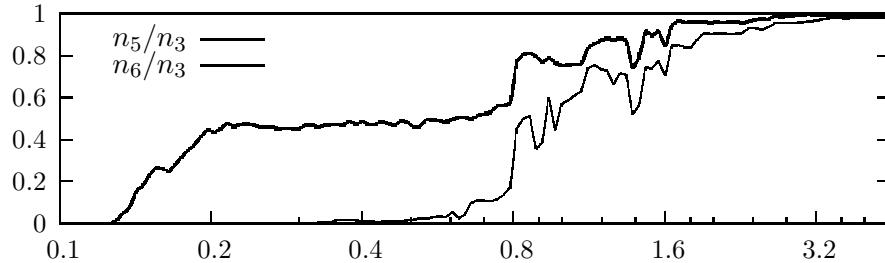
Nous devons ainsi analyser 7 combinaisons de ' $'$ et ' $)$ ', de longueur maximale 4, qui peuvent facilement correspondre à certaines particularités anatomiques plus ou moins bruitées. On ne saurait pas interpréter correctement des combinaisons plus complexes, pour lesquelles il n'y aurait pas d'évidence d'une structure particulière. En règle générale, on va privilégier les appariements les plus internes, dans la mesure où ils satisfont la contrainte de distance minimale. Ce principe est essentiel pour assurer un traitement linéaire, "au vol", cohérent et durant lequel on ne veut pas reconsiderer les décisions antérieures. Pour une analyse des différents cas et plus de détails sur la mise en œuvre de l'appariement, on se reportera à l'annexe A.2.

À l'issue de cette étape, on a obtenu un ensemble de points d'axes potentiels. Chacun de ces points garde des références sur les deux pixels de contours auxquels il est associé, et il est étiqueté par la distance de l'appariement correspondant. En corollaire, on pourrait dès maintenant estimer la "largeur" des doigts, en fonction des distances des appariements réalisés, mais il est préférable de le faire après un peu plus de nettoyage. Une cohérence structurelle globale sera établie lors de la construction des axes, dans un nouveau processus complet de groupement perceptuel [Low85, MN92, SB93] qui partira des points médians, pris en tant que primitives. Un certain nombre de points médians aberrants seront alors écartés.

Auparavant, nous allons analyser quantitativement l'ensemble actuel des points médians pour décider de la qualité des contours sources et d'une correction éventuelle.

2.3.4 Validation et adaptation de l'échelle de détection

L'échelle de détection qui a été déterminée a priori en fonction de l'image (voir 2.3.1) est très généralement satisfaisante, c'est-à-dire qu'elle permet de détecter l'essentiel des 5 axes, sans trop de parasites. Il en est ainsi avec toutes les images bien cadrées et pas trop bruitées, nous verrons au chapitre 3 que cela représente 94 % de nos "bonnes" images.

FIG. 2.10 – Nombre de points médians en fonction de l'échelle α

Pour des images plus difficiles, nous devons parfois procéder à une correction de l'échelle. Rappelons qu'il ne s'agit pas de déterminer une échelle optimale mais simplement d'obtenir l'essentiel des axes des doigts. Pour quantifier la qualité de la détection nous nous basons sur un comptage des pixels médians. Ce comptage est réalisé, ligne par ligne, au fur et à mesure de l'appariement et le résultat est exprimé sous la forme d'une suite de valeurs (n_i) qui donnent le nombre de lignes de balayage comptant au moins i pixels médians. Les valeurs n_1 et n_2 ont l'inconvénient de prendre en compte les os de l'avant-bras et d'être donc très dépendantes du cadrage. La première valeur significative est n_3 qui peut être considérée comme une bonne estimation du nombre de lignes intersectant l'axes de l'index ou de l'annulaire. Les valeurs n_i suivantes sont relativisées à n_3 pour s'affranchir des dimensions et de l'orientation de la main.

Ainsi, le rapport $\frac{n_5}{n_3}$ est une estimation du rapport de longueur entre le pouce et l'index. Ce rapport devrait s'approcher d'une valeur théorique de l'ordre de 0.64, d'après les tables de normalité [Poz74], c'est-à-dire le rapport $\frac{\mu_1}{\mu_2}$ selon le tableau 2.1, page 76. Il faut tout de même se donner une marge de tolérance importante pour accepter des différences d'inclinaison entre les axes, un certain repliement du pouce, l'absence de points médians au niveau des articulations, une détection plutôt chaotique des os du carpe, la présence d'artefacts, etc. Quand au rapport $\frac{n_6}{n_3}$, nul en théorie, c'est un bon indicateur de l'importance du bruit. La figure 2.10 montre l'évolution de ces deux rapports, pour l'image d'exemple, en fonction de l'échelle de calcul des contours.

Pour des échelles très grossières, le rapport $\frac{n_5}{n_3}$ est très faible, car la détection des contours des doigts n'est pas réalisée correctement, puis il augmente doucement. Le lissage confond alors plusieurs doigts en une seule structure, sur une longueur plus ou moins importante selon l'échelle. Ce rapport est ensuite relativement stable, avec une valeur de l'ordre de 0.45 (contre 0.64 en théorie), sur un intervalle d'échelles comprises entre 0.2 et 0.8, approximativement. Le rapport $\frac{n_6}{n_3}$ reste très faible tant que le bruit est bien effacé durant la détection des contours. Au-delà d'une certaine échelle limite, de l'ordre de 0.8 pour cette image, les deux rapports augmentent rapidement en raison du bruit et leur évolution devient plus chaotique. Avec une autre image, on pourrait faire des observations similaires, avec seulement des valeurs d'échelles différentes.

On peut ainsi définir deux valeurs de seuil, $s_{53} \leq s_{63}$ qui déterminent la validation et la recherche d'une échelle de détection satisfaisante. La double condition $\frac{n_5}{n_3} \geq s_{53}$ et $\frac{n_6}{n_3} \leq s_{63}$ définit l'intervalle de validité. Nous avons retenu comme seuils $s_{53} = s_{63} = 0.15$, pour lesquels cet intervalle est, en général, assez large. La correction de l'échelle de détection se fait alors de manière itérative, par encadrement et bisection :

- Si $\frac{n_5}{n_3} < s_{53}$, la détection des contours des doigts est considérée comme insuffisante et on doit prendre une échelle plus fine de manière à l'améliorer.

- Si $\frac{n_6}{n_3} > s_{63}$, on détecte trop de détails ou de bruit et on doit prendre une échelle plus grossière.

Il est à noter que, par définition, $\frac{n_5}{n_3} \geq \frac{n_6}{n_3}$. Le cas ambigu $\frac{n_5}{n_3} < s_{53}$ et $\frac{n_6}{n_3} > s_{63}$ est donc écarté par un choix de seuils tels que $s_{53} \leq s_{63}$.

Les échelles considérées restent toujours relativement grossières et le problème est alors suffisamment régularisé pour autoriser cette technique itérative. Il faut noter que ces mêmes rapports changent peu s'ils sont mesurés seulement sur les pixels validés par la construction des axes. Il est donc préférable de faire l'adaptation d'échelle tout de suite. En phase d'encadrement, il est logique d'utiliser une suite de valeurs α qui suivent une progression géométrique. Ensuite la bisection se réalise en prenant les racines carrées successives de la raison initiale.

Nous avons retenu une raison initiale de l'ordre de $\sqrt[8]{2}$, (ou son inverse), ainsi 4 itérations d'encadrement représentent une variation d'échelle d'un facteur $\sqrt{2}$, ce qui est déjà considérable pour corriger un défaut de cadrage de l'image. Quelques itérations de bisection conduisent à des variations d'échelle insignifiantes. Si l'on n'obtient pas de résultat satisfaisant après quelques itérations, on conclut à un échec de construction des axes. Cela se produit sur quelques images, pour lesquelles l'intervalle de validité s'avère trop réduit, notamment lorsque le bruit qui affecte n_6 est trop important ou lorsque le pouce est trop écarté, ce qui abaisse nettement la valeur de n_5 . Dans de tels cas, l'encadrement peut "déraper" vers des échelles totalement inadaptées ou alors la bisection a un comportement chaotique. La relation constatée entre le nombre d'itérations nécessaires à une convergence et la qualité du résultat correspondant sera discutée lors de l'analyse des résultats expérimentaux (voir chapitre 3). Il en ressort qu'il n'est pas utile de s'autoriser un nombre élevé d'itérations et qu'il est souvent préférable de conclure rapidement à un échec pour de telles images. Cela concerne moins de 2% des images que nous avions considérées comme "bonnes", selon un examen visuel a priori, comme décrit en 1.3.

L'échelle qui est finalement validée permet ainsi de détecter l'essentiel des 5 axes et elle réalise un bon compromis entre l'adaptation aux dimensions effectives de la main et l'élimination du bruit. Cette valeur d'échelle servira aussi de référence pour estimer le rapport signal/bruit qui interviendra dans l'expression de la précision de la localisation (voir 2.5). Il nous reste à construire effectivement ces axes.

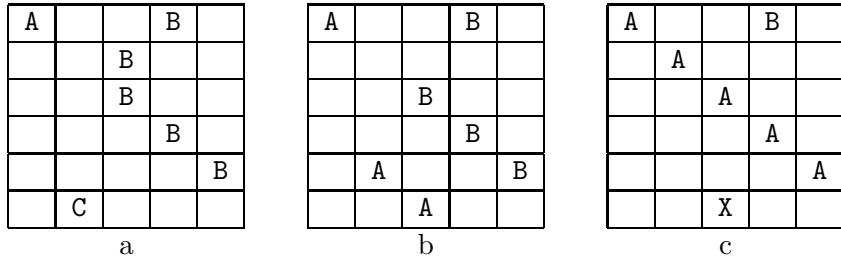
2.4 Construction des axes des doigts

Étant donnés des pixels médians, il s'agit maintenant de les regrouper pour mettre en évidence leurs contributions à des structures cohérentes. Le processus suit une trame classique de groupement perceptuel [Low85, MN92, SB93] qui se pratique usuellement sur des pixels de contour.

Une première phase consiste à chaîner les pixels suffisamment proches les uns des autres. Par construction, il s'agira bien de chaînes dont tout pixel, sauf ceux d'extrémités, a deux voisins. Après un peu de nettoyage, les chaînes résultantes sont assemblées, les unes à la suite des autres, selon une relation particulière qui traduit l'alignement des chaînes constitutives d'un même axe. Notre méthode considère qu'il existe une direction générale d'alignement, à $\pm \frac{\pi}{4}$, supposée ici être la verticale. Un certain nombre d'heuristiques permettent finalement de ne conserver que 5 axes et de les affecter chacun à un doigt. Les étapes successives de cette construction sont illustrées par la figure B.1, page 179 en annexe, avec un coloriage arbitraire qui permet de distinguer les différentes structures.

2.4.1 Chaînage par connexité

Le chaînage par connexité peut être réalisé au fur et à mesure de la génération des pixels médians. Durant le balayage ligne par ligne et pour chaque nouvel appariement, il suffit de regarder si le pixel médian correspondant, de coordonnées (i, j) , peut prolonger une chaîne existante. Dans le cas d'une connexité directe, cela signifie que le dernier pixel actuel d'une



On schématise ici de petits exemples de cartes des pixels médians. Chaque case est un pixel et les pixels médians sont ceux marqués d'une lettre. Les pixels marqués d'une même lettre satisfont les règles qui permettent de les chaîner entre eux.

FIG. 2.11 – Exemples de chaînage par connexité

chaîne se trouve en l'une des trois positions⁶ qui réalisent la 8-connexité : $(i-1, j-1)$, $(i, j-1)$ ou $(i+1, j-1)$. La chaîne est alors augmentée de ce nouveau pixel, on construit ainsi la chaîne des B en figure 2.11-a. Dans le cas contraire, le pixel est le départ d'une nouvelle chaîne, par exemple, C. Les chaînes qui ne peuvent plus être prolongées sont retirées de l'ensemble à examiner.

Pour qu'il n'y ait pas d'ambiguïté dans cette construction, il suffit d'un écart minimal de 3, autrement dit deux pixels intercalés, entre deux pixels médians consécutifs sur une même ligne de balayage. Il est alors clair que tout pixel médian ne peut pas être 8-connexe à plus d'un pixel médian de la ligne précédente (et aussi de la suivante). L'algorithme d'appariement des contours nous permet d'assurer l'espacement minimal requis entre deux pixels médians consécutifs sur une même ligne de balayage (voir annexe A.2). Notons par ailleurs que, le long d'une chaîne ainsi construite, les ordonnées sont toujours strictement croissantes.

Pour rattraper des petits défauts de construction des contours ou d'appariement, on doit étendre la relation de connexité. Si on ne trouve rien en $(i-1, j-1)$, $(i, j-1)$ et $(i+1, j-1)$, on regarde les trois positions de mêmes abscisses et d'ordonnée $j-2$, puis $j-3$, ..., jusqu'à $j-k$, pour un certain k (voir figure 2.11-b). Nous avons fixé $k=4$, une valeur qui exclut les situations ambiguës et qui améliore encore légèrement le résultat final par rapport au choix d'un k plus faible. Pour $k=5$, ou plus, on peut rencontrer des configurations qui conduiraient au croisement de deux chaînes. Par exemple, sur la figure 2.11-c, le pixel X serait alors chaîné à B.

Par ailleurs, la cohérence structurelle est renforcée en s'assurant que la largeur d'appariement ne varie pas trop brusquement entre deux pixels consécutifs d'une même chaîne. On peut voir cette contrainte de cohérence comme une propriété de connexité en trois dimensions, la troisième étant la largeur d'appariement. Une variation brutale de la largeur est souvent la conséquence d'une erreur, lors de la sélection des contours ou lors de l'appariement, et il est préférable d'écartier tout de suite ces défauts. Il y a aussi de fortes variations de largeur aux extrémités des os. Nous rencontrons alors un défaut inhérent à l'appariement horizontal : aux extrémités d'un os nettement incliné, il se produit une déviation des pixels médians par rapport à l'axe principal de cet os, voir figures 2.9-a et 2.9-b. Le fait d'isoler ces pixels est une première étape du processus de correction qui nous permettra d'obtenir une bonne estimation de l'orientation de chaque os.

Nous avons fixé la limite de connexité à 8 pixels de variation de largeur entre deux pixels médians consécutifs. Pour justifier ce choix, on remarquera que, pour deux contours "parallèles", les fluctuations extrêmes dues à la discréétisation peuvent induire, entre deux lignes, une variation de largeur d'appariement qui peut atteindre 4 pixels. Par ailleurs, les variations naturelles de largeur de la diaphyse sont inférieures à un pixel par ligne. Le seuil de 8 que nous avons retenu,

⁶La convention usuelle d'adressage dans les images induit l'usage des ordonnées croissantes vers le bas.

		4	4	4		
	7	3	3	3	7	
6	6	2	2	2	6	6
5	5	1	1	1	5	5
		0				

Les pixels marqués de 1 à 7, examinés dans cet ordre, peuvent être l'extrémité d'une chaîne que l'on cherche à prolonger par le pixel qui est en case 0.

FIG. 2.12 – Priorités de chaînage pour un pixel médian

tolère ainsi les écarts de largeur “ordinaires”. Par contre, dans le cas qui se traduit par une déviation de l'axe, on observe des variations de largeur nettement plus rapides et qui conduisent donc à un changement de composante connexe.

Ainsi, on collecte tous les pixels médians pour former des composantes connexes qui sont des chaînes et qui vérifient de plus une certaine régularité de leur attribut (la largeur d'appariement). Le nombre de composantes connexes peut aller de quelques dizaines pour de “bonnes” images⁷ à quelques centaines pour des images plus bruitées. Pourtant, le nombre de chaînes intersectant une ligne de balayage quelconque est toujours très limité. Sur l'ensemble de nos images, il reste inférieur à 10 et il est de l'ordre de 3 en moyenne. L'échelle de calcul des contours a été justement déterminée dans ce but. De plus, les chaînes candidates à un prolongement peuvent être facilement maintenues par abscisses croissantes et, en particulier, elles ne se croisent pas.

Cette technique de chaînage met donc en œuvre des lois de proximité et de similitude qui interviennent souvent dans un groupement perceptuel. La proximité des pixels d'axe traduit à la fois la continuité et la bonne symétrie des contours. La similitude de largeur ajoute de la cohérence en privilégiant l'appariement de contours de même nature.

On exploite implicitement l'existence d'une direction privilégiée et d'une distance minimale entre chaînes. Ces particularités permettent une mise en œuvre par balayage qui est nettement plus rapide que les solutions plus générales comme celles basées sur la recherche de chemins optimaux dans un graphe [Mon71, Mar76]. Globalement, le coût du chaînage est approximativement linéaire dans le nombre de pixels médians (quelques milliers), c'est-à-dire très rapide, notamment en regard du coût des balayages de toute l'image nécessaires au calcul des contours. On remarquera surtout que les composantes connexes sont connues par ordonnées (puis par abscisses) croissantes, selon leur pixel de début et aussi selon leur pixel de fin. On mémorise ces deux ordres qui permettront d'accélérer les opérations de fusion qui suivent.

La construction que nous venons de décrire, favorise volontairement un prolongement assez proche de la verticale. On notera que le chaînage d'un axe orienté à $\pm \frac{\pi}{4}$ n'est réalisable que par connexité directe ($\Delta j = 1$ et $|\Delta i| = 1$), ce qui rend sa construction relativement plus fragile.

2.4.2 Extension de la proximité et filtrage

Pour rattraper des aléas de discréétisation sur les os relativement inclinés, il faut aussi accepter des voisins avec un décalage de 2, voire 3, pixels en abscisse. On voit bien ce cas, par exemple, pour le troisième os de la figure 2.9-b, page 45. De tels prolongements ne sont cependant envisagés que s'il n'y a pas de prolongement possible satisfaisant $|\Delta i| \leq 1$.

La figure 2.12 indique l'ordre de priorité que nous suivons pour la recherche d'une chaîne candidate à un prolongement par le pixel courant, celui qui est en case marquée 0. Il est plus simple de recourir à un traitement en deux passes.

⁷Pour notre image d'exemple, il y a 20 chaînes dont 12 parasites et 3 des axes sont coupés (voir annexe B.1).

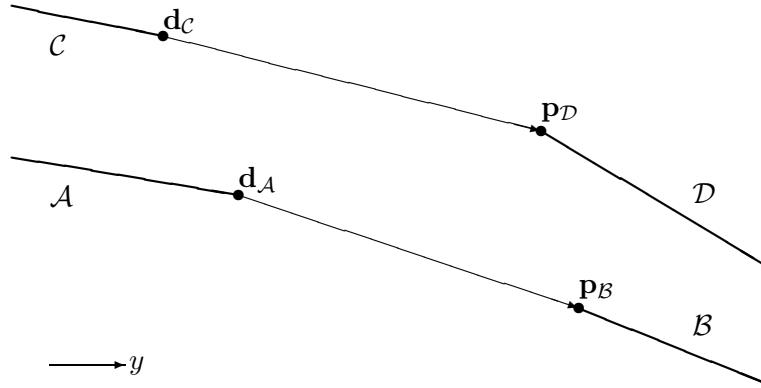


FIG. 2.13 – Raccordement naturel des chaînes d’axes

La première passe est la collecte “au vol” des pixels médians qui satisfaisent les contraintes $|\Delta i| \leq 1$ et $1 \leq \Delta j \leq 4$, ce qui s’obtient comme exposé en 2.4.1, en examinant successivement les cases marquées de 1 à 4. On favorise ainsi un chaînage vertical.

La seconde passe reprend les chaînes ainsi produites et fusionne celles qui satisfont les contraintes $2 \leq |\Delta i| \leq 3$, $\Delta j > 0$ et $|\Delta i| + \Delta j \leq 5$, entre les deux pixels respectivement à l’extrémité de l’une, cases marquées 5 à 7, et au début de l’autre, case 0. Cette seconde passe utilise de nouveau un balayage par ordonnées croissantes et le nombre de paires à examiner est ainsi approximativement linéaire en le nombre de chaînes.

On maintient bien sûr la contrainte sur les variations de la largeur d’appariement pour éviter de reconnecter des pixels incohérents de ce point de vue. Notons que, pour écarter toute ambiguïté de choix d’un prolongement, il faut maintenant s’assurer d’un espacement minimal de 7 entre les pixels médians d’une même ligne, soit 6 pixels intercalés. Il n’y a ainsi, au plus, qu’une seule extrémité de chaîne candidate sur chaque ligne de la figure 2.12.

Ensuite, on procède à la suppression des pixels isolés ($\Delta j > 1$ ou $|\Delta i| > 1$) et situés aux extrémités des chaînes. Un pixel isolé est pourtant conservé lorsqu’il contribue à faire le lien entre deux sous-chaînes consistantes. Ensuite, les chaînes considérées comme trop courtes (ayant arbitrairement moins de 10 pixels) sont éliminées. Ces opérations sont très efficaces pour écarter l’essentiel des petits parasites et on arrive ainsi à supprimer jusqu’à 10 % des pixels médians initiaux.

Le nombre des chaînes qui résultent des opérations de fusion et de filtrage est nettement réduit (on passe de 20 à 12 pour l’image d’exemple, voir figure B.1, page 179 en annexe) et ce d’autant plus que le nombre de composantes connexes initiales était élevé. Dans le cas d’images relativement bruitées, on passe ainsi de quelques centaines de composantes à nettement moins de la centaine.

Les opérations de filtrage contribuent aussi à écarter les pixels médians aux extrémités des os, puisque ceux-ci ont été isolés en petites composantes par la contrainte sur les variations de la largeur d’appariement. On peut alors considérer que les médians qui sont conservés sont proches des axes et on peut mesurer une largeur moyenne sur ces appariements. Bien qu’elle soit sensible à l’inclinaison, cette mesure, notée L_{ref} , va nous servir d’estimation de la “largeur” des doigts et elle sera affinée après restriction aux seuls axes des doigts (voir 2.4.7).

Par contre, ces opérations ne permettent pas de relier les chaînes d’un même axe, lorsque l’espacement entre celles-ci est relativement important, comme cela se produit fréquemment aux articulations. Pour décider de raccordements à “grande” distance, on ne doit plus raisonner sur les pixels, mais considérer une notion plus globale de “bonne” continuité entre les chaînes. En fait, on a déjà effectué ce changement de niveau conceptuel lors du filtrage sur la longueur.

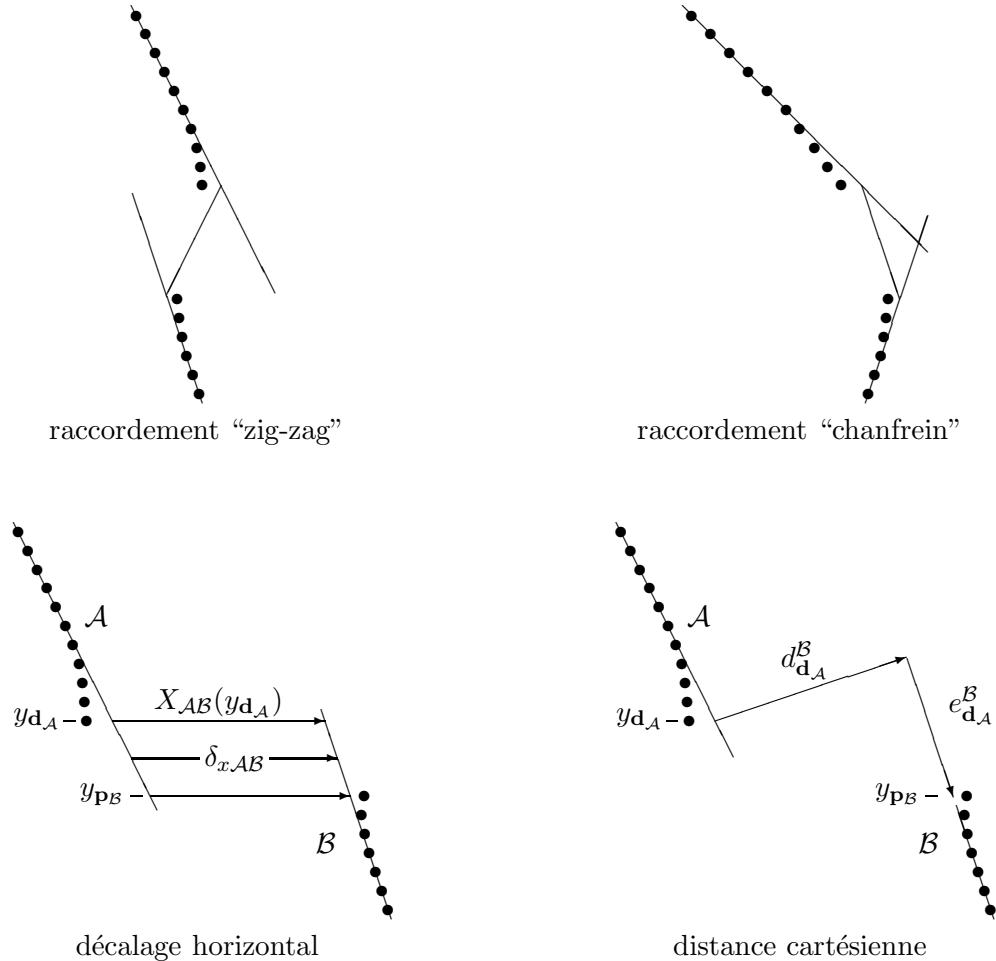


FIG. 2.14 – Continuité de deux chaînes d’axe

2.4.3 Continuité de deux chaînes

Avant de s’intéresser plus en détail à l’algorithme de raccordement, ce qui sera l’objet de 2.4.4 et 2.4.5, on a besoin d’un ordre de préférence entre des paires candidates et, pour cela, d’une mesure particulière de la continuité entre deux chaînes, au sens usuel en groupement perceptuel. Pour la suite, étant donnée une chaîne \mathcal{A} , on va noter $\mathbf{p}_\mathcal{A}$ son pixel proximal et $\mathbf{d}_\mathcal{A}$ son pixel distal, selon la terminologie anatomique.

En se référant à l’ordre de construction, par un balayage qui va du carpe vers le bout des doigts, $\mathbf{p}_\mathcal{A}$ et $\mathbf{d}_\mathcal{A}$ sont aussi, respectivement, le premier et le dernier pixel de la chaîne. Enfin, on rappelle que les ordonnées sont croissantes vers le bas de l’image et donc le long d’une chaîne.

Pour deux chaînes \mathcal{A} et \mathcal{B} , cela n’aurait guère de sens d’utiliser simplement la distance euclidienne $\|\mathbf{p}_\mathcal{B} - \mathbf{d}_\mathcal{A}\|$ ou, d’ailleurs, toute autre distance entre les pixels d’extrémité $\mathbf{d}_\mathcal{A}$ et $\mathbf{p}_\mathcal{B}$ qui sont en vis-à-vis. Par exemple, sur la figure 2.13, cela conduirait certainement à raccorder \mathcal{A} et \mathcal{D} . Pour obtenir de bons résultats, il est nécessaire de tenir compte également de la direction d’ensemble de chaque chaîne [Low87, MN92, DR92]. Néanmoins, le critère usuel de concordance de la direction n’est pas le plus pertinent et, comme le montre la figure 2.14, un raccordement “zig-zag” peut apparaître comme angulairement plus abrupt [Sau93, SU88] qu’un raccordement “chanfrein”.

Dans notre cas particulier, il s’agit aussi de faire abstraction des déviations causées aux extrémités des os par l’appariement horizontal (la figure 2.9, page 45, en donne quelques exemples concrets). Cela peut s’obtenir facilement à partir d’un calcul de régression linéaire sur les pixels

d'une chaîne. Comme la direction est assez proche de la verticale, on peut toujours exprimer la droite associée à une chaîne \mathcal{A} par une fonction linéaire de la forme $x = X_{\mathcal{A}}(y) = \bar{x}_{\mathcal{A}} + \lambda_{\mathcal{A}} \cdot (y - \bar{y}_{\mathcal{A}})$ (voir détail en annexe A.4). A priori, ce calcul est faussé par les pixels qui s'écartent de l'axe général, mais on a déjà isolé puis éliminé une bonne partie de ceux-ci et notamment les plus extrêmes. On peut donc considérer que l'on obtient ainsi une estimation convenable de la droite représentative de la chaîne \mathcal{A} , sans recourir à des techniques plus robustes [RL87].

Étant données deux chaînes \mathcal{A} et \mathcal{B} , on utilise une pseudo-distance de la forme :

$D(\mathcal{A}, \mathcal{B}) = \sqrt{\delta_{x\mathcal{AB}}^2 + \delta_{y\mathcal{AB}}^2}$, où $\delta_{x\mathcal{AB}}$ est une mesure du "décalage" horizontal entre les deux chaînes (voir figure 2.14, à gauche) et où $\delta_{y\mathcal{AB}}$ exprime leur espacement vertical. On va voir maintenant comment les directions estimées $\lambda_{\mathcal{A}}$ et $\lambda_{\mathcal{B}}$ interviennent dans ce calcul.

On peut exprimer le décalage horizontal comme une fonction : $X_{\mathcal{AB}}(y) = X_{\mathcal{B}}(y) - X_{\mathcal{A}}(y)$, évidemment variable le long de l'axe. La valeur $\delta_{x\mathcal{AB}} = \frac{1}{2} (X_{\mathcal{AB}}(y_{\mathbf{d}_{\mathcal{A}}}) + X_{\mathcal{AB}}(y_{\mathbf{p}_{\mathcal{B}}}))$ donne un rôle symétrique à \mathcal{A} et \mathcal{B} et on peut la considérer comme une mesure du décalage effectif moyen, bien sûr si $\mathbf{d}_{\mathcal{A}}$ et $\mathbf{p}_{\mathcal{B}}$ sont en vis-à-vis, autrement dit si \mathcal{B} est "en-dessous" de \mathcal{A} . Par linéarité, on peut aussi l'écrire sous la forme $\delta_{x\mathcal{AB}} = X_{\mathcal{AB}}\left(\frac{y_{\mathbf{d}_{\mathcal{A}}} + y_{\mathbf{p}_{\mathcal{B}}}}{2}\right)$, ce qui simplifie encore son calcul.

On notera que cette mesure de décalage est nulle quand les deux chaînes sont parfaitement "alignées" ($X_{\mathcal{B}} = X_{\mathcal{A}}$) ou si leurs droites sont convergentes en $y = \frac{y_{\mathbf{d}_{\mathcal{A}}} + y_{\mathbf{p}_{\mathcal{B}}}}{2}$. Plus généralement, cette mesure est faible dans des raccordements de type "chanfrein". Par exemple, pour \mathcal{A} et $X_{\mathcal{B}}(\mathbf{p}_{\mathcal{B}})$ fixés et en faisant varier le paramètre $\lambda_{\mathcal{B}}$ (la direction de \mathcal{B}), on voit facilement comment $\delta_{x\mathcal{AB}}$ traduit la notion intuitive du "bon raccordement" des deux chaînes. D'autre part, dans le cas de chaînes "parallèles", on notera que le décalage $\delta_{x\mathcal{AB}}$ peut rester relativement important. C'est ce qui permet d'obtenir une séparation correcte entre les chaînes relatives à des doigts différents et qui justifie donc que l'on fasse de $\delta_{x\mathcal{AB}}$ le terme prépondérant dans $D(\mathcal{A}, \mathcal{B})$.

Enfin, une prise en compte de l'espacement vertical est nécessaire, au moins pour déterminer une relation d'ordre cohérente entre les chaînes constitutives d'un même axe. On utilise pour cela l'expression $\delta_{y\mathcal{AB}} = k_y \cdot (y_{\mathbf{p}_{\mathcal{B}}} - y_{\mathbf{d}_{\mathcal{A}}})$, avec un coefficient $0 < k_y < 1$ qui pondère son importance comparée à $\delta_{x\mathcal{AB}}$. La valeur de k_y a été fixée empiriquement à 0.4, ce qui permet de tolérer des "coupures" d'axes assez importantes. Le prédictat $\delta_{y\mathcal{AB}} > 0$ sert aussi naturellement à caractériser " \mathcal{B} en-dessous de \mathcal{A} ".

Dans le cas de chaînes nettement inclinées, il serait plus rigoureux de remplacer le décalage horizontal $X_{\mathcal{AB}}(y_{\mathbf{d}_{\mathcal{A}}})$ par la distance $d_{\mathbf{d}_{\mathcal{A}}}^{\mathcal{B}}$ entre le point $(X_{\mathcal{A}}(y_{\mathbf{d}_{\mathcal{A}}}), y_{\mathbf{d}_{\mathcal{A}}})$ et la droite d'équation $x = X_{\mathcal{B}}(y)$ (voir figure 2.14, à droite). De même, au lieu de l'espacement vertical $y_{\mathbf{p}_{\mathcal{B}}} - y_{\mathbf{d}_{\mathcal{A}}}$, on devrait considérer la distance $e_{\mathbf{d}_{\mathcal{A}}}^{\mathcal{B}}$, entre, d'une part, la projection orthogonale de $(X_{\mathcal{A}}(y_{\mathbf{d}_{\mathcal{A}}}), y_{\mathbf{d}_{\mathcal{A}}})$ sur la droite $x = X_{\mathcal{B}}(y)$ et, d'autre part, le point $(X_{\mathcal{B}}(y_{\mathbf{p}_{\mathcal{B}}}), y_{\mathbf{p}_{\mathcal{B}}})$. Enfin, pour symétriser, il faudrait faire de même en échangeant les rôles de \mathcal{A} et \mathcal{B} et considérer la demi-somme de ces deux expressions. Ce calcul est, de toute évidence, nettement plus lourd que celui de $\delta_{x\mathcal{AB}}$ et $\delta_{y\mathcal{AB}}$ et on constate qu'il n'apporte pas d'amélioration notable dans la construction des axes. On trouvera une justification de cette équivalence approximative en annexe A.4.

En ce qui concerne les largeurs d'appariement, que ce soient les valeurs prises aux extrémités ou des moyennes, il semble difficile de les utiliser dans l'expression de la distance entre deux chaînes. En effet, on observe souvent des variations très importantes de la largeur d'appariement entre deux chaînes consécutives le long d'un même axe. Cela provient de la grande diversité des formes des contours et c'est ce qui a parfois causé une césure indésirable pendant le chainage. Prendre en compte à nouveau les largeurs d'appariement, même légèrement, aurait finalement tendance à contrarier l'évidence géométrique des alignements.

On retient donc l'expression de la forme $D(\mathcal{A}, \mathcal{B}) = \sqrt{\delta_{x\mathcal{AB}}^2 + \delta_{y\mathcal{AB}}^2}$, définie comme ci-dessus, sachant qu'en pratique on pourra éviter le calcul de la racine carrée.

2.4.4 Regroupement des chaînes

La formulation du regroupement perceptuel par continuité comme un problème de graphe est assez classique [Low85, SU88]. Dans un tel graphe, chaque sommet est une primitive (ici une chaîne) et les arcs, valués par une mesure de la continuité perceptive entre deux primitives, sont des regroupements à envisager.

Dans notre cas, on peut ainsi considérer le graphe orienté $G = (S, A)$ défini comme suit, pour un ensemble de chaînes C :

- les sommets du graphe sont les chaînes $S = C$,
- les arcs décrivent les raccordements potentiels entre une extrémité distale et une extrémité proximale : $A = \{(\mathcal{A}, \mathcal{B}), \forall \mathcal{A}, \mathcal{B} \in C \text{ tq } \delta_{y,\mathcal{AB}} > 0 \text{ et } D(\mathcal{A}, \mathcal{B}) < D_{\max}\}$, pour un certain D_{\max} .

La construction des axes peut alors s'envisager de manière assez naturelle comme la recherche d'un sous-ensemble de A satisfaisant une certaine forme d'optimalité.

Auparavant, il convient de discuter du choix du seuil D_{\max} qui va jouer sur la qualité effective du résultat. C'est un problème indépendant du choix de l'algorithme, même s'il détermine aussi $|A|$ et donc le coût de construction des axes.

Le choix d'une constante D_{\max} relativement faible a comme avantage notable de contribuer à réduire le nombre de couples $(\mathcal{A}, \mathcal{B})$ à examiner et donc le coût de l'algorithme de raccordement. Du fait des contraintes sur le nombre de pixels médians par ligne et du fait aussi d'une certaine régularité entre les lignes, le nombre de chaînes \mathcal{B} qui, pour \mathcal{A} fixée, satisfont $\delta_{y,\mathcal{AB}} > 0$ et $D(\mathcal{A}, \mathcal{B}) < D_{\max}$, sera en général très limité. Sur l'ensemble de nos images, avec les valeurs de paramètres que nous avons choisies, on observe $|A| < 4|S|$, au lieu de $O(|S|^2)$ sans la contrainte par D_{\max} , sachant que $|S|$ peut être de l'ordre de la centaine.

Enfin, notons que par construction (voir 2.4.1), on dispose des chaînes à la fois dans l'ordre des y_{d_A} croissants, dans l'ordre des y_{p_B} croissants et aussi par abscisses croissantes car elles ne se croisent pas. Ces particularités sont mises à profit, lors d'un balayage vertical de l'ensemble des chaînes, pour énumérer rapidement les couples $(\mathcal{A}, \mathcal{B})$ potentiels. Le coût effectif de ce balayage approche ainsi un comportement quasi linéaire en le nombre de chaînes.

La valeur D_{\max} limite la distance acceptable entre deux chaînes consécutives d'un même axe et son choix est donc constraint essentiellement par la recherche d'une certaine robustesse envers les interruptions "naturelles" de cet axe. On va relativiser cela par rapport à la "largeur" d'un doigt, la seule mesure un peu pertinente et disponible à ce stade du traitement. L'expression retenue est alors de la forme $D_{\max} = k_D L_{\text{ref}}$ où k_D est un paramètre à fixer et où L_{ref} est l'estimation de la largeur moyenne d'appariement que l'on a déterminée en 2.4.2. On voit que dans le cas de deux chaînes bien alignées ($\delta_{x,\mathcal{AB}} = 0$), l'égalité limite $D(\mathcal{A}, \mathcal{B}) = D_{\max}$ nous donne $\delta_{y,\mathcal{AB}} = D_{\max} = k_D L_{\text{ref}}$, soit $y_{p_B} - y_{d_A} = \frac{k_D}{k_y} \cdot L_{\text{ref}}$.

La valeur de k_D a été fixée empiriquement à 1.5, soit $\frac{k_D}{k_y} = 3.75$, ce qui permet de récupérer des coupures d'axe relativement conséquentes, comme on peut en rencontrer entre deux diaphyses sur des mains très jeunes (voir les figures 3 et 4 de l'introduction, pages 8 et 9). La valeur $k_D = 1$ s'avère trop faible, on constate effectivement qu'elle conduit au maintien d'un certain nombre de coupures d'axes. À l'inverse, avec des valeurs comme $k_D = 2$ ou plus, on observe trop souvent des prolongements exagérés des axes, vers des artefacts du bas de l'image et vers les os du carpe, voire de l'avant-bras. On évite donc de tels prolongements qui vont nuire à l'identification.

L'autre cas extrême, $D(\mathcal{A}, \mathcal{B}) = D_{\max}$ avec $\delta_{y,\mathcal{AB}} = 0$, nous donne $|\delta_{x,\mathcal{AB}}| = k_D L_{\text{ref}}$. Or, dans le cas de doigts serrés, le décalage horizontal entre deux axes voisins peut s'approcher très près de L_{ref} . Autrement dit, avec k_D de l'ordre de 1.5, la contrainte $D(\mathcal{A}, \mathcal{B}) < D_{\max}$ n'exclut pas le raccordement potentiel entre deux chaînes appartenant respectivement à l'un et à l'autre de ces deux axes. Il faut se rendre à l'évidence qu'un seuillage de D ne suffit pas pour assurer à la fois une séparation correcte entre deux axes voisins et la continuité de chaque axe.

2.4.5 Algorithme de construction d'axes

Une technique de regroupement reconnue [SB98, SM97, PF98] consiste à itérer la recherche d'une coupe optimale du graphe pour isoler une composante à chaque fois. On dispose pour cela d'un algorithme approché efficace [PSL90, GL96] qui est basé sur le calcul des principales valeurs propres et des vecteurs propres associés de la matrice d'adjacence binaire (ou encore de sa matrice laplacienne) obtenue par le seuillage de la distance entre les primitives. Le vecteur propre principal (le second pour la matrice laplacienne) donne alors la composante principale, un principe que l'on peut itérer. Néanmoins, une telle approche détermine seulement un groupe de primitives connexes, sans imposer de contrainte de topologie linéaire. Il en est de même pour des solutions par programmation dynamique [Mon71, SU88] ou d'autres techniques de construction de chemins [Mar76] qui ne gèrent pas spécifiquement l'existence de "fourches". Dans notre cas, il s'agit de construire un ensemble de chemins concurrents que l'on doit maintenir disjoints.

Pour contraindre la structure linéaire des axes, nous allons considérer le graphe biparti $G = (S, A)$ défini comme suit, pour un ensemble de chaînes C :

- les sommets du graphe sont, d'une part les extrémités distales des chaînes et, d'autre part leurs extrémités proximales : $S = S_d \cup S_p$, $S_d = \{\mathbf{d}_A, \forall A \in C\}$, $S_p = \{\mathbf{p}_A, \forall A \in C\}$,
- les arcs décrivent les raccordements potentiels entre une extrémité distale et une extrémité proximale : $A = \{(\mathbf{d}_A, \mathbf{p}_B), \forall A, B \text{ tq } \delta_{y,AB} > 0 \text{ et } D(A, B) < D_{\max}\}$, pour un certain D_{\max} .

La construction des axes peut alors s'envisager comme la recherche, dans le graphe G , d'un couplage optimal entre S_d et S_p . Il existe pour cela divers algorithmes [GM95, JM93] relativement efficaces (polynomiaux en $|S|$ et $|A|$) et une adaptation particulière à notre cas sera développée.

Une première idée pour construire effectivement des axes est de raccorder les couples de chaînes (A, B) qui sont réciproquement les plus proches, autrement dit, qui vérifient :

$$\begin{aligned} D(A, B) &< D_{\max}, \quad \delta_{y,AB} > 0, \\ D(A, B) &< D(C, B), \quad \forall C \in C \text{ tq } C \neq A \text{ et } \delta_{y,CB} > 0, \\ D(A, B) &< D(A, D), \quad \forall D \in C \text{ tq } D \neq B \text{ et } \delta_{y,AD} > 0. \end{aligned}$$

Par exemple, dans un cas comme celui de la figure 2.13, page 51, la "distance" $D(A, B)$ est inférieure aux deux distances croisées $D(A, D)$ et $D(C, B)$, ce qui valide le raccordement de A vers B . Généralement, cette infériorité sera même très nette, D étant définie dans ce but, et cela confère une grande robustesse aux décisions de raccordement. Les raccordements ainsi détectés sont en pratique toujours satisfaisants, au sens où ils relient effectivement des chaînes consécutives sur un même axe, du moins en ce qui concerne les métacarpiens et les phalanges. Des cas plus ambigus s'observent seulement dans le carpe ou avec les artefacts périphériques.

Mais ce type de raccordement ne suffit pas pour construire complètement les axes. Par exemple, toujours dans le cas de la figure 2.13, si l'on a $D(A, D) < D(C, D)$, cela interdit le raccordement de C vers D . Cependant, ce raccordement peut s'obtenir par une seconde passe sur le graphe G privé des deux sommets \mathbf{d}_A et \mathbf{p}_B et des arcs liés à ces sommets, c'est-à-dire :

$$\begin{aligned} G_{/AB} &= (S_{/AB}, A_{/AB}), \text{ avec} \\ S_{/AB} &= S - \{\mathbf{d}_A\} - \{\mathbf{p}_B\} \text{ et} \\ A_{/AB} &= A - \{(\mathbf{d}_A, \mathbf{p}_D) \in A, \forall D \in C\} - \{(\mathbf{d}_C, \mathbf{p}_B) \in A, \forall C \in C\}. \end{aligned}$$

Plus généralement, on doit itérer et le critère d'arrêt est qu'il ne reste plus d'arc "inférieur" à D_{\max} . On aboutit ainsi à la construction des axes intuitivement attendus en quelques itérations, pour toutes les images de notre base de test. Il faut noter que si le nombre d'itérations devenait trop important, ce serait le signe de données confuses et de raccordements loin d'être évidents.

Pour optimiser ce processus itératif, il faut mémoriser, pour chaque sommet \mathbf{d}_A , l'ensemble $\{(\mathbf{d}_A, \mathbf{p}_D) \in A, \forall D \in C\}$ et, pour chaque sommet \mathbf{p}_B , l'ensemble $\{(\mathbf{d}_C, \mathbf{p}_B) \in A, \forall C \in C\}$. Cela peut se faire pendant l'énumération des arcs. Chaque ensemble doit être ensuite ordonné selon la distance D croissante, ce qui représente $|S|$ petits tris, de quelques éléments à chaque fois. Cela peut se résoudre encore plus efficacement par une adaptation de l'algorithme de Gale et Shapley, dédié au problème classique des “mariages stables” [GS62, GI89].

Notre adaptation, très simple, consiste à ne considérer que les arcs de A , c'est-à-dire à utiliser des “listes de préférences” incomplètes, au lieu des $\left(\frac{|S|}{2}\right)^2$ arcs dans l'algorithme de couplage standard. Sans compter les tris préliminaires, on aboutit ainsi à un coût en $O(|A| + |S|)$, puisque dans cet algorithme chaque arc n'est considéré qu'une seule fois avant d'être éventuellement écarté. En pratique, $|A|$ est très nettement inférieur à $\left(\frac{|S|}{2}\right)^2$ qui serait le coût de cet algorithme sans la contrainte par D_{\max} .

D'autre part, il faut rappeler que l'algorithme de Gale et Shapley est en général assymétrique, c'est-à-dire que le choix d'un sens de résolution (par ex. *les hommes proposent, les femmes disposent*) détermine une solution alors que le sens inverse détermine une autre solution. Dans notre cas, le problème est clairement symétrique (en échangeant les rôles de $S_{\mathbf{d}}$ et $S_{\mathbf{p}}$) car les listes de préférences respectent l'ordre global induit par la mesure D sur les arcs, en supposant un ordre strict (ie. D ne donne pas deux fois la même valeur pour deux paires distinctes). Ainsi les premiers choix sont les mêmes pour les deux sens de résolution et ainsi de suite.

La taille des tris préliminaires ou, ce qui est à peu près équivalent, le nombre d'itérations dans la version naïve sont très faibles en pratique. Pour estimer une borne supérieure, on notera simplement que le coût global des $|S|$ petits tris est théoriquement majoré par le coût d'un tri de l'ensemble de tous les arcs. Pourtant, compte tenu des valeurs respectives de $|A|$ et $|S|$, la comparaison sera, en pratique, plutôt à l'avantage du “gros” tri.

Notons maintenant que, si l'on réalise effectivement le tri des arcs selon D croissant, alors on peut facilement construire les axes en s'inspirant de l'algorithme glouton de Kruskal, destiné usuellement au calcul d'un arbre ouvrant de poids minimal [Kru56]. On commence par les arcs les plus “courts” et un arc n'est retenu que s'il relie deux sommets \mathbf{d}_A et \mathbf{p}_B non encore reliés à d'autres. En pratique, on n'a pas besoin de distinguer les sommets \mathbf{p}_A et \mathbf{d}_A et on peut travailler directement sur des chaînes ayant chacune une référence à une chaîne précédente et une référence à une chaîne suivante. Il faut aussi noter que, par la définition des arcs qui impose une croissance stricte en y , nous sommes dispensés de gérer l'acyclicité comme dans l'algorithme de Kruskal.

On a ainsi un algorithme très simple :

Les deux extrémités \mathbf{p}_A et \mathbf{d}_A de chaque chaîne \mathcal{A} sont marquées initialement “libres”.

Construire une file avec les arcs ordonnés par distance croissante.

Tant que la file n'est pas vide :

- sortir l'arc en tête de file, soit $(\mathbf{d}_A, \mathbf{p}_B)$ cet arc,
- si les extrémités \mathbf{d}_A et \mathbf{p}_B sont toutes deux “libres”
 - relier \mathbf{d}_A et \mathbf{p}_B (ces extrémités ne seront alors plus “libres”),
 - sinon cet arc est ignoré.

Les tests des extrémités et les opérations de raccordements se font en $O(1)$, soit $O(|A|)$ pour vider la file. Il faut noter que c'est un “vrai” $O(1)$, notre cas est plus simple que “union-find”, car nous construisons des structures purement linéaires. Le coût de l'algorithme complet est donc en $O(|A| \cdot \log |A|)$, du fait du tri préliminaire. L'équivalence avec les algorithmes précédents se prouve facilement par l'absurde, dans le même style que la preuve de l'algorithme de Kruskal. Compte tenu des dimensions de notre problème $|A| < 200$ et de sa simplicité intrinsèque (la

structure de graphe est inutile, un tri, une boucle), ce dernier algorithme est au moins aussi efficace que les précédents. Un autre intérêt est son principe clairement incrémental qui permet, si le regroupement est insuffisant, d'augmenter dynamiquement D_{\max} . Nous verrons en 2.4.6 qu'il est aussi utile, parfois, de pouvoir prendre en compte des arcs qui ne respectent pas la stricte monotonie en y .

Une autre approche envisageable serait de chercher un couplage maximal (en poids et en cardinal) dans G , après avoir fixé un poids pour chaque arc, par exemple, $D_{\max} - D(\mathcal{A}, \mathcal{B})$ pour le poids de l'arc $(\mathbf{d}_{\mathcal{A}}, \mathbf{p}_{\mathcal{B}})$. C'est alors équivalent à la recherche d'un flot maximal et on dispose pour cela d'algorithmes du type Ford et Fulkerson et dérivés [FF62, Kuh55], au mieux en $O(|S| \cdot |A|)$ par la méthode hongroise. Compte tenu des dimensions de notre problème, ces algorithmes sont assez coûteux, sans garantie de convergence. Enfin, la mise en œuvre est relativement complexe.

Mais l'inconvénient essentiel de cette dernière approche est que l'on n'obtient pas le résultat escompté, car l'objectif du couplage maximal détermine seulement une optimisation globale. Par exemple, dans le cas de la figure 2.13, page 51, le raccordement naturel n'est assuré que si les “distances” (non euclidiennes) vérifient en plus : $D(\mathcal{A}, \mathcal{B}) + D(\mathcal{C}, \mathcal{D}) < D(\mathcal{A}, \mathcal{D}) + D(\mathcal{C}, \mathcal{B})$. Dans le cas contraire, on obtiendrait un croisement indésirable des axes.

L'algorithme “à la Kruskal” est donc celui que nous avons effectivement retenu. Les axes sont ainsi assemblés progressivement, tant qu'il existe un raccordement localement préférable. Cet algorithme, avec une définition de la “distance” D bien adaptée, réalise une séparation correcte entre les axes. Le choix du seuil de distance D_{\max} est ensuite tel que toutes les chaînes d'un doigt sont normalement raccordées les unes à la suite des autres. Enfin, l'échelle de détection ayant été éventuellement ajustée pour cela, les pixels médians sont en nombre suffisant pour composer 5 axes de bonne longueur et une bonne organisation structurelle est alors inhérente au contenu de l'image.

Il arrive pourtant qu'un axe reste coupé en deux morceaux et, d'autre part, qu'une chaîne parasite reste isolée ou, au contraire, prolonge exagérément un axe. Nous allons voir maintenant comment retoucher la construction des axes pour obtenir effectivement 5 axes bien identifiables.

2.4.6 Cas des axes coupés

Le cas d'un axe “en deux morceaux” reste assez fréquent⁸ et l'origine de ce problème se trouve dans divers “défauts” de l'image source. Il peut s'en suivre des aléas de détection des contours qui se propagent à la construction des points médians et qui perturbent finalement le bon alignement des chaînes.

La figure 2.15 illustre bien ce genre de problèmes. Dans cet exemple, la présence d'une chaîne \mathcal{B} “décalée” est due à une interruption du contour gauche de l'os, ce qui s'explique par un amincissement de la calcification tubulaire. L'algorithme d'appariement construit alors des points médians qui s'appuient à gauche sur le contour extérieur, disons la “peau”. Pour les chaînes \mathcal{A} et \mathcal{C} , le contour “os” est bien présent et il alors logiquement préféré. D'autre part, on constate que le contour gauche de \mathcal{D} passe progressivement, en remontant, d'un contour “os” à un contour “peau”. En effet, le contour “os” disparaît au niveau de l'articulation et une transition continue est réalisée par la régularisation qui est intégrée au calcul du gradient.

Lors de la construction des axes, les raccordements respectifs de \mathcal{C} à \mathcal{A} et de \mathcal{B} à \mathcal{D} sont alors naturels et incontestables du fait des alignements et ce, pour n'importe quel choix raisonnable de D et D_{\max} . Pour chacun de ces deux raccordements, on notera aussi la parfaite cohérence de type des contours et des règles d'appariements qui déterminent les points médians qui sont

⁸Cela concerne environ 7 % des bonnes images.

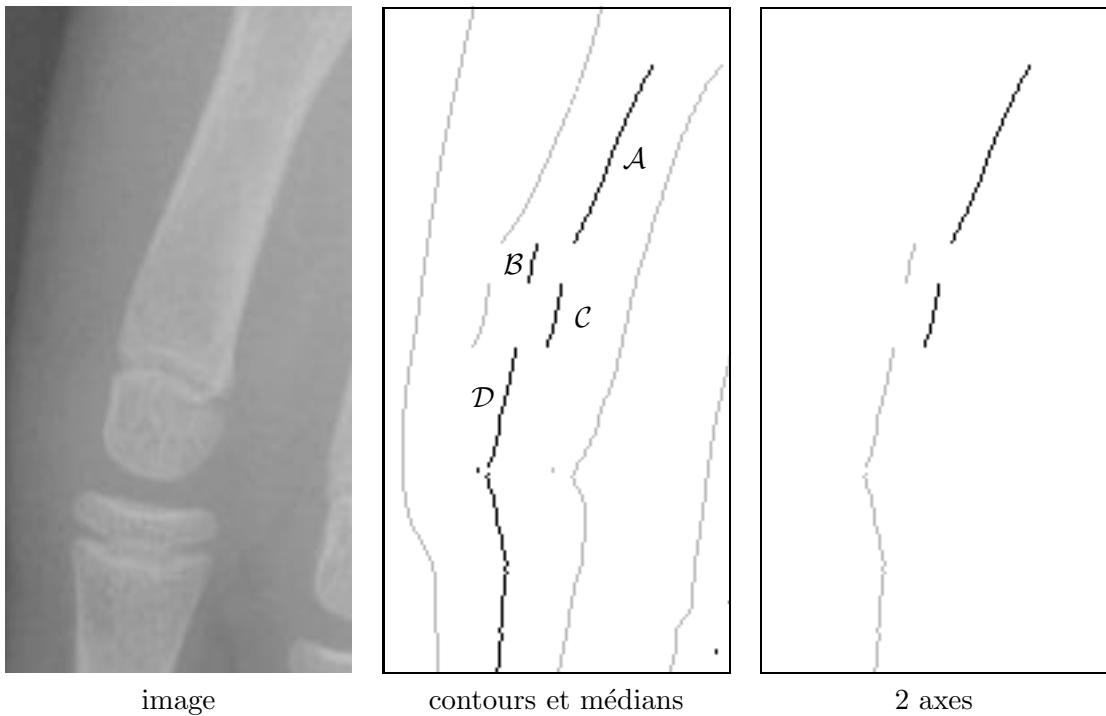


FIG. 2.15 – Exemple de raccordement difficile

en vis-à-vis. On voit ainsi toute l’ambiguïté d’une telle situation qui conduit à un blocage de la construction d’un axe et que l’on ne peut pas prévenir, sans remettre en cause une procédure qui s’avère, en général, tout à fait appropriée et très robuste.

Il est en fait relativement simple de corriger ce problème par un traitement spécifique a posteriori. Dans un premier temps, on va devoir décider du raccordement de \mathcal{B} “en dessous” de \mathcal{C} . Notons que c’est le seul raccordement encore possible, compte tenu des extrémités libres. Ensuite, il faudra rétablir la monotonie en y le long de l’axe.

Notre algorithme de raccordement doit donc considérer aussi les arcs “rétrogrades”, comme $(\mathcal{C}, \mathcal{B})$ de l’exemple qui vérifie bien $D(\mathcal{C}, \mathcal{B}) < D_{\max}$ mais tel que $\delta_{yCB} \leq 0$. Le nombre des arcs rétrogrades est a priori comparable à $|A|$, celui des arcs “ordinaires”. Ce dernier, en pratique de l’ordre de 4 fois le nombre de chaînes, est lié au choix de D_{\max} . Pour ne pas perturber la construction “normale”, les arcs rétrogrades ne sont pris en compte qu’après avoir considéré tous les arcs ordinaires, ce qui revient à leur donner une priorité de la forme $D(\mathcal{C}, \mathcal{B}) + D_{\max}$, par exemple. Ainsi, presque tous les arcs rétrogrades sont finalement ignorés, car au moins une de leur extrémités n’est plus libre au moment de leur prise en compte. Une alternative est de ne tester que les extrémités des axes obtenus avec les arcs ordinaires. On a alors un coût quadratique dans le nombre d’axes, en pratique peu différent du coût linéaire de l’ordre de $|A|$.

Il faut noter qu’en relâchant la contrainte de monotonie en y , on a réintroduit le risque de construire des cycles. Dans notre exemple, on voit que le problème se pose si l’on a également $D(\mathcal{D}, \mathcal{A}) < D_{\max}$. En s’appuyant sur une relation d’ordre strict entre les axes, on évite facilement la construction d’un cycle. On ne pourra donc raccorder l’axe $\mathcal{B} \dots \mathcal{D}$ à la suite de l’axe $\mathcal{A} \dots \mathcal{C}$ que si les extrémités des chaînes constitutives satisfont la double contrainte d’ordre $y_{p_A} < y_{p_B}$ et $y_{d_C} < y_{d_D}$, contrainte à laquelle s’ajoute la caractérisation d’un arc rétrograde $y_{p_B} \leq y_{d_C}$.

Autrement dit, on ne s’autorise un raccordement rétrograde entre deux axes que si le premier commence avant l’autre et se termine aussi avant. La contrainte d’ordre interdit notamment le raccordement d’un axe plus court qui commencerait après et se finirait avant. Par exemple, s’il n’y avait pas la chaîne \mathcal{D} à la suite de \mathcal{B} , on ne chercherait pas à raccorder \mathcal{B} seul à $\mathcal{A} \dots \mathcal{C}$.

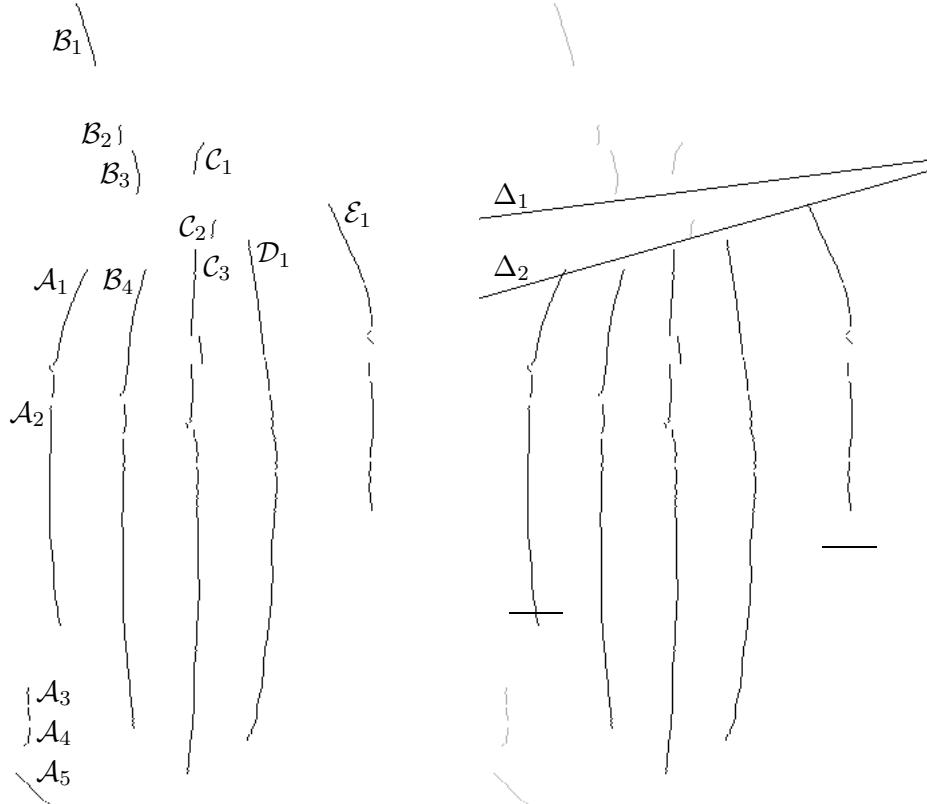


FIG. 2.16 – Découpage des chaînes parasites

C'est effectivement l'allongement permis par la présence de \mathcal{D} , et aussi réciproquement \mathcal{A} , qui motive le raccordement.

Pourachever un raccordement rétrograde, il reste à fusionner le détail des deux axes de manière à rétablir la monotonie stricte en y . Il s'agit d'une opération de fusion très classique qui ne présente pas de difficulté majeure. Le cas délicat et ambigu, celui de deux pixels d'axes de même ordonnée (un sur chaque chaîne), est généralement évité par les contraintes d'appariement. Les exceptions notables qui peuvent affecter l'axe d'un doigt, concernent moins de 1% des bonnes images. Il s'agit du cas de deux appariements de part et d'autre du canal médullaire d'un métacarpien qui serait trop large pour valider l'appariement central. Il n'y a alors pas de règle absolue qui permettrait de préférer l'un ou l'autre des deux pixels. Une heuristique acceptable consiste à privilégier une certaine forme d'alignement en se basant sur les calculs de régression qui permettent d'exprimer la distance des pixels aux axes.

2.4.7 Détection des axes des doigts

Pour finaliser la détection des axes, on procède ensuite à diverses opérations de nettoyage. Cela consiste tout d'abord à éliminer d'éventuels axes parasites en se basant sur des critères simples de consistance, par exemple, trop bas dans l'image, pas assez de pixels, etc.

Il n'y a pas de difficulté particulière pour fixer les seuils de ces critères, puisque les traitements qui précèdent ont justement favorisé l'évidence d'axes longs. S'il n'était pas possible de conserver au moins 5 axes, ce serait l'indication d'un grave manque de cohérence des pixels médians car ceux-ci sont a priori en nombre approprié pour former ces axes. Sur de "bonnes" images, cela ne se produit effectivement pas. À l'inverse, il est possible de conserver plus de 5 axes, notamment ceux de l'avant-bras. La figure B.1, page 179 en annexe, montre les 6 axes obtenus pour notre

exemple. Le choix final des 5 axes des doigts est relativement simple, si on ne cherche pas encore à les identifier un à un.

D'autre part, on constate que quelques chaînes parasites (les axes des os du carpe, voire de l'avant-bras ou des artefacts) viennent prolonger exagérément les axes des doigts. La figure 2.16 donne un exemple de ces divers problèmes. On y voit que l'axe $\mathcal{A}_1 \dots \mathcal{A}_2$ (l'auriculaire) se prolonge par trois chaînes parasites $\mathcal{A}_3 \dots \mathcal{A}_5$. L'axe $\mathcal{B}_1 \dots \mathcal{B}_4 \dots$ passe successivement par le cubitus, l'os pyramidal, l'os crochu puis l'annulaire. Le cas de l'axe du grand os, ici \mathcal{C}_1 , qui prolonge l'axe du majeur $\mathcal{C}_3 \dots$ est très fréquent, on le voit aussi sur la figure B.1 en annexe.

Pour corriger ces problèmes, il s'agit de détecter les axes trop longs et de les recouper de manière appropriée. Comme la structure fine d'une chaîne traduit une très forte cohérence des contours, on conserve l'intégralité d'une chaîne qui serait touchée par une limite a priori, et on ne s'autorise à recouper qu'entre deux chaînes consécutives sur un axe. Le but est alors d'établir une cohérence globale entre les extrémités des 5 axes. Notons que c'est la première utilisation d'un tel niveau de structure, les étapes précédentes⁹ ne traitaient que de propriétés relativement locales (contours, appariements, connexité des médians, chaînes proches).

Pour réaliser la séparation avec le carpe, on va considérer que les extrémités proximales des métacarpiens sont alignées, ce qui constitue une assez bonne approximation de la réalité. Il s'agit donc de déterminer une droite telle que les axes coupant cette droite sont manifestement trop longs et doivent être recoupés. Pour plus de robustesse, on détermine cette droite en deux itérations.

Dans un premier temps, on cherche à placer une droite passant dans le carpe et approximativement perpendiculaire à l'axe de la main. On utilise pour cela quelques heuristiques assez voisines de celles qui ont servi pour l'orientation initiale (voir 2.1.1). Mais la construction présente est plus précise puisque l'on peut maintenant se baser sur l'extrémité distale du majeur¹⁰ et, éventuellement, les axes de l'avant-bras¹¹. Dans l'exemple de la figure 2.16, la droite Δ_1 qui est ainsi obtenue, permet déjà de couper entre \mathcal{B}_3 et \mathcal{B}_4 et entre \mathcal{C}_1 et \mathcal{C}_2 . Notons aussi qu'elle intersecte \mathcal{E}_1 , cette chaîne est alors conservée entière.

Une seconde droite est alors obtenue par régression linéaire sur les extrémités proximales des axes ainsi raccourcis. Dans l'exemple de la figure 2.16, la droite Δ_2 est donc basée sur les extrémités proximales des chaînes \mathcal{A}_1 , \mathcal{B}_4 , \mathcal{C}_2 , \mathcal{D}_1 et \mathcal{E}_1 . Cette droite permet de recouper entre \mathcal{C}_2 et \mathcal{C}_3 . Une telle construction est très robuste envers les prolongements parasites puisque ceux-ci ne concernent, au pire, qu'un ou deux axes qui ont été déjà recoupés et qui ne s'écartent donc pas trop de l'alignement recherché. Cette construction est aussi relativement robuste dans le cas où l'on aurait un axe trop court, dans la mesure où la connexité des chaînes permet de rattraper, pour les autres axes, le mauvais placement de la droite Δ_2 qui en résulte.

Il faut également traiter le cas de chaînes parasites qui pourraient se trouver à l'extrémité distale des axes, comme $\mathcal{A}_3 \dots \mathcal{A}_5$ dans l'exemple. Avec des images relativement bien cadrées, on peut considérer que de telles chaînes ne peuvent pas se trouver dans le prolongement du majeur. En se basant sur la longueur que l'on a mesurée pour le majeur, on peut ainsi estimer celle d'un autre axe. Le tableau 2.1, page 76, tiré de données de normalité [Poz74], nous donne des rapports entre les longueurs des axes.

Notons qu'à ce stade, il n'est pas encore possible d'orienter l'image, c'est-à-dire, de décider de quel côté se trouvent le pouce, l'index, etc. Pour l'index et l'annulaire, sans les distinguer, on prend une longueur de référence égale à celle du majeur. On fait donc une légère surestimation ($\frac{\mu_2}{\mu_3} = \frac{\mu_4}{\mu_3} \approx 0.92$), normalement sans importance.

⁹Mis à part le comptage des pixels médians que l'on a utilisé en 2.3.4 pour valider l'échelle, mais il s'agissait seulement de cohérence statistique.

¹⁰Une identification robuste du majeur est que son extrémité distale est celle d'abscisse médiane.

¹¹On tente pour cela d'identifier une ou deux chaînes relativement consistantes dans le haut de l'image.

Pour le pouce et l'auriculaire, de longueurs différentes, on doit faire un compromis entre les rapports respectifs ($\frac{\mu_1}{\mu_3} \approx 0.58$ et $\frac{\mu_5}{\mu_3} = 0.75$). Un rapport de valeur 0.70 conduit à une surestimation déjà très importante de la longueur du pouce, avec un risque notable de conserver des chaînes parasites. La longueur de l'auriculaire est alors sous-estimée, mais l'important est de se placer au-delà de la dernière articulation. En effet, même si l'axe est interrompu à ce niveau, la connexité le long de la phalange distale assure finalement un découpage correct. Compte tenu de la longueur relative de cette phalange ($\lambda_{35} = 13\%$), il faut prendre un rapport d'au moins 0.65. Un valeur de rapport intermédiaire entre les deux limites 0.65 et 0.70, c'est-à-dire de l'ordre de 0.67 ou 0.68, donne effectivement de bons résultats.

Une identification complète des axes, c'est-à-dire, savoir de quel côté de l'image se trouve le pouce, sera finalement nécessaire pour guider la suite du traitement avec l'apport de paramètres spécifiques à chaque axe. Le pouce est a priori l'axe le plus court. Pourtant la longueur n'est pas un critère suffisamment robuste, car l'axe de l'auriculaire peut être incomplet. L'expérience montre qu'il est préférable de considérer les distances mesurées entre les extrémités distales des axes, d'une part, entre le pouce et l'index et, d'autre part, entre l'auriculaire et l'annulaire. Ce critère est nettement plus robuste, mais il échoue pourtant sur quelques cas. L'orientation définitive de l'image se fera donc plus tard, après un nouveau calcul des axes à l'échelle fine.

2.4.8 Exemple d'un cas plus difficile

Comme nous le verrons au chapitre 3, l'image qui nous sert d'exemple courant (figures 1.1, 2.8 et annexe B.1) est relativement ordinaire parmi les images dites de "bonne" qualité. Les quelques autres exemples que nous avons choisis pour illustrer les problèmes de construction des axes (figures 2.9, 2.15 et 2.16), sont eux aussi traités de manière globalement correcte. Nous allons regarder maintenant le cas d'une image beaucoup plus difficile (voir figure 2.17). Le problème principal pour cette image est un manque de contraste, comme on peut le voir figure 2.18.

L'histogramme des maxima du gradient (voir figure 2.19), qui est à comparer à celui de la figure 2.5, page 41, montre qu'il est difficile ici d'obtenir une séparation correcte entre signal et bruit. Le premier pic qui correspond au bruit est presque entièrement fusionné avec un second pic d'importance comparable qui correspond à des contours. La détermination du seuil bas, selon la méthode décrite en 2.3.2, est alors relativement fragile. Elle nous donne pourtant la valeur s_b attendue pour séparer ces deux pics. Le problème est qu'un seuillage à ce niveau supprime les contours les plus faibles, ceux des phalanges distales, et conserve une part importante du bruit. À titre indicatif, si la détection de s_b avait échoué, le seuil se situerait au niveau de s'_b , ce qui supprimerait l'essentiel du bruit, mais encore plus de signal.

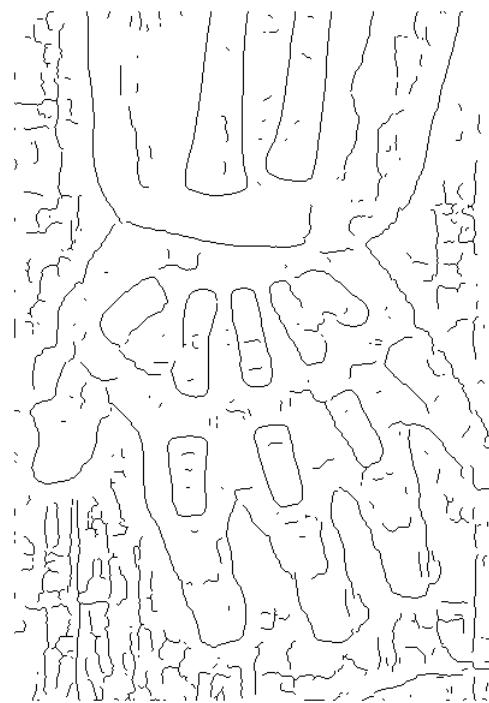
Notons que ceci concerne pourtant un gradient calculé à une échelle proche de l'optimal. L'aspect chaotique de la figure 2.20 montre l'extrême difficulté du choix d'une échelle de détection appropriée, en comparaison avec les courbes plus régulières de la figure 2.10, page 47.

Pour des échelles supérieures à 0.25, le signal et le bruit ont des valeurs de gradient qui se confondent entièrement et un minimum relatif du type de s_b n'est pas détectable. Avec le choix d'un seuil plus élevé, du type de s'_b , on perd l'essentiel du signal, à un point tel que le dénominateur n_3 est souvent nul (aucune ligne de balayage ne contient au moins 3 pixels médians). On a alors $n_6 = n_5 = n_3 = 0$ et donc l'indétermination des rapports $\frac{n_6}{n_3}$ et $\frac{n_5}{n_3}$. Pour visualiser clairement ce cas sur la figure 2.20, on a attribué la valeur arbitraire 1.2 à ces rapports. Rappelons que l'on a, par définition, $0 \leq n_6 \leq n_5 \leq n_3$ et donc, normalement, $\frac{n_6}{n_3} \leq \frac{n_5}{n_3} \leq 1$. Pour des échelles au-delà de 1, le bruit domine, avec de nombreux pixels médians parasites et des rapports qui tendent alors vers 1.

La plage d'échelles "propres" se limite ainsi à un intervalle entre 0.18 et 0.25. L'échelle initiale, déterminée d'après les dimensions de l'image, est de l'ordre de 0.21 et donne un score $\frac{n_5}{n_3}$ de l'ordre 0.07, soit moins de la moitié du seuil minimal fixé à 0.15. Le deuxième essai, avec



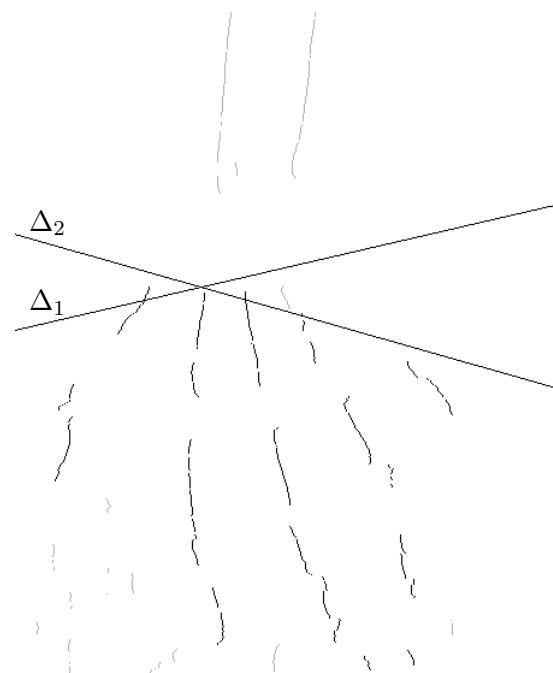
image (éclaircie pour l'impression)



contours



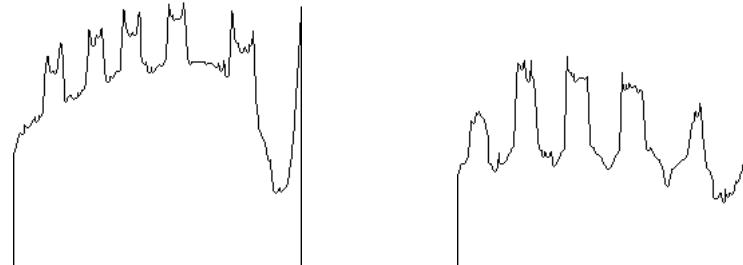
points médians



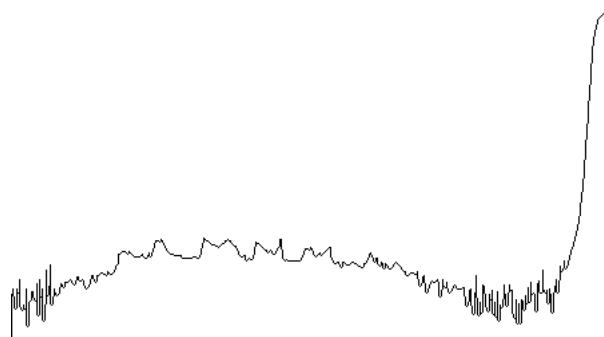
axes détectés

FIG. 2.17 – Résultats obtenus avec une image insuffisamment contrastée

Dans cette figure, l'ordonnée indique le niveau de gris relevé le long d'une coupe horizontale de l'image. On observe ainsi l'amplitude des variations (le signal) qui déterminent les contours longitudinaux. Le pic de droite correspond à une marge blanche sur le cliché.



pour l'exemple courant, figures 1.1, etc.
coupe à hauteur des métacarpiens coupe à hauteur des phalanges



pour l'image de la figure 2.17, coupe à hauteur des métacarpiens

FIG. 2.18 – Exemples de profils des niveaux de gris dans une image

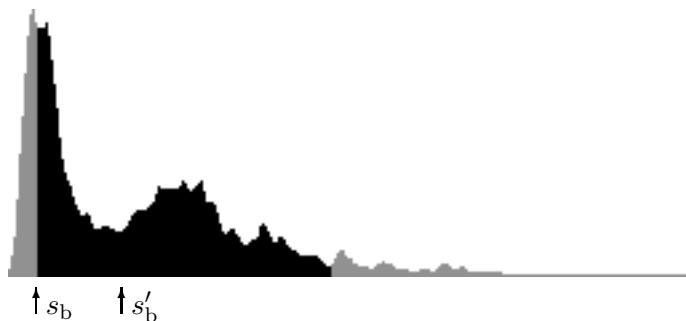
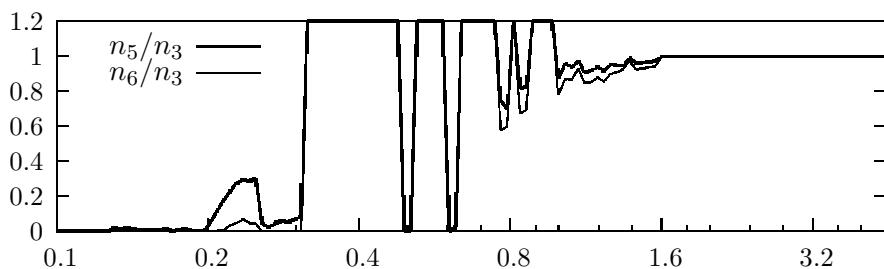


FIG. 2.19 – Seuils de détection des contours, pour l'image de la fig. 2.17



La valeur “plafond” 1.2 traduit les cas où $n_3 = 0$.

FIG. 2.20 – Nombre de points médians en fonction de l'échelle

une échelle de l'ordre de 0.23, donne un score “acceptable” de l'ordre de 0.28 (à comparer à 0.45 pour notre exemple usuel et au 0.64 théorique). Malgré le bruit important, le score $\frac{n_6}{n_3}$ reste très faible, de l'ordre 0.05, ce qui montre¹² l'efficacité des contraintes d'appariement.

Finalement, la détection des points médians se fait relativement bien, à l'exception de l'axe de l'auriculaire pour lequel il manque des contours, notamment le contour gauche de ce métacarpien. Les points médians s'organisent en 185 chaînes connexes, de longueur moyenne inférieure à 10. Ces chaînes sont donc, pour la plupart, éliminées par le filtrage. Il reste alors 47 chaînes relativement consistantes qui sont ensuite regroupées en 12 axes dont 6 parasites.

Le choix des 5 axes est correct. Notons pourtant que le raccourcissement important de l'axe de l'auriculaire a pour conséquence de faire “basculer” la droite de régression. D'ordinaire ce serait sans grande conséquence, mais ici cela conduit à couper une partie de l'axe du métacarpien de l'annulaire dont la connexité est trop faible. Notons d'autre part qu'une chaîne parasite est conservée à l'extrémité distale du majeur, ce qui conduit à maintenir aussi des parasites aux extrémités de l'index et de l'annulaire.

2.5 Détermination de l'échelle de localisation

Il s'agit maintenant d'obtenir une localisation plus précise des contours, en utilisant une échelle nettement plus fine que celle qui a été utilisée pour la détection des axes. On va commencer par discuter du choix de cette échelle, en se donnant un objectif de localisation a priori suffisant.

2.5.1 Objectif de précision

En ce qui nous concerne, il est raisonnable et suffisant de viser un objectif de précision de l'ordre du demi-pixel, ce qui correspond à un contour bien placé à la résolution de l'image numérique, c'est le résultat obtenu par la NMS. D'un point de vue métrique, cela correspond à moins de 0.2 mm à la résolution basse de 75 dpi. Cela permet notamment une mesure de distance de contour à contour avec une précision au $\frac{1}{2}$ mm, donc comparable à celle d'une mesure “manuelle” qui serait réalisée directement sur la radiographie avec une règle graduée. Comme on l'a vu en introduction, la technique radiographique permettrait de gagner un peu sur la taille des pixels, jusqu'à 400 dpi, mais avec une augmentation de bruit. Nous allons voir que, dans notre cas, cela n'a en fait guère de sens de chercher une meilleure précision.

Des précisions supérieures, dites sub-pixeliques, sont pourtant réalisables, quand la qualité des images le permet. Cela revient à estimer plus finement la position des maxima du gradient, au voisinage de la position détectée au demi-pixel près. Cela peut se faire par interpolation [Har84, JKS95, Dev95] ou, de manière plus précise, en utilisant des batteries de filtres “décalés” d'une fraction de pixel. Les filtres de Deriche, en particulier, se prêtent bien à un tel usage [MD97].

Mais, il ne faut pas oublier que cela permet simplement d'atteindre une précision sub-pixelique dans le calcul de la position d'un maximum du gradient. L'écart entre un maximum du gradient et la position “réelle” du contour peut être nettement plus important, parfois de l'ordre de quelques pixels. Nous allons examiner les principales contributions à de tels écarts que sont les diverses sources de bruit et de flou. Par contre, on va ignorer la déformation perspective (voir l'introduction) dont on a dit qu'on peut en faire abstraction en respectant un certain cadrage.

¹²Si l'on avait la probabilité $p^n(1-p)$ d'avoir exactement n pixels médians sur une ligne, on aurait alors $\frac{n_5}{n_3} = p^2$ et $\frac{n_6}{n_3} = p^3 = (\frac{n_5}{n_3})^{\frac{3}{2}}$, soit pour cet exemple, $\frac{n_6}{n_3} \approx 0.15$, à comparer au 0.05 qui est mesuré.

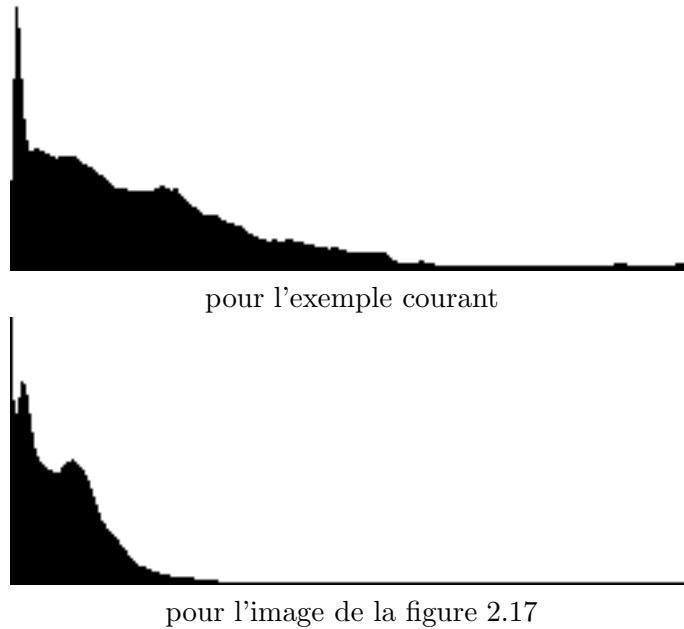


FIG. 2.21 – Histogrammes des valeurs du module du gradient

2.5.2 Influence du bruit

Le bruit, textures anatomiques et causes techniques confondues, induit une erreur aléatoire dans l'estimation de la position de chaque maximum du gradient. Il est possible de quantifier statistiquement ces erreurs par leur écart-type.

Selon Canny, c'est l'inverse du critère de localisation, soit $\frac{1}{L(\alpha)} = \frac{\sigma_b}{A} \frac{1}{\sqrt{2\alpha}}$, pour les filtres que nous utilisons (filtres de Deriche) et pour le modèle “classique” du bruit blanc gaussien additif (voir annexe A.1). Ces hypothèses sur le bruit sont requises du point de vue analytique. En pratique, il est surtout important d'avoir une distribution relativement compacte, c'est-à-dire plus formellement, un kurtosis¹³ modéré, inférieur ou égal à celui de la loi normale, et, à défaut de bruit blanc, une relative importance des hautes fréquences spatiales pour assurer l'efficacité du lissage.

Le problème est que $\frac{A}{\sigma_b}$, le rapport signal/bruit dans l'image source, n'est pas directement mesurable. En effet, la séparation entre signal et bruit est parfois loin d'être évidente, on s'en rend compte sur les exemples de la figure 2.18. À titre indicatif, en moyenne sur nos images, le rapport $\frac{A}{\sigma_b}$ est de l'ordre de 7 (soit 17 dB), mais pour quelques images, le rapport est inférieur à 3 (9.5 dB), valeurs estimées indirectement par le procédé qui est décrit ci-dessous. Ces valeurs relativement mauvaises s'expliquent par le fait que la “hauteur” du signal A est très inférieure à la dynamique de l'image, comme on peut le voir figure 2.18.

Il est relativement plus facile de mettre en évidence le rapport signal/bruit dans le champ du module du gradient. La différentiation est un moyen classique d'accéder aux caractéristiques spectrales (c'est aussi une forme de décorrélation) et la régularisation qui l'accompagne et qui est indispensable au calcul des dérivées, réalise une atténuation relativement sélective du bruit. Autrement dit, la moyenne et la variance du module du gradient du bruit sont nettement plus faibles, relativement au signal “contour”, que ne le seraient les mêmes mesures sur les valeurs des pixels dans l'image source.

¹³Le kurtosis ou coefficient d'aplatissement se calcule à partir du moment d'ordre 4, sa valeur est d'autant plus élevée que les “queues” de distribution sont importantes. Par construction, sa valeur pour une distribution gaussienne est une constante et, notamment, elle ne dépend pas de la variance.

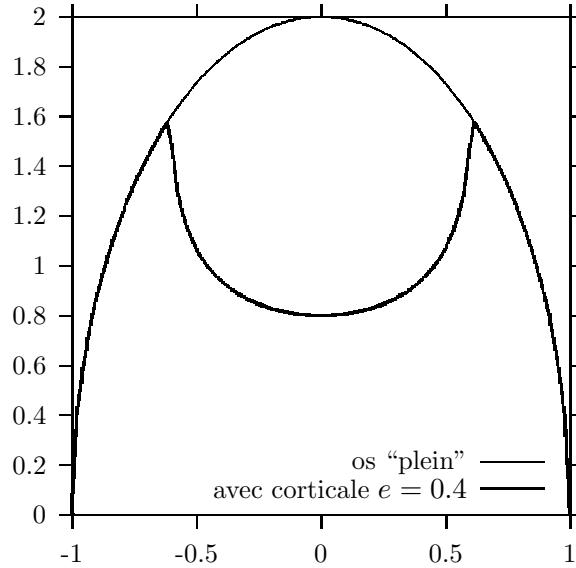


FIG. 2.22 – Profil théorique de l'épaisseur d'absorption d'un os

C'est particulièrement notable aux échelles que nous avons utilisées pour la détection. Cela suffit normalement pour que le bruit se concentre sur un pic très pointu dans un histogramme du module du gradient, comme on le voit figure 2.21, en haut¹⁴. Du fait de la forte pente de ce pic, l'erreur possible de classification entre signal et bruit ne concerne qu'un nombre relativement faible de pixels et elle a peu d'influence sur les mesures des variances respectives. On dispose ainsi d'une mesure fiable du rapport signal/bruit du gradient. Ce rapport est de l'ordre de 20, en moyenne sur nos images, et de 21 pour notre exemple courant. Pour l'image de la figure 2.17, l'estimation est beaucoup moins fiable, du fait de l'ambiguïté entre signal et bruit, et la méthode présentée donne un rapport signal/bruit de l'ordre de 6.

Ce rapport signal/bruit du gradient correspond au critère de détection de Canny, c'est-à-dire $D_2(\alpha_{\text{det}}) = \frac{A}{\sigma_b} \sqrt{\frac{1}{\alpha_{\text{det}}}}$ pour l'échelle de détection α_{det} . La disparition du facteur $\frac{1}{\sqrt{2}}$ s'explique par le passage en dimension deux : on a considéré que la variance du gradient du bruit est le double de la variance de chacune des dérivées partielles¹⁵, alors que le signal reste mono-dimensionnel (dans la direction du gradient). Ainsi le rapport signal/bruit mesuré est $D_2 = \frac{1}{\sqrt{2}} \cdot D_1$, avec D_1 le critère usuel de Canny, tel qu'il est exprimé en annexe A.1. Pour les deux exemples donnés ci-dessus, cela donne des valeurs respectivement de l'ordre de 10 et 3.

En posant $k_{\text{det}} = \frac{1}{\sqrt{2}} \frac{1}{D_2(\alpha_{\text{det}})\sqrt{\alpha_{\text{det}}}} = \frac{1}{\sqrt{2}} \frac{\sigma_b}{A}$, l'écart-type de l'erreur de localisation pour une échelle α_{loc} peut ainsi s'écrire : $\frac{1}{L(\alpha_{\text{loc}})} = \frac{1}{D_2(\alpha_{\text{det}})\sqrt{2\alpha_{\text{det}}\alpha_{\text{loc}}}} = \frac{k_{\text{det}}}{\sqrt{\alpha_{\text{loc}}}}$.

Le paramètre k_{det} traduit la qualité de l'image et il est connu à l'issue de la phase de détection. Les mesures de $D_2(\alpha_{\text{det}})$, obtenues sur notre base de "bonnes" images, donnent ainsi des valeurs de k_{det} en général inférieures à 0.25 ($\sqrt{\text{pixel}}$), avec une moyenne de l'ordre de 0.10 et un écart-type de l'ordre de 0.04. Les erreurs de localisation pour des valeurs usuelles de α_{loc} sont donc facilement compatibles avec notre besoin de précision. Par exemple pour $\alpha_{\text{loc}} = 1$, un écart de localisation de l'ordre de 2 fois l'écart-type, soit $2k_{\text{det}}$, est encore acceptable pour notre objectif du demi-pixel.

¹⁴On notera que les histogrammes représentés ici diffèrent nettement de ceux présentés aux figures 2.5, page 41, et 2.19, page 63. Ici, on comptabilise les valeurs du gradient pour tous les pixels de l'image et pas seulement pour les maxima, comme dans ces autres figures.

¹⁵En utilisant $\sigma_{|\nabla|}^2 = \sigma_{\frac{\partial}{\partial x}}^2 + \sigma_{\frac{\partial}{\partial y}}^2$ et $\sigma_{\frac{\partial}{\partial x}}^2 = \sigma_{\frac{\partial}{\partial y}}^2$.

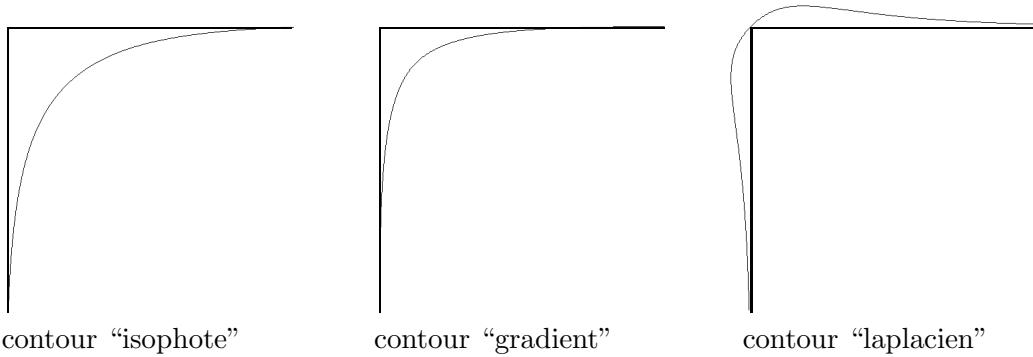


FIG. 2.23 – Différences dans le contour d'un "coin" selon la caractérisation employée

Notons que ces résultats sont établis pour un modèle très théorique de contour du type "marche". En réalité, un contour est plus "flou" et il est d'usage d'assimiler le contour réel à un contour marche qui serait déjà lissé [EZ98].

Dans notre cas, si on considère un os de section circulaire, de diamètre 2, l'épaisseur de cet os qui est traversée par un rayon X passant à la distance d de l'axe, est égale à $2\sqrt{1-d^2}$, ceci dans le cas d'un os "plein" (voir figure 2.22). Pour un os avec un canal médullaire de largeur $2(1-e)$, où e est l'épaisseur de la couche corticale, on obtient l'épaisseur d'os traversée en soustrayant $2\sqrt{(1-e)^2-d^2}$ sur l'intervalle $d \in [e-1, 1-e]$. On obtient ainsi des profils d'épaisseur qui se comparent aux profils des niveaux de gris relevés sur une image, comme ceux en haut de la figure 2.18, page 63. Cela reste très approximatif, car en réalité la densité osseuse n'est pas homogène et la conversion entre l'absorption des rayons X et le niveau de gris est loin d'être linéaire. Dans une moindre mesure, il s'ajoute à cela diverses autres sources de flou, dues au procédé radiographique et à la numérisation de l'image, mais on peut considérer que l'origine principale est anatomique.

Une mesure effective de la "largeur" σ_{src} de ce flou serait délicate [EZ98] et relativement imprécise. Elle est estimée empiriquement de l'ordre de 2 à 4 pixels, valeur qui dépend surtout des dimensions de l'os et bien sûr de la résolution de l'image.

Ce flou intrinsèque a pour effet d'atténuer l'amplitude des dérivées du signal, donc le rapport signal/bruit qui intervient dans les critères D et L de Canny. Mais le paramètre k_{det} qui est mesuré, est basé sur le rapport signal/bruit effectif, c'est-à-dire qu'il prend en compte les effets de ce flou et il n'y a donc pas lieu de corriger explicitement l'expression $\frac{1}{L(\alpha_{loc})} = \frac{k_{det}}{\sqrt{\alpha_{loc}}}$.

La dépendance des critères D et L en fonction de l'échelle (et du flou intrinsèque) est une particularité de certains filtres, notamment ceux de Deriche et les filtres gaussiens [Fau93, KSM06]. Plusieurs auteurs [Td90, BS94, KG94, ÅH96, Ste98] remettent aussi plus ou moins en cause le bien fondé du critère de localisation de Canny. En fait, toutes ces approches donnent des valeurs de $\frac{1}{L}$ de l'ordre de $\frac{\sigma_b}{A}$, peu différentes de celles que nous obtenons pour les échelles α_{src} et α_{loc} usuelles.

Pour un rapport signal/bruit de l'ordre de 7, la moyenne sur nos images, on a ainsi un écart-type légèrement inférieur à $\frac{1}{10}$ de pixel. En général, avec une probabilité de 99 %, l'écart effectif de localisation reste donc inférieur au quart de pixel. Il faut aussi penser que le signal s'atténue en allant vers les phalanges distales mais une part du bruit, celle due au bruit quantique et aux textures anatomiques, décroît elle aussi, ce qui limite la dégradation du rapport signal/bruit.

Pour un rapport signal/bruit de l'ordre de 3, l'écart-type atteint $\frac{2}{10}$ de pixel et on peut seulement "garantir" une localisation au demi-pixel près. Mais de telles images posent déjà des problèmes de détection et de construction des axes, ce qui compromet fortement la réalisation des mesures.



FIG. 2.24 – Mesure de la largeur d'une métaphyse

2.5.3 Erreur de caractérisation

Indépendamment de l'influence du bruit, la caractérisation des contours comporte une erreur de localisation systématique [Ber84, Vv94, DG93]. L'écart est négligeable pour des contours relativement rectilignes, comme le long des diaphyses pour ce qui nous concerne, mais il peut être relativement important lorsque la courbure de l'isophote augmente (voir figure 2.23 et annexe A.1). De plus, l'écart augmente avec le lissage “global” défini par $\sigma^2 = \sigma_{\text{src}}^2 + \sigma_{\text{loc}}^2$. Quantitativement, le problème concerne autant les maxima du gradient que les passages par zéro du laplacien, mais les effets respectifs sont de sens opposés. Notons pourtant que le calcul du laplacien demande généralement des valeurs de σ_{loc}^2 plus fortes, en raison du bruit.

Le contour calculé est donc notamment “arrondi” et cela pose naturellement un problème de justesse des mesures, par exemple pour la largeur d'une métaphyse (voir figure 2.24). On ne sait quantifier l'erreur de caractérisation que sur des modèles très simples, utiles simplement pour se donner une majoration. Sur un cas réel, une correction explicite n'aurait de sens que si l'on pouvait connaître avec précision le profil anatomique du contour.

Dans le cas d'une caractérisation par le gradient, le contour calculé se trouve “à l'intérieur” du “vrai” contour. En partant des résultats de Verbeek et van Vliet et al. [Vv94], l'écart radial absolu peut être majoré par $1.3 \frac{\sigma^2}{R}$, où R est le rayon de courbure du contour. Par exemple pour $\sigma^2 = 20$, un cas défavorable avec $\sigma_{\text{src}} = 4$ et $\sigma_{\text{loc}} = 2$, on obtient un écart inférieur au quart de pixel pour R supérieur à 104. Cette valeur limite pour R est déjà perçue comme nettement courbe. Dans un cas plus favorable, avec $\sigma_{\text{src}} = 2$ et $\sigma_{\text{loc}} = 1$, on peut même descendre jusqu'à $R = 26$. Sur des portions de contours relativement rectilignes et avec un choix raisonnable de σ_{loc} , l'erreur de caractérisation est donc facilement maintenue à des valeurs très faibles, notamment en dessous du quart de pixel.

Mais, on doit aussi réaliser des mesures qui s'appuient sur des portions de contours ayant des courbures plus importantes. C'est le cas par exemple pour les épiphyses, pour des métaphyses arrondies, avec des rayons de l'ordre de 10 ou 20 et aussi pour les os du carpe. L'écart peut alors facilement dépasser le pixel. Enfin, on a des rayons très faibles, de l'ordre de 2 ou 3, aux coins des métaphyses et sur des épiphyses lenticulaires. La majoration en $\frac{\sigma^2}{R}$ n'est valable que pour des contours suffisamment réguliers, soit $\frac{\sigma}{R}$ relativement petit, avec la limite $\frac{\sigma}{R} < \frac{1}{2}$ selon [Vv94].

Pour analyser le cas des petits rayons, il est plus réaliste de considérer un modèle de contour anguleux [DG93, Roh94, KSM06]. L'écart maximal δ qui se mesure le long de la bissectrice de l'angle, est d'autant plus important que l'angle θ est “pointu” et il est proportionnel à σ . Plus précisément, dans le cas d'une caractérisation par les maxima du gradient, Deriche et Giraudon [DG93] obtiennent des valeurs de l'ordre de $\delta = 0.715\sigma$ pour $\theta = \frac{\pi}{2}$, de 1.329σ pour $\theta = \frac{\pi}{4}$, etc. Köthe et al. [KSM06] aboutissent à des valeurs comparables. Même si cette estimation est moins défavorable que celle en $\frac{\sigma^2}{R}$, elle donne des valeurs d'écart de l'ordre de plusieurs pixels.

On connaît donc le sens de l'erreur et on pourrait recentrer la localisation et ainsi diviser l'incertitude par 2. Mais, en pratique, seul l'ordre de grandeur semble pertinent. Une correction plus précise supposerait la mesure de σ_{src} et de R ou θ , ce qui resterait très approximatif. Dans certains cas, on peut aussi utiliser la spécificité de la mesure, par exemple, sur la figure 2.24, on peut considérer que la mesure se fait selon une direction à $\frac{\pi}{4}$ par rapport aux bissectrices des angles et donc diviser l'incertitude par $\sqrt{2}$.

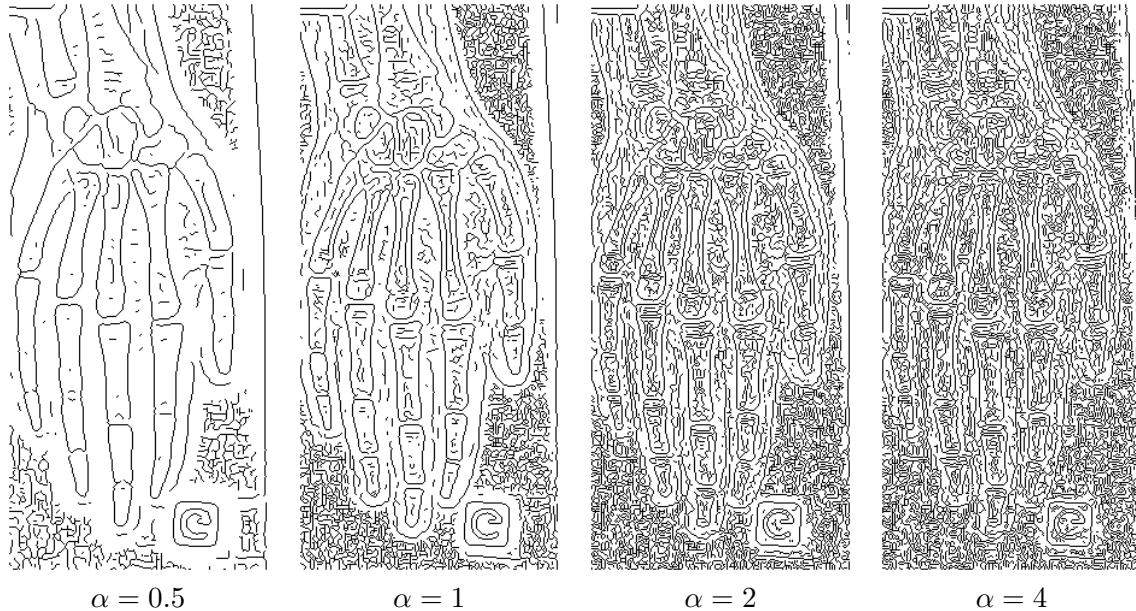


FIG. 2.25 – Maxima du gradient en fonction du paramètre d'échelle

Au mieux, tout cela permet de se rapprocher d'une erreur de l'ordre d'un pixel à chaque extrémité.

2.5.4 Choix de l'échelle de localisation

On a vu que l'erreur de localisation imputable au bruit était relativement faible et pouvait être considérée comme inférieure au quart de pixel. Sur des portions de contour relativement rectilignes, il en est de même pour l'erreur de caractérisation. Dans ce cas, la localisation au demi-pixel est donc très sûre et l'échelle de localisation importe peu. Par des techniques un peu robustes¹⁶ on peut ainsi s'assurer d'une erreur qui ne dépasse pas le pixel pour les mesures de distance entre deux contours.

Dans le cas de courbures plus prononcées, une localisation au demi-pixel près, voire au pixel près, n'est plus assurée. Le seul paramètre que l'on peut contrôler est l'échelle de filtrage σ_{loc} ou $\alpha_{loc} = \frac{2}{\sigma_{loc}}$ et, pour réduire l'erreur de caractérisation, il faut augmenter α_{loc} .

Mais on approche de la limite asymptotique dès que σ_{src} domine dans $\sigma^2 = \sigma_{src}^2 + \sigma_{loc}^2$. À noter aussi que l'on approche des limites du filtrage discret et, pour $\alpha_{loc} = 4$, $\sigma_{loc} = \frac{1}{2}$, la régularisation est quasi-inexistante. Sur la figure 2.25, on peut voir encore quelques différences notables entre les contours obtenus respectivement pour $\alpha = 1$ et $\alpha = 2$, notamment pour les épiphyses. En passant de $\alpha = 2$ à $\alpha = 4$, il n'y a plus de différence dans les vrais contours, mais on note une densité accrue du bruit.

En effet, un autre critère pour le choix de l'échelle est la densité des maxima du gradient du bruit ou, inversement, la distance moyenne entre ces maxima parasites. Une valeur théorique pour cette distance nous est donnée par le critère d'unicité de Canny (voir annexe A.1). On retiendra que ce critère ne dépend pas du niveau du bruit mais plutôt de ses caractéristiques spectrales. Canny s'est basé sur l'hypothèse d'un bruit blanc, assez réaliste en ce qui concerne le bruit de haute fréquence qui nous concerne ici.

Pour les filtres que nous utilisons, ce critère s'exprime par $x_{max}(\alpha) = \frac{2\pi}{\alpha\sqrt{5}}$. Un seuil très significatif, $x_{max} = 2$, soit un pixel intercalé, est déjà atteint pour α de l'ordre de 0.7. Au-delà, on va effectivement observer des "contacts" entre les maxima parasites et les vrais contours.

¹⁶Par exemple, en prenant la médiane de quelques mesures similaires.

Ces contacts s'observent en nombre croissant avec α et ils vont naturellement compliquer la validation des contours en créant des ambiguïtés dans le chaînage par connexité. Des contraintes de cohérence, basées sur le calcul préalable des axes, ainsi que l'apport d'un peu d'information a priori, vont nous permettre de guider la validation des contours et de résoudre correctement une bonne part de ces ambiguïtés. Il convient pourtant de ne pas augmenter exagérément α_{loc} et l'expérience montre que les résultats se dégradent pour des valeurs de α_{loc} au-delà de 2 ou 3.

La mauvaise localisation est surtout un problème pour des comparaisons entre d'une part nos mesures, faites sur les contours de type “gradient”, et d'autre part des mesures “classiques” (recueils de références anatomiques ou mesures “manuelles” sur l'image). En effet, si ces dernières font l'objet d'une erreur de caractérisation, celle-ci serait plutôt comparable au cas des contours du type “laplacien” qui correspondent au modèle de l'œil [Har84, Mar82]. La figure 2.23, page 67, illustre ce problème qui concerne toutes les mesures qui sont prises en se basant sur des coins. Lorsqu'il s'agit de comparer entre elles des mesures prises sur des contours obtenus par la même caractérisation (gradient), avec des courbures (ou angles) d'isophotes et des profils d'intensités identiques, on peut considérer que les mesures seront très voisines, avec une même erreur de caractérisation et une influence du bruit qui est comparativement négligeable.

On a déjà admis que le flou intrinsèque σ_{src} qui est d'origine principalement anatomique mais qui est relativement mal connus, dépend linéairement de la largeur des os, en unité métrique, et de la résolution des images, pour la conversion en nombre de pixels. On dispose d'ailleurs à ce stade du traitement d'une mesure robuste de la largeur moyenne d'appariement, notée L_{ref} . Cette mesure est assez représentative de la largeur des doigts (en pixels) et on peut ainsi poser $\sigma_{\text{src}} = \lambda L_{\text{ref}}$. Formellement, il s'agit d'une caractéristique locale mais on peut considérer que λ est une constante pour une image. On a alors intérêt à se donner des propriétés similaires lors du choix de σ_{loc} , c'est-à-dire, poser $\sigma_{\text{loc}} = \mu L_{\text{ref}}$, pour un certain μ .

Ainsi, comme $\sigma^2 = \sigma_{\text{src}}^2 + \sigma_{\text{loc}}^2$, on a $\sigma = \sqrt{\lambda^2 + \mu^2} L_{\text{ref}}$. Par ailleurs, pour ce qui concerne le rayon de courbure R d'un contour, on peut aussi poser $R = \nu L_{\text{ref}}$ et ainsi $\frac{\sigma^2}{R} = \frac{\lambda^2 + \mu^2}{\nu} L_{\text{ref}}$. Les deux expressions des écarts de caractérisation, celle en σ et celle en $\frac{\sigma^2}{R}$, suivent donc une dépendance linéaire en fonction de L_{ref} , c'est-à-dire en fonction des dimensions effectives dans l'image. Ainsi les erreurs relatives peuvent être vues comme des constantes propres à chaque définition d'une mesure.

L'échelle de localisation est donc choisie comme $\alpha_{\text{loc}} = \frac{30}{L_{\text{ref}}}$. Le paramètre 30 est empirique et donne des valeurs de α_{loc} qui ne dépassent pas 2 sur l'ensemble de nos images. On a ainsi des valeurs de σ_{loc} de l'ordre de la moitié de celles de σ_{src} et on ne gagne pas beaucoup à utiliser des valeurs plus faibles. Pour notre exemple courant, on obtient $\alpha_{\text{loc}} \approx 1.6$, sans différence notable dans les vrais contours par rapport à $\alpha = 2$. Le problème est maintenant la sélection des “bons” contours dans une carte des maxima du gradient qui s'avère très dense du fait du bruit.

2.6 Localisation et identification de chaque os

Les axes que l'on a construits, ont maintenant un rôle essentiel puisque l'on va chercher des points de contours approximativement symétriques de part et d'autre de chacun de ces axes. Une marge de tolérance permet de rattraper les erreurs éventuelles d'estimation de la position des axes et les variations de la largeur d'appariement. On arrive ainsi à construire des portions de contours longitudinaux relativement longues et à écarter l'essentiel des maxima parasites. Après un peu de nettoyage, les contours que l'on a sélectionnés sont normalement les contours des os et, par construction, on sait les rattacher à leurs axes respectifs.

On en déduit ensuite une meilleure localisation des axes et il reste à procéder à un découpage le long de chaque axe pour déterminer les séparations entre les os successifs. Cette dernière

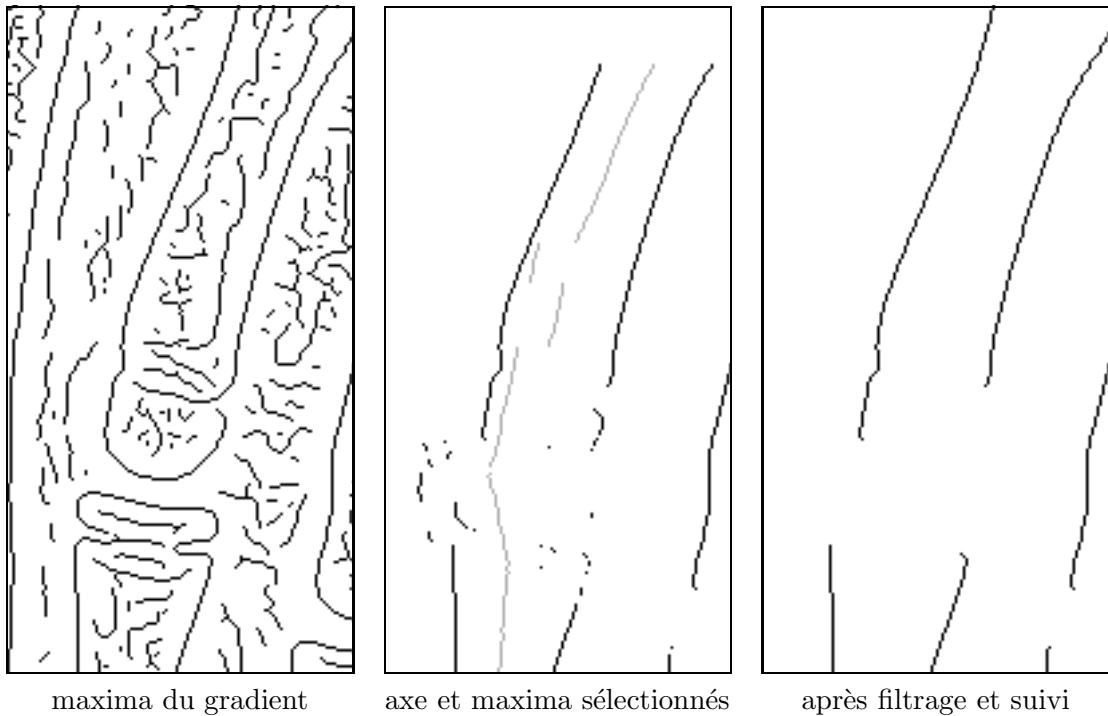


FIG. 2.26 – Sélection des contours à partir des axes

analyse est guidée à la fois par le contenu de l'image et par référence à des longueurs "normales" qui déterminent a priori la détection des limites de chaque os.

2.6.1 Construction des contours longitudinaux

Le calcul de la carte des maxima du gradient se fait essentiellement comme décrit en 2.2, en utilisant l'échelle fine de localisation α_{loc} . Mais le bruit est trop important pour permettre de caractériser les contours en n'utilisant que des critères sur l'intensité du gradient. Des contraintes de structure globale sont alors nécessaires. Pour cela, de part et d'autre de chaque point d'axe, on va chercher un contour gauche et un contour droit, à la fois à une distance acceptable et correctement "orientés".

Le critère d'orientation s'exprime par le cosinus de l'angle entre le vecteur gradient et une estimation locale de la normale à l'axe. On ne retient un maximum de gradient que si le cosinus correspondant est supérieur à un certain seuil¹⁷. Une valeur de 0.9 pour ce seuil, soit un angle limite d'environ 26° , donne suffisamment de tolérance vis-à-vis de la courbure naturelle d'un contour longitudinal. Et, si on considère que le gradient du bruit a une distribution uniforme en direction, ce critère suffit pour en éliminer 85 %.

Le critère de position conduit à rechercher un maximum distant du point d'axe d'environ la moitié de la largeur d'appariement, c'est-à-dire, à peu près au même emplacement que le pixel de contour (celui obtenu à l'échelle de détection) qui a déterminé ce point d'axe. La largeur d'appariement n'est généralement pas sous-estimée car, à l'échelle de détection, il y a peu de détails intérieurs à l'os qui pourraient être validés dans un appariement. Comme il s'agit de contours relativement rectilignes, on peut considérer qu'une marge extérieure de deux pixels suffit pour rattraper les diverses petites erreurs (caractérisation, bruit, discréétisation).

Par contre, le maximum recherché peut se trouver très nettement en dedans de la position estimée. Il est en effet assez fréquent que la largeur d'appariement soit surestimée, lorsqu'elle

¹⁷Le sens de la normale est naturellement inversé selon que cherche à caractériser un contour gauche ou droit.

s'appuie sur un contour “peau” ou sur une fusion entre le contour “os” et le contour “peau”¹⁸. Dans le cas où l'on a apparié deux contours de nature différente, par exemple la chaîne \mathcal{B} de la figure 2.15, page 58, l'axe estimé est nettement décalé et se trouve parfois très proche du contour recherché (voir figure 2.26). Une distance inférieure au dixième de la largeur d'appariement n'est alors pas rare. Toutefois, compte tenu des largeurs respectives “os” et “peau”, on considère qu'il n'est pas utile de regarder de l'autre côté de l'axe.

En pratique, on balaye donc de la marge extérieure jusqu'à la position de l'axe. On parvient ainsi à rattraper le cas d'un axe nettement décalé. Si on trouve plusieurs maxima correctement orientés sur cet intervalle, on retient le plus fort en module. On peut en effet considérer que les contours les plus marqués, à proximité d'un axe et correctement orientés, sont ceux des os. On ne conserve donc, au plus, qu'un pixel de contour par côté d'un axe et par ligne de balayage.

Cette phase de sélection demande, dans sa globalité, de balayer exhaustivement de l'ordre de 10 à 30 % des pixels de la carte des maxima, pourcentage qui dépend de la largeur moyenne d'appariement et de la longueur des axes. Le coût est relativement faible et tous ces pixels doivent être examinés. Notamment, on n'exploite pas d'hypothèse a priori qui serait basée sur la connexité longitudinale, car cela privilégierait un allongement arbitraire des chaînes de contour. Cela introduirait ainsi une corrélation qui n'est pas compatible avec les étapes suivantes de mise en évidence par accumulation qui privilégie la cohérence structurelle.

Ce n'est qu'une fois que les pixels ont été sélectionnés indépendamment les uns des autres, qu'ils sont assemblés en chaînes connexes. La technique est similaire à celle qui est décrite pour les axes en 2.4.1, mais on exploite en plus le fait que les maxima sont déjà étiquetés par l'axe auquel ils se rapportent et qu'ils sont aussi identifiés comme gauches ou droits. Les deux critères de position et d'orientation excluent alors toute incohérence entre le chaînage et l'étiquetage structurel. Ce dernier sera facile à propager lors des étapes suivantes.

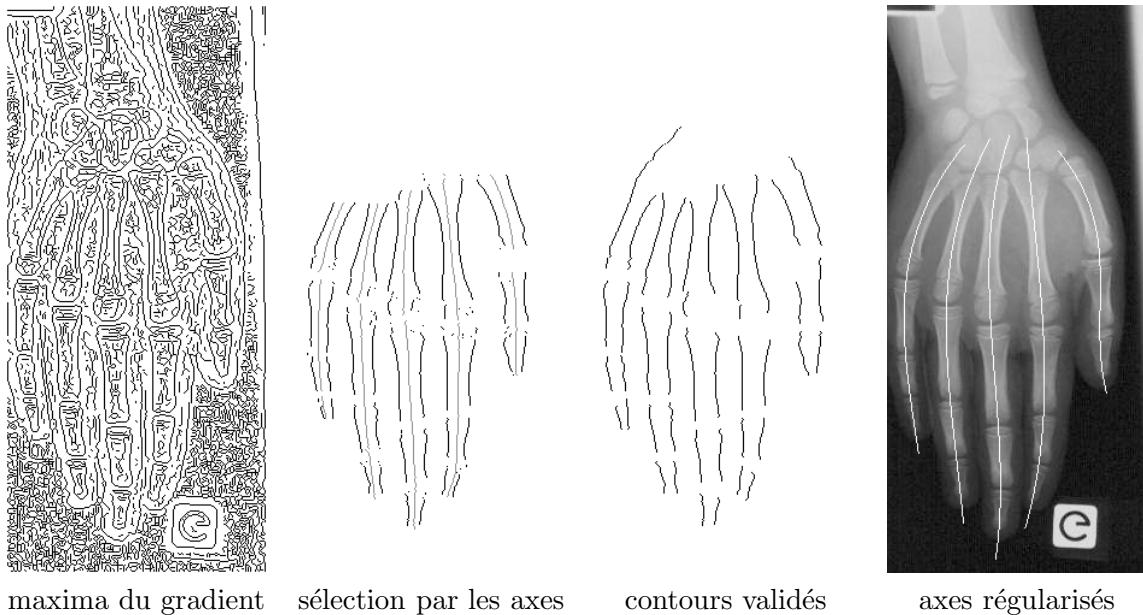
On peut néanmoins sélectionner des maxima erronés, visibles par exemple sur les figures 2.26 et 2.27. Il s'agit de parasites ponctuels qui sont dus à du bruit de forte intensité, ou de chaînes relativement courtes qui correspondent à des détails anatomiques. Ces parasites sont détectés par défaut lorsque le contour “os” est inexistant ou ne satisfait pas les critères. Nous pratiquons donc un filtrage sur la longueur des chaînes.

Le seuil de longueur minimale peut être facilement fixé car les parasites sont toujours de petite taille et le fait de supprimer un fragment de bon contour, par exemple celui d'une épiphyse, est rattrapé par la suite. On s'est fixé un seuil de l'ordre de quelques pourcents de la longueur des axes (estimée par la moyenne des longueurs de l'index et l'annulaire) et on peut vérifier que des petites variations de ce pourcentage ou, de manière équivalente, une erreur d'estimation de la longueur des axes ne changent pas significativement le résultat final.

Ensuite, une extension des chaînes permet de compléter ce résultat en bouchant des trous, pour raccorder ou prolonger les composantes longitudinales. On procède par l'incorporation de pixels de maxima, initialement non sélectionnés mais connexes aux précédents. On rattrape ainsi les défauts de continuité des axes et une extrapolation de la direction d'axe permet de conserver une cohérence par la contrainte d'orientation. Cela évite notamment le raccordement des contours entre deux doigts voisins à l'extrémité proximale des métacarpiens. Notons pourtant que, lorsque les pixels générateurs pour un même côté d'un os ne sont pas tous dans la même composante connexe, on peut aboutir à la croissance de chaînes “parallèles” concurrentes, avec plusieurs pixels de contour pour une même ligne de balayage.

Dans ce cas, pour être cohérent avec une sélection directe à partir d'un point d'axe, il faut conserver le pixel ayant la plus grande valeur de gradient. On introduit cependant un peu

¹⁸Si de telles fusions peuvent se produire à l'échelle de détection (voir 2.4.6 et la figure 2.15, page 58), l'échelle fine permet de distinguer et de localiser correctement les deux contours (voir figure 2.26).

FIG. 2.27 – Contours longitudinaux à échelle fine ($\alpha \approx 1.6$) et nouveaux axes

d'hystérésis dans cette comparaison. En cas de valeurs très voisines, on garde le premier pixel trouvé et, comme la croissance des extrémités de chaînes est réalisée en parallèle, c'est-à-dire en largeur d'abord, cela revient à privilégier le pixel le plus proche d'un générateur. La régularité du gradient suffit alors pour éviter une alternance chaotique entre les contours concurrents.

2.6.2 Localisation et identification des axes

Les contours ainsi obtenus font ensuite l'objet d'un nouvel appariement pixel à pixel. La technique est plus simple et beaucoup plus légère que celle qui est utilisée pour l'appariement initial, en 2.4. Dans le cas présent, on sait par construction quels contours sont appariés et on connaît l'axe auquel ils se rapportent. L'axe initial est peut être mal localisé, comme on l'a vu. En fait, il est alors un peu décalé mais sa direction est relativement bien estimée et elle permet de diriger correctement les appariements.

Pour engendrer les nouveaux points médians, il suffit de suivre les deux contours appariés, en gérant au fur et à mesure le décalage entre les ordonnées des pixels pour maintenir une direction qui correspond le mieux (localement) à la perpendiculaire à l'axe. La figure B.2, page 179 en annexe, montre les points médians ainsi obtenus. En utilisant la direction estimée de chaque axe, on évite ainsi les artefacts qui provenaient de l'appariement arbitrairement horizontal. Les précautions introduites en 2.4, notamment pour corriger les déviations en bout d'os, sont alors inutiles. On notera aussi que ces nouveaux axes sont mieux localisés que les contours qui les déterminent, car les erreurs de caractérisation réalisées sur des contours symétriques sont de sens opposés et se compensent. Il ne reste alors essentiellement que les erreurs dues au bruit qui peuvent être considérées comme négligeables¹⁹.

Maintenant, on peutachever l'identification des axes, c'est-à-dire décider de quel côté de l'image se trouve le pouce, en reprenant le principe exposé en fin de 2.4.7. Cela va permettre de guider la suite du traitement avec l'apport de paramètres spécifiques à chaque axe.

Dans un premier temps, on va construire une représentation très régulière des axes, calculée par régression polynomiale, de la forme $x = P(y)$, sur l'ensemble des points médians d'un

¹⁹L'écart-type de l'erreur de localisation d'un axe est alors $\frac{1}{\sqrt{2L}}$, c'est-à-dire le facteur $\frac{1}{\sqrt{2}}$, puisqu'il s'agit d'une moyenne sur 2 valeurs, appliqué à l'écart-type $\frac{1}{L}$ vu en 2.5.2.

axe. En pratique, on constate qu'un polynôme de degré minimal 3 est nécessaire et suffisant pour assurer une bonne localisation et pour respecter la courbure d'un axe, c'est aussi ce qui est retenu dans [PGP⁺01, MDDS01]. La dernière image de la figure 2.27, page 73, montre le résultat obtenu sur notre exemple. On peut notamment traduire les formes en S , cas du majeur et du pouce sur notre exemple, ce que ne permettrait pas un polynôme de degré 2. On peut ainsi interpoler correctement un axe, en position et en direction, et recréer une continuité au niveau des épiphyses et des articulations.

Pour des degrés supérieurs à 3, l'évaluation des coefficients est trop facilement sujette à des instabilités numériques, surtout dans le cas d'axes presque rectilignes. Le degré 3 s'avère assez stable pour permettre des extrapolations fiables, à courte distance. On aura besoin en effet de prolonger un peu les axes, au-delà des extrémités du métacarpien et de la phalange distale.

Notons que ces extrémités ne sont souvent pas atteintes par les contours appariés, du fait notamment du critère de direction et des difficultés intrinsèques, respectivement ambiguïtés et manque de contraste. Il faudra donc estimer ces extrémités, d'une manière compatible avec les limites de la technique de segmentation, on verra cela à la fin de 2.6.4.

2.6.3 Problème de l'échelle pour la segmentation

La segmentation longitudinale consiste à localiser les limites des os le long de chaque axe. Techniquement, cela revient à identifier les bons extrema de la dérivée première de l'image dans la direction locale de l'axe [CGE94, CNHS99, PGP⁺01, MDDS01, HJT07, MFMFAL03].

Cette dérivée se calcule facilement, puisqu'elle s'exprime comme $\mathcal{D}_\mathbf{a} = \mathbf{a}^T \mathbf{g}$, avec \mathbf{a} le vecteur unitaire tangent localement à l'axe et \mathbf{g} le gradient de l'image. Cette caractérisation est préférable à l'utilisation des maxima issus de la NMS, car on conserve le signe de la dérivée et on peut écarter des bruits de fort gradient mais mal orienté. Pour cette dernière raison, $\mathcal{D}_\mathbf{a}$ est également préférable à la simple composante en y du gradient, dès que l'axe est notablement incliné. Le vecteur tangent \mathbf{a} se calcule directement et de manière robuste à partir de la forme polynomiale de l'axe. Le champ \mathbf{g} utilisé est celui déjà calculé, à l'échelle de localisation α_{loc} dont le choix résulte des compromis discutés en 2.5.

La figure 2.28 donne un exemple des variations de la dérivée $\mathcal{D}_\mathbf{a}$ le long d'un axe (il s'agit du cas de l'index dans notre image usuelle). On voit que les extrema sont nombreux, très rapprochés et qu'il y a souvent des ambiguïtés, en raison du bruit, des séparations plus ou moins nettes entre diaphyses et épiphyses, de détails internes à l'os, etc. On a extrapolé l'axe, a priori, pour être sûr de considérer entièrement le métacarpien et la phalange distale et il faut donc aussi écarter des extrema qui correspondent à des structures du carpe ou à des objets parasites en bas de l'image (à droite sur la figure).

D'autre part, il n'est pas possible d'utiliser une échelle mieux adaptée à la détection. L'échelle qui a permis la détection initiale des axes est beaucoup trop grossière et, comme on le voit figure 2.25, page 69, des échelles intermédiaires effacent encore trop de détails le long de l'axe. Pour gagner un peu de robustesse de détection, on pourrait considérer un lissage non isotrope, en utilisant l'échelle de localisation dans la direction de l'axe et une échelle plus grossière dans la direction transversale. Notons aussi que cette échelle transversale devrait être adaptée au long de l'axe, pour limiter les effets négatifs d'un lissage sur des extrémités arrondies. En particulier, il conviendrait de désactiver le lissage transversal dans le cas des petites épiphyses qui sont caractérisées par de fortes courbures de leur contour.

On peut envisager un schéma de diffusion non linéaire, relativement classique [CLMC92, Wei97] si on considère que la direction de l'axe est aussi celle du gradient aux points qui nous intéressent. Cette approche exploite, par définition, le gradient et la courbure de l'isophote. Mais, avec une mise en œuvre itérative, elle est relativement coûteuse et, de par sa nature diffuse, elle ne permet pas vraiment d'utiliser la direction de l'axe qui est plus pertinente et plus robuste

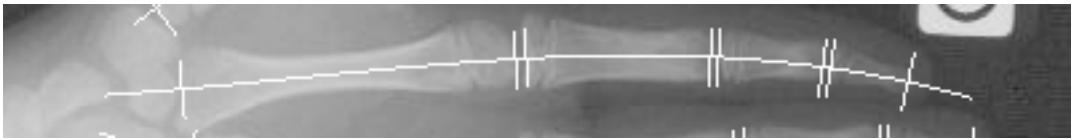
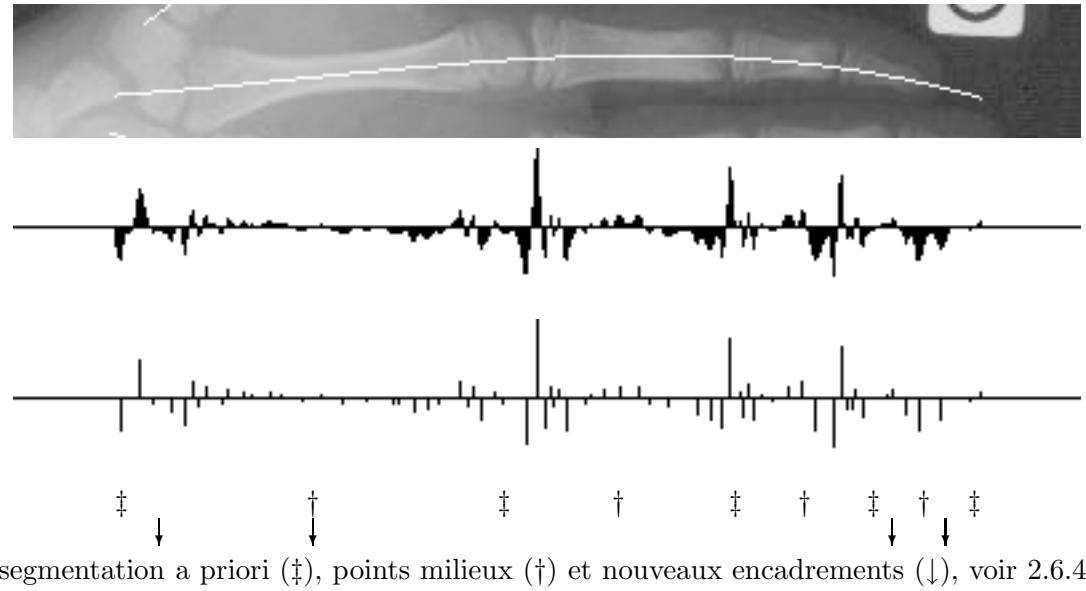


FIG. 2.28 – Variations de la dérivée le long de l’axe et segmentation résultante

que la direction locale du gradient. Ici, il conviendrait plutôt de diffuser la direction de l’axe, dans un champ artificiel, et de l’utiliser à la place du gradient. On peut également envisager une approche plus directe, avec un lissage anisotrope dont la mise en œuvre peut se faire de manière relativement efficace, par convolution ou filtrage récursif [GSv03]. Pourtant ce calcul, qui doit être orienté localement selon l’axe, reste relativement compliqué et demande de nombreuses interpolations entre pixels, ce qui est une source potentielle d’artefacts.

Nos expérimentations ont montré que ces diverses techniques de régularisation ne permettent pas d’améliorer utilement la qualité de la dérivée et un apport de contraintes structurelles est alors indispensable pour déterminer correctement les limites effectives des os.

2.6.4 Segmentation longitudinale d’un axe

Pour guider la segmentation, nous utilisons des rapports qui s’expriment entre les longueurs des différents os de chaque axe. Les valeurs de référence pour ces rapports sont données dans le tableau 2.1. C’est à ce stade que l’identification des axes, c’est-à-dire savoir de quel côté se trouve le pouce, devient indispensable. Nous réalisons alors une première segmentation arbitraire, au prorata des longueurs normales relatives pour chaque os (les marques ‡ de la figure 2.28), mais les frontières entre os ainsi estimées ne peuvent pas être utilisées directement, par exemple, pour sélectionner le contour le plus proche. En effet, leurs positions sont trop sensibles aux diverses erreurs qui peuvent affecter les extrémités et la longueur d’un axe. Ce choix aurait également le gros inconvénient de forcer une interprétation de “la” normalité, déterminée par les constantes λ_{ij} du tableau 2.1.

L’utilisation de ces constantes demande donc une certaine souplesse d’interprétation et il faut introduire une phase de relaxation qui permet de remettre en cause la segmentation a priori. Les extrémités estimées de l’axe et les paramètres de normalité ne servent qu’à déterminer des points milieux (les marques † de la figure 2.28) qui sont supposés se trouver à l’intérieur de

axe j	rang i				μ_j
	λ_{0j}	λ_{1j}	λ_{2j}	λ_{3j}	
1	0.44	0.32	0.24		0.14
2	0.45	0.27	0.16	0.12	0.22
3	0.40	0.30	0.18	0.12	0.24
4	0.39	0.30	0.19	0.12	0.22
5	0.43	0.28	0.16	0.13	0.18

On note le rang $i = 0$ pour un métacarpien, $\dots, i = 3$ pour une phalange distale et on note l'axe $j = 1$ pour le pouce, $\dots, j = 5$ pour l'auriculaire. On considère que le pouce n'a pas de troisième phalange. Cette interprétation n'est pas pertinente d'un point de vue anatomique, il serait plus juste de dire que le pouce n'a pas de métacarpien, mais elle s'avère la plus pratique pour les calculs.

Pour l'os de rang i et d'axe j , nous notons par λ_{ij} la longueur normale relative de cet os dans le cumul des longueurs normales pour ce doigt et nous notons par μ_j la part de ce cumul dans le cumul général pour les 5 doigts. Autrement dit, les longueurs mesurées devraient normalement vérifier les égalités suivantes $L_{ij} = \lambda_{ij} \sum_k L_{kj}$ et $\sum_k L_{kj} = \mu_j \sum_{kl} L_{kl}$.

Les valeurs respectives des λ_{ij} et μ_j données ici, sont tirées des tables de normalité usuelles [Poz74]. Il s'agit des longueurs épiphyses comprises, comme dans l'exemple figure 2.29. Notons que les valeurs effectives peuvent varier selon l'individu, avec des écarts-types de l'ordre de 5 à 10 % des valeurs nominales. Pourtant, les valeurs normales changent peu en fonction de l'âge et elles peuvent être utilisées pour une segmentation a priori. Enfin, on considère que ces valeurs prennent en compte la déformation perspective qui résulte d'un cadrage normal de la radiographie.

TAB. 2.1 – Valeurs des coefficients “normaux” de longueur relative des os et des axes

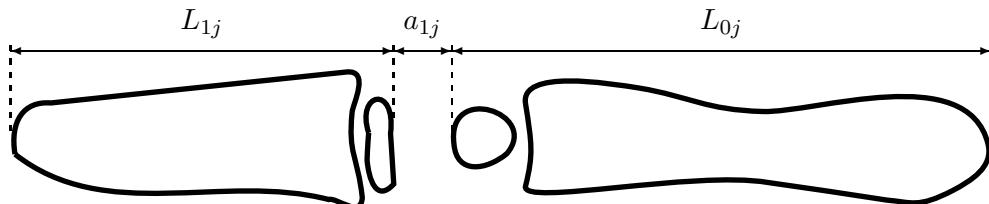


FIG. 2.29 – Longueurs de la première phalange et du métacarpien, pour un axe $j > 1$

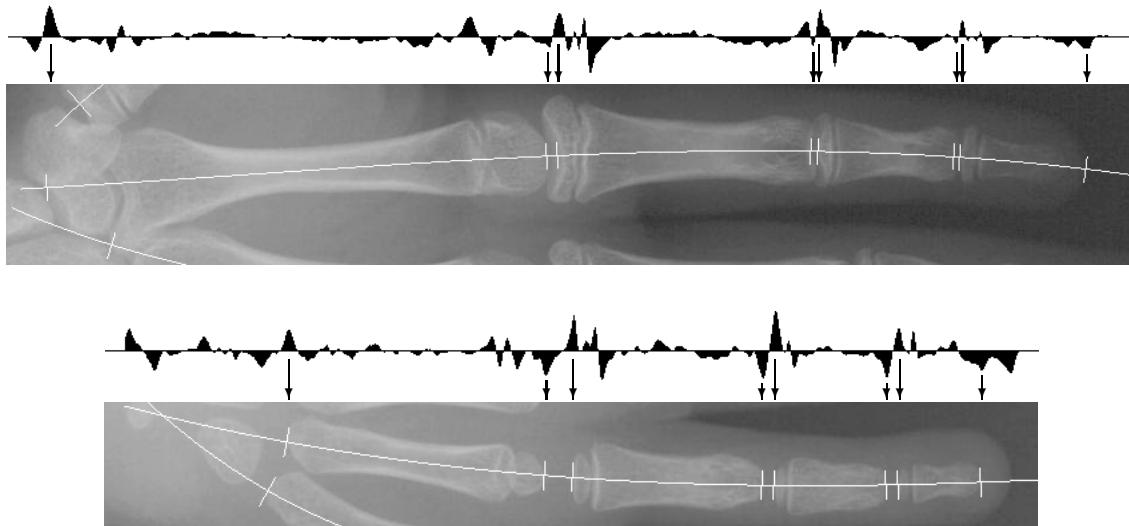


FIG. 2.30 – Exemples de signatures plus complexes

chaque os. Ensuite, dans un intervalle d'analyse qui est limité par les milieux estimés pour deux os consécutifs, on peut rechercher le bord distal de l'un et le bord proximal du suivant, en se basant sur les variations de la dérivée le long de l'axe. Assez souvent, on détecte assez facilement la coupure comme le meilleur couple, formé des deux extrema consécutifs de la dérivée \mathcal{D}_a , de sens opposés, dans l'ordre négatif puis positif.

La figure 2.30 donne des exemples où la caractérisation des coupures est relativement plus délicate. Le choix d'un extremum absolu de \mathcal{D}_a , sur l'intervalle d'analyse, ne conduit pas toujours à la bonne réponse. On doit tenir compte de la distance entre deux extrema consécutifs, qui constitue une mesure de la dimension d'une structure ou de l'espacement entre deux structures. On considère également la “largeur” d'un extremum de \mathcal{D}_a qui est estimée par l'écart entre les deux extrema de sa dérivée²⁰ qui l'encadrent. Ces contraintes structurelles permettent de filtrer correctement les variations qui se rapportent aux brillances corticales, ainsi que les coupures entre diaphyse et épiphysse qui donnent toutes des signaux relativement très étroits, plus facilement détectables par la dérivée seconde. Les détails intérieurs à l'os ne satisfont pas les contraintes d'ordre, d'espacement ou de couplage. On notera que tout cela revient à considérer que la caractérisation d'une articulation repose principalement sur sa structure spectrale.

Nous avons mis cela en œuvre par une recherche séquentielle de type automate qui présente donc une certaine similitude avec la construction des points médians. En amont, on utilise un seuillage faible, adapté sur l'intervalle, mais qui ne sert essentiellement qu'à réduire le nombre de signaux injectés dans l'automate. En pratique, il semble préférable de balayer chaque intervalle d'analyse dans le sens distal vers proximal (de droite à gauche dans les figures 2.28 et 2.30) pour détecter d'abord l'extrémité de l'os distal. La recherche de l'extrémité de l'os proximal se fait ensuite, avec un cas particulier pour les 4 métacarpiens qui seuls peuvent présenter une épiphysse distale plus ou moins séparée de la diaphyse.

La justesse de cette méthode repose sur le double fait que, premièrement, les erreurs sur les extrémités d'un axe restent dans certaines limites, ainsi chaque point milieu se situe bien à l'intérieur de l'os correspondant, et que, deuxièmement, les critères de caractérisation sont bien ciblés et pertinents sur un intervalle de recherche qui est supposé contenir une seule articulation. Pour quantifier la tolérance la méthode, on peut regarder pour quels écarts sur les extrémités de l'axe, un point milieu estimé reste dans les limites de l'os. Pour chaque axe, on obtient ainsi un cas

²⁰C'est-à-dire une dérivée seconde du signal. Cette mesure est parfaitement fondée dans le cas où \mathcal{D}_a a un profil gaussien, en réponse à un contour de type “marche”, parfait et isolé.

particulier de simplex en dimension 2 dont les variables sont ces écarts. Le fait qu'un point milieu, calculé avec ces écarts, reste dans les limites effectives se traduit par une paire d'inéquations, paramétrées par les valeurs du tableau 2.1 et le problème est parfaitement symétrique vis-à-vis d'écarts positifs ou négatifs. On vérifie ainsi que la phalange distale pose le plus de problème car, du fait de sa faible longueur, son point milieu estimé se retrouve rapidement décalé en dehors des limites effectives. La méthode est naturellement deux fois plus robuste sur le pouce. Ce résultat est assez heureux, car c'est l'axe du pouce qui est le plus délicat à construire.

Les mains très jeunes pourraient aussi poser un problème a priori, au regard des espaces articulaires importants entre des os relativement courts (voir les exemples figure 3, page 8 de l'introduction). L'intervalle de tolérance pour le placement d'un point milieu est alors réduit, mais les extrémités des axes sont en fait bien estimées, car il y a peu d'ambiguités de contours, et les points milieux restent donc dans les limites de leurs os respectifs.

De manière générale, le placement de l'extrémité distale d'un axe est donc le plus critique, puisqu'il affecte fortement la dernière phalange, mais des erreurs de l'ordre de 5 % de la longueur totale du doigt sont toujours acceptables. Avec des extrémités distales correctement estimées, la méthode tolère des écarts pouvant aller jusqu'à 25 % sur les extrémités proximales.

Ces résultats ont déterminé les heuristiques qui conduisent à l'extrapolation des axes. On peut considérer que les extrémités sont toujours sous-estimées a priori. Pour commencer, on peut donc procéder arbitrairement à l'allongement des axes qui sont considérés comme trop courts, relativement aux autres et dans des limites qui permettent une extrapolation correcte. Notons qu'un axe nettement trop court pose des problèmes de précision pour l'estimation de son approximation polynomiale et les portions extrapolées peuvent alors dévier notablement de l'axe réel. Une telle erreur rend également aléatoire l'identification des axes qui est basée sur leur longueurs respectives. C'est pourquoi, on décide de l'échec de la détection d'un axe si sa longueur, avant correction, est estimée comme étant trop courte. Le seuil est fixé expérimentalement à 20 % en dessous de la normale²¹. Pour plus de robustesse, la méthode ne tient pas compte d'un axe qui serait le seul manifestement trop long, pour décider de l'allongement des autres.

En considérant l'alignement des extrémités proximales des métacarpiens (voir le principe en 2.4.7), on décide déjà si certaines de ces extrémités doivent être "remontées". Ensuite, avec les ratios μ_j du tableau 2.1, on décide si certaines extrémités distales doivent être "descendues", pour obtenir des rapports de longueurs proches des valeurs normales. Enfin, les 5 axes sont systématiquement étendus aux deux extrémités. Compte tenu de la sous-estimation initiale, une extension empirique de 10 % de leur longueur, à chaque extrémité, conduit normalement à une surestimation qui reste dans les limites acceptables pour placer correctement les points milieux. Dans le cas d'axes majoritairement trop courts, cette correction n'est pas suffisante. En fait, cela résulte d'un grave défaut dès la détection des axes (voir par exemple le cas de la figure 2.17, page 62). Il n'est alors pas utile de poursuivre le traitement car une qualité d'image qui est déjà délicate pour la détection, ne permet pas une localisation appropriée à des mesures.

Dans le cas général, on parvient ainsi à détecter et localiser correctement les articulations, de manière précise et robuste. On en déduit les longueurs effectives pour les phalanges proximales et médianes et, avec les valeurs du tableau 2.1, on peut estimer la longueur normale du métacarpien et donc déterminer, avec un peu de marge, un intervalle de recherche pour l'extrémité proximale de ce métacarpien (les marques ↓ de la figure 2.28). L'autoadaptation de l'intervalle d'analyse permet généralement de s'affranchir d'une estimation approximative de l'extrémité de l'axe, mais la présence de plusieurs maxima (carpe, bruit, etc.) exige un intervalle relativement étroit et bien localisé, ce qui peut poser des problèmes du fait de la variabilité naturelle. Le premier

²¹C'est très variable, selon l'os considéré, mais une déviation de la longueur de 20 % par rapport à la normale correspond à au moins 2 ou 3 fois l'écart-type, d'après [Poz74]. Ce seuil de 20 % sera également justifié au chapitre 3, par l'analyse de nos résultats expérimentaux.

exemple de la figure 2.30 montre un cas relativement courant d'erreur sur la limite proximale du métacarpien. Pour les 4 "vrais" métacarpiens, l'extrémité proximale n'a pas d'épiphyse et elle est définie simplement comme un maximum de la dérivée. Pour le pouce, il convient plutôt de reprendre la caractérisation qui a été utilisée pour les phalanges.

On procède de manière assez similaire pour l'extrémité distale de la dernière phalange. On voit sur les exemples la difficulté à distinguer les extrémités "os" et "peau" par l'intensité du gradient. On utilise donc essentiellement la position sur l'axe. Le problème devient difficile quand ces contours sont faibles et se confondent avec le bruit. Un *a priori* de normalité est alors déterminant.

La définition de la normalité conditionne donc, à plusieurs reprises, la segmentation le long de chaque axe. Il est alors naturel de penser que cette influence n'est pas négligeable et elle sera étudiée expérimentalement au chapitre suivant.

2.7 Bilan, mise en œuvre et généralisation

Les figures 2.28, page 75, et B.3, page 180, montrent que le résultat obtenu sur notre exemple usuel est globalement satisfaisant. En particulier, les mesures de longueur de chaque os sont alors directement accessibles. En analysant statistiquement les longueurs ainsi obtenues, nous verrons au prochain chapitre que l'on peut compter sur des résultats de qualité comparable pour la plupart des "bonnes" images.

Notre méthode parvient ainsi à réaliser les opérations de détection et de localisation qui sont délicates en l'absence de définition préliminaire des zones d'intérêt et d'*a priori* sur la qualité de l'image. Nous utilisons essentiellement un seul *a priori*, celui de "bon" cadrage qui nous permet d'estimer des dimensions caractéristiques qui déterminent le paramétrage. Notre méthode peut ainsi constituer la première phase d'une chaîne de traitements automatiques en permettant l'acquisition de mesures et la localisation qui seraient requises pour la mise en œuvre de méthodes particulières ou plus sophistiquées. Nous en donnons quelques exemples en 2.7.2.

Les techniques qui sont décrites dans ce chapitre sont très nettement adaptées au problème particulier des radiographies de la main, par le choix des procédures et de leur paramétrage. La variabilité du contenu et de la qualité technique des images demande pourtant une certaine généralité et une grande souplesse d'autoadaptation, d'autant plus que l'on s'interdit toute intervention au cas par cas. Les principes qui guident le paramétrage sont rappelés en 2.7.1.

On peut donc envisager la mise au point d'autres réglages, pour des images assez différentes, par la qualité ou le contenu. On peut aller aussi plus loin et considérer la possibilité de réemployer certaines procédures, relativement génériques, dans d'autres contextes. C'est le cas notamment pour la construction des axes, dès lors qu'il existe une direction privilégiée (voir 2.7.3).

2.7.1 Réglage des paramètres

On suit une méthodologie classique de groupement perceptuel [SB93] qui enchaîne un nombre relativement important d'étapes. À chaque étape, on essaie donc de regrouper certains éléments primitifs, sélectionnés par un seuillage, dans des structures un peu plus cohérentes. Les éléments trop isolés sont alors écartés et les structures conservées sont des primitives pour l'étape suivante.

Dans la plupart des cas, un paramètre intervient dans un seuillage et, pour assurer une autoadaptation, on a recours à un paramétrage indirect. Les paramètres que l'on doit fixer servent alors à déterminer l'adaptation de paramètres effectifs à partir de mesures faites sur l'image. Il s'agit, soit de déterminer un seuil à partir d'éléments déjà acquis, soit de conditionner les valeurs qui sont seuillées, et il s'agit de conférer à chaque paramètre assez d'efficacité dans l'élimination d'un certain type de parasites. Une modélisation partielle permet généralement d'estimer une valeur théorique, proche de celle effectivement requise, mais sans plus. Ensuite,

une adaptation expérimentale reste nécessaire pour corriger les approximations du modèle et doser une certaine tolérance dans l'interprétation.

Fixer la valeur d'un paramètre revient souvent à arbitrer un compromis, dans ce contexte approximatif, entre la probabilité d'occurrence des fausses sélections et celle des faux rejets. Les mécanismes de mise en évidence par accumulation qui sont mis en place après un seuillage, reposent sur l'hypothèse [Low85] que la co-occurrence de plusieurs fausses sélections dans une même structure est relativement improbable, ce qui relâche des contraintes sur la valeur du paramètre. Cela demande, par ailleurs, de veiller à l'indépendance statistique entre les éléments construits lors des étapes précédentes. Dans le même temps, la validation *a posteriori* par l'appartenance à une structure cohérente jouent un rôle de "réparation" et permet de limiter les problèmes liés aux faux rejets.

Le nombre de paramètres est relativement important, de l'ordre d'une bonne quarantaine, mais la conception générale de l'algorithme et le rôle bien localisé de chacun de ces paramètres nous permettent un réglage progressif. Le découpage en étapes bien distinctes, par les objectifs et par les techniques mises en œuvre, facilite aussi grandement la détermination des valeurs effectives des paramètres. Notamment, on peut juger de la qualité intrinsèque de chaque étape et procéder par une suite de réglages que l'on peut considérer comme indépendants les uns des autres. En pratique, il est même utile de "dérégler" les étapes précédentes pour mieux apprécier l'adéquation générique d'un réglage. Le paramétrage peut être ainsi déterminé après quelques réglages expérimentaux et son adéquation est vérifiée *a posteriori* sur un large échantillon.

Ainsi, quelques essais suffisent pour déterminer une valeur satisfaisante pour un paramètre, au sein d'un processus d'ensemble relativement séquentiel, et, même si quelques retours sont nécessaires, le tout est très léger en comparaison avec la dimension combinatoire du problème. On peut vérifier *a posteriori* que, pour l'ensemble des "bonnes" images, des variations assez sensibles des paramètres ne changent pas significativement le résultat final, notamment la localisation et la segmentation des axes. Cela permet de retenir des valeurs relativement "rondes", comme celles qui sont proposées au fil de ce chapitre.

Par contre, pour de mauvaises images, le problème est très instable et, pour un lot de quelques images déjà, il n'est pas possible de trouver un paramétrage commun satisfaisant.

Dans les premières étapes, l'apport de connaissance anatomique est très faible. En fait, cela revient considérer que la main est formée de 5 doigts et qu'elle est relativement bien cadrée. On utilise ainsi implicitement le fait qu'il existe plusieurs niveaux de structures (main/doigts/détails d'ossification/textures) et que les dimensions caractéristiques d'un certain niveau sont relativement constantes dans une image et nettement différentes de celles des autres niveaux. Les rapports de dimensions entre niveaux sont des invariants anatomiques très grossiers.

Les paramètres qui déterminent l'obtention de l'orientation de l'image (voir 2.1.1) se limitent à définir les deux bandes transversales de façon relativement arbitraires. Avec un cadrage usuel, la définition de ces bandes est peu contrainte et c'est ce qui fait la robustesse de cette technique. Le recours à des relations d'ordre, en comparant les dimensions de l'image et les valeurs moyennes sur les deux bandes, nous évite d'avoir à fixer des seuils. Des techniques plus sophistiquées (voir 2.1.2) requièrent elles aussi des contraintes de cadrage et elles seraient plus délicates à régler. Elles font notamment intervenir des réglages d'échelles et des seuils et certaines sont très sensibles à la présence d'éléments étrangers dans l'image.

Finalement, l'orientation n'est pas un problème dans la mesure où notre méthode permettrait d'explorer rapidement quelques orientations. Le point essentiel est l'hypothèse de cadrage qui permet l'adaptation de divers paramètres en fonction des dimensions de l'image, principalement l'échelle de détection (voir 2.3.1) et les seuils sur les largeurs d'appariement (voir 2.3.3). Des paramètres basés sur les dimensions caractéristiques des objets à analyser, semblent nécessaires

dans toute technique et, si on ne peut pas les estimer *a priori*, on doit procéder à l'essai de plusieurs valeurs pour obtenir une réponse suffisante.

Les seuils sur les maxima du gradient à l'échelle de détection sont adaptés en fonction d'un histogramme (voir 2.3.2). Les autres paramètres qui interviennent dans la validation des contours, y compris ceux de cette adaption, sont peu critiques car l'appariement apporte ensuite une cohérence structurelle suffisante. La validation du nombre de points médians (voir 2.3.4) se réfère à des considérations anatomiques, mais sans grande contrainte sur la valeur effective des paramètres associés.

La construction des axes (voir 2.4) comporte de nombreuses étapes et elle utilise de ce fait de nombreux paramètres. Un certain nombre de ces paramètres se rapportent au groupement des pixels en chaînes et servent à définir la relation de connexité locale. Même si cette connexité n'est pas isotrope, ces paramètres sont relativement indépendants de l'application et ne demandent pas de réglage particulier. En fait, les valeurs de ces paramètres traduisent certaines propriétés intrinsèques comme, par exemple, la séparation minimale entre les chaînes qui est assurée par la technique de génération des points médians.

Les paramètres qui interviennent ensuite pour regrouper les chaînes et faire les axes sont, par contre, assez dépendants de la structure anatomique d'une main. Ils sont ajustés en tenant compte de la longueur possible d'une coupure au niveau d'une articulation et de l'écart entre deux axes voisins. On fournit en fait des facteurs de forme qui permettent de déterminer ces paramètres à partir de la largeur moyenne d'appariement mesurée. Cependant la quantification de la continuité de deux chaînes, même paramétrée au mieux, laisse subsister une certaine ambiguïté et c'est ce qui fonde le choix de l'algorithme de groupement pour cette étape. La logique de décision des regroupements à effectuer est ainsi prépondérante dans la construction des bons axes, là où ils existent. La paramétrisation est simplement ajustée pour ne pas les interdire et intervient plutôt pour limiter le raccordement d'éléments superflus.

Par ailleurs, l'opération la plus coûteuse est le calcul d'une carte de gradient. Le nombre d'itérations requises, lors de l'ajustement de l'échelle de détection (voir 2.3.4), a donc un rôle déterminant sur le temps de calcul. Deux points particulièrement critiques sont, d'une part, le choix de l'échelle initiale en fonction des dimensions de l'image, et, d'autre part, la loi de correction de cette échelle. Il s'agit de la problématique classique de recherche d'une solution approchée dans les algorithmes d'optimisation itérative. En utilisant la robustesse des étapes suivantes, nous avons préféré paramétriser pour l'estimation d'une échelle initiale sûre, dans les limites où la mise en évidence des axes restait suffisante. On verra en 3.1 que, de ce point de vue, il n'est généralement pas utile d'affiner le choix de cette échelle. Il s'agit donc là de l'exception où le paramétrage ne vise pas l'optimisation locale d'un critère de qualité.

On a passé ainsi en revue tous les paramètres qui interviennent dans la construction des axes à l'échelle de détection. Les quelques critères anatomiques pris en compte jusqu'à ce stade ne font que traduire la forme générale de la main, avec une tolérance suffisante pour s'adapter aux variations de maturation et de pose. Le paramétrage de la construction des axes à l'échelle de localisation se fait sur des principes similaires. Mais on dispose maintenant d'une mesure de la largeur effective d'un doigt et ensuite, plus précisément, de celle des os et on utilise ainsi beaucoup moins d'estimations empiriques. Le paramétrage est également facilité car les contraintes structurelles bien dirigées sont alors plus efficaces que les seuillages. Cela se vérifie notamment dans la validation des contours longitudinaux.

Le paramétrage prend un aspect plus délicat dans la phase de segmentation longitudinale des axes (voir 2.6.4). Notre choix d'une résolution dirigée par des contraintes structurelles permet de limiter le paramétrage technique, notamment il n'y a pas à fixer de seuils critiques. Par contre, la méthode repose sur un jeu de paramètres qui traduisent les dimensions anatomiques attendues. On doit alors planter des valeurs très précises qui fixent *a priori*, et qui sont tirées de

statistiques sur les mensurations. Ces valeurs dirigent fortement le choix des coupures entre les os et favorisent une interprétation conforme à la normalité. L'influence d'un paramétrage particulier sur les résultats sera étudiée expérimentalement au chapitre suivant.

2.7.2 Association avec d'autres techniques de traitements des images

Ainsi, sans requérir aucun traitement en amont, avec simplement une hypothèse de cadrage peu contraignante, notre méthode parvient à une compréhension de l'image qui est très largement suffisante pour initialiser et paramétrier correctement diverses techniques complémentaires.

La segmentation à laquelle nous aboutissons, permet d'identifier, pour chaque os, la portion d'axe et les contours longitudinaux qui lui sont associés. Cela permet la définition de "formes englobantes" convenablement localisées, orientées et dimensionnées. Ce résultat est bien adapté à la mise en œuvre de traitements complémentaires, éventuellement spécialisés à chaque os, comme des mesures de longueurs, la détection et localisation des épiphyses, etc. En particulier, la simple mesure des longueurs est la base de la technique d'analyse par le profil métacarpo-phalangien [PGNG72] qui intervient dans le diagnostic de nombreuses pathologies [PG97].

L'analyse de la forme des épiphyses est une composante essentielle des diverses méthodes de détermination de l'âge osseux [GP83, RCT88, TWC⁺83, THGC01, ER67]. L'approche par les contours n'est alors pas la meilleure, mais on dispose maintenant des éléments de localisation et d'orientation qui permettent d'envisager diverses techniques d'analyse soit par des signatures spectrales, soit par recalage de modèles de forme ou d'apparence. On notera que la localisation permet aussi de procéder à des mesures de densité significatives et donc à une normalisation des intensités. Ainsi, on a le moyen de faire abstraction des divers facteurs techniques de la radiographie et la seule variable restante est alors le stade de maturation de l'articulation considérée.

Plus globalement, on dispose aussi d'un nombre suffisant de points clefs fiables pour envisager des techniques de recalage sur l'ensemble de la main. On verra au chapitre 4, comment cela peut contribuer à la détection et à l'identification des os du carpe. Dans le même ordre d'idée, si on considère des modèles de forme, il est maintenant possible de faire le choix préalable entre plusieurs variantes topologiques selon le stade de maturation, même si celui-ci n'est pas explicité, et de construire un ensemble de points clefs pertinents.

Notre méthode peut ainsi constituer une base pour des protocoles d'apprentissage automatique, en soulignant cependant que le principe d'un apprentissage statistique restera contestable pour la caractérisation d'anomalies rares. Si l'on vise un apprentissage de la normalité, on peut considérer que l'on dispose d'une segmentation généralement bonne sur un grand nombre d'images. Nous verrons au chapitre suivant qu'il est possible de détecter les erreurs les plus flagrantes, ce qui permet d'exclure de l'apprentissage (ou du recalage) les quelques éléments les plus déstabilisants. Les erreurs plus fines, souvent indiscernables de la variabilité anatomique, seront ensuite moyennées par la régression inhérente à des techniques statistiques.

2.7.3 Applications à d'autres contextes

De nombreux objets naturels ou fabriqués présentent des symétries axiales et, dans un certain nombre de configurations, de telles symétries en 3D donnent lieu à des symétries approximatives dans l'image, notamment les symétries dites obliques [Kan81]. L'existence de symétries permet d'envisager la détection des objets par leurs axes. À défaut d'être toujours linéaires, ces axes sont très réguliers et cela facilite leur mise en évidence par accumulation et regroupement. L'appariement des contours contribue à fiabiliser leur détection car on peut mettre en œuvre des critères de cohérence plus forts que les critères locaux utilisés lors d'un chaînage des contours.

Une contrainte de symétrie est toutefois moins forte qu'une contrainte de forme et elle est particulièrement pertinente lorsque les axes de symétrie sont les seuls éléments structurants

qui soient relativement invariants, au regard de la variabilité des images à traiter. C'est le cas en particulier quand, au-delà des aspects techniques d'illumination, de pose ou de bruit, la variabilité concerne aussi la forme des objets et que l'on doit ainsi combiner deux modèles complexes, celui de formation des images et celui des objets eux-mêmes.

Dans notre cas, les axes sont invariant vis-à-vis du développement osseux, au contraire des contours et, plus généralement, de la forme des os dont l'évolution serait délicate à modéliser avec assez de souplesse et de précision pour guider la reconnaissance. En pratique, notre méthode s'accorde avec des symétries assez imparfaites que présentent les os longs et leur voisinage.

Il faut aussi remarquer que, pour une paire de contours, il y a invariance de la position de leur axe de symétrie en fonction de l'échelle, comparativement aux contours qui se déplacent de manière importante, voire qui peuvent fusionner. Les effets de l'échelle sur un champ symétrique s'annulent et l'axe de symétrie ne bouge pas (ou, en pratique, relativement peu pour une symétrie approximative). La propriété se vérifie toute une plage d'échelles, avec une limite supérieure qui dépend de la présence d'objets voisins et une limite inférieure qui dépend du niveau de bruit. Cette invariance permet d'assurer directement la cohérence entre deux échelles extrêmes. Ainsi, par exemple, des axes détectés à une échelle très grossière sont directement utilisables pour valider des paires de pixels de contours obtenus à une échelle nettement plus fine.

La mise en évidence et la sélection des axes par des caractérisations usuelles est un processus relativement lourd et complexe à paramétrier (voir annexe A.3). C'est le cas plus particulièrement quand il existe de nombreuses symétries, plus ou moins locales, dont l'évidence est relativement subjective et dont un certain nombre ne sont pas intéressantes. On peut cependant envisager de telles techniques isotropes pour un prétraitement qui permettrait de mettre en évidence une direction d'ensemble.

Notre algorithme d'appariement des contours est utilisable dès que l'on sait déterminer une direction de balayage et le traitement est alors très rapide, en regard des techniques usuelles. Dans notre cas, cette direction doit être voisine de la perpendiculaire aux axes et l'utilisation d'une direction différente pour chaque axe est envisageable si on peut l'estimer a priori. Dans le cas d'une symétrie oblique, la direction d'axe compte peu, il convient d'estimer la direction particulière d'appariement. Dans tous les cas, pour des variations lentes de la largeur d'appariement, cet algorithme tolère une certaine erreur sur la direction de balayage.

Pour un coût qui reste essentiellement dicté par le balayage pixel par pixel, l'automate d'appariement permet de traiter spécifiquement des cas non triviaux, comme par exemple les contours emboîtés du type '()' ou l'ambiguïté du motif '()'. Avec les caractérisations classiques des symétries, dans le meilleur des cas, on ne considère que des paires de la forme '()' et on ne voit pas de tels aspects structurels. En changeant les règles d'évaluation, il est possible de spécifier différemment quelle paire est prépondérante dans un cas '()' ou dans quelles conditions '()' sera interprété comme l'expression d'une ou de deux symétries.

En changeant l'alphabet d'entrée, les états significatifs ou les règles de transition, il est aussi possible de s'adapter à d'autres modèles structurels. Par exemple, on utilise la même méthode pour traiter les articulations, avec des extrema des dérivées première et seconde en entrée et plusieurs variantes de l'automate selon que l'on accepte ou non la présence d'une épiphyse.

Une telle approche se justifie chaque fois qu'un filtrage structurel est nécessaire, soit parce que, utilisées seules, les caractérisations numériques seraient trop ambiguës, soit parce qu'il n'est pas possible de mettre en œuvre des techniques alternatives qui demanderaient, par exemple, un apprentissage.

Chapitre 3

Validation expérimentale

Nous allons maintenant expérimenter les limites de la méthode que nous avons détaillée au chapitre précédent. Pour un rappel sur la terminologie anatomique, on se reportera aux premières pages de l'introduction, notamment les figures 1 et 2.

L'essentiel de cette étude est basé sur des statistiques sur plusieurs centaines d'images. En 3.1, nous allons vérifier la validité apparente de nos résultats, vis-à-vis des références de normalité morphométrique. Le besoin d'une qualité minimale des images apparaît nettement. Il faut également considérer l'influence des paramètres qui induisent une localisation *a priori* des articulations. Comme nous le verrons en 3.2, des différences de sensibilité sont alors nettes selon les os. D'autre part, en 3.3, sur un petit échantillon, nous pourrons effectivement comparer les mesures de longueurs issues de notre segmentation à celles réalisées par les radiologues.

Cette validation est convaincante pour les phalanges proximales et médianes, c'est-à-dire pour les articulations qui les délimitent. Il est ainsi possible d'envisager des applications diagnostiques, comme esquissé en 3.4, mais des traitements complémentaires sont nécessaires pour fiabiliser la détection et améliorer la précision. Il apparaît notamment comme nécessaire de traiter le carpe, pour des informations intrinsèquement utiles aux diagnostics, mais aussi pour améliorer la localisation des métacarpiens.

3.1 Validation de la chaîne de traitement

3.1.1 Robustesse de l'algorithme

Le tableau 3.1 résume le déroulement général de notre algorithme sur les 861 images de la base de test. Le détail par catégorie de qualité visuelle nous permet de valider le bien-fondé des critères énoncés en 1.3. Ainsi, pour un peu plus de 96 % des images dites de “bonne” qualité, la construction et la segmentation des axes aboutissent sans problème majeur. On constate que le taux de succès décroît nettement lorsque les images ne satisfont pas tous les critères de qualité.

De manière générale, un échec est prononcé le plus souvent lors de la validation de l'échelle de détection (colonne `méd.` du tableau 3.1), c'est-à-dire très tôt dans le processus de traitement. On utilise pour cela une mesure de la “quantité d'axes détectés” qui est basée sur le comptage

qualité visuelle	nombre d'images	échecs				succès	
		méd.	axes	seg.	total		
bonne	601	11	3	10	24	577	96 %
intermédiaire	146	14	3	9	26	120	82 %
mauvaise	114	39	3	10	52	62	54 %

TAB. 3.1 – Aboutissement du traitement des axes des doigts

qualité visuelle		nombre d’itérations						total
		1	2	3	4	échecs		
bonne	nb im.	563	8	10	9	11	601	
intermédiaire	nb im.	114	8	3	7	14	146	
mauvaise	nb im.	65	3	4	3	39	114	

TAB. 3.2 – Nombre d’itérations lors de la localisation des axes

des points médians (voir 2.3). Lorsqu’il n’est pas possible, après quelques itérations, de détecter une trace suffisante de 5 axes, sans trop de parasites, nous déclarons un échec. C’est le cas pour 8 images qui ne sont pas des images de main et pour un certain nombre de mains très mal cadrées ou insuffisamment contrastées. Le choix d’un nombre limite d’itérations sera justifié en 3.1.2.

Les échecs qui peuvent être déclarés à l’issue de la construction des axes à l’échelle fine (colonne **axes** du tableau 3.1), sont relativement rares. Le contrôle à ce stade consiste à vérifier que les longueurs relatives des axes, avant corrections, sont sensiblement conformes à la normale, selon les valeurs μ_j du tableau 2.1, page 76. Compte tenu de la robustesse de la phase de segmentation, on reste assez tolérant en ce qui concerne les axes trop longs¹. Par contre, on considère un échec dès qu’un axe apparaît comme trop court de 20 %, relativement aux autres, car son extrapolation pourrait être trop imprécise. Les erreurs de longueur rendent également aléatoire l’identification des axes, c’est-à-dire savoir de quel côté se trouve le pouce. On notera que l’on ne peut pas détecter d’anomalie de longueur dans le cas où tous les axes sont trop courts, mais restent relativement bien proportionnés, ce qui se produit généralement lors d’un manque de contraste aux extrémités distales.

De manière plus fine, après la segmentation, on peut contrôler la mesure de longueur relative pour chaque os, vis-à-vis des valeurs λ_{ij} du tableau 2.1, page 76. Ceci fera l’objet d’une étude détaillée en 3.1.4. On ne reporte ici (colonne **seg.** du tableau 3.1) que les mesures manifestement trop courtes ou trop longues, respectivement inférieure au dixième ou supérieure au double de la longueur normale. Nous déclarons alors un échec car l’identification des os le long de l’axe est certainement fausse, avec un décalage possible dans l’identification des os, soit du fait d’un axe trop court et confusion avec une épiphyse, soit du fait d’un axe trop long et confusion avec un os du carpe ou un artefact.

Dans les deux cas, **axes** et **seg.**, l’interprétation est sévère puisque les autres os peuvent être malgré tout correctement segmentés. Nous justifions ce rejet de l’image par l’incertitude sur l’identification et les difficultés que l’on aurait à gérer et interpréter les données manquantes lors de l’analyse des résultats. Le peu d’images qui sont ainsi rejetées, n’altère pas la validité statistique de l’étude qui suit.

Globalement, en ce qui concerne les 24 cas d’échec constatés pour de “bonnes” images, un nouvel examen visuel a posteriori nous a montré que 23 de ces images présentaient des défauts techniques (cadrage, artefacts, contraste, etc.) dont l’importance avait été sous-estimée lors de la classification initiale. Dans le cas restant, la cause principale de l’échec est un défaut anatomique, l’image ayant été prise suite à un traumatisme. De tels défauts justifiaient le reclassement des 24 images dans la catégorie de qualité “intermédiaire”, pour laquelle le taux de succès tomberait alors aux environs de 70 %, ce qui reste un relativement bon résultat.

3.1.2 Influence du nombre d’itérations

Nous avons limité à 4 le nombre maximum d’itérations qui peuvent être réalisées durant la phase d’ajustement de l’échelle de détection. Comme le montre le tableau 3.2, la détection est généralement satisfaisante dès la première itération. C’est le cas, par exemple, pour environ 94 %

¹Notamment, le cas d’un seul axe trop long, avant segmentation, n’affecte pas le traitement des autres axes.

des “bonnes” images. Quelques itérations sont justifiées lorsque les dimensions de la main ont été surestimées en raison d’un cadrage trop large.

Dans les autres cas, le fait d’autoriser un grand nombre d’itérations n’améliore pas les résultats de manière significative. Le plus souvent, la qualité ou le contenu de l’image ne permettent pas de déterminer une détection suffisante des 5 axes et cela quelle que soit l’échelle.

Par exemple, si l’on considère les 11 images dites de “bonne” qualité qui donnent lieu à un échec de détection, on constate que, pour 3 d’entre elles, une échelle pourrait être validée en 5 itérations. Mais, dans l’un des cas, on aboutit alors à un échec de construction des axes (passage de la colonne **méd.** à la colonne **axes** dans le tableau 3.1) et, dans les deux autres cas, on aboutit à une segmentation, acceptable selon 3.1.1, mais de relativement mauvaise qualité selon 3.1.3 et 3.1.4. Enfin, dans la plupart des cas d’échec, il n’est pas possible de détecter suffisamment de points d’axes, quelle que soit la valeur échantillonnée pour le paramètre d’échelle, et le mécanisme d’itération ne converge donc pas.

3.1.3 Erreurs de segmentation

Pour analyser les résultats de la segmentation, nous avons effectué une comparaison, pour chaque os, entre sa longueur mesurée d’après la segmentation automatique et sa longueur théorique estimée à partir des tables de normalité. Comme dans la légende du tableau 2.1, page 76, pour l’os de rang i et d’axe j , nous notons L_{ij} la longueur effectivement mesurée après la segmentation des axes. Une longueur normale \mathcal{L}_{ij} peut être estimée, pour chaque os, par :

$$\mathcal{L}_{ij} = \lambda_{ij}\mu_j \sum_{kl} L_{kl}$$

La somme (de 19 termes) est une estimation de la longueur totale des 5 axes et les valeurs λ_{ij} et μ_j sont celles données dans le tableau 2.1, page 76. Elles vérifient la relation $\sum_{ij} \lambda_{ij}\mu_j = 1$ et on considère qu’elles prennent en compte les effets d’une déformation projective standard.

Le calcul des valeurs \mathcal{L}_{ij} est donc fait, dans un premier temps, en tenant compte de la somme pour tous les os. Il est clair que ces valeurs peuvent être affectées de manière notable par quelques mesures L_{kl} aberrantes, quand cela conduit à une mauvaise estimation de la longueur des axes correspondants². Ainsi, une “petite” variation de 10 % sur un métacarpien induit une variation de l’ordre de 1 % sur la somme et, donc, sur chaque \mathcal{L}_{ij} . Pour limiter ces perturbations, on va écarter les os dont la longueur mesurée semble anormale, en procédant comme suit.

Pour chaque os, à partir de la valeur L_{ij} mesurée directement, et de l’estimation \mathcal{L}_{ij} , obtenue par pondération de la somme pour tous les os, nous calculons l’écart relatif “brut” $e_{ij} = \frac{L_{ij} - \mathcal{L}_{ij}}{\mathcal{L}_{ij}}$. Cet écart constitue un critère simple pour détecter les mesures probablement “fausses” et nous utilisons un seuil qui est fixé arbitrairement à ± 0.2 ; on verra pourquoi cette valeur en 3.1.5 et 3.3. Une mesure L_{ij} est donc considérée comme aberrante si $|e_{ij}| > 0.2$ et on considère alors que l’os correspondant est mal segmenté. Cette tolérance de 20 % est à comparer aux écarts-types normaux qui sont, en valeurs relatives, de l’ordre de 10 % pour les phalanges distales et de 7 % pour les métacarpiens, d’après [Poz74].

Pour les 577 “bonnes” images, il y a 1136 telles mesures aberrantes, tous os confondus, soit un peu plus de 10 % sur un total de 10 963 mesures (577×19). C’est en moyenne un peu moins de 2 “erreurs” par image, mais le tableau 3.3 montre qu’elles sont nettement regroupées. Ainsi, environ 9 % des images (53 sur 577) concentrent un tiers des erreurs (379 sur 1136), avec au moins 5 erreurs chacune. À l’inverse, un peu plus de 27 % des images (158 sur 577) ne font l’objet d’aucune mesure erronée et le taux de réussite approche les 52 %, si l’on tolère une erreur par image. On mesure encore une fois l’importance de la qualité d’image.

²Alors qu’une erreur de localisation pour une articulation n’affecte pas cette somme mais seulement les valeurs \mathcal{L}_{ij} pour les deux os concernés.

qualité visuelle		nombre de mesures aberrantes						total
		0	1	2	3	4	plus	
bonne	nb im.	158	141	106	72	47	53	577
	% im.	27 %	24 %	18 %	12 %	8 %	9 %	
	nb er.	0	141	212	216	188	379	1136
intermédiaire	nb im.	10	25	23	19	11	32	120
	% im.	8 %	21 %	19 %	16 %	9 %	27 %	
	nb er.	0	25	46	57	44	236	408
mauvaise	nb im.	1	11	9	10	7	24	62
	% im.	2 %	18 %	15 %	16 %	11 %	39 %	
	nb er.	0	11	18	30	28	202	289

TAB. 3.3 – Nombre de mesures aberrantes par image

rang	axe					total	taux
	1	2	3	4	5		
méta	59	30	19	25	59	192	7 %
prox.	51	17	7	8	35	118	4 %
méd.		29	14	20	84	147	6 %
dist.	167	102	65	132	213	679	24 %
total	277	178	105	185	391	1136	10 %
taux	16 %	8 %	5 %	8 %	17 %	10 %	

Contrairement aux conventions d'indices adoptées pour le tableau 2.1, page 76, la présentation et l'analyse des résultats qui sont faites dans ce chapitre, considèrent que le pouce n'a pas de phalange médiane. Il est en effet préférable ici de mettre en évidence les problèmes spécifiques aux phalanges distales, pouce inclus.

Pour des raisons similaires, on va aussi considérer que le pouce comporte un métacarpien^a. Le choix restant entre une phalange proximale ou médiane est alors arbitraire.

^aDu point de vue anatomique, du fait de la présence d'une épiphyshe proximale sur son premier os, il serait plus logique de considérer que le pouce comporte trois phalanges, comme tous les doigts, et pas de métacarpien.

TAB. 3.4 – Nombre de mesures aberrantes, os par os, pour les images de “bonne” qualité

os rang	axe	qualité d'image					
		bonne	intermédiaire	mauvaise			
méta	1	87.4	15.7	73.2	30.3	73.2	40.4
méta	2	91.5	8.6	90.1	11.9	87.2	12.5
méta	3	94.2	7.5	92.1	10.5	83.6	26.7
méta	4	94.1	9.5	88.8	20.0	88.7	18.9
méta	5	90.3	14.1	84.8	23.0	84.4	21.9
prox.	1	91.2	16.2	77.7	29.0	74.8	37.7
prox.	2	93.1	8.4	89.9	10.4	89.2	11.3
prox.	3	95.1	5.2	95.5	3.7	84.8	27.1
prox.	4	95.4	6.1	92.5	14.5	90.9	15.7
prox.	5	93.5	12.3	87.4	21.1	84.7	29.3
méd.	2	92.5	12.7	86.5	26.4	81.0	29.3
méd.	3	94.7	8.8	93.8	8.7	64.5	37.0
méd.	4	93.9	9.7	90.3	15.6	89.6	16.9
méd.	5	87.3	20.0	77.3	31.7	76.8	40.0
dist.	1	81.8	19.2	75.7	28.0	76.3	25.4
dist.	2	87.1	14.5	82.5	24.1	77.1	24.1
dist.	3	89.7	10.6	80.8	18.6	67.1	28.9
dist.	4	84.6	18.2	81.5	27.4	80.1	21.1
dist.	5	78.1	22.1	76.2	29.3	78.9	20.7

TAB. 3.5 – Moyenne (en gras) et écart-type du score de segmentation \tilde{q}_{ij} , os par os

Le tableau 3.4 montre la distribution des mesures aberrantes, os par os, pour ce qui concerne les “bonnes” images. On constate ainsi que presque 60% (679 sur 1136) des ces mesures concernent les phalanges distales. Vu autrement, environ 24% (679 sur 577×5) des mesures de phalanges distales sont “erronées”, soit en moyenne plus d’une erreur par image. Nous avancerons des explications à cela en 3.1.5.

L'estimation des longueurs normales est alors refaite, en écartant de la somme de référence les os dont les mesures effectives sont jugées aberrantes et en pondérant correctement selon les valeurs conservées. On calcule donc $L_{\text{valides}} = \sum_{ij, |e_{ij}| \leq 0.2} L_{ij}$ et $\mathcal{W}_{\text{valides}} = \sum_{ij, |e_{ij}| \leq 0.2} \lambda_{ij} \mu_j$ et, ensuite :

$$\tilde{\mathcal{L}}_{ij} = \lambda_{ij} \mu_j \frac{L_{\text{valides}}}{\mathcal{W}_{\text{valides}}}$$

ce qui nous donne de nouveaux écarts relatifs “corrigés” $\tilde{e}_{ij} = \frac{L_{ij} - \tilde{\mathcal{L}}_{ij}}{\tilde{\mathcal{L}}_{ij}}$.

On réalise ainsi ce qui constitue d'ordinaire une étape dans un processus d'estimation robuste, en notant que le petit nombre de valeurs initiales (ie. 19) et l'usage ultérieur de cette correction ne justifient pas de recourir à des techniques itératives et plus sophistiquées [RL87] où, par exemple, les valeurs aberrantes ne seraient pas directement écartées mais feraient l'objet d'une pondération décroissante en fonction des $|e_{ij}|$ et réévaluée à chaque itération. n

3.1.4 Scores de segmentation

À partir des mesures introduites ci-dessus, on définit pour chaque os un score de segmentation “brut” $q_{ij} = 100.(1 - |e_{ij}|)$ et un score “corrigé” $\tilde{q}_{ij} = 100.(1 - |\tilde{e}_{ij}|)$. Le score corrigé réduit l'influence des valeurs aberrantes et c'est celui qui est utilisé ci-dessous pour les analyses de résultats os par os.

groupe d'os	qualité d'image				
	bonne		intermédiaire		mauvaise
méta	91.5	11.8	85.8	21.5	83.4
prox.	93.6	10.6	88.6	18.9	84.9
méd.	92.1	13.8	87.0	23.3	78.0
dist.	84.3	17.9	79.3	25.9	75.9
axe 1	86.8	17.5	75.5	29.1	74.7
axe 2	91.1	11.6	87.3	19.7	83.6
axe 3	93.4	8.5	90.5	13.0	75.0
axe 4	92.0	12.5	88.3	20.4	87.3
axe 5	87.3	18.5	81.4	27.0	81.2
tous	90.3	14.3	85.1	22.8	80.7
centr.	93.9	7.8	91.5	12.9	87.4
score Q	88.2	10.3	78.7	19.7	71.9
					20.8

Par exemple, la ligne **méta** donne les statistiques de scores “corrigés” sur l’ensemble des \tilde{q}_{0j} , c’est-à-dire pour tous les métacarpiens.

La ligne **tous** concerne l’ensemble des \tilde{q}_{ij} , c’est-à-dire une moyenne arithmétique des scores “corrigés”, sur tous les os de toutes les images, alors que la ligne **score Q** correspond à la moyenne arithmétique, sur toutes les images, des moyennes géométriques Q définies à partir des scores “bruts” q_{ij} .

La ligne **centr.** correspond aux 6 os dits “centraux” pour lesquels la segmentation s’avère la plus robuste et caractérisés par : $(\text{rang}, \text{axe}) \in \{\text{méta, prox}\} \times \{2, 3, 4\}$.

TAB. 3.6 – Moyenne (en gras) et écart-type du score de segmentation, par groupe d’os

groupe	25 %	50 %	75 %	90 %	95 %
méta	97	95	90	83	75
prox.	98	96	93	89	83
méd.	98	96	92	84	75
dist.	96	91	81	55	43
axe 1	97	93	85	69	42
axe 2	97	94	89	82	75
axe 3	98	96	92	87	80
axe 4	98	96	92	84	69
axe 5	98	95	88	59	43
tous	98	95	90	79	61
centr.	98	96	93	88	83
score Q	94	91	86	78	72

Lire, par exemple : pour les os de l’axe 1, au moins 25 % des scores de segmentation \tilde{q}_{i1} atteignent ou dépassent la valeur 97.

TAB. 3.7 – Quantiles des scores, par groupe d’os, pour les images de “bonne” qualité

rang axe	méta					prox.					méd.					dist.				
	1	2	3	4	5	1	2	3	4	5	2	3	4	5	1	2	3	4	5	
score	89	96	97	96	94	95	97	93	91	96	93	96	90	96	94	92	93	93	96	75
moy.	87	92	94	94	90	91	93	95	95	94	93	95	94	94	87	82	87	90	85	78

TAB. 3.8 – Scores de segmentation, os par os, pour l’image d’exemple

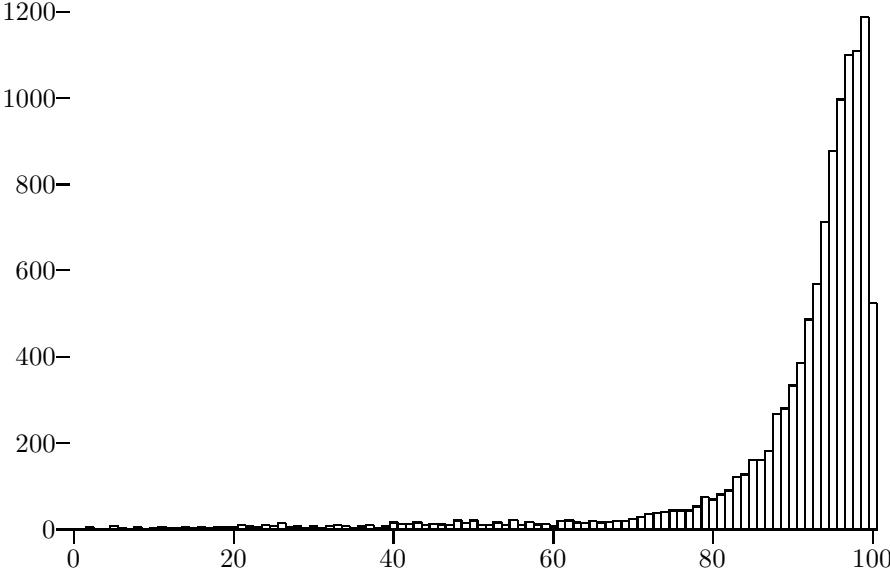


FIG. 3.1 – Histogramme des scores, tous os confondus, pour les images de “bonne” qualité

Pour l’ensemble de la base, les moyennes et les écarts-types des scores de segmentation sont indiqués dans le tableau 3.5, pour les trois qualités, soit respectivement pour les 577, 120 et 62 images du tableau 3.1, page 85.

Si l’on considère que la distribution des \tilde{e}_{ij} est centrée sur 0 en médiane (et en moyenne), la quantité $D = \frac{1}{N} \sum |\tilde{e}_{ij}|$ est la *déviation absolue moyenne* des longueurs (relatives) pour l’os ij . C’est une mesure de dispersion usuelle en analyse de données [Hub81] et son avantage par rapport à la variance classique $\sigma^2 = \frac{1}{N} \sum |\tilde{e}_{ij}|^2$ est d’être plus robuste vis-à-vis des déviations anormales. Dans le cas d’une distribution parfaitement gaussienne des \tilde{e}_{ij} , avec une variance σ_G^2 , on aurait $D = \sqrt{\frac{2}{\pi}} \cdot \sigma_G \approx 0.8 \sigma_G$. La moyenne des scores de segmentation, c’est-à-dire $100.(1 - D)$, nous permet donc d’estimer ce σ_G d’une distribution supposée normale.

D’autre part, la variance des $|\tilde{e}_{ij}|$ s’écrit $\frac{1}{N} \sum |\tilde{e}_{ij}|^2 - (\frac{1}{N} \sum |\tilde{e}_{ij}|)^2 = \sigma^2 - D^2$, en supposant toujours que $\sum \tilde{e}_{ij} = 0$ et avec σ qui note la variance effective des \tilde{e}_{ij} . Ainsi l’écart-type des scores de segmentation, c’est-à-dire $100.\sqrt{\sigma^2 - D^2}$, nous permet d’estimer l’importance des valeurs aberrantes.

Le tableau 3.6 fournit les mêmes résultats avec une agrégation sur certains groupes d’os caractéristiques. La dernière ligne de ce tableau donne des statistiques comparables pour le score global $Q = (\prod_{ij} q_{ij})^{\frac{1}{19}}$ qui est une mesure de la normalité de l’ensemble des mesures L_{ij} pour une image considérée. L’utilisation dans ce score de la valeur absolue $|e_{ij}|$ évite les compensations systématiques, notamment quand une erreur de localisation pour une articulation donne deux écarts e_{ij} proches en valeurs absolues et de signes contraires. Le choix des scores bruts et d’une moyenne géométrique permet d’accentuer l’influence des mauvais scores.

Pour une analyse plus fine, on peut utiliser des tables de quantiles ou des histogrammes. Le tableau 3.7 donne quelques quantiles pour les 577 “bonnes” images et la figure 3.1 montre un histogramme des scores de segmentation, pour ces mêmes images, tous os confondus. On peut vérifier que 70 à 75 % des scores atteignent ou dépassent le score moyen.

Pour l’image que nous avons utilisée comme exemple courant dans les chapitres précédents, les scores \tilde{q}_{ij} sont indiqués sur la première ligne (**score**) du tableau 3.8. La seconde ligne (**moy.**) reprend, pour comparaison, les moyennes correspondantes observées sur l’ensemble des “bonnes” images (voir tableau 3.5). La figure 3.2 donne une représentation des écarts \tilde{e}_{ij} correspondants.

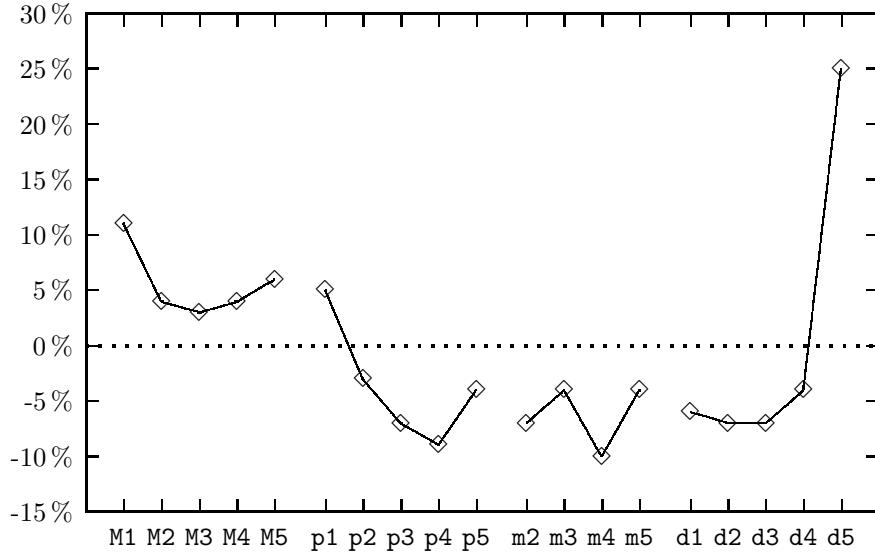


FIG. 3.2 – Visualisation des écarts relatifs pour l'image d'exemple

Ainsi, pour le premier métacarpien, le score obtenu sur cette image est 89 et le score moyen sur l'ensemble des “bonnes” images est 87. Ce score moyen assez faible pourrait traduire les difficultés observées en général sur le pouce. Pour notre exemple, un écart de 11 % pour le premier métacarpien (M1) n'est pourtant pas évidente, voir la figure B.3, page 180 (en annexe), et cet os semble mieux détecté que le dernier métacarpien (M5) qui obtient le score 94. En fait, ce premier métacarpien est effectivement nettement plus long que la normale. D'autre part, si on considère les phalanges proximale et médiane du quatrième axe (p4 et m4), on obtient deux scores comparables, respectivement 91 et 90. Mais sur la figure B.3, une erreur technique n'est manifeste que sur la phalange proximale, alors que sur la figure 3.2 le “profil” pour les proximales semble pourtant plus régulier que celui des médianes. Par contre, le score 75 sur la dernière phalange distale (d5) est significatif d'une erreur, on voit sur la figure B.3 que la segmentation a retenu le contour “peau”.

3.1.5 Premières conclusions

Pour interpréter de telles discordances entre une interprétation visuelle du résultat de la segmentation et les scores obtenus, il ne faut pas oublier que ces scores ne font que quantifier la normalité des mesures. Celle-ci traduit la normalité biométrique des os, vis-à-vis de “la” norme de référence, mais observée avec le biais technique de la segmentation. Les statistiques présentées ci-dessus ne permettent pas de faire clairement la part entre les erreurs de notre méthode et la variabilité naturelle. Il faut donc rester prudent dans l'analyse de ces premiers résultats. En attendant des expérimentations mieux ciblées, en 3.2 et 3.3, on peut avancer que :

- La longueur des os semble présenter une variabilité naturelle importante. Pour les scores supérieurs à 80, l'histogramme de la figure 3.1 correspond relativement bien avec celui d'une distribution gaussienne (repliée), avec un écart-type de l'ordre de 10, en se basant sur le score moyen ou l'inflexion de la courbe aux alentours du score 90. De tels écarts, entre une mesure et sa valeur normale estimée, sont conformes à ceux donnés par les observations morphométriques [Poz74] et ne sont pas révélateurs d'erreurs de segmentation.
- Pour les scores inférieurs à 80, les fréquences observées dépassent nettement les valeurs des résidus d'une gaussienne et il s'agit très probablement d'erreurs de segmentation.

- Dans environ 75 % des cas, les scores individuels satisfont $\tilde{q}_{ij} \geq 90$ (et 90 % pour $\tilde{q}_{ij} \geq 80$). Dans ce cas, on peut considérer que notre algorithme aboutit à une localisation probablement correcte de l'os et que celui-ci est mesuré avec une erreur technique (projection, discréétisation, localisation des contours, direction de l'axe, définition de la mesure, etc.) assez faible en regard de la variabilité naturelle. On pourra préciser cela en 3.3.

Une étude détaillée par tranche d'âge ne révèle pas de différence systématique dans ces résultats. Par contre, on constate facilement l'influence de la qualité de l'image sur la réussite de la segmentation, ce qui justifie les exigences techniques énoncées en 1.3. De toute évidence, des radiographies mal acquises (contraste, saturation, cadrage, etc.) donnent des résultats moins fiables.

Les meilleurs scores sont obtenus pour le troisième axe, les phalanges proximales et, plus particulièrement, le groupe d'os dits "centraux". Pour ces derniers, on a $\tilde{q}_{ij} \geq 90$ dans près de 88 % des cas et $\tilde{q}_{ij} \geq 80$ dans 97 % des cas. On a des résultats de moindre qualité par ailleurs et, sans exclure des différences de variabilité biométrique, on peut apporter un certain nombre d'explications techniques aux différences constatées entre certains os ou groupes d'os :

- Les problèmes pour les axes extérieurs peuvent s'expliquer par le fait que l'algorithme de construction des axes privilégie l'orientation verticale. Pourtant, on devrait alors observer des différences importantes entre le pouce et l'auriculaire, car le premier est généralement nettement plus écarté, voire replié. Une autre explication serait la dissymétrie des contours au niveau des métacarpiens de ces deux axes, avec la présence unilatérale de la peau, ce qui peut induire des erreurs de détection de l'axe. Enfin, il ne faut pas oublier simplement l'influence de la longueur d'un axe sur sa mise en évidence, son approximation polynomiale, son extrapolation, etc.

- Une légère perte de qualité pour les phalanges médianes peut se justifier par la perte de précision relative car les longueurs mesurées sont plus faibles.

- Les scores relativement moyens qui sont obtenus pour les métacarpiens, illustrent la difficulté de la segmentation entre ces os et ceux du carpe. On verra en 3.2 que les contraintes de normalité ont alors un rôle non négligeable et certainement préjudiciable. Une solution à ce problème passe par le traitement conjoint des os du carpe (voir chapitre 4).

- Enfin, les mauvais résultats sur les phalanges distales peuvent facilement s'expliquer par le manque de contraste qui perturbe déjà la construction des axes et qui, même lorsque l'axe est correct, crée beaucoup d'ambiguïté pour la segmentation. On notera que les extrémités distales sont souvent à la limite du perceptible et que la détection visuelle est très délicate, elle-aussi.

3.2 Influence des paramètres de normalité

Les rapports λ_{ij} et μ_j du tableau 2.1, page 76, définissent "la" norme biométrique. Ces rapports déterminent les scores de segmentation et ainsi la confiance que l'on peut accorder au résultat d'une segmentation et aux mesures qu'elle détermine. De manière encore plus critique, ces paramètres sont utilisés comme a priori pour guider le processus de segmentation des os (voir 2.6) et, malgré un mécanisme d'adaptation relativement efficace, des valeurs λ_{ij} et μ_j différentes peuvent conduire à des segmentations différentes.

Il est donc important de regarder l'influence de ces paramètres et nous avons procédé par comparaison des segmentations qui résultent de quelques paramétrages particuliers. En effet, faute de mesures de référence³, on ne peut pas caractériser les erreurs de segmentation et on ne peut qu'observer les différences qui sont induites par des variations des paramètres.

³On verra en 3.3 la possibilité d'une comparaison avec quelques mesures "manuelles", mais celles-ci ne sont pas en nombre statistiquement déterminant pour la présente étude.

jeux à comparer				nombre d'images
-5 %	0 %	5 %	10 %	568
	0 %	5 %	10 %	5
-5 %	0 %			4
		5 %	10 %	1
			10 %	1

TAB. 3.9 – Nombre d'images exploitables selon les jeux de paramètres à comparer

Idéalement, de telles variations artificielles ne devraient pas affecter la segmentation et les longueurs mesurées devraient être respectivement identiques. Ainsi, pour un même os d'une même image, une différence entre deux mesures obtenues en utilisant des paramètres différents correspond à au moins une erreur de segmentation et savoir laquelle nous importe peu, d'ailleurs ces deux mesures peuvent être erronées. Par contre, on ne détecte pas les cas de deux erreurs identiques. On suppose que ces derniers cas sont relativement peu nombreux et, du moins, qu'ils ne biaissent pas l'interprétation qualitative.

Notons que cette expérimentation permet aussi de simuler le comportement du programme sur un lot d'images “anormales” qui présenteraient toutes une même déviation morphologique, opposée à celle qui est imposée artificiellement à la norme par la variation des paramètres.

3.2.1 Variations des paramètres

Un nouveau jeu de paramètres, dit 5 %, a été construit en augmentant les valeurs des λ_{0j} (métacarpiens) de 0.02 (soit environ 5 % en valeur relative), en diminuant d'autant les valeurs des λ_{1j} , en augmentant de 0.01 les valeurs des λ_{2j} et en diminuant de 0.01 les valeurs des λ_{3j} (phalanges distales) et un cas particulier pour le pouce (+0.02, -0.01 et -0.01). Les longueurs relatives sur chaque axe vérifient ainsi toujours les identités de définition : $\sum_i \lambda_{ij} = 1, \forall j$.

Ceci ne permet pas d'explorer spécifiquement le rôle des paramètres μ_j , mais on s'intéresse ici à des variations individuelles des longueurs estimées des os et la modification des λ_{ij} est un moyen suffisant pour cela. Le fait de varier de manière identique sur tous les axes permet alors d'augmenter la consistance statistique de l'expérience. En 3.2.3, on reviendra sur le problème d'un biais éventuel qui serait induit par ces choix expérimentaux.

Le jeu de paramètres 5 % va donc favoriser, notamment, une surestimation de la longueur a priori des métacarpiens, une sous-estimation de celle des phalanges distales et un déplacement de la position a priori des deux articulations qui concernent ces os.

L'articulation médiane (de plus inexistante pour le pouce, comme considéré ici) n'est pas concernée car on a préféré cibler un effet maximal sur les deux autres articulations. On notera que les articulations médianes et distales sont de structures similaires et ne comportent qu'une seule épiphyshe lenticulaire. Les articulations distales peuvent être considérées comme les plus délicates en raison du contraste et des dimensions plus faibles. Les articulations proximales sont très nettement différentes, avec deux épiphyses dont une plutôt ronde et importante. Comme vu en 2.6, la détection des articulations utilise deux modèles (automates) différents et ils sont ainsi tous deux éprouvés ici. Le choix d'ignorer l'articulation médiane ne devrait donc pas enlever de généralité à l'étude expérimentale et aux conclusions qui suivent.

De manière similaire, par des variations inverses, nous avons construit un jeu dit -5 % et, avec des variations ± 0.05 et ± 0.02 , nous avons construit un autre jeu dit 10 %. La référence (valeurs du tableau 2.1, page 76) sera dite 0 %. Rappelons qu'une variation des paramètres λ_{ij} influe aussi sur l'estimation de la qualité de la segmentation.

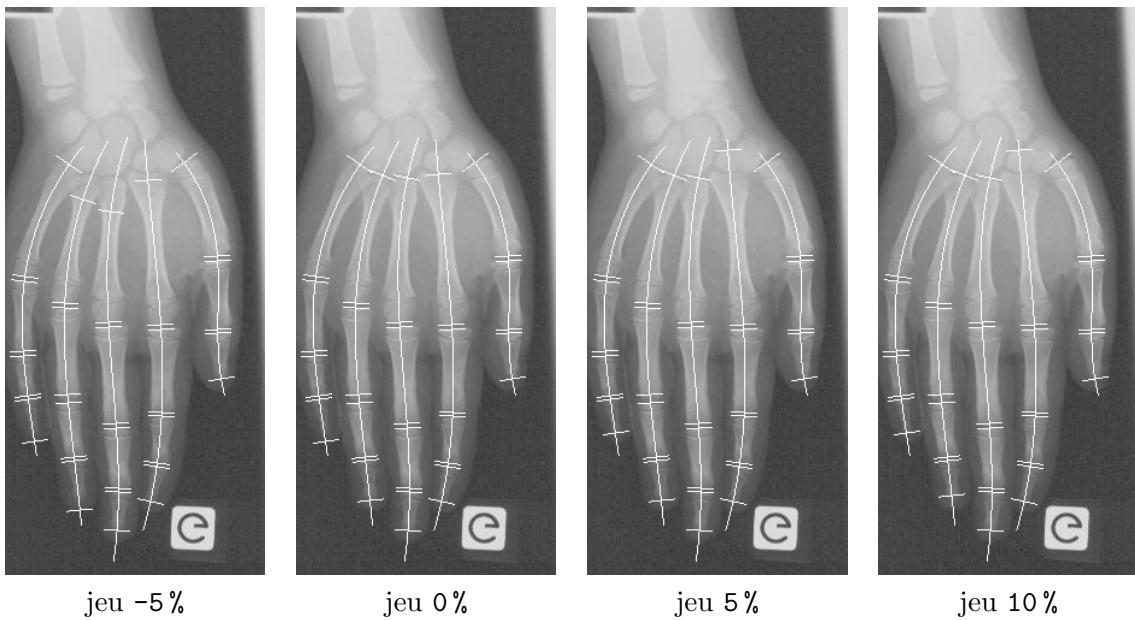


FIG. 3.3 – Exemple de variations dans la segmentation, en fonction des paramètres de normalité

	méta					prox.					méd.					dist.				
	1	2	3	4	5	1	2	3	4	5	2	3	4	5	1	2	3	4	5	
-5 %	83	120	92	87	100	57	69	80	72	58	39	50	45	33	38	29	32	41	35	
0 %	83	125	120	108	100	57	69	80	72	58	39	50	45	33	38	29	32	33	35	
5 %	83	146	120	108	100	57	69	80	72	58	39	50	45	33	38	29	32	33	28	
10 %	83	146	120	108	100	57	69	80	72	58	39	50	45	33	38	29	32	29	28	

TAB. 3.10 – Variation des mesures résultantes (en pixels), os par os, pour l'image d'exemple

Dans les cas extrêmes, cela peut se traduire par quelques différences dans la caractérisation d'une segmentation erronée, du fait de mesures manifestement trop courtes ou trop longues (voir 3.1.1 et colonne seg. du tableau 3.1). Le lot des images qui sont effectivement utilisables dans la présente expérimentation, ne correspond donc pas exactement aux 577 “bonnes” images considérées jusqu’ici. Comme indiqué dans le tableau 3.9, seulement 568 de ces images “résistent” à toutes les variations. Pour les autres, dans 4 cas, les deux jeux 5 % et 10 % conduisent à une segmentation non valide et, dans 5 autres cas, c'est le jeu -5 % qui pose un problème. Enfin, pour 2 images, la segmentation obtenue avec le jeu de référence est rejetée, alors que celles obtenues avec d'autres paramètres sont acceptables. Le dernier cas laisse notamment perplexe sur la stabilité du traitement pour cette image.

La figure 3.3 permet de visualiser les différentes segmentations obtenues sur notre image usuelle, selon le jeu de paramètres utilisé, et le tableau 3.10 permet de comparer numériquement les longueurs effectivement mesurées, celles affectées d'une erreur manifeste étant en gras.

Les remarques particulières que l'on peut faire sur cet exemple, sont représentatives des conclusions que l'on tirera de l'expérimentation globale :

- En ce qui concerne les métacarpiens, le paramétrage standard convient le mieux, le jeu -5 % introduit 3 erreurs alors que, pour les jeux 5 % et 10 %, on ne constate qu'une erreur sur le second métacarpien.
- La segmentation des phalanges proximales et médianes n'est pas affectée par les modifications des paramètres, notamment l'erreur sur la quatrième phalange proximale (un défaut de contour au voisinage de l'axe) est persistante.
- Enfin, on a une réponse très variable sur deux phalanges distales.

jeux de paramètres	-5 %	0 %	5 %	10 %
nombre moyen de mesures aberrantes par image	2.1	1.9	2.5	3.0
score Q moyen	93.0	94.1	93.3	91.9

TAB. 3.11 – Estimation de la qualité de la segmentation selon les jeux de paramètres

3.2.2 Analyse des résultats

Le tableau 3.11 indique, pour chaque jeu de paramètres et sur les 568 images comparables, les moyennes des mesures de qualité définies en 3.1. Les mêmes paramètres interviennent donc, a priori, pour guider la segmentation et, a posteriori, pour quantifier la normalité du résultat vis-à-vis de ces paramètres. Le jeu 0 % peut être effectivement considéré comme le meilleur des 4, mais la baisse de score observée pour les autres jeux est très inférieure à celle qui est attendue dans l'hypothèse où la réalisation de la segmentation donnerait des résultats identiques quelle que soit la paramétrisation. On devrait alors observer, respectivement, des baisses de l'ordre de 3 points de score pour ±5 % et de 6 points pour 10 %. Autrement dit, la segmentation tend à “suivre” la paramétrisation.

Pour préciser au cas par cas la sensibilité vis-à-vis du paramétrage, les tableaux 3.12 à 3.15 donnent la distribution du nombre de différences dans les mesures, entre deux jeux de paramètres, os par os et sur l'ensemble des images qui permettent des mesures communes à ces deux jeux (voir tableau 3.9).

Entre les jeux 0 % et 5 %, on dispose des mesures pour 573 images. Le tableau 3.12 indique que, tous os confondus, il y a 2270 différences de longueurs sur 10 887 paires de mesures (573×19), soit près de 21 %. Deux tiers de ces différences (1 503 sur 2 270) concernent les phalanges distales et le tiers restant concerne essentiellement les métacarpiens, notamment le second. Cela représente en moyenne près de 4 (exactement 2270/573) différences par image. Il y a très peu d'effet sur les phalanges proximales et médianes.

Les figures 3.4 et 3.5 montrent les histogrammes des différences relatives⁴, respectivement pour les métacarpiens et les phalanges distales. On constate que le jeu 5 % conduit effectivement à rechercher des métacarpiens a priori plus longs et des phalanges distales a priori plus courtes. Sur l'histogramme concernant les métacarpiens, le pic très net, vers l'abscisse -14, correspond au cas assez fréquent, illustré par la figure 3.3, où la segmentation du second métacarprien englobe le trapézoïde, ce qui donne effectivement lieu à une erreur relative de l'ordre de 14 %, en moyenne. Les résultats sont beaucoup plus confus pour ce qui concerne les phalanges distales.

Si l'on compare le tableau 3.12 avec le tableau 3.13, celui des différences entre les jeux 0 % et 10 %, on constate une augmentation modérée du nombre de différences, avec une distribution globale similaire. Entre les jeux 5 % et 10 % (voir tableau 3.14), il y a beaucoup moins de différences sur les métacarpiens, ce qui laisse penser que les effets néfastes sont le plus souvent acquis dès une variation de 5 % de la longueur a priori. Par contre, le nombre de différences sur les phalanges distales reste élevé, ce qui confirme le comportement plus aléatoire de leur segmentation.

Avec le jeu -5 % (voir tableau 3.15), il y a une augmentation très nette du nombre de différences sur les métacarpiens et une petite amélioration sur les phalanges distales. Ce jeu de paramètres conduit à sous-estimer la longueur a priori des métacarpiens, ce qui favorise le calage de la segmentation sur des parasites intérieurs à l'os. Inversement, la longueur a priori des phalanges distales est surestimée, mais l'effet positif n'est pas net. On a en effet des problèmes avec le contour “peau” et le manque de contraste rend toujours les choses très sensibles au bruit.

⁴Si les deux mesures à comparer sont notées respectivement m_0 et m_5 , on considère la différence relative qui s'exprime par $\frac{2(m_5 - m_0)}{m_0 + m_5}$. En abscisse des figures 3.4 et 3.5, cette valeur est multipliée par 100. Pour conserver une bonne lisibilité, la tranche 0 ne comptabilise pas les différences nulles (ie. les mesures identiques) mais seulement les très petites différences non nulles.

rang	axe					total	taux
	1	2	3	4	5		
méta	156	241	122	87	128	734	26 %
prox.	12	2	1	0	5	20	1 %
méd.		2	1	0	10	13	1 %
dist.	11	350	378	413	351	1503	52 %
total	179	595	502	500	494	2270	21 %
taux	10 %	26 %	22 %	22 %	22 %	21 %	

TAB. 3.12 – Nombre de mesures différentes, os par os, entre les jeux de paramètres 0 % et 5 %

rang	axe					total	taux
	1	2	3	4	5		
méta	197	286	128	123	167	901	31 %
prox.	18	2	1	2	9	32	1 %
méd.		4	1	2	16	23	1 %
dist.	15	439	483	438	385	1760	61 %
total	230	731	613	565	577	2716	25 %
taux	13 %	32 %	27 %	25 %	25 %	25 %	

TAB. 3.13 – Nombre de mesures différentes, os par os, entre les jeux de paramètres 0 % et 10 %

rang	axe					total	taux
	1	2	3	4	5		
méta	66	71	13	53	67	270	9 %
prox.	15	1	0	2	6	24	1 %
méd.		4	0	2	8	14	1 %
dist.	13	319	423	411	320	1486	52 %
total	94	395	436	468	401	1794	16 %
taux	5 %	17 %	19 %	20 %	17 %	16 %	

TAB. 3.14 – Nombre de mesures différentes, os par os, entre les jeux de paramètres 5 % et 10 %

rang	axe					total	taux
	1	2	3	4	5		
méta	414	292	489	429	388	2012	70 %
prox.	17	1	2	0	11	31	1 %
méd.		3	5	2	15	25	1 %
dist.	14	218	263	423	346	1264	44 %
total	445	514	759	854	760	3332	31 %
taux	26 %	22 %	33 %	37 %	33 %	31 %	

TAB. 3.15 – Nombre de mesures différentes, os par os, entre les jeux de paramètres 0 % et -5 %

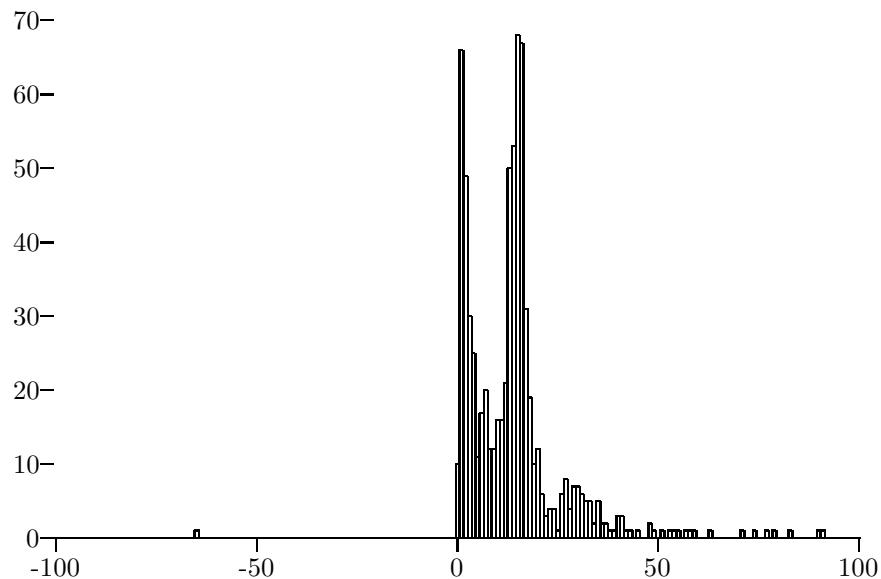


FIG. 3.4 – Histogramme des différences relatives, entre les jeux de paramètres 0 % et 5 %, sur les métacarpiens

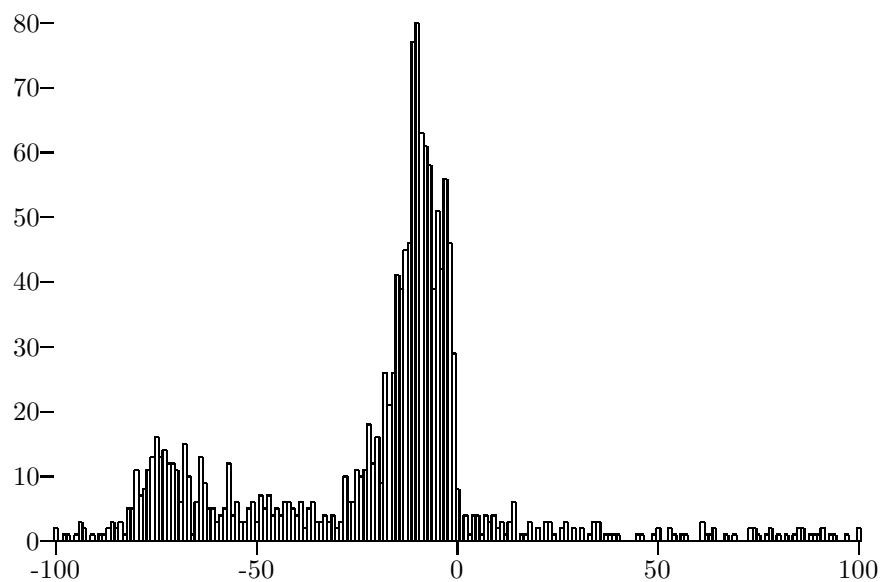


FIG. 3.5 – Histogramme des différences relatives, entre les jeux de paramètres 0 % et 5 %, sur les phalanges distales

	nombre de mesures différentes									
	0	1	2	3	4	5	6	7	plus	total
nb im.	5	32	68	125	138	99	72	26	8	573
% im.	0.87 %	5.6 %	11.9 %	21.8 %	24.1 %	17.3 %	12.6 %	4.5 %	1.40 %	
% th.	0.29 %	2.8 %	11.0 %	23.2 %	28.6 %	21.3 %	9.7 %	2.7 %	0.48 %	

TAB. 3.16 – Nombre de mesures différentes par image, entre les jeux de paramètres 0 % et 5 %

	nombre de mesures différentes									
	0	1	2	3	4	5	6	7	plus	total
nb im.	0	9	34	90	129	130	98	61	22	573
% im.	0 %	1.57 %	5.9 %	15.7 %	22.5 %	22.7 %	17.1 %	10.6 %	3.8 %	
% th.	0.04 %	0.59 %	3.9 %	13.2 %	25.4 %	28.4 %	18.9 %	7.5 %	2.0 %	

TAB. 3.17 – Nombre de mesures différentes par image, entre les jeux de paramètres 0 % et 10 %

	nombre de mesures différentes									
	0	1	2	3	4	5	6	7	plus	total
nb im.	11	54	111	177	153	46	14	6	2	574
% im.	1.92 %	9.4 %	19.3 %	30.8 %	26.7 %	8.0 %	2.4 %	1.05 %	0.35 %	
% th.	0.82 %	6.9 %	22.0 %	33.4 %	25.2 %	9.4 %	1.9 %	0.23 %	0.02 %	

TAB. 3.18 – Nombre de mesures différentes par image, entre les jeux de paramètres 5 % et 10 %

	nombre de mesures différentes									
	0	1	2	3	4	5	6	7	plus	total
nb im.	0	4	12	40	56	121	137	115	87	572
% im.	0 %	0.70 %	2.10 %	7.0 %	9.8 %	21.2 %	24.0 %	20.1 %	15.2 %	
% th.	0.00 %	0.10 %	0.86 %	4.1 %	12.2 %	22.9 %	27.6 %	20.8 %	11.4 %	

TAB. 3.19 – Nombre de mesures différentes par image, entre les jeux de paramètres 0 % et -5 %

3.2.3 Corrélations entre les erreurs de segmentation

On va maintenant envisager l’existence de corrélations entre les différences de segmentation qui affectent une même image. On s’intéresse alors seulement à l’occurrence de ces différences, sans considération de leur importance numérique.

Une corrélation d’origine technique est envisageable à plusieurs titres. En particulier, les heuristiques d’extrapolations des axes conduisent à lier leurs longueurs a priori, en ajustant celles des plus courts par rapport aux plus longs. De manière plus générale, la mauvaise construction ou segmentation d’un axe peut avoir des causes qui affectent aussi les autres axes. On devrait ainsi observer un biais global, provenant de la qualité de l’image (cadrage, contraste, etc.), ou plus localisé (direction particulière, présence d’un artefact, etc.). Enfin, on notera que la construction des jeux de paramètres, notamment le fait d’imposer des variations similaires sur tous les axes, devrait aussi induire des effets perceptibles.

Dans le tableau 3.16, la ligne **nb im.** indique le nombre d’images qui sont concernées par exactement 0, 1, 2, … différences dans les mesures de longueurs, entre les jeux de paramètres 0 % et 5 %. La ligne **% im.** indique les fréquences correspondantes, par exemple seulement 5 images sur 573, soit environ 0.87 %, ne donnent lieu à aucune différence.

C_{ab}	d 1	M 2	d 2	M 3	d 3	M 4	d 4	M 5	d 5
M 1 d 1	0.06	0.10	0.09	0.13	0.06	0.04	-0.04	0.00	-0.03
		0.06	0.01	-0.01	-0.01	0.08	0.00	0.08	0.01
M 2 d 2			0.08	0.06	0.05	-0.04	0.09	0.05	0.04
				-0.04	0.29	-0.06	0.06	0.01	0.04
M 3 d 3					0.08	0.22	-0.03	0.04	0.10
						-0.01	0.03	0.03	-0.02
M 4 d 4							-0.05	0.16	-0.03
								0.02	0.10
M 5									-0.01

TAB. 3.20 – Corrélation entre deux mesures pour les jeux de paramètres 0 % et 5 %

C_{ab}	d 1	M 2	d 2	M 3	d 3	M 4	d 4	M 5	d 5
M 1 d 1	0.18	0.12	0.04	0.11	0.04	0.08	0.03	0.05	-0.04
		0.08	-0.06	-0.01	-0.02	0.07	-0.06	0.11	0.07
M 2 d 2			0.04	0.13	0.01	0.05	0.07	0.04	0.04
				-0.04	0.15	-0.01	0.05	-0.04	-0.03
M 3 d 3					0.04	0.17	0.08	0.01	0.05
						-0.05	0.09	0.02	0.08
M 4 d 4							-0.01	0.10	-0.08
								-0.04	0.13
M 5									0.00

TAB. 3.21 – Corrélation entre deux mesures pour les jeux de paramètres 0 % et 10 %

La ligne % th. donne des fréquences théoriques qui sont calculées avec l'hypothèse que les différences résultent de processus indépendants d'un os à l'autre. Ainsi la probabilité d'observer exactement k différences sur une image, correspond au coefficient de x^k dans $\prod_{i=1}^{19} (p_i \cdot x + 1 - p_i)$, où chaque probabilité p_i , particulière à un os i , est estimée à partir des données du tableau 3.12. Par exemple, la probabilité d'avoir une différence sur le premier métacarpien est considérée comme égale à 156/573 , etc.

Les tableaux 3.17 à 3.19 font de même pour les autres paires de jeux de paramètres. On constate, de manière générale, une grande similitude entre les fréquences observées et les fréquences théoriques, avec une tendance à un certain étalement des premières. Ainsi, quels que soient les jeux de paramètres utilisés, on peut noter que les fréquences observées autour du nombre moyen de différences, sont de l'ordre de seulement 80 à 85 % des valeurs théoriques correspondantes. Inversement, sur les cas extrêmes, c'est-à-dire pour la co-occurrence de très peu ou de beaucoup de différences, les fréquences observées sont relativement plus fortes que celles calculées sous l'hypothèse d'indépendance. Ces cas extrêmes sont cependant plus délicats à quantifier statistiquement car les nombres d'occurrences comptabilisées sont très faibles.

Cette analyse tend à montrer que l'hypothèse d'indépendance entre plusieurs occurrences de différences sur une même image est relativement bien vérifiée, avec pourtant une corrélation légère mais perceptible. On aboutit à des constations similaires, du point de vue qualitatif et quantitatif, quand on se limite à un groupe d'os, pourvu qu'il donne lieu à suffisamment d'occurrences pour valider des statistiques. C'est le cas notamment si l'on considère seulement les métacarpiens ou seulement les phalanges distales et l'on n'observe pas d'effet particulier à l'un ou l'autre de ces groupes d'os.

C_{ab}	d 1	M 2	d 2	M 3	d 3	M 4	d 4	M 5	d 5
M 1 d 1	0.31	0.06	-0.03	0.09	0.00	0.04	-0.09	0.06	-0.06
		0.01	0.04	0.06	0.01	-0.01	0.04	0.09	0.09
M 2 d 2			0.00	-0.02	0.03	0.08	-0.05	0.01	-0.07
				-0.03	0.17	0.03	0.07	-0.01	0.11
M 3 d 3					-0.07	-0.05	-0.03	-0.02	-0.08
						0.00	0.03	0.00	0.11
M 4 d 4							-0.08	0.00	-0.08
								0.02	0.22
M 5									-0.01

TAB. 3.22 – Corrélation entre deux mesures pour les jeux de paramètres 5 % et 10 %

C_{ab}	d 1	M 2	d 2	M 3	d 3	M 4	d 4	M 5	d 5
M 1 d 1	-0.03	0.10	-0.09	0.08	-0.02	0.09	0.11	0.09	0.04
		0.00	0.02	-0.03	-0.03	0.04	-0.03	0.04	0.01
M 2 d 2			-0.02	-0.07	0.02	0.06	0.11	0.13	0.06
				-0.07	0.11	-0.02	0.03	-0.01	0.11
M 3 d 3					0.01	0.22	0.06	0.08	0.02
						0.06	0.02	0.02	0.00
M 4 d 4							0.05	0.23	0.01
								0.11	0.15
M 5									0.04

TAB. 3.23 – Corrélation entre deux mesures pour les jeux de paramètres 0 % et -5 %

Pour une étude plus fine, on va maintenant étudier la corrélation entre les occurrences des différences concernant seulement deux os, considérés en particulier. Si l'on note ces os **a** et **b**, on dispose de p_a et p_b , les fréquences respectives d'une différence sur ces os, fréquences données par les tableaux 3.12 et suivants, comme déjà vu. On notera $p_{\bar{a}}$ et $p_{\bar{b}}$ les fréquences complémentaires. De même, on peut obtenir par comptage p_{ab} qui correspond à la fréquence du cas des deux différences conjointes sur une même image, ainsi que $p_{\bar{a}\bar{b}}$ qui correspond au cas “ni sur **a** ni sur **b**”, et $p_{a\bar{b}}$ et $p_{\bar{a}b}$ qui correspondent aux cas avec une et une seule différence sur les deux os.

Pour deux os **a** et **b**, on considère alors le coefficient de corrélation normalisée, aussi dit coefficient r de Pearson. Dans notre cas, par définition des fréquences, ce coefficient r s'écrit : $C_{ab} = \frac{p_{ab} - p_a p_b}{\sqrt{(p_{aa} - p_a^2)(p_{bb} - p_b^2)}}$, en remarquant que $p_{aa} = p_a$ et $p_{bb} = p_b$. Pour mieux mettre en évidence ses symétries, on peut facilement le réécrire sous la forme : $C_{ab} = \frac{p_{ab} p_{\bar{a}\bar{b}} - p_{\bar{a}b} p_{a\bar{b}}}{\sqrt{p_a p_{\bar{a}} p_b p_{\bar{b}}}}$.

Les tableaux 3.20 à 3.23 donnent les valeurs calculées de C_{ab} , pour toutes les paires (**a**, **b**) entre métacarpiens ou phalanges distales, pour nos 4 combinaisons de jeux de paramètres. Les noms des os sont abrégés par **M** pour un métacarprien et **d** pour une phalange distale, suivi du numéro d'axe.

Dans le cas théorique d'évènements **a** et **b** indépendants, le coefficient C_{ab} vaut 0. Des valeurs strictement positives traduisent une corrélation au sens usuel (un couplage **a** ↔ **b**, **ā** ↔ **ā̄**) et des valeurs strictement négatives traduisent une anti-corrélation (un couplage **a** ↔ **ā̄**, **ā** ↔ **b**). Les valeurs extrêmes sont respectivement 1 et -1. Des valeurs absolues inférieures à 0.3, voire 0.5, sont généralement considérées comme traduisant de “faibles” corrélations, le critère subjectif étant basé sur $|C_{ab}|^2$ plutôt que sur $|C_{ab}|$.

La transformée de Fisher $z' = \frac{\ln(1+r) - \ln(1-r)}{2}$ permet de passer du coefficient r (ie. C_{ab}) à la variable z' qui suit une loi normale avec une variance $\sigma_{z'}^2 = \frac{1}{N-3}$. Dans notre cas $N \approx 573$ et $\sigma_{z'}$ est donc de l'ordre de 0.04. Pour les valeurs de r qui nous intéressent, la transformée de Fisher est très proche de l'identité et on peut assimiler un intervalle de confiance autour d'une valeur r à l'intervalle de confiance de même largeur autour de la valeur z' correspondante. Cela revient à considérer que l'on a un pseudo-écart-type de 0.04 sur C_{ab} , notre estimation du coefficient de corrélation entre les deux os **a** et **b**.

Notons que, quand une phalange proximale ou médiane est impliquée, les comptages sont souvent nuls ou ne dépassent pas quelques unités, ce qui donne des estimations de la corrélation non définies ou peu significatives. Par exemple, si $p_a = p_b = \frac{1}{N}$, le facteur C_{ab} ne peut valoir que $\frac{-1}{N-1}$, si $p_{ab} = 0$, ou alors 1, si $p_{ab} = \frac{1}{N}$. En fait, un problème sur un de ces os est souvent lié à un problème général à tout son axe.

À la lecture des tableaux 3.20 à 3.23, quelques constatations générales et aussi quelques particularités se dégagent :

- Dans la plupart des cas, les valeurs absolues $|C_{ab}|$ sont faibles et comparables à l'écart-type, donc statistiquement peu différentes de 0.
- Globalement, les faibles⁵ valeur positives l'emportent pourtant sur les faibles valeurs négatives, avec une moyenne légèrement positive, inférieure à l'écart-type. Cela peut correspondre à un petit biais induit par la qualité générale de l'image, par l'algorithme ou encore par des variations anatomiques. Chacune de ces influences apparaît ainsi comme très modérée.
- En particulier, il n'y a pas de corrélation forte ($|C_{ab}| > 0.5$) entre métacarpiens et phalanges distales, y compris sur le même axe. Cela confirme le fait que la segmentation des phalanges proximales et médianes est très robuste et qu'elle détermine ensuite la segmentation aux deux extrémités comme des processus nettement indépendants, fonctions de leurs paramètres respectifs et des particularités locales de l'image.
- On peut cependant remarquer que les cas de corrélation négative s'observent le plus souvent entre un métacarprien et une phalange distale. Cela peut s'expliquer par la construction des jeux de paramètres, avec l'hypothèse d'un comportement assymétrique vis-à-vis d'une sous-estimation ou surestimation. On confirmerait ainsi que, selon le sens de la variation, les différences se produisent, soit un peu plus sur les métacarpiens, soit un peu plus sur les distales, et un plus rarement sur les deux en même temps, tout ceci par rapport aux résultats attendus pour des processus totalement indépendants.

Les plus fortes valeurs de corrélation sont positives. On peut les ranger en deux catégories :

- Il y a le cas du métacarpien et de la phalange distale du pouce avec le jeu de paramètres 10 % (voir tableaux 3.21 et 3.22). On notera cependant qu'il y a peu de cas de différences de segmentation sur la phalange distale du pouce (voir tableaux 3.13 et 3.14, page 97) et on peut vérifier qu'elles sont presque toutes liées à une différence sur le métacarpien et aussi sur la phalange centrale. Il s'agit alors, ni plus ni moins, que d'un gros problème sur tout le pouce pour une dizaine ou une quinzaine d'images.
- Il y a le cas de deux métacarpiens ou de deux distales, le plus souvent sur des axes voisins. Cette localisation tend à exclure des problèmes liés à la qualité générale de l'image, à notre algorithme ou encore au choix des jeux de paramètres. Les effets seraient alors répartis plus uniformément sur toutes les paires. Des explications plus plausibles peuvent venir de la structure du carpe, de la présence d'artefacts au bas de l'image ou de variations anatomiques particulières.

⁵Disons, pour être précis, que faible signifie ici de valeur absolue inférieure à 0.1, mais ce seuil importe peu.

3.2.4 Nouvelles conclusions

Il apparaît donc que la sensibilité de l'algorithme de segmentation vis-à-vis des longueurs a priori est très variable selon les os. On va admettre que les observations faites ci-dessus à propos des différences de mesures valent aussi pour les erreurs qui semblent ainsi très localisées. En complément des premières conclusions (voir 3.1.5), on peut maintenant préciser que :

- La segmentation des phalanges proximales et médianes est très stable, surtout pour ce qui concerne les 3 axes centraux. On voit ainsi que la localisation est très robuste, alors que les variations expérimentées prédisposaient notamment à placer des coupures entre une épiphyshe et la diaphyshe d'un même os. Cela montre l'efficacité de l'utilisation d'un modèle structurel pour réaliser la détection d'une articulation par le balayage d'un intervalle. Cette efficacité permet de définir un intervalle assez large qui résiste effectivement à des différences relativement importantes entre les positions réelles et celles estimées à partir des longueurs a priori.
- Pour les autres os, métacarpiens et phalanges distales, on voit que les longueurs a priori déterminent fortement la segmentation résultante. On peut penser que les longueurs normales (jeu 0 %) donnent en moyenne les meilleurs résultats, du moins pour les métacarpiens, mais la tolérance sur la longueur a priori s'avère malheureusement inférieure à ce que semble être la variabilité naturelle. Ainsi, une déviation de 5 % par rapport à la normale suffit à donner un fort taux d'erreurs. Notons aussi qu'une surestimation semble généralement préférable à une sous-estimation mais le compromis est fragile, compte tenu de la variabilité naturelle et d'un certain nombre de particularités anatomiques. Nous avons intégré cela au mieux dans les heuristiques d'extrapolation des axes et dans la définition des intervalles de recherche pour la segmentation des extrémités. Les résultats présentés ci-dessus montrent les limites de ces corrections a priori.
- En ce qui concerne les métacarpiens, et surtout le second, il semble impossible de réaliser un compromis a priori entre le risque de valider un contour parasite interne et celui de déborder dans le carpe. Une segmentation satisfaisante des métacarpiens demande de traiter aussi le carpe. L'approche développée dans le chapitre 4 devra permettre justement de localiser les os du carpe et, ainsi, d'encadrer les articulations des métacarpiens sur ces os. On pourra alors envisager des résultats comparables à ceux que nous avons obtenus pour les phalanges proximales et médianes.
- En ce qui concerne les phalanges distales, en raison du manque de contraste, il s'avère que les bruits, le contour distal "os" et le contour "peau" du bout du doigts ont des valeurs très proches et que leur ordre d'intensité est imprévisible, ce qui rend la détection aléatoire. On notera que ces extrémités distales sont également très difficilement lisibles "à l'œil". À cela s'ajoute les problèmes de localisation de l'axe lorsqu'il doit être extrapolé suite à un défaut de détection. Par contre, nous pouvons localiser correctement la dernière articulation qui est utilisée dans certains examens classiques, comme TW2 [TWC⁺83].

3.3 Validation des mesures sur un échantillon

Une validation effective de quelques mesures a été réalisée avec le concours des radiologues [HSN⁺97]. Le but ici est de pouvoir faire abstraction de la variabilité biométrique en vérifiant si les longueurs calculées sur les images sont comparables aux mesures obtenues de manière classique, avec une règle graduée, sur le cliché radiographique. Le cas échéant, les mesures sont réalisées épiphyses comprises, ce qui correspond au résultat attendu de notre segmentation. Le côté fastidieux des mesures "manuelles" explique que cette vérification soit limitée à quelques images et à quelques mesures par image seulement.

Nous ne considérons donc que 6 os, les métacarpiens et les premières phalanges des trois doigts centraux, c'est d'ailleurs une pratique admise par les radiologues pour des diagnostics "rapides". En ce qui nous concerne, il s'agit des os qui obtiennent généralement les meilleures



FIG. 3.6 – Exemple d’extrémités concaves sur une phalange proximale

scores de segmentation (voir tableaux 3.6 et 3.7, page 90). Enfin, ces os sont parmi les plus longs, ce qui prédispose aux erreurs relatives les plus faibles.

Les mesures “manuelles” sont arrondies au mm le plus proche, soit une erreur relative de 1 à 5 % sur les longueurs qui nous intéressent. On notera que $\frac{1}{2}$ mm correspond à environ 3 pixels pour la résolution de 150 dpi. D’autre part, l’erreur de localisation des contours, comme on l’a vu en 2.5, peut perturber la mesure sur l’image de l’ordre de un ou deux pixels, dans certains cas. Enfin, dans le cas d’extrémités concaves, comme sur la figure 3.6, notre mesure le long de l’axe diffère de la mesure “manuelle” qui, selon la pratique traditionnelle, serait plutôt une mesure “hors-tout” (avec certaines exceptions [Poz74]).

Le test a porté sur un échantillon de 28 images, représentatives de presque toutes les tranches d’âge⁶. Les images ont été choisies parmi celles dites de “bonne” qualité, à l’exception de deux qui sont de qualité “intermédiaire”, identifiées 170.1 et 749.1. Un certain nombre de ces images présentent des anomalies morphométriques qui sont significatives d’un point de vue médical. Le choix, relativement aléatoire, en fait aussi un échantillon assez représentatif de la variabilité technique sur l’ensemble des images à traiter. On a ainsi des cadrages et des contrastes très variés, deux résolutions et une grande diversité d’éléments périphériques, parfois très complexes.

Les résultats obtenus sur ces images sont illustrés aux figures B.4, page 181 en annexe, et suivantes. Il apparaît que les limites des os sont, dans l’ensemble, plutôt bien détectées. Les défauts visibles s’expliquent le plus souvent par un mauvais placement des axes, notamment aux extrémités distales. Cela peut provenir d’une erreur d’appariement ou d’un manque de cohérence entre plusieurs symétries, comme “os” et “peau”, plus particulièrement pour le pouce du fait de sa pose latérale. Une explication concomitante est naturellement le manque de contraste systématique au niveau de la dernière phalange. Si l’approximation polynomiale de l’axe n’est pas suffisamment déterminée, les portions extrapolées peuvent alors dévier notablement.

Le tableau 3.24 donne, pour les 6 os centraux, les scores a priori \tilde{q}_{ij} définis en 3.1.3 et 3.1.4. Pour plus de clarté, les noms des os sont notés ici par M pour un métacarpien et p pour une phalange proximale, suivi du numéro d’axe. La dernière colonne indique la moyenne des 6 scores. Les deux dernières lignes permettent de comparer, respectivement, les moyennes sur ces 28 images et celles sur l’ensemble des “bonnes” images, ces dernières valeurs étant reprises des tableaux 3.5, page 89, et 3.6, page 90. Ce petit échantillon est donc, globalement et selon les scores a priori, de qualité légèrement supérieure à la moyenne mais il présente pourtant un certain nombre de scores individuels nettement inférieurs à leur moyenne respective.

Pour évaluer numériquement la justesse des mesures de longueur, nous avons procédé à des calculs de scores, similaires à ceux introduits en 3.1.4, mais obtenus en remplaçant la longueur estimée a priori \mathcal{L}_{ij} par la longueur effectivement mesurée \mathbf{L}_{ij} . La résolution de chaque image est connue et permet la conversion des mm en pixels. La mesure obtenue par notre programme étant toujours notée L_{ij} , on définit ainsi une différence relative $d_{ij} = \frac{L_{ij} - \mathbf{L}_{ij}}{\mathbf{L}_{ij}}$, puis un score effectif $r_{ij} = 100.(1 - |d_{ij}|)$. Ces scores figurent dans le tableau 3.25. On vérifie ainsi que les os sont plutôt bien segmentés, avec des scores effectifs généralement meilleurs que les scores a priori.

⁶Nous voulions une image par tranche d’âge, selon Greulich et Pyle, soit 15 tranches pour chaque sexe, mais deux tranches ne sont pas représentées par manque d’image.

image	âge	score <i>Q</i>	M2	M3	M4	p2	p3	p4	moy.
142.1	0.50	86	93	92	91	85	97	93	91.8
144.1	12.00	95	96	97	95	100	90	92	95.0
155.1	13.50	95	100	100	100	97	93	93	97.2
156.1	14.00	95	95	95	100	99	91	94	95.7
161.1	9.00	92	97	97	98	96	98	97	97.2
164.1	18.00	93	96	93	95	99	93	95	95.2
170.1	13.50	88	98	86	90	99	96	95	94.0
173.1	13.00	92	99	92	97	100	95	95	96.3
188.1	12.50	94	95	97	99	98	95	95	96.5
203.1	0.25	65	94	93	89	87	97	93	92.2
221.1	14.00	94	97	99	100	99	91	93	96.5
227.1	11.50	93	84	98	100	99	91	96	94.7
230.1	5.75	82	98	96	97	99	96	100	97.7
232.1	4.17	92	97	94	96	96	96	97	96.0
258.1	5.00	95	97	96	100	98	94	97	97.0
292.1	7.00	95	89	99	97	99	92	96	95.3
424.1	3.00	94	98	97	96	99	100	100	98.3
434.1	2.67	92	100	95	94	90	98	95	95.3
533.1	16.00	95	96	99	92	100	92	96	95.8
666.2	15.00	95	97	98	96	100	92	95	96.3
671.1	5.00	95	97	96	97	95	100	100	97.5
727.1	17.00	92	94	97	94	97	97	99	96.3
731.1	16.00	95	86	100	99	99	93	96	95.5
748.1	1.00	93	100	94	98	89	98	90	94.8
749.1	6.00	93	96	96	97	98	90	90	94.5
780.1	1.50	93	91	98	97	87	96	93	93.7
810.1	1.00	89	96	92	96	88	99	98	94.8
819.1	0.25	87	94	94	91	93	95	99	94.3
moy.		91	95	95	96	95	94	95	95.6
bonnes		88.2	91.5	94.2	94.1	93.1	95.1	95.4	93.9

TAB. 3.24 – Scores a priori sur un échantillon de 28 images

Au total, sur 168 mesures, seulement une quinzaine, les scores en gras dans le tableau 3.25, dépassent les 5 % d'écart. On peut faire les remarques suivantes :

- Pour l'image 156.1, la mesure de M4 est fausse en raison d'une erreur sur l'extrémité proximale. Ce problème est assez fréquent et il s'explique par la présence de maxima de gradient parasites, correctement structurés et relativement plus "forts" que le contour recherché. On a des cas similaires avec les images 173.1 en M3, 533.1 en M4, 666.2 en M3 et 727.1 en M4. Ces cas ne sont pas systématiquement détectables par le score a priori.
- Pour les images 227.1, 292.1 et 731.1, on a une erreur nettement plus importante sur l'extrémité proximale de M2. Ce cas est celui qui est détaillé en haut de la figure 2.30, page 77. L'intervalle d'analyse est trop large et il permet d'accepter le contour du trapézoïde.
- Pour l'image 170.1, il y a un problème en p4 du fait d'une mauvaise interprétation de la brillance corticale à son extrémité distale. Ce problème affecte aussi l'image 727.1 en p2.
- Pour les images 748.1 et 780.1, la segmentation est relativement correcte, mais les épiphyses de l'articulation entre métacarpien et phalange sont à la limite du discernable, ce qui peut affecter différemment les mesures "manuelles" et automatiques. Notons aussi que les coefficients

image	âge	dpi	M2	M3	M4	p2	p3	p4	moy.
142.1	0.50	150	99	98	98	97	96	99	98.1
144.1	12.00	150	99	99	99	95	97	97	97.8
155.1	13.50	75	99	99	100	100	98	98	99.2
156.1	14.00	150	100	100	94	95	96	95	96.8
161.1	9.00	150	99	99	98	100	100	99	99.1
164.1	18.00	150	98	99	99	97	99	98	98.4
170.1	13.50	150	95	99	100	98	98	94	97.3
173.1	13.00	150	99	94	100	99	98	98	98.1
188.1	12.50	150	100	100	99	99	99	98	99.1
203.1	0.25	150	98	97	99	98	97	99	98.3
221.1	14.00	150	100	100	95	95	96	97	97.0
227.1	11.50	150	84	100	99	99	99	100	96.8
230.1	5.75	150	100	98	99	98	100	99	99.1
232.1	4.17	150	100	99	100	99	100	100	99.6
258.1	5.00	150	99	99	98	98	98	100	98.7
292.1	7.00	150	87	99	100	99	99	98	97.0
424.1	3.00	150	100	98	100	99	98	100	99.1
434.1	2.67	150	99	99	98	99	98	99	98.7
533.1	16.00	150	99	100	94	95	95	98	97.0
666.2	15.00	150	98	94	97	95	98	98	96.7
671.1	5.00	150	100	100	99	100	99	98	99.4
727.1	17.00	75	99	99	94	90	99	99	96.9
731.1	16.00	75	87	99	98	99	100	99	97.1
748.1	1.00	150	97	99	98	87	95	89	94.2
749.1	6.00	150	97	97	97	99	98	97	97.6
780.1	1.50	150	99	93	92	97	98	97	96.1
810.1	1.00	150	100	100	99	99	98	97	99.0
819.1	0.25	150	96	98	98	99	95	100	97.8
moy.			97.4	98.5	98.0	97.3	98.0	97.9	97.9

TAB. 3.25 – Scores effectifs sur l'échantillon

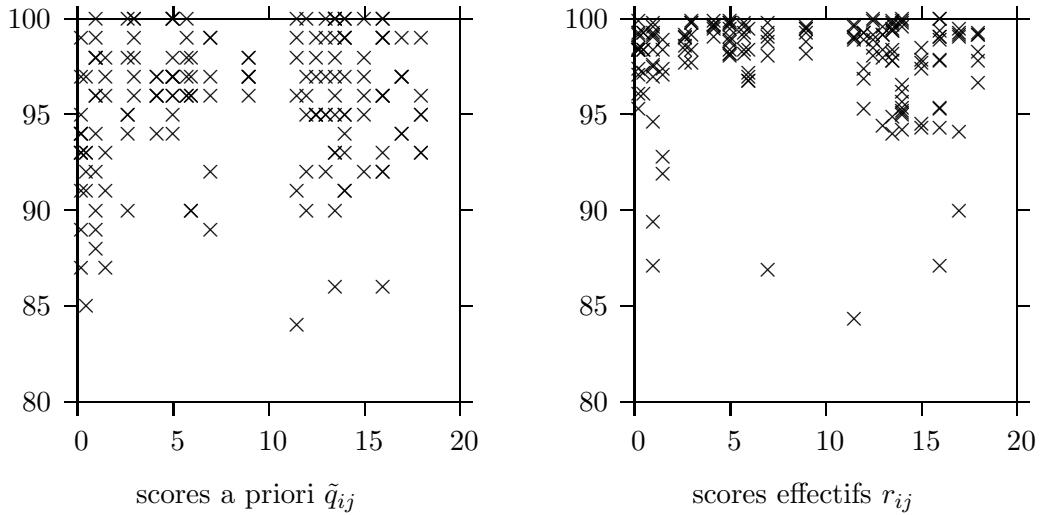
de normalité correspondent à l'apparition simultanée de toutes les épiphyses, mais il peut y avoir des particularités. Ainsi, pour l'image 748.1, M2 serait le seul métacarpien avec une épiphyse.

- À l'exception d'une erreur perceptible pour l'image 170.1 en M2, les scores effectifs de 95 ou plus ne peuvent pas être considérés comme des erreurs de segmentation.

Ces résultats sont obtenus en utilisant les paramètres du jeu standard 0 % pour diriger la segmentation. Avec le jeu 5 %, on ne corrige aucun des cas de sous-estimation des longueurs de métacarpiens et deux de ces cas sont maintenant des surestimations. On introduit par ailleurs 7 autres cas de surestimation, dont 5 qui concernent M2 et le trapézoïde. Les autres effets notables portent sur les phalanges distales, avec la correction de 5 surestimations et l'introduction de 2 sous-estimations.

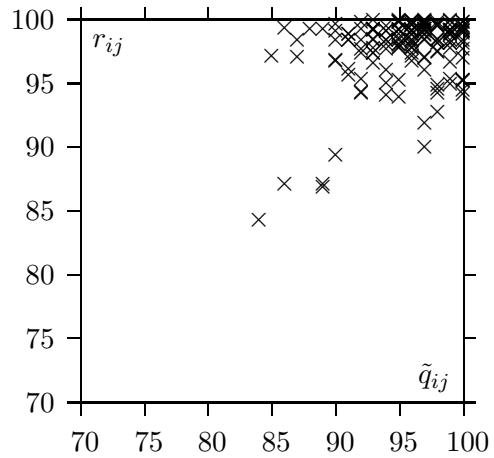
En passant du jeu 5 % au jeu 10 %, les seuls effets notables portent sur les phalanges distales, avec la correction de 5 surestimations et l'introduction de 32 sous-estimations.

En utilisant le jeu -5 %, on corrige 3 des 5 cas de surestimation des longueurs de métacarpiens, les deux autres sont maintenant sous-estimées, mais on introduit 26 nouveaux cas de sous-estimation (sur un total de 3×28 mesures de métacarpiens). Les autres effets notables sont 6 sous-estimations sur les métacarpiens des axes 1 ou 5 et, pour les phalanges distales, 5 corrections et 8 surestimations.



L'axe des abscisses correspond à l'âge osseux et l'axe des ordonnées correspond au score \tilde{q}_{ij} ou r_{ij} calculé.
Chaque marque \times correspond à la mesure d'un os d'une image de l'échantillon. Il s'agit d'un report des valeurs données dans les tables 3.24 et 3.25. Les scores effectifs ne sont pas arrondis à l'entier afin de mieux restituer la densité du graphique.

FIG. 3.7 – Scores obtenus selon l'âge



Il s'agit d'un croisement entre les valeurs données dans les tables 3.24 et 3.25. L'axe des abscisses correspond aux scores a priori \tilde{q}_{ij} et l'axe des ordonnées correspondent aux scores effectifs r_{ij} .

Chaque marque \times correspond à la mesure d'un os d'une image de l'échantillon.

FIG. 3.8 – Relation entre scores effectifs et scores a priori

Ces observations vont dans le sens des conclusions plus générales faites en 3.2.4 sur l'influence du paramétrage. Notamment, la moyenne des moyennes des scores effectifs sur les 6 os, c'est-à-dire la valeur dans la case en bas à droite du tableau 3.25, passe de 97.9 à respectivement 97.2, 97.2 et 96.2 pour les jeux 5 %, 10 % et -5 %.

La figure 3.7 montre qu'il n'y a pas de relation particulière entre l'âge osseux et les scores, a priori ou effectifs. Cela demandait confirmation car les problèmes sont très différents selon les âges : le manque de contraste en bas âge et l'ambiguïté des structures ensuite. On voit aussi que les différences entre nos mesures et celles "manuelles" sont généralement inférieures à 5 %. Ces nombreuses petites différences s'expliquent le plus souvent par diverses incertitudes techniques de part et d'autre, comme rappelé ci-dessus. Des différences supérieures à 5 % sont des erreurs flagrantes de notre technique de segmentation, sauf des cas discutables comme les images 748.1 et 780.1 pour lesquelles la présence des épiphyses est très ambiguë.

Les scores a priori sont relativement plus mauvais, ce qu'on explique essentiellement par la déviation des longueurs anatomiques réelles par rapport aux moyennes de normalité. La figure 3.8 montre la relation entre les scores a priori, en abscisse, et les scores effectifs, en ordonnée. On voit qu'un score $\tilde{q}_{ij} \geq 90$ correspond généralement à un score $r_{ij} \geq 95$, c'est-à-dire une segmentation correcte. Mais dans quelques cas, un score $\tilde{q}_{ij} \geq 90$, a priori acceptable, est associé à un mauvais score r_{ij} , c'est-à-dire à une erreur de segmentation. Inversement, un score $\tilde{q}_{ij} < 90$ peut être associé à une segmentation relativement correcte. Ceci limite la pertinence des scores a priori comme critères de fiabilité.

3.4 Applications

Nous allons évaluer maintenant l'utilisation des résultats disponibles dans deux pratiques diagnostiques des plus courantes, l'âge osseux et le profil métacarpo-phalangien. Nous verrons ensuite comment nos résultats permettent d'envisager diverses variantes de recalage qui seraient la base de traitements localisés, avec une application particulière qui consiste en une étude statistique de la forme des axes. Cette dernière confirme la grande variabilité anatomique et la difficulté à utiliser des a priori pour diriger les premières phases de l'analyse des images.

3.4.1 Relation à l'âge osseux

Nous allons regarder si la segmentation obtenue peut déterminer directement l'âge osseux. L'ossification de l'articulation d'un doigt est reliée au rapport $\frac{L}{L+a}$ entre, d'une part, la longueur L d'un os contigu et, d'autre part, la somme de cette longueur et de l'espace articulaire que l'on notera a . Quand l'âge augmente, $\frac{a}{L}$ décroît et le rapport approche de 1.

Les mesures L et $L+a$ résultent directement de la segmentation. Si la détection est correcte, l'erreur de localisation est alors faible en valeur relative, au maximum quelques pourcents pour les os les plus long. Ainsi, pour l'articulation entre une phalange proximale et son métacarpien, en reprenant les notations de la figure 2.29, page 76, on peut considérer le rapport $\frac{L_{1j}}{L_{1j}+a_{1j}}$, pour les 3 axes centraux, $j \in \{2, 3, 4\}$. On utilise ici la longueur de la première phalange dont la segmentation est plus fiable que celle du métacarpien.

Nous avons procédé à ce calcul pour 509 "bonnes" images pour lesquelles l'âge osseux est connu et la segmentation est réalisable (au sens vu en 3.1.1), quelle que soit la qualité a priori de cette segmentation. Il apparaît que, pour une même main, les trois rapports pour $j \in \{2, 3, 4\}$ ont des valeurs similaires. La figure 3.9 donne ces valeurs en fonction de l'âge osseux. Dans cette figure, chaque marque \times correspond à la mesure de l'ossification d'une articulation entre métacarpien et phalange proximale pour l'un des trois axes centraux d'une image parmi les 509. Dans leur ensemble, elles semblent donc bien suivre une loi empirique qui les relie à l'âge osseux,

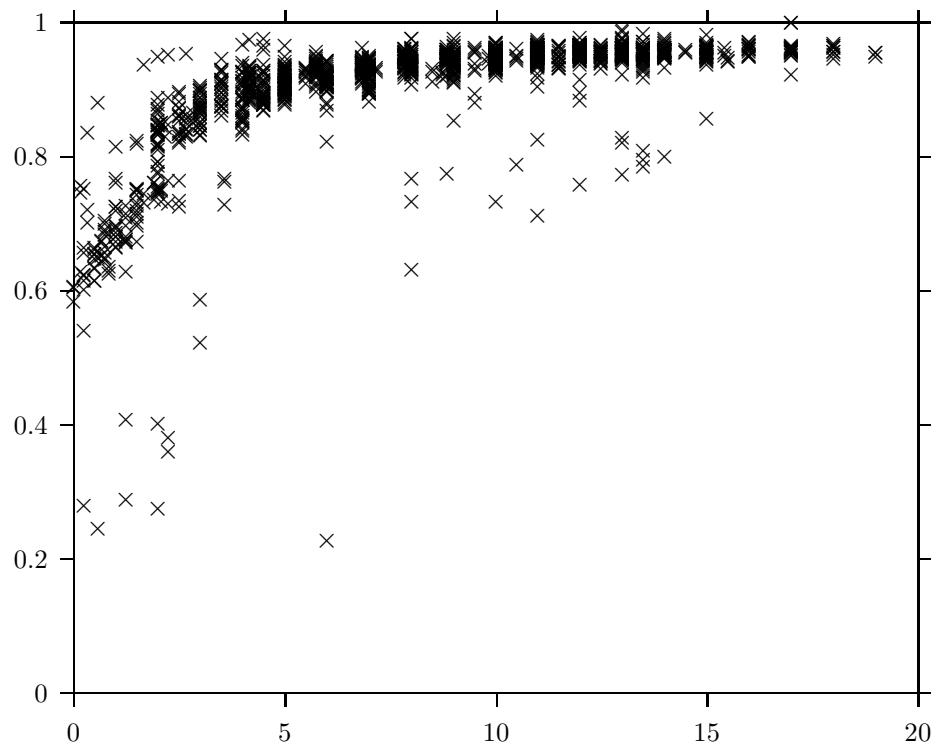


FIG. 3.9 – Ossification des articulations en fonction de l'âge osseux

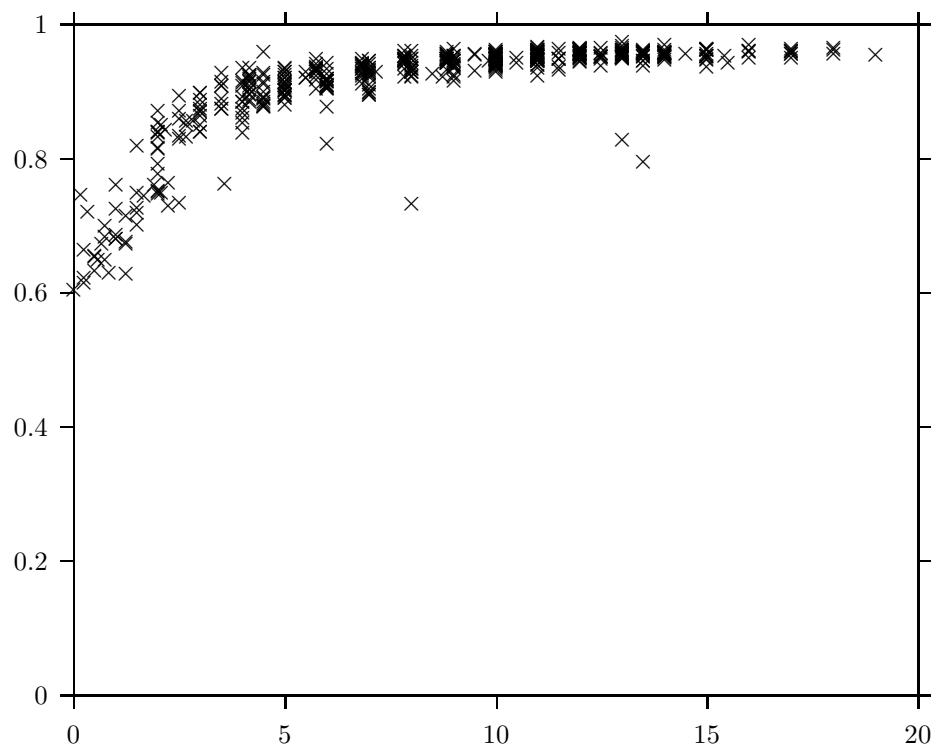


FIG. 3.10 – Médianes de l'ossification des articulations en fonction de l'âge osseux

avec une discontinuité vers 2 ans, ce qui correspond à l'apparition des épiphyses. Les quelques valeurs manifestement aberrantes correspondent à des erreurs de segmentation, avec des scores a priori généralement inférieurs à 30, ce qui permettrait de les écarter facilement.

La figure 3.10 est obtenue de manière similaire, mais en ne marquant que la valeur médiane des 3 rapports calculés pour chaque image. On voit que cela suffit à éliminer l'essentiel des points aberrants, preuve encore que les erreurs sont relativement indépendantes et n'affectent généralement, au plus, qu'un seul axe d'une image. La redondance sur 3 mesures est alors presque suffisante pour constituer une mesure robuste, mais on peut difficilement faire mieux. Le pouce est structurellement différent et la segmentation des axes extérieurs est moins fiable.

On obtient des résultats voisins avec la seconde articulation et les rapports $\frac{L_{2j}}{L_{2j}+a_{2j}}$ ou $\frac{L_{1j}}{L_{1j}+a_{2j}}$, mais les lois sont un peu différentes et il n'est pas évident de combiner toutes ces valeurs numériques dans une même régression. De toute façon, la dispersion des valeurs est trop grande et la pente de la courbe est trop faible pour permettre de déduire un âge osseux précis, à partir de la simple mesure de ces rapports. En fait, le contact articulaire se réalise rapidement et, par la suite, la valeur du rapport n'évolue plus que du fait de l'allongement relatif de l'os.

Une mesure plus progressive de la maturation d'une articulation devrait prendre en compte le remplissage de tout l'espace inter-diaphysaire et, entre autres, d'une manière plus ou moins explicite, la forme des épiphyses. On a assez d'éléments pour définir, pour chaque articulation, une "boîte englobante" convenablement localisée, orientée et dimensionnée. Notamment pour une application à l'âge osseux, il était important de pouvoir traiter correctement l'articulation de la phalange distale, même si cette dernière est mal contrastée. Cette localisation précise peut servir de base pour la mise en œuvre de traitements complémentaires, adaptés à chaque articulation, pour des mesures de longueur, de largeur ou encore de densité. Des mesures de largeurs relatives de l'épiphyse et de la diaphyse, perpendiculairement à l'axe, donnent déjà des résultats plus facilement exploitables [Pie95] que les mesures longitudinales.

D'une manière plus générale et plus sophistiquée, on peut procéder à des recalages et ainsi envisager des approches dirigées par un modèle propre à chaque articulation. On a également le moyen de réaliser automatiquement l'apprentissage statistique qui est requis comme préliminaire pour la mise en œuvre de telles approches.

Enfin, rappelons que l'analyse des articulations des doigts n'est pas suffisante dans les méthodes classiques [GP83, TWC⁺83] de détermination de l'âge osseux. L'analyse du carpe est également nécessaire.

3.4.2 Profil métacarpo-phalangien

La figure 3.11 donne une représentation des écarts \tilde{e}_{ij} pour deux images d'exemple. On a utilisé pour cela la définition $\tilde{e}_{ij} = \frac{L_{ij} - \bar{L}_{ij}}{\bar{L}_{ij}}$, avec L_{ij} la longueur mesurée et \bar{L}_{ij} une longueur de référence adaptée à chaque image (voir 3.1.3).

Certains troubles de la croissance et malformations congénitales se traduisent par de fortes anomalies morphométriques, de plusieurs fois l'écart-type, sur quelques os particuliers [PG97]. La technique diagnostique du profil métacarpo-phalangien se base sur l'analyse de graphiques semblables à la figure 3.11, mais utilise des références de longueur absolue \bar{L}_{ij} données par des tables et adaptées à une sous-population particulière (sex, age, etc.). On peut utiliser diverses expressions de l'écart, notamment le Z-score : $\frac{L_{ij} - \bar{L}_{ij}}{\sigma_{ij}}$, normalisé par l'écart-type propre à chaque os [PGNG72], ou le Q-score : $\log\left(\frac{L_{ij}}{\bar{L}_{ij}}\right)$ [DV91, DV92]. En fait, les éléments du diagnostic ne sont pas les scores individuels mais plutôt les profils, par exemple croissant ou décroissant, concave ou convexe, observés dans et entre certains groupes d'os. Les différentes mesures de l'écart se distinguent alors par des propriétés particulières d'amplification et de mise en évidence.

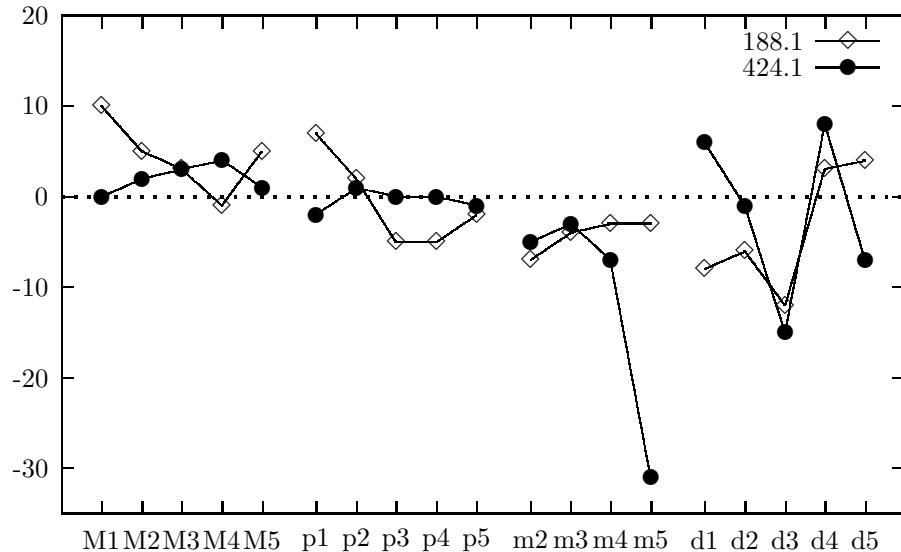


FIG. 3.11 – Profil métacarpo-phalangien pour les images 188.1 et 424.1

Pour l'image 188.1, il n'y a pas d'erreur technique perceptible et l'anomalie possible sur d3 est anatomique. Pour l'image 424.1, la valeur aberrante pour m5 est une forte anomalie anatomique et on peut vérifier que cet os est détecté et mesuré correctement par notre méthode. La seule erreur technique est une mesure trop courte de d3 pour l'image 424.1. On voit sur cet exemple la difficulté à distinguer automatiquement les anomalies anatomiques des erreurs techniques. On notera aussi le profil naturellement chaotique pour les phalanges distales.

Pour développer une telle application, notre algorithme de segmentation semble déjà assez robuste pour les phalanges proximales et médianes et il le sera d'autant plus si on peut fiabiliser les mesures par une analyse complémentaire des articulations, comme vu pour l'âge osseux. Par contre, il est nécessaire d'améliorer la localisation des extrémités proximales des métacarpiens et il faut se poser la question de l'utilité des phalanges distales, d'ailleurs rarement utilisées de manière déterminante dans les examens classiques. Notre méthode permet aussi de réaliser, en préalable, un apprentissage automatique de la normalité sur un ensemble d'images conséquent et approprié.

Il reste à établir et à valider le fonctionnement dans le cas de fortes déviations qui seraient spécifiques à une pathologie particulière. Mais de tels cas sont trop rares pour permettre une analyse statistique directe. La méthodologie qui est employée en 3.2, peut être reprise pour une expérimentation indirecte, au prix d'une paramétrisation spécifique à chaque pathologie. Ainsi, par exemple, on peut simuler le comportement sur des déviations de 20 à 40 % en utilisant les images, beaucoup plus courantes, qui présentent des déviations dans l'intervalle $\pm 10\%$ et en imposant une déviation complémentaire de 30 % par une variation artificielle de la normalité.

Enfin, pour l'usage diagnostic, il est concevable de laisser au praticien la possibilité de choisir une paramétrisation adaptée à une certaine hypothèse de pathologie.

3.4.3 Recalage et étude statistique de la forme des axes

La segmentation que nous obtenons permet de placer de manière relativement fiable les points bien identifiés que sont les extrémités des métacarpiens et des phalanges. Cela permet d'initier un recalage. Une application directe de ce résultat est une analyse statistique de la cohérence de forme des axes. En ne traitant que la forme des axes, on n'a pas à considérer la grande variabilité de structure des os qui compliquerait un modèle de forme, basé sur les contours, ou

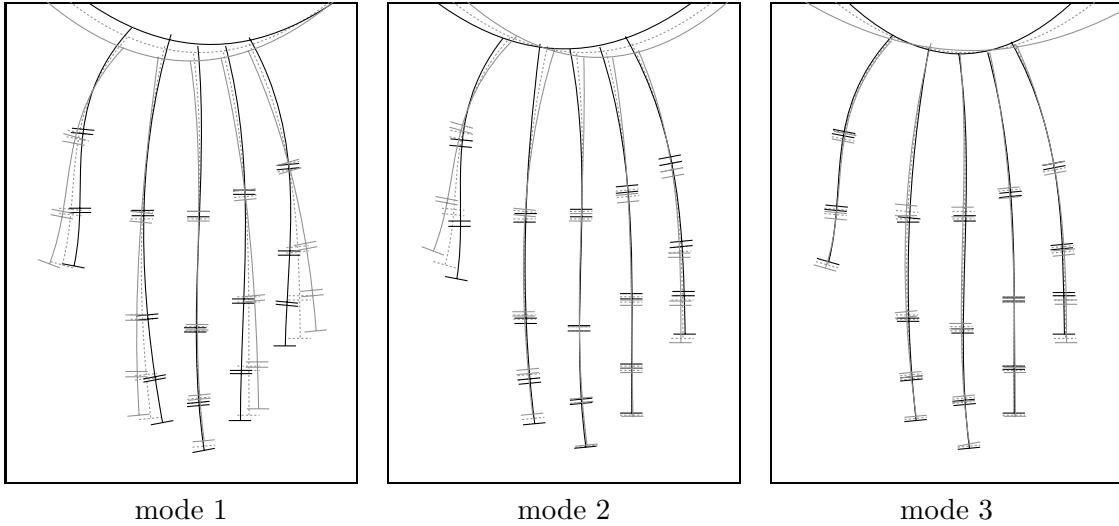


FIG. 3.12 – Trois premiers modes de covariance des axes, avec un recalage standard

un modèle d'apparence, basé sur les intensités. On peut ainsi envisager d'acquérir des résultats statistiquement significatifs sur quelques centaines d'images.

Le modèle de forme que nous utilisons ici, se base sur 38 points clefs qui sont les extrémités des 19 os. Des points supplémentaires, placés le long de chaque axe, sont utiles pour prendre en compte la courbure de cet axe sur la longueur d'un os. Au total, ce modèle utilise 117 points. On a choisi de découper la longueur d'un os en, respectivement, 8, 6, 4 et 2 intervalles relativement réguliers, en allant des métacarpiens aux phalanges distales, ce qui suit approximativement les dimensions relatives de ces os et permet d'approcher une densité curviligne uniforme. Les positions longitudinales de ces nouveaux points sont interpolées à partir de celles des points clefs et elles sont naturellement affectées par la variabilité anatomique de la longueur de chaque os et par les erreurs de segmentation. Le choix du nombre de points interpolés est donc aussi une manière de pondérer ces perturbations.

Nous avons utilisé la technique classique [CTCG95, DM98] qui consiste en une analyse en composantes principales de la matrice de covariance des points de définition, après une normalisation par recalage sur le modèle. Selon la méthode dite de Procrustes [Goo91], usuelle pour cette application, la transformation de recalage consiste en une similitude. Une symétrie axiale est décidée au préalable si besoin, selon le côté du pouce.

Les 3 premiers modes de covariance ainsi obtenus, sont représentés figure 3.12. Ces modes principaux portent respectivement de l'ordre de 38 %, 29 % et 6 % de la variance globale et ce résultat est obtenu sur une base d'un peu plus de 200 images choisies sans erreur manifeste, selon les scores a priori. Pour chaque mode, à partir des informations de moyenne et de variance pour chaque point du modèle, on a retracé les axes (par interpolation polynomiale cubique), les extrémités des os et une courbe de régression (polynomiale de degré 2) pour les extrémités proximales des 5 métacarpiens. Le tracé en pointillé correspond aux axes moyens et les traits continus correspondent à des variations de 1 écart-type de part et d'autre. On voit que cela donne déjà de fortes amplitudes. Une illustration plus standard, avec des variations de l'ordre de 2 à 3 fois l'écart-type, montrerait des configurations irréalistes avec des croisements potentiels.

En procédant comme ci-dessus à un recalage sur l'ensemble des points de définition, on constate que le résultat est très sensible aux variations d'écartement des doigts. Un écartement des doigts augmente en effet la distance moyenne des points de définition avec leur barycentre et cette distance est la mesure qui détermine l'échelle de la similitude. Même avec la pondération qui résulte du choix de répartition des points interpolés, l'influence des phalanges distales est

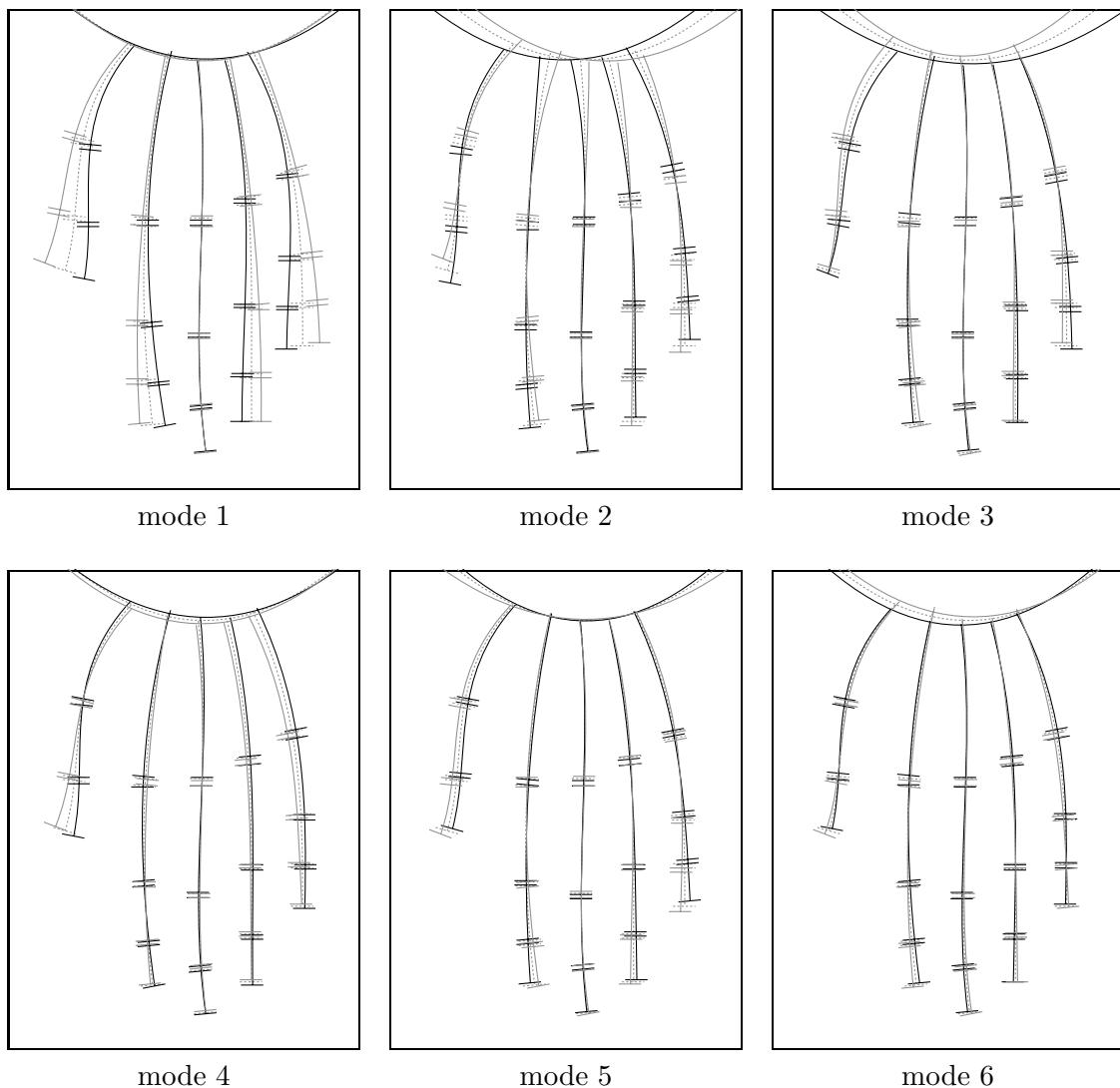


FIG. 3.13 – Premiers modes, avec recalage sur les 3 phalanges proximales centrales

importante et ceci perturbe la normalisation des longueurs, comme on le voit avec le mode 1.

Cela brouille considérablement la pertinence d'une analyse en composantes principales. Si le recalage ne tient pas compte des phalanges distales, on évite aussi qu'il soit perturbé par les erreurs de segmentation les plus fréquentes. Le recalage doit de même ignorer le pouce qui dispose d'une très grande liberté articulaire et on pourra alors retrouver ses variations dans les composantes de forme. Enfin, on peut veiller à ce que l'ajustement du facteur d'échelle ne dépende pas de l'avancement de l'ossification, de manière à pouvoir observer correctement d'éventuelles corrélations avec ce dernier. En repensant ainsi la séparation classique entre pose et forme, on peut espérer mieux distinguer d'éventuelles cohérences utiles.

Nous avons donc retenu le principe d'un recalage effectué sur un ensemble restreint de points, ceux qui sont attachés aux phalanges proximales des trois axes centraux, ainsi que les extrémités proximales correspondantes sur les phalanges médianes. On a ainsi des références de longueur qui prennent en compte une zone d'ossification plus un espace articulaire, cela pour des os dont la détection est très fiable et qui sont à peu près au centre du domaine géométrique. Une utilisation alternative des seuls trois métacarpiens centraux permet d'illustrer différemment certaines corrélations mais le résultat s'avère en général plus bruité, du fait des erreurs assez fréquentes sur les extrémités proximales de ces os.

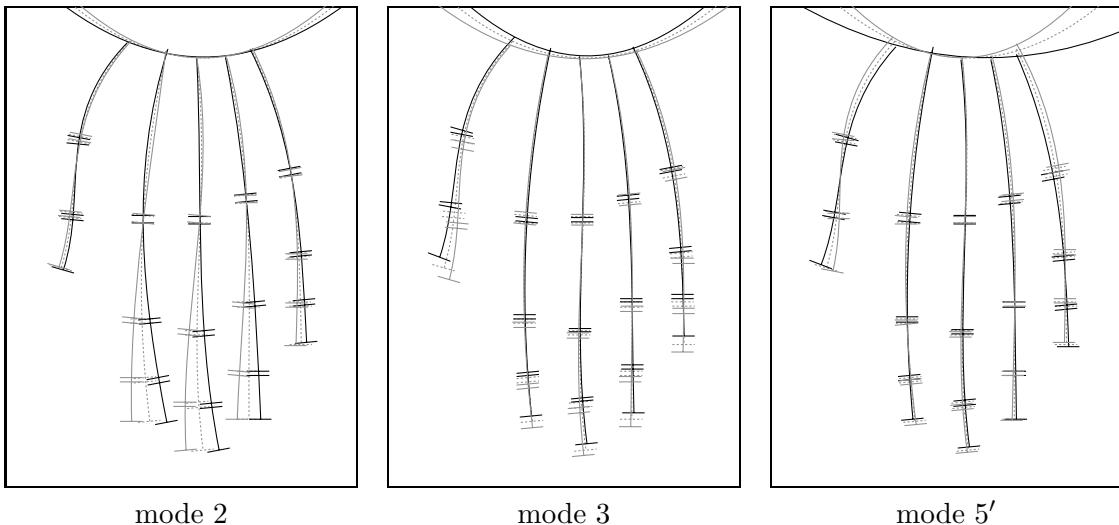


FIG. 3.14 – Quelques modes particuliers, avec recalage sur les 3 métacarpiens centraux

Les 6 premiers modes obtenus par l'analyse en composantes principales, après un recalage sur les trois phalanges proximales centrales, sont représentés figure 3.13. Ces modes portent respectivement de l'ordre de 39 %, 24 %, 8 %, 7 %, 4 % et 3 % de la variance globale observée pour ce recalage. La figure 3.14 reprend trois de ces modes pour lesquels il est instructif de considérer le recalage défini sur les métacarpiens centraux.

- Le mode 1 décrit l'écartement des doigts “en éventail”. Comparativement à ce que pouvait montrer la figure 3.12, il n'y a donc pas de relation perceptible avec la longueur des axes, mais on observe un couplage assez net entre l'écartement des doigts et l'avancement de l'ossification des espaces articulaires. Une explication est la variation avec l'âge du rapport d'aspect entre la longueur et l'écartement, alors que la longueur des phalanges prime dans l'estimation de la similitude. D'autres causes, anatomiques ou liées à la pratique radiographique, sont possibles.

- Le mode 2 correspond essentiellement à une rotation entre deux blocs relativement rigides, d'une part, les métacarpiens et, d'autre part, les phalanges des 3 axes centraux. La figure 3.14 montre plus clairement la rigidité du bloc des métacarpiens et le mouvement respectif des doigts.

La combinaison de ces deux premiers modes autorise une grande liberté de rotation de chaque doigt autour de son articulation métacarpo-phalangienne. Le nombre de points concernés et l'amplitude des déplacements latéraux suffisent à expliquer que cela constitue les deux modes principaux de variance. Dans ces deux premiers modes, il n'y a pas de variation significative des longueurs des os qui sont donc clairement décorrélées de la forme des axes. Cela justifie la séparation du problème en deux étapes comme nous le faisons, c'est-à-dire, d'abord la localisation transversale des axes puis leur segmentation longitudinale.

- Le mode 3 est une variation simultanée de la longueur de tous les métacarpiens, relativement aux phalanges. Avec la figure 3.14 qui utilise les métacarpiens comme référentiel, on confirme que cela s'accompagne d'une variation de courbure de la ligne d'articulation entre carpe et métacarpe, c'est-à-dire, d'une variation des longueurs relatives entre métacarpiens.

- Le mode 4 montre une variation de la largeur du carpe, relativement aux phalanges. Il y a ainsi une possibilité de variation indépendante du rapport d'aspect du bloc des métacarpiens.

- Pour le mode 5, il est déjà plus difficile de donner une interprétation à ce qui semble être surtout un “résidu” de variations, complémentaire des premiers modes. Le mode 5', figure 3.14, est différent et correspond à une variation particulièrement sensible des longueurs relatives entre métacarpiens.

À partir de ce rang, on observe un certain nombre de variations qui sont en fait indépendantes, avec des variances du même ordre de grandeur. La définition de directions principales dans ce sous-espace des variations relativement isotropes est alors mal déterminée et elle fournit des combinaisons aléatoires, selon les petites fluctuations de covariance induites par le mode de recalage et par le choix de la base d'apprentissage.

- Le mode 6 montre une autre possibilité de variation des longueurs relatives des métacarpiens.

Comme le montrent les expérimentations précédentes, on ne peut pas lier les longueurs relatives des métacarpiens entre elles ou avec celles des phalanges. La combinaison des seuls modes 3, 5' et 6 permet déjà d'engendrer diverses variations importantes de position et de courbure pour la ligne d'articulation entre carpe et métacarpe. Un modèle pour cette ligne doit donc avoir au moins 3 degrés de liberté et son recalage sur seulement 5 points est relativement instable.

Pour le pouce on peut observer la grande liberté articulaire de chaque phalange. Pour les autres doigts, même s'il n'existe pas à proprement parler de liberté articulaire inter-phalangienne dans le plan de l'image, on doit considérer qu'il en est ainsi d'un point de vue statistique. Par exemple, après un recalage qui se base seulement sur une phalange proximale, l'analyse en composantes principales montre l'existence d'un mode propre de variation de la direction des deux phalanges qui sont dans son prolongement, sans corrélation avec la forme des autres axes, ni l'avancement de l'ossification, ni quelque longueur. Le couplage qui est ainsi observé entre les deux phalanges médiane et distale est purement mécanique et un recalage sur la phalange médiane met en évidence une liberté supplémentaire de rotation de la phalange distale seule. On peut de même identifier, axe par axe, des modes indépendants de variation des longueurs, dont une variation couplée des deux phalanges médiane et distale, relativement à la proximale, et une variation de longueur de la phalange distale seule, relativement à la médiane.

En fait, il s'agit clairement d'un ensemble articulé avec, pour chaque os, une liberté angulaire et une liberté de variation de sa longueur, comme envisagé dans [MFMMMFAL05]. Mais une localisation effective de chaque articulation, ou une détermination équivalente, est indispensable pour initier le recalage d'ensemble d'un tel modèle, avant d'envisager une mise en correspondance plus précise et particulière à chaque zone d'intérêt. Notre méthode est alors un moyen rapide et fiable pour obtenir cette localisation préalable, en utilisant très peu d'a priori de forme.

Un tel ensemble articulé peut être naturellement paramétré par le système de coordonnées polaires relatives qui utilise la longueur des os (et des espaces articulaires) et les angles entre deux segments consécutifs sur un axe. En procédant à une analyse en composante principale dans ce système de coordonnées, on met en évidence l'indépendance de forme des axes. Les trois premiers modes restent similaires à ceux de la figure 3.13 et ce sont les seuls modes qui expriment une certaine cohérence globale de forme. On observe ensuite un grand nombre de modes qui correspondent à une variation individuelle de longueur ou d'angle.

3.4.4 Difficultés pour un modèle actif de forme des axes

Une des étapes d'un modèle actif de forme consiste à déplacer chaque point de définition de manière à optimiser la correspondance entre son voisinage dans l'image et un profil modèle. Dans notre cas, ce principe pourrait être réalisé sur une carte des symétries, voir A.3, avec des profils qui traduisent la recherche des maxima dans cette carte. Le calcul de cette carte demande une estimation préalable de la largeur d'appariement et son coût incite à une évaluation paresseuse basée sur une première localisation grossière des axes. Le principe serait alors d'utiliser le modèle actif à une échelle relativement plus fine pour améliorer la localisation. La difficulté provient de ce que les symétries potentielles sont en nombre important et très rapprochées, alors que la variabilité des formes se prête mal à des a priori statistiques.

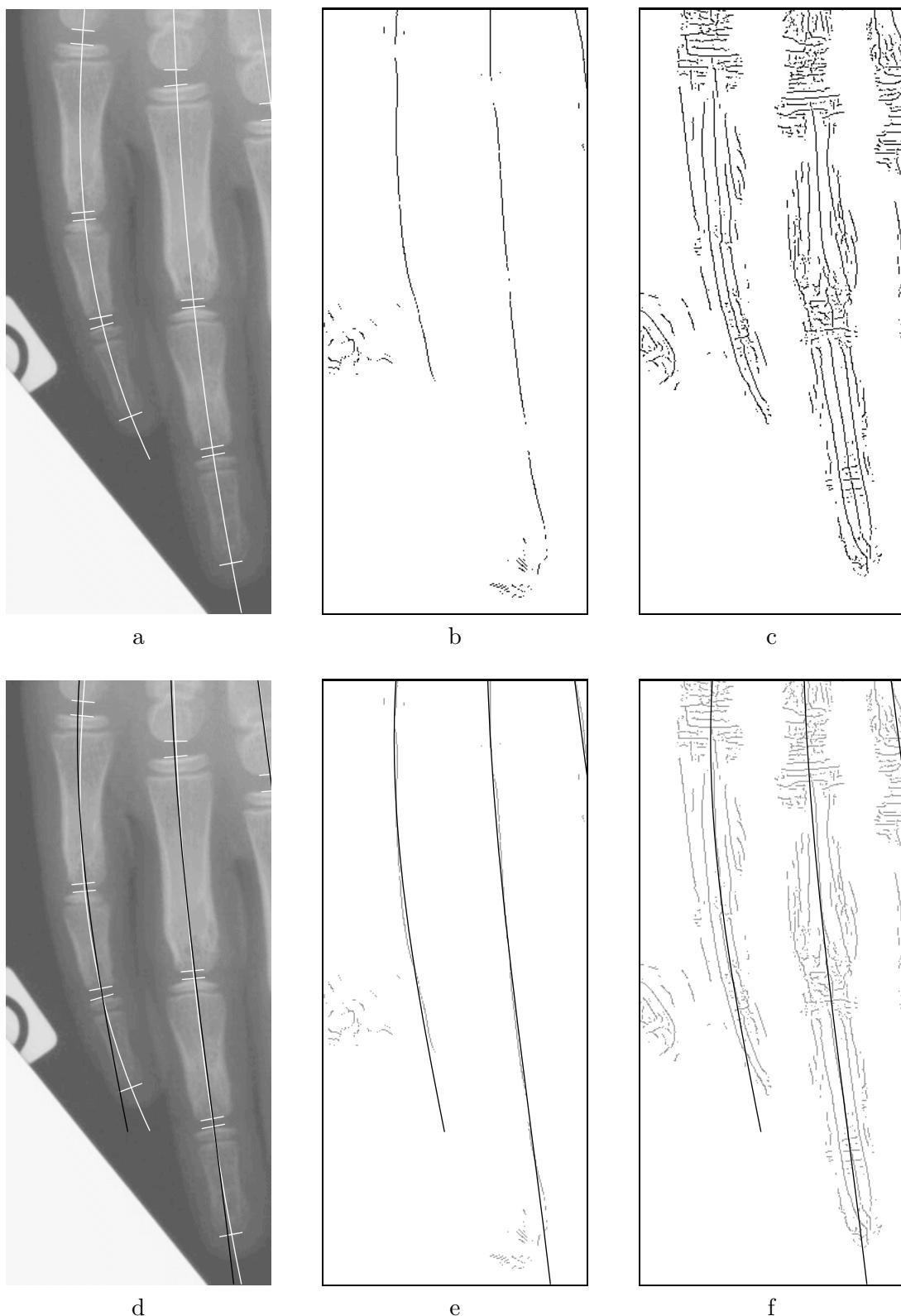


FIG. 3.15 – Exemple de difficulté de mise en œuvre d'un a priori statistique

La figure 3.15 donne un exemple de problèmes rencontrés couramment sur les phalanges distales. En haut de cette figure, de gauche à droite, on a représenté :

- en 3.15-a, l'extrait d'une radiographie avec, tracés en blanc, les axes obtenus par la méthode décrite au chapitre 2,
- en 3.15-b, la carte des symétries calculée à une échelle de détection,
- en 3.15-c, la carte des symétries calculée à une échelle de localisation.

Les résultats sur l'image complète sont visibles respectivement sur les figures B.16, page 187, à gauche, A.5, page 154 et A.6, page 155. Pour calculer les cartes de symétries, on a utilisé les échelles et la largeur d'appariement retenues lors du calcul des axes.

Au bas de la figure 3.15, on a superposé en noir le tracé qui correspond à l'axe moyen du modèle statistique. Ce tracé a été recalé axe par axe sur la phalange médiane obtenue par notre méthode. En ignorant les problèmes qui relèveraient d'une autre initialisation, on peut considérer que cela correspond à un résultat intermédiaire pour un modèle actif qui aurait déjà résolu les modes principaux correspondant aux libertés de l'articulation métacarpo-phalangienne. L'angle d'articulation entre les deux premières phalanges est très proche de la valeur normale.

Par contre, en 3.15-d, on voit une nette différence sur une phalange distale entre l'axe calculé par appariement (en blanc) et l'axe moyen (en noir). Un critère de normalité statistique inciterait alors à douter de notre résultat, pourtant correct. En 3.15-b et A.5, on voit que la convergence d'un modèle de forme serait facile à une échelle de détection. Mais la symétrie est mal déterminée sur l'extrémité distale et 3.15-e montre une mauvaise localisation, peu différente de l'axe moyen. En 3.15-f, on voit qu'il serait difficile de guider le modèle de forme sur la bonne convergence, car il existe une symétrie concurrente basée sur un appariement entre peau et os. Cette autre symétrie est plus proche de l'axe moyen et elle présente une meilleure cohérence en bout d'axe. Ce sont donc ses points qui seront choisis dans la phase "du modèle vers l'image", ce qui détermine le maintien du modèle sur une fausse solution.

Il existe bien sûr des cas contraires où l'utilisation d'un modèle statistique permettrait de corriger un axe qui serait mal estimé par la méthode du chapitre 2. En fait, il n'y a pas de règle générale pour pondérer la pertinence relative des deux méthodes. Notre méthode de mise en évidence et de construction de l'axe utilise aussi des hypothèses de régularité et l'apport d'un a priori de forme ne semble pas exploitable pour décider de la localisation transversale d'un axe.

3.5 Conclusion

L'ensemble des observations particulières apportées en 3.3 confirment les hypothèses émises en 3.1.5 et 3.2.4. Les interprétations positives sont même renforcées puisque la précision effective est généralement meilleure que celle estimée par les scores a priori. En fait, ces scores qui sont les seuls critères de qualité disponibles a priori, ne font que traduire la normalité des mesures de longueur et manquent de pertinence pour quantifier la justesse d'une segmentation. Ils sont finalement utiles pour détecter les cas aberrants, sans toutefois expliciter s'il s'agit d'une erreur technique ou d'une anomalie anatomique. Cela montre aussi toute la difficulté à utiliser des a priori de longueur lors de la segmentation.

La mesure des phalanges proximales et médianes est ainsi très précise et aussi très robuste, vis-à-vis des variations biométriques courantes et pour tous les stades de la croissance. Cela découle naturellement de la bonne localisation des articulations à laquelle nous parvenons. Ce processus est dirigé essentiellement par l'image et des contraintes structurelles adaptées servent à résoudre les ambiguïtés, sans influence notable par les a priori de normalité qui sont faiblement utilisés.

La méthode du chapitre 2 peut donc être vue comme une première étape essentielle dans la localisation des structures. En particulier, on a suffisamment d'information pour procéder à un recalage global ou local avec des images d'un atlas ou un modèle statistique. Dans les deux cas, il semble pourtant nécessaire de distinguer plusieurs stades de développement, pour s'adapter à la complexité d'évolution de forme et d'apparence de chaque os. Il est alors préférable d'incorporer dès ce stade l'expertise développée dans les méthodes classiques, en respectant notamment les étapes les plus caractéristiques du développement propre à chaque os. Le caractère automatique de notre méthode de localisation permet aussi d'envisager l'apprentissage sur de grandes bases qui est requis pour tenir compte de la diversité anatomique.

À partir de cette localisation, on peut également envisager des techniques complémentaires d'analyse des articulations. Cela permettrait par ailleurs de corriger la localisation des extrémités des os et de confirmer d'éventuelles anomalies morphométriques. On pourrait fiabiliser ainsi les mesures de longueur qui sont requises pour des techniques diagnostiques reconnues, comme le profil métacarpo-phalangien, ou effectuer d'autres mesures particulières utilisées dans certaines méthodologies alternatives pour l'âge osseux.

Par contre, la méthode du chapitre 2, à elle seule, manque de robustesse pour localiser l'extrémité proximale des métacarpiens et, plus nettement, l'extrémité des phalanges distales. Cela résulte notamment de la grande variabilité anatomique, avec une très forte indépendance morphométrique pour chaque os. Cela interdit la mise en œuvre de contraintes de normalité ou de cohérence suffisamment précises pour diriger correctement le choix d'un contour d'extrémité.

L'autre source d'erreur est la mauvaise localisation transversale des extrémités des axes. Cela s'explique par des ambiguïtés dans la caractérisation des symétries qui proviennent, soit des formes particulières aux extrémités proximales des métacarpiens, soit du faible contraste pour les phalanges distales. Le choix d'une fausse symétrie induit une déviation de l'axe et un manque de symétrie donne un axe trop court. Notons que les diverses techniques classiques de caractérisation des symétries sont affectées de manière similaire, que les approches dirigées par un modèle s'adaptent mal à la variabilité et que notre technique reste relativement efficace par la mise en œuvre de contraintes structurelles d'appariement facilement paramétrables.

L'estimation des paramètres de forme d'un axe n'est pas suffisamment déterminée pour assurer, selon le cas, une correction automatique ou une extrapolation correcte, surtout quand l'image prédispose localement à de multiples faux appariements. L'approximation polynomiale que nous utilisons est aussi satisfaisante pour régulariser les directions tangente et normale le long de l'axe détecté, tout en conservant de faibles déviations transversales sur la partie interpolée. Mais elle est aussi mal déterminée pour assurer la localisation transversale sur une partie extrapolée.

Pour les métacarpiens, du fait de l'utilité de ces os dans les diagnostics, nous devons envisager des techniques complémentaires. L'analyse statistique en 3.4.3 laisse entrevoir une possibilité d'exploiter la cohérence de direction et d'espacement des axes des métacarpiens pour améliorer leur localisation transversale. Il reste le problème délicat du calage longitudinal pour ancrer des points clefs sur les extrémités proximales et cela pose toujours le problème de l'usage de contraintes de normalité pour y parvenir. Au prochain chapitre, nous verrons comment utiliser les résultats acquis ici, en tenant compte des faiblesses constatées, pour guider la segmentation du carpe et, en retour, mieux localiser les extrémités des métacarpiens.

La localisation des extrémités des phalanges distales est très aléatoire, essentiellement par manque de contraste et l'inadéquation d'a priori de longueur ou de direction. On peut même douter de la faisabilité d'une solution générale satisfaisante, dès lors qu'une lecture visuelle est elle aussi parfois très délicate. De fait, la longueur des phalanges distales est peu utilisée dans les diagnostics. Par contre, leurs articulations sont très utiles [Kir84] et nous parvenons à les localiser correctement.

Chapitre 4

Définition d'un modèle géométrique pour le carpe

La région du carpe présente de nombreuses particularités qui compliquent la détection des contours des os. Notamment la contiguïté et la superposition des régions osseuses ont tendance à atténuer l'amplitude des “vrais” contours, alors que divers détails anatomiques vont générer des parasites relativement forts.

Cela complique notamment les opérations de détection par seuillage. La densité des maxima de gradient parasites pose alors des difficultés dans le suivi des contours. Les figures 1.1, page 18, et 2.3, page 38, montrent que le problème de la mise en évidence des structures est relativement plus délicat qu'en ce qui concerne les os longs. Les contours obtenus ne sont pas complètement fermés, ni même constitués d'une seule chaîne et il y a également des bifurcations. La validation d'une chaîne de maxima comme un fragment de contour est alors loin d'être évidente a priori. Il arrive notamment qu'une chaîne soit constituée de la mise bout à bout de pixels des contours de deux os. Un apport de règles spécifiques n'est envisageable que si on a résolu le problème de l'identification, c'est-à-dire, si l'on sait déterminer à quel(s) os une chaîne peut se rattacher.

D'autre part, lorsqu'un os n'a pas atteint le stade à partir duquel il est visible sur l'image, il faut éviter les fausses détections (interprétations de bruits comme des contours) et les fausses identifications (confusions entre deux os). Notons que, même si l'on disposait déjà d'une bonne estimation de l'âge osseux, il n'existe pas d'ordre chronologique général qui pourrait aider à déterminer ces problèmes. Comme on peut le voir avec le tableau 2 de l'introduction, les os du carpe présentent une grande variabilité dans leur ordre d'apparitions respectives. En particulier, la cardinalité de l'ensemble des os du carpe qui sont visibles sur une image, ne détermine pas les os qui composent cet ensemble.

Comme il s'agit de formes plutôt rondes, une caractérisation utilisant des symétries centrales semble bien appropriée. On va ainsi reprendre des principes similaires à ceux mis en œuvre pour les os longs, avec un point central à la place d'un axe.

Pour initier le processus de détection et d'identification, il faut un premier placement a priori des centres de référence. Une mise en évidence de ces centres de symétries particulières, directement à partir de l'image, n'est pas envisageable, en raison de la multiplicité des symétries plus ou moins locales (voir notamment l'annexe A.3). Plus généralement, les techniques de mise en évidence par accumulation [Bal81, Big88, RWY90] sont toujours relativement coûteuses, le filtrage des nombreuses symétries parasites est délicat à paramétrier et, de toute façon, cela ne règle pas les problèmes d'identification évoqués ci-dessus.

Nous avons choisi d'utiliser la localisation des os longs, obtenue par les techniques développées au chapitre 2, pour déterminer un placement a priori d'un centre de référence pour chaque os du carpe. Nous verrons en 4.2 comment des constructions géométriques permettent d'y parvenir

d'une manière relativement robuste vis-à-vis de la variabilité biométrique et des éventuelles erreurs de segmentation constatées au chapitre 3. Pour affiner la caractérisation des contours, la position de chaque centre doit être ensuite recalée sur les contours déjà identifiés.

Nous allons commencer par regarder comment la donnée d'un point central de référence peut effectivement contribuer à détecter et à identifier le contour d'un os du carpe et, inversement, comment recaler la position d'un tel point à partir de contours, éventuellement incomplets.

4.1 Détection et identification des contours

4.1.1 Utilisation d'un point de référence pour la caractérisation des contours

Dans un premier temps, chaque os peut être considéré indépendamment. Il s'agit de quantifier l'appartenance d'un pixel de maxima du gradient au contour d'une région osseuse donnée. Dans une détection de contour classique, on utiliserait simplement le module du gradient, mais nous voulons exprimer ici le fait que le contour cherché s'organise autour d'un point central, dit point de référence. Contrairement aux techniques plus générales [Low85, Sau93], nous supposons que ce point est connu a priori.

Nous noterons \mathbf{C} ce point de référence, \mathbf{p} les coordonnées du pixel candidat considéré et, par abus, ce pixel lui-même, \mathbf{g} le vecteur gradient (fonction de \mathbf{p}) calculé en ce pixel et \mathbf{r} le vecteur $\overrightarrow{\mathbf{C}\mathbf{p}}$. La caractérisation d'un maximum du gradient comme pixel de contour va prendre en compte trois critères, assez comparables dans le principe à ceux utilisés pour les os longs (voir 2.6) :

- l'intensité du gradient $|\mathbf{g}|$ en ce pixel, critère usuel purement local qui permet déjà d'écartier par seuillage une part importante des parasites.

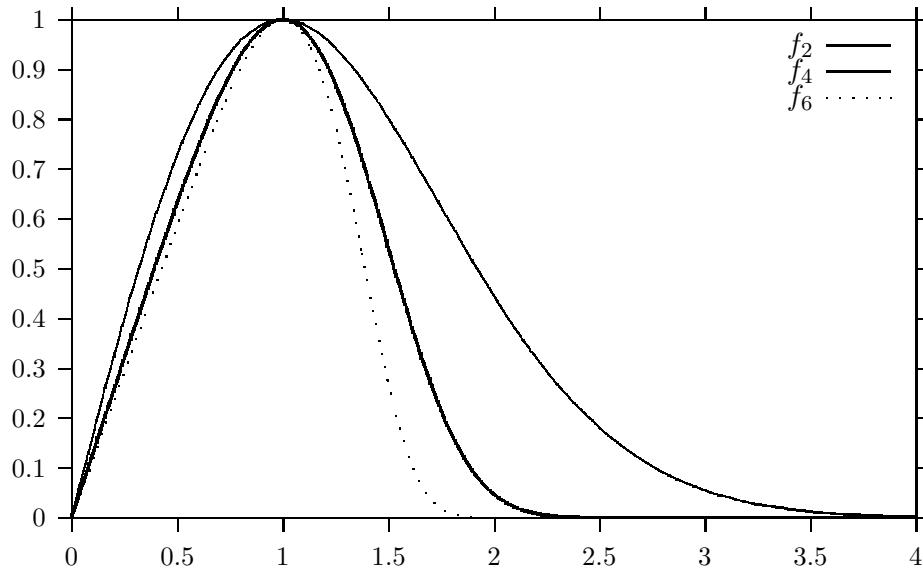
- le cosinus $\cos(\mathbf{g}, \mathbf{r})$ qui exprime la direction du gradient, par rapport au point de référence ; on veut que le gradient soit dirigé vers le point de référence, c'est-à-dire que ce cosinus doit être proche de -1 . En pratique, une tolérance assez large est requise pour prendre en compte des formes non circulaires et un mauvais centrage du point de référence. Remarquons que la condition $\cos(\mathbf{g}, \mathbf{r}) < 0$ est théoriquement suffisante pour rejeter un pixel d'un éventuel contour adjacent, celui d'un os voisin et où le gradient est alors de direction opposée.

- une fonction $f(|\mathbf{r}|)$ de pondération selon la distance du pixel considéré au point de référence. Cette fonction sera maximale à une distance "idéale" : elle doit contribuer à pénaliser les contours des autres os qui sont relativement loin et ne doit pas favoriser des détails internes qui seraient à faible distance du point de référence. Pour les os longs, on avait simplement utilisé un intervalle de validité de la distance à l'axe, ici il faut un critère plus flou car la taille d'un os est moins bien connue a priori. Et, de nouveau, on doit accepter des formes non circulaires et un mauvais centrage du point de référence.

Ces trois critères sont à valeurs bornées et, au lieu de les traiter individuellement, on peut considérer leur produit $\mathcal{K}_{\mathbf{C}}(\mathbf{p}) = -|\mathbf{g}| \cos(\mathbf{g}, \mathbf{r}) f(|\mathbf{r}|)$, ce qui est plus pratique pour mettre en œuvre un seuillage. Le changement de signe permet d'avoir une interprétation plus intuitive.

Pour la commodité des calculs, il est alors très intéressant de faire apparaître le produit scalaire $\mathbf{g} \cdot \mathbf{r} = |\mathbf{g}| |\mathbf{r}| \cos(\mathbf{g}, \mathbf{r})$ dans l'expression de $\mathcal{K}_{\mathbf{C}}(\mathbf{p})$, c'est-à-dire, d'avoir un facteur $|\mathbf{r}|$ dans l'expression de $f(|\mathbf{r}|)$. En ajoutant la contrainte $f(0) = f(\infty) = 0$ et le souci d'une certaine régularité, ainsi qu'une décroissance rapide, on pense assez naturellement à une forme en $x e^{-x}$.

Pour calibrer la réponse, on ajoute les contraintes $f'(R_{\text{ref}}) = 0$ et $f(R_{\text{ref}}) = 1$, avec R_{ref} le "rayon de référence" de l'os considéré et, après le changement de variable $\rho = \frac{|\mathbf{r}|}{R_{\text{ref}}}$, on a une solution $f_1(\rho) = \rho \frac{e^{-\rho}}{e^{-1}}$. On remarque que ces contraintes de normalisation suppriment toute possibilité de réglage de la tolérance, par exemple le choix de la valeur de f pour $\rho = 2$.

FIG. 4.1 – Fonctions f_k de pondération selon la distance au point de référence

On retrouve une certaine souplesse en considérant la famille des fonctions $f_k(\rho) = \rho \frac{e^{-\frac{\rho^k}{k}}}{e^{-\frac{1}{k}}}$ (voir figure 4.1). Pour la commodité des calculs, nous avons retenu k entier pair et le critère composite $\mathcal{K}_C(\mathbf{p})$ s'exprime alors facilement à partir du produit scalaire $\mathbf{g} \cdot \mathbf{r}$, de \mathbf{r}^2 et des constantes k et R_{ref} .

En pratique, R_{ref} peut être adapté os par os, soit d'après une éventuelle estimation de l'âge osseux, soit par la mesure du contour obtenu lors d'une première détection. On peut commencer par lui donner une valeur standard de l'ordre de $\frac{L_{\text{ref}}}{2}$, la moitié de la largeur moyenne des os des doigts. Une telle valeur correspond assez bien aux dimensions maximales d'un os du carpe.

Cette simplification est également pertinente pour tous les os dans leur forme "jeune". Même si ρ est dans ce cas nettement inférieur à 1, avec des valeurs de $f_k(\rho)$ relativement faibles, il y a alors peu de maxima parasites concurrents. Notons pourtant que le grand os, le scaphoïde et les deux épiphyses radiale et cubitale, s'allongent progressivement avec l'âge. La caractérisation par la distance n'est plus aussi homogène, mais le seuillage conserve suffisamment de bons maxima, à distance correcte du point de référence, pour fournir les germes du contour.

Le choix $k = 4$ apparaît comme un bon compromis. Le critère est alors assez efficace pour pénaliser les contours des os voisins et assez tolérant dans le cas où le point de référence serait excentré. Dans le cas où le point \mathbf{C} serait placé en dehors de l'os, la partie de contour la plus proche est invalidée par le signe du cosinus, mais \mathcal{K}_C répond positivement sur l'autre partie (voir la figure 4.2, détection par \mathbf{C}_A) et le contour peut être ensuite prolongé par les techniques usuelles. Un tel placement n'arrive que pour des os de petite taille et la distance reste dans des limites qui donnent des valeurs de $f_k(\rho)$ qui sont suffisantes pour la détection par un seuillage.

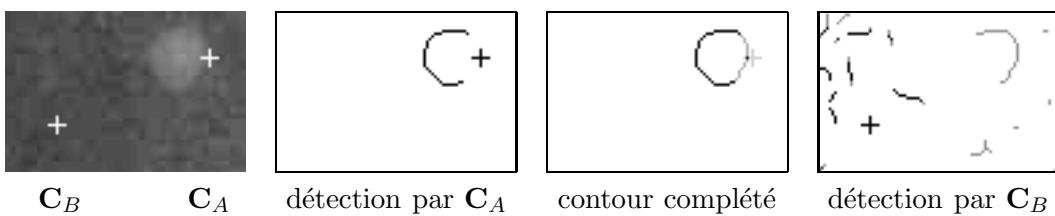


FIG. 4.2 – Cas de points de référence en dehors de l'os

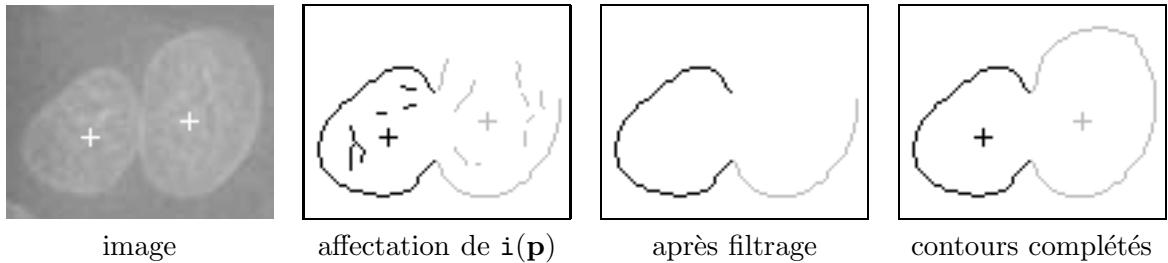


FIG. 4.3 – Cas d'une chaîne de maxima partagée par deux contours

Cela a aussi des inconvénients. Un point de référence peut générer une réponse positive sur le contour d'un os voisin, ainsi dans l'exemple de la figure 4.2, le point \mathbf{C}_B peut détecter le contour de l'os A . Quand un os est absent, comme l'os B qui devrait se trouver en bas à gauche dans l'exemple, une technique de seuillage adaptatif peut être mise en défaut par l'absence de "vrai" signal, comme on le voit sur la figure. Si, dans un tel cas, on demande au seuillage seul d'éviter les fausses détections, cela rend son réglage automatique assez délicat.

Il est donc préférable de traiter ensemble les deux points de référence et d'utiliser le fait que l'un va générer une réponse nettement plus forte que celle de l'autre sur le seul contour présent.

4.1.2 Utilisation conjointe de plusieurs points de référence

On considère maintenant n points de référence, un par os, et on définit autant de critères $\mathcal{K}_i(\mathbf{p})$, $i = 1, \dots, n$. Pour tout pixel \mathbf{p} de maximum du gradient, on peut alors considérer un entier $i(\mathbf{p})$ tel que $\mathcal{K}_{i(\mathbf{p})}(\mathbf{p}) = \max_{i \in \{1, \dots, n\}} \mathcal{K}_i(\mathbf{p})$. Cet entier indique l'os au contour duquel sera rattaché le pixel \mathbf{p} considéré.

Formellement, il n'y a pas unicité de $i(\mathbf{p})$. Étant donnés un pixel \mathbf{p} , la direction de son gradient et un point de référence, il est très facile de construire des lieux remarquables où placer un autre point de référence donnant la même réponse. En pratique pourtant, l'unicité est généralement vérifiée et avec une bonne marge, du fait de l'orientation et de la distance. Par régularité, le long d'une chaîne de maxima, la fonction $i(\mathbf{p})$ est alors constante par morceaux. Les pixels pour lesquels la définition de $i(\mathbf{p})$ peut être ambiguë, sont à la jonction de ces morceaux.

La figure 4.3 montre le cas le plus courant. Cela correspond à la construction remarquable où les deux points de référence sont symétriques par rapport à la droite portée par le gradient au pixel de rebroussement. En fait, avec les aléas de discréétisation, il n'y a pas nécessairement d'ambiguïté sur un pixel et, même dans ce cas, le choix du contour de rattachement pour ce pixel serait finalement peu important.

Il pourrait y avoir des problèmes d'affectation beaucoup plus ambigus dans le cas où les deux régions osseuses se chevauchent. Par exemple, si un point de référence se trouve dans l'intersection des deux régions et que le second point de référence se trouve à l'extérieur de l'une et de l'autre et, de même, si les deux points se trouvaient dans l'intersection.

En fait, le recouvrement des régions osseuses reste relativement limité et on considère que le placement des points de référence permet d'éviter les cas envisagés ci-dessus. La figure 4.4 montre un exemple correctement résolu avec un placement ordinaire des points de référence.

La figure 4.4 illustre un autre problème lié aux chevauchements. Au voisinage d'un point quadruple, à l'intersection théorique de deux contours, le gradient et les critères \mathcal{K}_i s'annulent. Les contours peuvent alors se trouver coupés en plusieurs fragments déconnectés. Or, la détection d'un fragment demande qu'il contienne suffisamment de pixels qui dépassent le seuil sur \mathcal{K}_i . Dans l'exemple, on a ainsi un défaut de détection du fait de deux intersections trop proches.

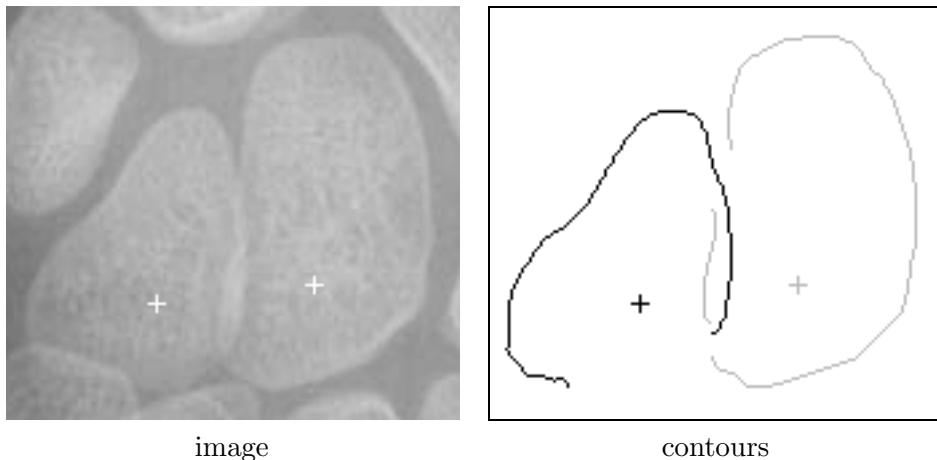


FIG. 4.4 – Cas du chevauchement des régions osseuses

4.1.3 Détails techniques

Les pixels qui sont validés par le seuillage sur \mathcal{K} sont assemblés pour former des chaînes, lors d'un balayage de l'image ligne par ligne. Notre algorithme est très voisin de celui vu en 2.4 pour construire les axes des os longs. La principale différence est que l'on relâche la contrainte de monotonie stricte en y , pour autoriser des suites horizontales de pixels. Une autre particularité importante est que l'on considère que les chaînes peuvent s'allonger par les deux bouts. Dans le cas d'un contour en forme de \cap ou \wedge , \sqcap , etc., l'algorithme commence donc par le haut et prolonge en parallèle les extrémités gauche et droite, au fil du balayage.

Tout ceci n'a un sens que si les pixels de maxima forment des structures curvilignes, sans ambiguïté, où tout pixel de contour n'a au plus que deux voisins au sens de la 8-connexité. La construction des maxima du gradient comporte un "amincissement", mais il peut y avoir des pixels qui génèrent des bifurcations¹. On supprime alors les quelques pixels strictement nécessaires pour obtenir la déconnexion des chaînes. En général, une des 3 chaînes incidentes est un parasite qui sera éliminé, et il est ensuite facile de refaire la continuité du contour, à condition que les bifurcations ne soient pas trop denses.

D'autre part, deux chaînes initialement distinctes seront regroupées lorsqu'elles se rejoignent sur une portion en forme de \cup ou \vee , \sqcup , etc. Ce cas se produit avec les deux branches d'un contour ouvert en haut ou si un contour comporte des concavités en partie haute. Dans ce dernier cas, on aura initié plusieurs chaînes de type \cap ou \wedge , \sqcap , etc., qui sont finalement fusionnées en parties basses. Le cas particulier du contour fermé revient à détecter que les deux extrémités d'une même chaîne se rejoignent. Le chaînage doit cependant respecter la cohérence de l'étiquetage par $i(p)$ et, dans un cas comme celui de la figure 4.3, on maintient donc la coupure en deux chaînes.

Ainsi, tout ensemble curvilinear de pixels d'un même contour peut être regroupé en une seule chaîne, directement au fil du balayage. C'est une différence notable par rapport aux algorithmes usuels qui génèrent de nombreuses chaînes sur certaines portions de contour, par exemple sur les bords gauches des \cap pour [Gir87]. On évite ainsi une seconde passe de fusion pour des cas dont la connexité et la structure curvilinear sont triviales.

La discussion concernant l'échelle de localisation (voir 2.5) reste pertinente pour les os du carpe. Le rapport signal/bruit est un peu moins bon que pour les métacarpiens mais il se compare à celui observé pour les premières ou secondes phalanges. Les contours ne sont pas rectilignes

¹Lors de la construction des axes des os longs, on savait éviter les bifurcations, en s'assurant d'un écart minimal entre deux pixels médians sur une même ligne de balayage, mais ce principe n'est pas adaptable ici.

mais les rayons de courbure sont généralement grands et l'erreur de caractérisation est assez faible, sauf pour les petits os et en quelques points singuliers. En fait, cela a peu d'influence pour des applications à l'âge osseux car l'évolution des formes est très importante. De plus, il s'agit d'une erreur systématique et la comparaison de mesures permet d'en faire abstraction.

L'échelle de localisation ne convient pas pour la détection, même si celle-ci est guidée par les points de référence, car il y a trop de maxima parasites et de bifurcations. L'échelle de détection utilisée pour les doigts est à l'inverse trop grossière et elle engendre la fusion des contours proches. Ce problème ne se posait pas pour les os longs, relativement plus espacés, et on privilégiait alors une détection sans apport d'information. Ici, on doit séparer des contours relativement plus proches mais la détection est guidée par les points de référence.

Une échelle de détection double de celle de localisation, soit $\alpha = \frac{\alpha_{loc}}{2}$, donne dans l'ensemble de bons résultats et la propagation entre deux échelles dans un rapport de 2 est alors relativement directe. Il n'est pas utile d'adapter ces échelles ; une amélioration de la détection passe plutôt par le recalage des points de référence et l'ajustement des valeurs de R_{ref} .

4.1.4 Remarques théoriques

Pour alléger les notations, on se replace dans le contexte d'un seul point de référence. On se place aussi dans le plan continu, un fragment de contours est une courbe et \mathbf{p} désigne un point de cette courbe. Nous noterons $\theta(\mathbf{p})$ l'angle polaire de \mathbf{p} par rapport au point de référence et pour une direction de référence arbitraire. Nous considérons aussi un paramétrage curviligne $\mathbf{p}(s)$ suffisamment régulier.

On s'intéresse alors à la variation angulaire : $\frac{d}{ds}\theta(\mathbf{p}(s)) = \frac{d\mathbf{p}}{ds} \cdot \nabla\theta(\mathbf{p}(s))$.

Si on considère que le contour est isophote, c'est-à-dire qu'il suit une ligne de niveau de l'image, sa tangente en tout point $\mathbf{p}(s)$ coïncide avec \mathbf{g}_\perp , le vecteur normal au gradient de l'image en ce point. On peut ainsi écrire : $\frac{d\mathbf{p}}{ds} = \lambda(s)\mathbf{g}_\perp$, avec $\lambda(s)$ qui est non nul et donc de signe constant², par régularité.

D'autre part, on peut toujours écrire $\nabla\theta(\mathbf{p}) = \frac{\mathbf{r}_\perp}{|\mathbf{r}|^2}$.

En conséquence, $\frac{d}{ds}\theta(\mathbf{p}(s)) = \frac{\lambda(s)}{|\mathbf{r}|^2}\mathbf{g}_\perp \cdot \mathbf{r}_\perp$ et comme, trivialement, $\mathbf{g}_\perp \cdot \mathbf{r}_\perp = \mathbf{g} \cdot \mathbf{r}$, on en conclut que si le fragment de contour vérifie $\cos(\mathbf{g}, \mathbf{r}) < 0$ en tout point, alors $\theta(\mathbf{p}(s))$ est strictement monotone sur ce segment.

Cette propriété est importante pour la mise en œuvre. En effet, elle permet de dire qu'un rayon issu du point de référence ne coupe le contour qu'en un seul point, ce qui valide le principe d'une paramétrisation angulaire du contour. Notamment, si un rayon rencontre deux maxima de gradient, l'un au moins est un parasite et on est en droit de le supprimer, en se basant sur \mathcal{K} , ses composantes $|\mathbf{g}|$, $\cos(\mathbf{g}, \mathbf{r})$ et $f(|\mathbf{r}|)$, ou tout autre critère de cohérence de son voisinage ou de consistance plus globale de sa chaîne. Pour cela, il n'est pas utile de se livrer à un balayage angulaire explicite et dense sur les pixels qui serait de mise en œuvre assez délicate. On peut se contenter de tester les extrémités des chaînes pour détecter si elles sont en conflit sur un certain secteur et procéder alors par échantillonnage du critère sur ce secteur.

Cette propriété permet aussi de donner une définition simple de la "longueur angulaire" d'un fragment de contour, notion qui sera beaucoup plus pertinente que la longueur curviligne pour en exprimer la consistance. Le calcul de cette mesure n'utilise que les valeurs angulaires des deux extrémités de la chaîne.

L'utilisation d'un point de référence, même si son placement est approximatif, permet donc d'écartier les parasites et de caractériser une portion relativement importante du contour. D'autre

²Le signe de λ ne dépend que du sens de la paramétrisation en s .

part, étant donné un contour, on peut se demander quel est le domaine potentiel où placer un point de référence qui le caractérise positivement dans son ensemble.

D'un point de vue formel, la contrainte $\cos(\mathbf{g}, \mathbf{r}) < 0$ en tout point du contour détermine le convexe intérieur au contour, c'est-à-dire l'intersection de tous les demi-plans tangents et du côté de \mathbf{g} . Le plus souvent, le contour est convexe et alors tout point à l'intérieur convient. Dans les autres cas qui nous concernent, le contour reste "étoilé" et il existe une partie convexe non vide, intérieure au contour. En pratique, cette région est relativement vaste et le plus difficile reste de placer le point de référence à l'intérieur d'un os très jeune, certes convexe, mais très petit. Mais dans ce dernier cas, même si le point de référence \mathbf{C} est placé légèrement en dehors, la caractérisation $\mathcal{K}_\mathbf{C}$ permet encore une détection suffisante et l'identification du contour.

Dans ce qui suit, on va s'intéresser plus en détail au problème du placement des points de référence. Avant de considérer des techniques de placement de points a priori, ce sera l'objet de 4.2, on va regarder comment il est possible d'améliorer le placement d'un tel point uniquement à partir d'informations extraites de l'image. Le but est double. Déjà, si besoin, le recalage d'un point de référence permet de mieux caractériser le contour de l'os correspondant, donc d'en améliorer itérativement la détection et la localisation. D'autre part, cela permet de déterminer par régression les paramètres du modèle de placement initial, de recaler certains points à partir d'autres ou plus généralement de pouvoir initialiser géométriquement divers modèles sur l'image.

Il faut écarter toute de suite l'idée de rechercher un point de référence comme un maximum du niveau de gris, même en partant d'une position relativement bien approchée. Déjà le lissage qui permettrait d'effacer juste ce qu'il faut des détails de structure n'est pas évident à déterminer. Ensuite, les zones les plus calcifiées sont souvent à la périphérie de l'os. Enfin, on peut même trouver, dans certaines images, des chemins d'intensité strictement croissante qui passent d'un os à l'autre. Les difficultés seraient comparables en procédant sur un critère de symétrie.

On va voir maintenant comment se baser sur une connaissance partielle du contour pour estimer une nouvelle position du point de référence.

4.1.5 Recalage d'un point de référence

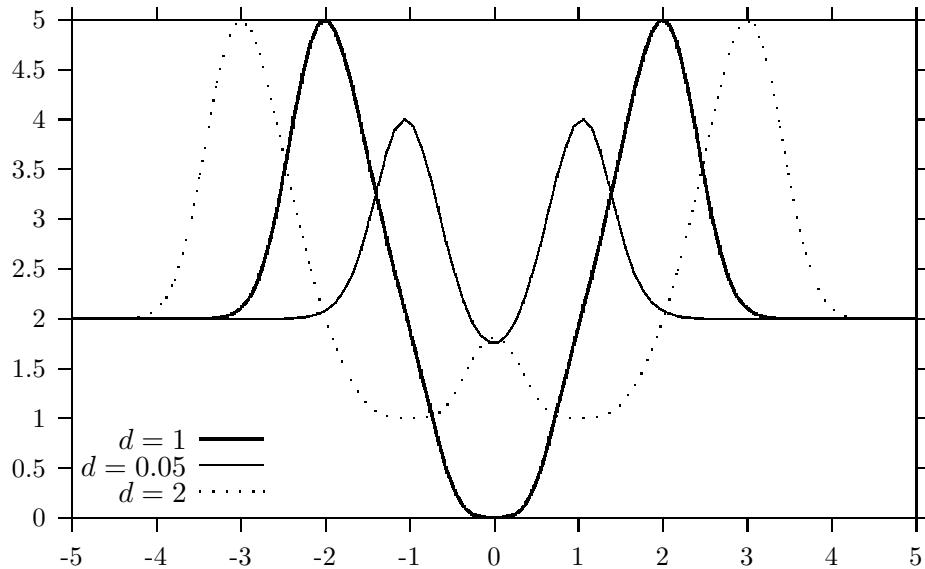
Dans ce qui suit, on notera \mathcal{C} le contour connu discréétisé et $|\mathcal{C}|$ son cardinal (le nombre de pixels). On va se placer dans le cadre classique du problème aux moindres carrés, en considérant la minimisation de critères de la forme générale : $E^2(\mathbf{X}) = \sum_{\mathbf{p} \in \mathcal{C}} e^2(\mathbf{X}, \mathbf{p})$, où $e(\mathbf{X}, \mathbf{p})$ est un terme relatif à la caractérisation de chaque pixel \mathbf{p} du contour par le point de référence \mathbf{X} qui est ici la variable.

Nous allons maintenant nous intéresser à des expressions particulières pour $e(\mathbf{X}, \mathbf{p})$. La principale difficulté est de convenir d'un critère de qualité réaliste pour un point de référence. On notera que le critère est défini sur le contour \mathcal{C} connu a priori, alors qu'il s'agit de l'optimiser pour le contour que le point de référence solution doit caractériser a posteriori. Un processus d'optimisation simultanée sur \mathbf{X} et \mathcal{C} est hors de propos, du fait de sa lourdeur et de la difficulté à établir des conditions de convergence. En particulier, une minimisation sur un ensemble variable aurait tendance à rejeter les pixels extrêmaux, alors que l'on peut penser que ce sont ces pixels qui comptent le plus pour "tirer" un point de référence initialement excentré.

Nous avons envisagé quelques méthodes de recalage basées sur des formulations relativement intuitives de $e(\mathbf{X}, \mathbf{p})$; nous explicitons ci-dessous celles qui nous ont semblé les plus pertinentes.

a) Une première méthode, très simple, revient à minimiser $E_r^2(\mathbf{X}) = \sum_{\mathbf{p} \in \mathcal{C}} \mathbf{r}^2$, dont la solution se calcule directement comme le barycentre des pixels de contour, soit $\mathbf{G} = \frac{1}{|\mathcal{C}|} \sum_{\mathbf{p} \in \mathcal{C}} \mathbf{p}$.

On notera que ce critère ne tient pas compte du gradient (en dehors de la caractérisation initiale de \mathcal{C}) et très peu de la forme du contour. On verra que le résultat, en tant que point de référence, est effectivement de qualité médiocre en général.

FIG. 4.5 – Variations de E_K^2 pour des dipôles

Pourtant, le point \mathbf{G} est utile à plusieurs titres. Déjà, il permet d'apprécier la longueur angulaire du contour, de manière généralement plus juste qu'en utilisant le point a priori qui est parfois placé à l'extérieur du contour³. D'autre part, $\sqrt{\frac{E_K^2(\mathbf{G})}{|\mathcal{C}|}}$ est une estimation du rayon moyen effectif de l'os. On pourra l'utiliser comme valeur de R_{ref} , en particulier pour la localisation, à condition qu'elle soit évaluée sur un contour déjà relativement complet. Cette mesure peut aussi être utile pour des applications de diagnostic, avec éventuellement une analyse plus complète des moments du contour.

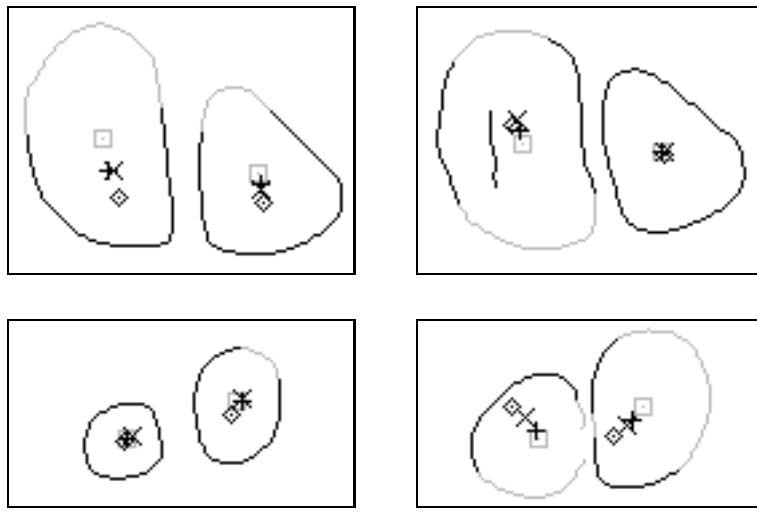
b) La seconde méthode, part de l'idée de chercher le point de référence \mathbf{X} qui maximise le critère $\mathcal{K}_{\mathbf{X}}$ sur l'ensemble du contour connu. Comme \mathcal{K} est borné (car \mathbf{g} , le gradient régularisé de l'image, est borné), cela revient à minimiser $E_K^2(\mathbf{X}) = \sum_{\mathbf{p} \in \mathcal{C}} (K - \mathcal{K}_{\mathbf{X}}(\mathbf{p}))^2$. Les dérivées partielles d'ordres 1 et 2 de E_K^2 selon les coordonnées de \mathbf{X} sont définies et on peut utiliser des techniques de minimisation relativement robustes comme la méthode de Levenberg-Marquardt [PTVF92a]. Néanmoins, cela reste une recherche itérative avec des évaluations coûteuses.

On peut étudier les conditions de la convergence sur un cas trivial, celui d'un dipôle parfait, c'est-à-dire un contour composé de deux points symétriques par rapport à l'origine, à la distance d de celle-ci, avec \mathbf{g} de module 1 et pointant vers l'origine. La figure 4.5 montre les variations de E_K^2 le long de la droite passant par ces deux points et l'origine. Les abscisses x sont normalisées pour avoir $R_{\text{ref}} = 1$ et on a tracé $E_K^2(x)$ pour trois valeurs significatives de d .

Le rayon de convergence décroît avec d (le rayon de l'os), mais on peut vérifier qu'il tend vers 1 quand d tend vers 0. Autrement dit, une convergence est toujours assurée si le point de référence est initialisé à une distance inférieure à R_{ref} du centre, c'est normalement le cas en prenant \mathbf{G} . On peut aussi vérifier que la valeur limite pour déterminer une solution unique est $d = 1$. Autrement dit, les points de contours ne doivent pas être à une distance supérieure à R_{ref} du centre. Ce n'est pas toujours vérifié et la solution \mathbf{K} n'est alors pas le centre de symétrie.

En pratique, avec un contour plus ou moins complet au lieu d'un dipôle, cela peut s'observer dans quelques rares exemples. Il est donc important de pouvoir estimer correctement R_{ref} .

³En théorie, \mathbf{G} pourrait aussi se trouver à l'extérieur du contour, mais ce cas suppose des formes concaves complètement irréalistes ici.



légende : $\diamond = \mathbf{G}$ $\times = \mathbf{K}$ $+=\mathbf{H}$ $\square =$ centre pour le contour complet

FIG. 4.6 – Comparaison des diverses méthodes de recalage d'un point de référence

c) Une troisième méthode, hybride, consiste à construire, pour chaque pixel \mathbf{p} du contour, son point de référence optimal $\hat{\mathbf{p}}$ qui, par définition de \mathcal{K} , est situé dans la direction du gradient à la distance R_{ref} , soit $\hat{\mathbf{p}} = \mathbf{p} + R_{\text{ref}} \frac{\mathbf{g}}{\|\mathbf{g}\|}$. On peut alors considérer le point \mathbf{H} , défini comme le barycentre de tous ces points $\hat{\mathbf{p}}$ et qui minimise par conséquent $E_h(\mathbf{X}) = \sum_{\mathbf{p} \in \mathcal{X}} (\hat{\mathbf{p}} - \mathbf{X})^2$.

Cette méthode est nettement plus légère que la précédente. Les deux tiennent compte de la direction du gradient et donc, avec l'approximation usuelle, de la tangente au contour et ainsi indirectement de sa forme. La prise en compte de la distance est très différente. Mais, intuitivement, on voit que par ces deux méthodes, un pixel de contour trop près peut "pousser" le point de référence et qu'un pixel trop loin peut le "tirer". Globalement, l'un ou l'autre de ces phénomènes sera plutôt en action selon le rapport entre R_{ref} et le rayon effectif.

Une différence notable est que la dernière méthode n'utilise pas le module du gradient, mais on peut considérer que celui-ci est relativement constant le long d'un contour et donc peu déterminant. Dans ce cas, la direction du gradient est la normale au contour, on a une définition purement géométrique et le point \mathbf{H} ne dépend que de la forme du contour. Enfin, par construction, il y a toujours unicité de la solution.

Toutes ces méthodes sont influencées par le placement du point de référence initial, dans la mesure où celui-ci détermine les pixels de contour qui sont détectés. La figure 4.6 permet de voir ce qu'il en est sur des cas concrets. Nous avons rassemblé dans cette figure les cas les plus typiques, notamment des contours plus ou moins fermés, voire en plusieurs morceaux, et le cas d'un fragment parasite. Pour amplifier les effets dans un but d'illustration, nous n'avons pas utilisé les contours obtenus après complétion par connexité (en gris dans la figure) car ceux-ci sont toujours fermés ou presque. Les contours \mathcal{C} utilisés ici pour calculer les points \mathbf{G} , \mathbf{K} et \mathbf{H} sont ceux représentés en noir et qui correspondent normalement à une étape intermédiaire.

On remarquera que, s'il existe un axe de symétrie du contour, au sens approprié à leur définition respective, alors les trois méthodes vont placer chacune un point sur cet axe (en supposant que les conditions d'unicité pour \mathbf{K} sont satisfaites). S'il y a plusieurs axes de symétrie, par exemple dans le cas du "triangle" figure 4.6, alors les points sont tous à l'intersection de ces axes. C'est donc le cas aussi, en particulier, pour un contour circulaire.

Plus généralement, on voit que dans le cas d'un contour fermé, les trois méthodes donnent des points très voisins du "centre". À l'inverse, lorsque le contour est notablement incomplet, les écarts par rapport à ce centre sont très nets et d'amplitude différente selon la méthode.

Le point **G**, marqué \diamond , est alors celui qui convient le moins comme point de référence. Les points **K**, marqué \times , et **H**, marqué $+$, sont moins affectés et sont, en général, très voisins. Pourtant **H** semble toujours légèrement mieux placé, vers “le” centre calculé sur le contour complet et marqué \square . Il n'y a pas d'interprétation évidente de l'influence respective de $|\mathbf{r}|$ dans les deux critères, notamment l'allongement d'un contour ne semble pas avoir d'effet déterminant.

La méthode qui s'impose donc, par son résultat et par sa facilité, est celle du calcul de **H**, mais on l'intègre dans un processus plus robuste. On calcule aussi **G** et, si les deux points diffèrent de trop, on recommence toute la phase détection, en se basant sur ce premier **H** comme point de référence. On obtient ainsi de nouveaux contours et un nouveau point **H**. En pratique, il n'est pas utile d'itérer plus. On dispose alors d'un contour relativement complet, avec une estimation correcte du centre de l'os et on peut aussi envisager d'adapter R_{ref} . Tout ceci contribue à une bonne caractérisation qui permet de résoudre les ambiguïtés de bifurcation dans le calcul à l'échelle de localisation. Cela sert aussi pour l'apprentissage d'un modèle de placement a priori.

4.2 Placement a priori des points de référence

Le problème de l'identification automatique des os du carpe a été peu traité. Les quelques travaux existants supposent des mains très jeunes car l'identification se base sur la détection préalable de ces os, par des techniques classiques, orientées contours ou régions. Cela demande que chaque os soit bien séparé de ses voisins. Or, on peut observer une superposition du trapèze et du trapézoïde dès parfois 4 ans (alors que sur d'autres images, pour le même âge, ces deux os sont encore à peine visibles) et le contact entre grand os et os crochu se fait généralement à partir de 5 ou 6 ans.

Ainsi, dans [PKKH93], la région du carpe est localisée grossièrement par un balayage puis les os sont détectés par une suite de seuillages et d'opérations morphologiques. Sur des mains jeunes, le grand os est facilement identifiable par sa position et ses dimensions. Les autres os sont ensuite identifiés par leur position dans un certain secteur angulaire autour du grand os. Quelques mesures élémentaires de forme et d'intensité sont utilisées comme signature pour régler les cas d'ambiguïté, c'est-à-dire quand plusieurs os se trouvent dans le même secteur. Dans [ZGL07], ce même principe est amélioré par une détection des contours qui utilise un lissage de type Perona et Malik et le détecteur de Canny. Dans [HJCT07], qui s'inspire aussi fortement de [PKKH93], l'identification se fait de proche en proche à partir de l'os crochu, ce qui suppose un ordre d'ossification particulier. Ces méthodes ne cherchent pas à distinguer le trapèze et le trapézoïde.

Dans [RCN⁺95], l'identification et la localisation sont réalisées à une échelle grossière par un réseau de neurones rétinien qui peut s'adapter aux cas de superposition entre les os voisins, mais qui demande une base d'apprentissage très conséquente pour incorporer les diverses combinaisons des stades de développements de chaque os.

Dans [SVV90], l'identification et la détection des contours sont réalisées conjointement à l'aide d'un atlas, très inspiré de TW2, qui pour chaque os et chaque stade, décrit la forme, l'apparence et des relations géométriques avec les os voisins. La résolution est guidée par un ordre particulier et un certain nombre de règles de cohérence sur les stades de développement.

Nous venons de voir que le placement approprié de points de référence permettait de résoudre l'identification et la détection des contours, sans a priori de forme. Nous avons vu au chapitre 2 une méthode qui permet de localiser les extrémités des phalanges et des métacarpiens le long de leur axe respectif. Maintenant, nous allons regarder comment des données de position, puis de longueur et d'orientation que l'on peut ainsi acquérir sur une image nous permettent de placer les points de référence. Il faut pour cela un modèle géométrique adapté aux données dont on dispose. La variabilité biométrique est relativement importante et des indicateurs de qualité calculés sur les mesures ne permettent de détecter que les erreurs techniques les plus grossières.

Nous avons vu au chapitre 3 que la localisation des os longs est en général très fiable, surtout pour les phalanges proximales et médianes. Les résultats sont un peu moins bons en ce qui concerne l'extrémité proximale des métacarpiens et cela constitue un problème quand on s'intéresse au carpe. Nous avons cherché tout d'abord à construire les points de référence comme des combinaisons linéaires des points d'extrémité localisés sur les axes des os longs. Le choix initial d'un modèle linéaire était motivé par la facilité de réalisation d'une phase d'apprentissage automatique, par une régression sur un grand nombre d'images. En fait, il s'avère qu'une telle solution n'est pas envisageable, même si l'on dispose également d'une localisation des extrémités des os de l'avant-bras. Il n'y a pas de modèle linéaire global qui puisse s'adapter à l'amplitude des variations morphométriques, aux erreurs techniques de localisation des métacarpiens et aux libertés de rotation du poignet.

Nous avons ensuite défini un modèle qui incorpore des composantes non linéaires, il s'agit en l'occurrence de longueurs et de directions. Avec une extraction robuste de ces données, le problème de la localisation des points de référence est alors convenablement déterminé.

4.2.1 Définition d'un modèle linéaire

Pour chaque phalange et métacarpien, nous disposons des coordonnées des deux points qui sont ses extrémités le long de son axe. Ces données sont extraites de l'image par la technique développée au chapitre 2. Ainsi, pour une image i , on peut former un vecteur \mathbf{U}_i qui regroupe les coordonnées de tous⁴ ces points. Toujours pour cette image et pour chaque os j du carpe, on considère le point de référence \mathbf{C}_i^j .

Un modèle linéaire pour l'os j est alors défini par une matrice \mathbf{M}^j . Une fois que cette matrice est connue, l'utilisation du modèle consiste à calculer $\mathbf{C}_i^j = \mathbf{M}^j \cdot \mathbf{U}_i$ pour obtenir le point de référence a priori pour l'os j dans une image i quelconque. En mode d'apprentissage, il s'agit de déterminer \mathbf{M}^j qui minimise $\sum_{i \in \mathcal{I}_j} (\mathbf{C}_i^j - \mathbf{M}^j \cdot \mathbf{U}_i)^2$, sur un ensemble d'images \mathcal{I}_j pour lesquelles \mathbf{U}_i et \mathbf{C}_i^j sont connus. C'est un problème classique. Nous allons nous attarder un peu sur les problèmes concernant l'acquisition et la pertinence de chaque \mathbf{C}_i^j .

Pour aboutir à un apprentissage sur un nombre d'images relativement important, nous avons procédé en deux étapes. D'abord, nous avons construit un modèle rudimentaire \mathbf{M}_0^j , en partant d'un petit sous-ensemble d'images \mathcal{I}_{j0} . Cela s'est fait de manière empirique, par régression, à partir de points de référence que l'on aura localisés manuellement, et aussi, plus directement, en "inventant" des relations linéaires qui semblent naturelles. Cette phase est donc relativement fastidieuse, qu'il s'agisse de relever des points de référence ou de vérifier visuellement que les combinaisons linéaires choisies donnent des résultats acceptables.

Nous nous sommes ensuite servis de ce premier modèle \mathbf{M}_0^j pour construire une base \mathcal{I}_j plus importante. Pour chaque image i destinée à former cette base, on calcule donc un point a priori $\mathbf{C}_{0i}^j = \mathbf{M}_0^j \cdot \mathbf{U}_i$. Ce point peut être relativement mal placé, mais la technique de recalage vue en 4.1.5 permet de se recentrer correctement sur l'image, en supposant ici que l'os j est bien détectable. Les limites d'un placement acceptable sont celles qui déterminent une identification correcte des contours qui servent au recalage. On obtient ainsi un nouveau point \mathbf{C}_i^j pour chaque image $i \in \mathcal{I}_j$. Cette phase est entièrement automatique et permet de construire assez facilement une base \mathcal{I}_j statistiquement suffisante pour calculer le modèle définitif \mathbf{M}^j .

Le calcul d'une matrice \mathbf{M}^j qui minimise $\sum_{i \in \mathcal{I}_j} (\mathbf{C}_i^j - \mathbf{M}^j \cdot \mathbf{U}_i)^2$ est un problème classique. Outre sa robustesse intrinsèque vis-à-vis des indéterminations, une technique de résolution par la décomposition en valeurs singulières [PTVF92b] apporte aussi de nombreuses informations sur la pertinence du modèle, la mise en évidence des redondances, les covariances, etc.

⁴En pratique, on convient facilement qu'il est possible d'écartier systématiquement des points non déterminants, à commencer par ceux des phalanges distales.

On utilise en fait une base \mathcal{I} commune pour tous les os du carpe car il serait fastidieux de gérer une base différente pour chaque os. Une part importante des calculs, la décomposition en valeurs singulières d'une grosse matrice dont les colonnes sont les vecteurs \mathbf{U}_i , $i \in \mathcal{I}$, est alors commune à tous les os j .

On doit alors tenir compte du fait que certains os seront absents de certaines images. D'autre part, pour un os j donné, ce sont les images dans lesquelles il est au début de son développement, qui fournissent les données les plus pertinentes pour l'estimation du modèle. Il s'agira, en phase d'utilisation, de pouvoir placer un point de référence très près d'un point de naissance de l'ossification, alors que dans le cas d'un os bien développé, l'obtention d'un placement acceptable à l'intérieur des contours effectifs est relativement facile. De plus, le centre d'un os bien développé ne correspond pas nécessairement au point de naissance de l'ossification, car la croissance n'est pas isotrope.

Ainsi, on cherche plutôt \mathbf{M}^j qui minimise $\sum_{i \in \mathcal{I}} w_i^j (\mathbf{C}_i^j - \mathbf{M}^j \cdot \mathbf{U}_i)^2$, un problème tout aussi classique qui se résout comme le précédent. Des valeurs pour les pondérations w_i^j (pertinence de l'image i pour l'estimation du modèle pour l'os j) sont déterminées lors du recalage qui permet de passer de \mathbf{C}_{0i}^j à \mathbf{C}_i^j . La valeur est nulle si l'os n'est pas détecté et, autrement, on prend une valeur positive, fonction décroissante du rayon effectif de l'os que l'on connaît alors. On pourrait ensuite envisager d'adapter ces valeurs en fonction des scores de segmentation des os longs (voir chapitre 3) et même de les affiner progressivement dans un processus d'estimation robuste [RL87]. Nous allons voir maintenant les limites de cette approche.

4.2.2 Défauts d'un modèle linéaire

À titre d'exemple, un problème relativement fréquent est celui de la mauvaise segmentation du second métacarpien qui englobe le trapézoïde (voir notamment la figure 4.7, image de gauche). Cela représente une erreur de l'ordre de 14 % de la longueur de ce métacarpien. Ce n'est qu'un peu plus que la limite du détectable, au regard de la variabilité biométrique, mais c'est déjà une erreur très importante, au regard des dimensions du carpe. De plus, on a une forte accumulation statistique autour de cet écart de 14 % et les mesures du second métacarpien tendent à former une loi bimodale.

Cela rend déjà la phase d'apprentissage un peu plus délicate. Les techniques de moindres carrés pondérés et d'estimation robustes n'utilisent que des pondérations du type des coefficients w_i^j introduits ci-dessus. C'est-à-dire, qu'il s'agit de pondérer globalement, pour l'os j du carpe et pour une image i , l'influence conjointe des mesures \mathbf{U}_i et du centre \mathbf{C}_i^j sur l'estimation des paramètres du modèle. Nous n'avons pas le moyen d'indiquer a priori une méfiance sur une composante particulière u_{ki} de \mathbf{U}_i . Dans notre exemple, ce serait ainsi la position de l'extrémité proximale du second métacarpien qui devrait être disqualifiée.

Par principe, c'est la technique d'estimation robuste qui décidera de la valeur de w_i^j en fonction de l'erreur $(\mathbf{C}_i^j - \mathbf{M}^j \cdot \mathbf{U}_i)^2$, autrement dit, en fonction de l'importance attribuée en général (termes m_{lk}^j de \mathbf{M}^j) à la k -ième composante dans le modèle pour l'os j . En fait, cela va fonctionner correctement, tant que l'on reste dans le cadre de principe de l'estimation robuste, c'est-à-dire notamment que la distribution des erreurs $(\mathbf{C}_i^j - \mathbf{M}^j \cdot \mathbf{U}_i)^2$, $\forall i \in \mathcal{I}$ reste unimodale.

La solution pour rester dans le domaine de bon fonctionnement est alors d'éjecter a priori de la base d'apprentissage \mathcal{I} toutes les images qui donnent lieu à des scores de segmentation inférieurs à 90. Si on ne tient pas compte des phalanges distales dans le modèle, on dispose facilement d'assez d'images satisfaisantes pour estimer correctement un modèle linéaire pour chaque os. Mais ce modèle est alors fortement calqué sur la normalité biométrique car on a écarté de nombreux cas de segmentation correcte mais présentant une forte déviation naturelle.

Ensuite, en phase d'utilisation, il s'avère que ces modèles sont inadaptés sur de nombreuses images. Les décalages des extrémités des os longs donnent des erreurs de position des points de

référence estimés qui peuvent facilement conduire à des erreurs d'identification entre deux os. En particulier, en revenant à notre exemple, une mauvaise segmentation du second métacarpien conduit à une erreur importante de localisation du trapézoïde. En effet, le modèle du trapézoïde, estimé uniquement sur des cas favorables, utilise largement ce métacarpien.

L'approche linéaire peut sembler un peu plus appropriée si on se donne la localisation des os de l'avant-bras. La technique employée pour détecter et localiser les os des doigts, peut aussi s'appliquer au radius et au cubitus, avec une paramétrisation qui tient compte des largeurs d'appariement respectives. On obtient ainsi la localisation de l'extrémité distale de ces deux os. Cette localisation est certes relativement imprécise, avec des problèmes similaires à ceux observés sur les extrémités proximales des métacarpiens, mais cela permet d'encadrer le carpe. La localisation longitudinale revient ainsi plus à une interpolation qu'à une extrapolation, ce qui contribue à limiter l'influence des erreurs sur les métacarpiens. Compte tenu des possibilités de recalage des points de référence, on obtient effectivement un peu plus souvent des résultats acceptables, mais il reste de nombreux cas d'échec.

Le problème le plus notable est lié à la liberté de rotation du poignet qui est un paramètre caché. Les os du carpe sont fortement liés au référentiel donné par les métacarpiens et ceux de la couche proximale, scaphoïde, semi-lunaire et pyramidal, glissent sur les épiphyses radiale et cubitale. On notera pour illustrer ce problème que, selon la rotation du poignet, l'axe du troisième métacarpien peut pointer, soit vers le radius, soit vers le cubitus, avec une amplitude qui peut le faire passer en dehors de l'épiphyse correspondante. Dans un tel cas, le décalage latéral des points de référence estimés en utilisant ces épiphyses peut facilement conduire à une erreur d'identification pour un os du carpe. On notera aussi que les épiphyses radiale et cubitale ne peuvent donc pas être localisées comme les os du carpe.

Cela revient à dire que les os de l'avant-bras contribuent de manière non linéaire et qu'ils n'apportent que des informations de distance et pas de position. De toute façon, le cadrage usuel d'une image est généralement trop serré et, dans ce cas, on ne dispose pas d'une longueur d'avant-bras suffisante pour la mise en évidence d'une fraction d'axe. Cette solution n'est donc pas applicable sur des images au "tout venant".

Par contre, les métacarpiens et les phalanges peuvent apporter des informations de longueur et de direction pertinentes pour une localisation. Mais les axes des métacarpiens ne sont pas parallèles et les axes d'une phalange et du métacarpien d'un même doigt ne sont pas plus alignés. Une opération comme le report de la longueur d'un os dans la direction d'un autre n'entre pas dans le cadre d'un modèle linéaire. Sur un tel modèle, il est ainsi exclu que les mécanismes de résolution parviennent à construire implicitement une quelconque forme d'estimation robuste des longueurs et des directions qui permettrait d'atténuer l'anormalité ou une erreur de localisation d'un os.

Cette expérimentation aura toutefois permis d'identifier les quelques relations géométriques qui sont essentielles. La méthodologie d'estimation d'un modèle de la forme $\mathbf{C}_i^j = \mathbf{M}^j \cdot \mathbf{U}_i$ reste pertinente, mais elle demande d'incorporer dans le vecteur \mathbf{U}_i des données de longueur et d'orientation acquises de manière relativement robuste. Cette démarche que nous allons maintenant développer, relève d'un schéma classique de généralisation de la régression linéaire, au même titre, par exemple, que les régressions polynomiales. Toute la difficulté d'une telle démarche réside alors dans le choix des données pertinentes. Nous avons choisi d'en réduire le nombre, en retenant les plus évidentes et en se basant sur la redondance entre plusieurs mesures élémentaires pour acquérir de la robustesse.

4.2.3 Solution non linéaire

Le placement du point de référence pour un os du carpe consiste alors à placer un point de coordonnées fixes dans un référentiel particulier à cet os et qui est calé sur l'image à partir

des données obtenues sur les phalanges et les métacarpiens. Le référentiel pour l'os j dans l'image i est déterminé par une origine \mathbf{o}_i^j et deux directions perpendiculaires, avec des unités particulières, soit deux vecteurs \mathbf{u}_i^j et \mathbf{v}_i^j . Le point de référence pour cet os s'exprime alors comme $\mathbf{o}_i^j + \delta^j \mathbf{u}_i^j + \epsilon^j \mathbf{v}_i^j$, où δ^j et ϵ^j sont les deux paramètres du modèle pour l'os j . Il est relativement facile de leur donner des valeurs empiriques δ_0^j et ϵ_0^j pour initier le processus d'apprentissage.

Pour chaque construction, on se base essentiellement sur un seul axe, celui qui semble le plus naturel pour le point de référence considéré. Il s'agit du deuxième axe pour le trapèze, le trapézoïde et le scaphoïde, du troisième pour le grand os, du quatrième pour l'os crochu, le pyramidal et le semi-lunaire. On évite d'utiliser les deux axes extrêmes qui sont généralement moins précis. On ne place pas de point pour l'os pisiforme qui se confond avec le pyramidal.

L'origine \mathbf{o}_i^j est le point milieu estimé du métacarpien de l'axe correspondant. En se plaçant au milieu du métacarpien, on évite notamment les erreurs éventuelles de localisation de l'axe en bout d'os. Le point \mathbf{o}_i^j est pris sur la courbe de régression polynomiale de l'axe qui est généralement bien localisée sur cette portion, car on détecte facilement les contours longitudinaux de la diaphyse du métacarpien. Pour le placement longitudinal, on n'utilise pas directement l'extrémité proximale du métacarpien qui peut être sujette à erreur. Il est plus stable de reporter la moitié de la longueur du métacarpien à partir de l'extrémité distale. Cette dernière est en général très précise et on va utiliser une estimation robuste de la longueur.

En reprenant les notations de la table 2.1, page 76, L_{kl} est la longueur mesurée de l'os de rang k de l'axe l et, par définition des coefficients λ_{kl} et μ_l , la longueur cumulée des os des 5 axes, estimée d'après ce seul os, peut s'exprimer comme $\bar{L}_{kl} = \frac{L_{kl}}{\lambda_{kl}\mu_l}$. Dans le cas d'une main "normale", cette longueur serait indépendante du choix de l'os, autrement dit, de k et l . En réalité, on obtient autant de valeurs différentes que de choix de k et l et il s'agit d'en tirer une mesure suffisamment robuste. En l'occurrence, on considère la valeur médiane de \bar{L}_{02} , \bar{L}_{03} et \bar{L}_{04} obtenues pour les trois métacarpiens centraux. Du seul fait de la variabilité naturelle, ces trois valeurs sont en général très voisines. De plus, cette estimation est très robuste vis-à-vis d'une erreur de localisation de l'extrémité proximale d'un métacarpien. On a vu que de telles erreurs sont indépendantes d'un axe à l'autre et que leur co-occurrence est très rare. Il ne semble pas pertinent de faire intervenir d'autres os. Notamment, les phalanges proximales sont mieux localisées et donnent des mesures fiables mais les rapports de longueurs effectives entre phalanges et métacarpiens font l'objet de variations assez importantes par rapport à la normale.

Autrement dit, pour placer le point \mathbf{o}_i^j , on prend comme longueur de référence la médiane des longueurs des trois métacarpiens centraux, correction faite des tailles normales relatives. Cela revient généralement à prendre une valeur très voisine de celle du métacarpien considéré, sauf si celui-ci a fait l'objet d'une grosse erreur de segmentation.

La direction de \mathbf{u}_i^j est alors donnée par la tangente à la courbe de régression polynomiale au point \mathbf{o}_i^j . La courbure de l'axe est faible, et si on a commis une erreur dans la localisation de ce point, cela ne provoque pas de variation significative de la direction. Ainsi, en simulant des déplacements relativement importants du point, on observe à peine un pixel d'écart transversal dans le placement du point de référence correspondant. La direction de \mathbf{v}_i^j est perpendiculaire, avec l'orientation choisie pour pointer dans l'ordre de numérotation des axes.

Enfin, pour compléter la définition du référentiel, on a besoin de fixer $|\mathbf{u}_i^j|$ et $|\mathbf{v}_i^j|$ qui seront caractéristiques des dimensions a priori du carpe. Pour la largeur $|\mathbf{v}_i^j|$, on prend simplement L_{ref} , la largeur moyenne des os des doigts dont l'estimation est très robuste. Pour avoir une longueur de référence $|\mathbf{u}_i^j|$, on ne peut pas utiliser le calcul précédent car on a besoin ici de la longueur totale de l'os, zones cartilagineuses incluses. La différence est importante pour les mains jeunes dont les épiphyses ne sont pas visibles et qui sont aussi les plus exigeantes dans le placement des points de référence. On procède donc à un calcul similaire, mais en considérant des

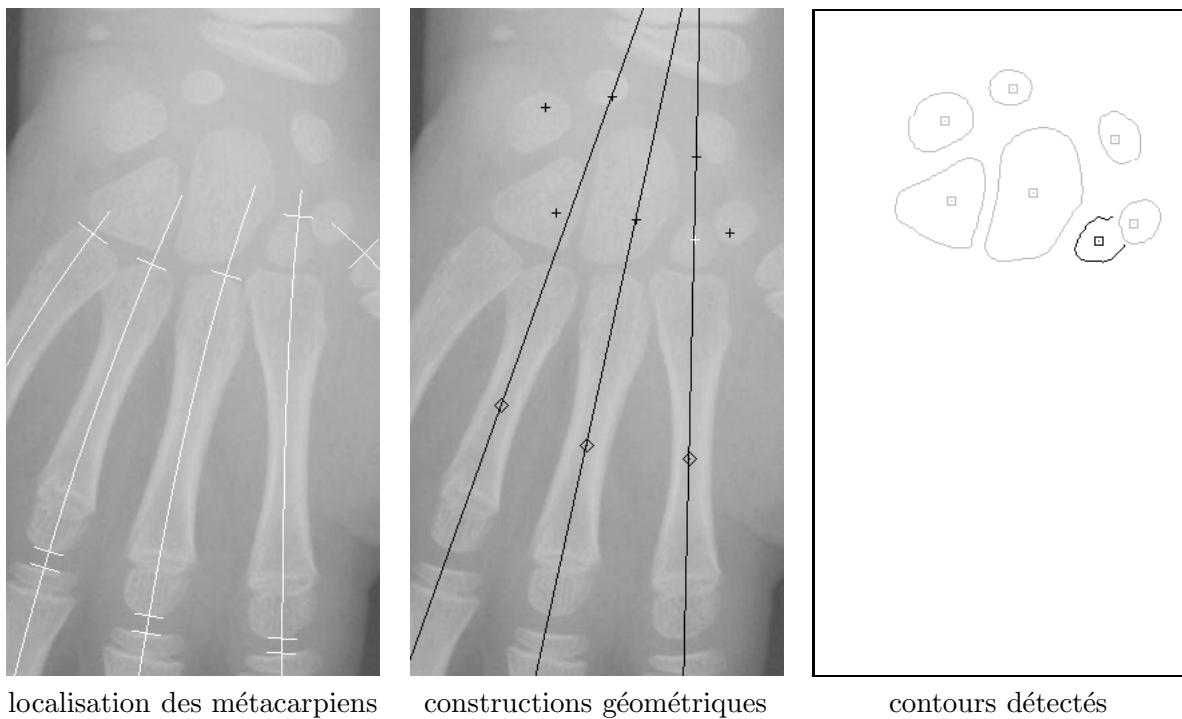


FIG. 4.7 – Exemple de placement des points de référence

longueurs égales à la distance entre les extrémités proximales du métacarpien et de la première phalange. En toute rigueur, la correction ainsi faite est imputable pour moitié au métacarpien et pour moitié à la phalange. Le résultat est néanmoins acceptable sans correction particulière qui utiliserait des coefficients de longueurs relatives hybrides.

4.3 Résultats et applications

La figure 4.7 montre les étapes du placement des points de référence a priori et de la détection des contours. Il s'agit là d'un exemple représentatif des difficultés courantes et, en particulier, du cas d'une erreur de segmentation sur le second métacarpien. On voit que la méthode est assez robuste pour estimer correctement un point milieu, marqué \diamond , pour chacun des trois métacarpiens centraux. Cela permet notamment de placer le point de référence du trapézoïde, marqué + en blanc sur l'image du centre, et ainsi d'identifier ses contours (en noir à droite).

On remarque aussi que les points de référence a priori, marqués +, ne sont pas toujours parfaitement situés, mais que le placement réalisé suffit pour une bonne détection des contours de leur région osseuse respective. Notamment, on réalise la séparation entre le trapèze et le trapézoïde. L'exemple de la figure 4.7 correspond à un âge osseux de presque 7 ans pour une fille. L'obtention de contours relativement complets et bien identifiés est en général possible pour des maturations un peu plus avancées avec une limite aux alentours de 10 ans. Le recalage sur ces contours permet d'obtenir de bons points de référence, marqués \square , pour guider les traitements ultérieurs.

La méthode peut être mise en échec, soit en cas d'erreurs de même sens sur la longueur d'au moins deux des trois métacarpiens centraux, soit en cas d'erreur sur l'extrémité distale d'un métacarpien. De tels cas sont très rares sur les images de bonne qualité. Dans le premier cas, l'estimation par la médiane est fausse et tous les points de référence sont affectés. Dans le second cas, seuls les points de référence qui sont basés sur ce métacarpien sont affectés. Il est difficile de prévoir a priori des solutions alternatives utilisables car seules les erreurs les plus grossières,

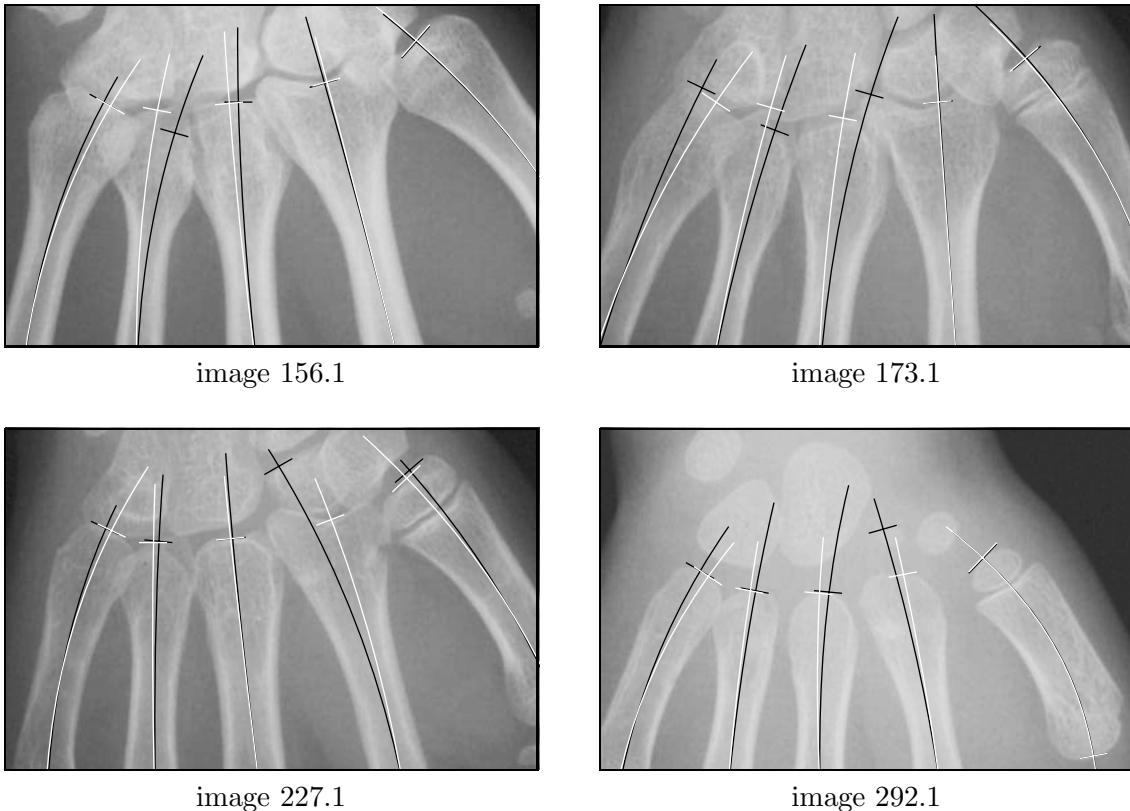


FIG. 4.8 – Exemples de correction des extrémités proximales des métacarpiens

celles qui dépassent largement la variabilité naturelle, sont détectables et pourraient donner lieu à une décision automatique entre deux méthodes.

Comme exemple de traitement ultérieur, au-delà des applications diagnostiques utilisant directement les os du carpe, comme l'âge osseux, on peut revenir sur la localisation de l'extrémité proximale du métacarpien. Une première amélioration provient d'une correction de la direction de l'axe qui est obtenue en renversant le modèle de placement du point de référence. Il s'agit plus exactement d'un modèle alternatif qui utilise les 5 axes. On corrige ainsi les problèmes liés à une mauvaise détermination des points médians à l'extrémité proximale du métacarpien. Ensuite, pour localiser l'articulation avec l'os du carpe en vis-à-vis, on utilise une technique similaire à celle utilisée pour les articulations des phalanges et la localisation du point de référence permet une contrainte de position assez précise. On corrige ainsi les problèmes liés à l'utilisation d'une contrainte de longueur guidée par un modèle de normalité.

Dans les cas où l'os du carpe n'est pas détecté, on peut encore utiliser le point de référence a priori. Il faut noter qu'il n'y a alors pas de recalage effectif sur l'image et que cela revient simplement à une régression sur les trois métacarpiens centraux, ce qui peut poser des problèmes pour les morphologies qui dévient notablement de la normalité.

Si l'on applique ceci à l'étude d'exemples vue en 3.3, tableau 3.25, page 106, et figures B.5, page 181 en annexe, et suivantes, on voit que l'on parvient ainsi à corriger les 4 cas 156.1, 173.1, 227.1 et 292.1, voir la figure 4.8, où l'ancienne segmentation est en noir et la nouvelle en blanc. Les 4 autres cas rencontrés dans 3.3 et que l'on ne parvient pas à corriger ainsi, correspondent à des ossifications très avancées (533.1, 666.2, 727.1 et 731.1). La limite d'efficacité d'une telle correction semble ainsi se situer aux alentours de 14 ou 15 ans, c'est-à-dire nettement au-delà de la limite d'obtention de bons contours pour le carpe. Ceci provient du bon recalage des points centraux, même sur des fragments de contours.

Au-delà de cet âge limite, la localisation d'une articulation, directement par ses contours ou par sa signature spectrale, est mal déterminée. La solution peut alors passer par la mise en correspondance de l'image analysée avec des images d'un atlas de référence ou un modèle d'apparence. Notre méthode suffisamment fiable de placement de points de référence permettrait d'initialiser le calage géométrique qui est nécessaire à la mise en œuvre de ces techniques. Notons que le problème est facilité pour des images dans lesquelles tous les os sont visibles. Pour des mains plus jeunes, on devrait gérer la multiplicité des chronologies d'apparition des os et notre technique d'identification par les points de référence permettrait de sélectionner le bon modèle chronologique.

Enfin, rappelons que la convergence des axes dans la région de l'épiphyse du radius est due à une orientation particulière des métacarpiens pour cette image et que cela ne pourrait être utilisé en toute généralité.

4.4 Problèmes non résolus

Pour compléter ce travail sur la détection et la localisation des os, il resterait essentiellement deux choses à faire :

1. la détection de l'os pisiforme. Cet os apparaît très tard et il se confond souvent avec l'os pyramidal. Sa détection ne peut pas se résoudre par l'adjonction d'un point de référence car on se trouve souvent dans le cas de deux contours imbriqués ou, du moins, englobant une large région commune. Un point de référence supplémentaire introduirait de nombreuses confusions d'identification qui affecteraient le recalage, la caractérisation à l'échelle de localisation et aussi l'apprentissage du modèle géométrique pour les deux os. En fait, on doit considérer que l'on traite la localisation de l'ensemble pyramidal+pisiforme, sans chercher à les distinguer et il faut envisager un traitement spécifique, par exemple un modèle d'apparence.
2. la localisation des épiphyses radiale et cubitale. Les techniques que nous avons développées peuvent y conduire si une des deux conditions suivantes est réalisée. Soit on dispose d'un cadrage assez large qui permet de détecter les axes du radius et du cubitus comme des os longs, mais cela peut nuire aux premières phases du traitement. Soit on impose une orientation particulière du poignet qui permet de traiter les épiphyses comme des os du carpe. Ce n'est pas toujours le cas sur les images dont on dispose. Pour une technique plus générale, la seule information utile est celle d'une distance, par exemple, par rapport au centre du troisième métacarpien. Cette distance est relativement bien connue après le recalage des points de référence et il faut ensuite envisager un balayage angulaire pour détecter et identifier les contours des épiphyses.

Nous n'avons pas expérimenté de modèle géométrique plus sophistiqué. Dans ce sens, on pourrait, par exemple, envisager d'introduire l'âge osseux estimé sur les os longs pour adapter le placement a priori des points de référence. La pertinence n'est pas établie et nous n'avons pas suffisamment d'images pour réaliser un apprentissage par tranche d'âge.

Conclusions

Notre problème était de pouvoir détecter, identifier et localiser les régions osseuses sur une radiographie de la main, de manière entièrement automatique, c'est-à-dire, sans ajustement de paramètres au cas par cas. Il s'agit ainsi d'éviter aussi bien le paramétrage lié à la qualité de l'image que le paramétrage morphométrique, pour faciliter une intégration indépendante en amont d'une application clinique. Dans notre approche, l'identification et la localisation précise de certains éléments anatomiques restent essentielles et constituent une base pour des traitements d'image plus spécifiques.

Une détection des contours par le gradient, sans autre information que celles extraites localement de l'image, nous semble la méthode la plus efficace et la plus neutre. À ce niveau, nous avons mis en œuvre des techniques relativement classiques. La difficulté dans la caractérisation des contours utiles provient essentiellement des cas de juxtapositions et de superpositions des régions osseuses, ainsi que de la variabilité dans leur développement respectif, qui conduisent à des ambiguïtés de détection ou d'identification. Nous devons introduire des connaissances a priori sur la morphologie du squelette de la main. Cela se traduit par un certain nombre de contraintes de relation qui lient les positions relatives des contours. Cet apport d'information, assimilable à un modèle de forme faible, est suffisamment souple pour assurer une interprétation fidèle de certaines morphologies qui s'écartent sensiblement de la normalité. Les contraintes sont introduites progressivement et elles sont paramétrées par des mesures acquises sur l'image lors des étapes précédentes.

Une étape essentielle de ce processus est la construction des axes des doigts. La symétrie axiale des doigts est suffisamment caractéristique pour permettre leur mise en évidence avec très peu d'information a priori. Il faut cependant filtrer l'information de symétrie pour n'accepter que certaines configurations de contours très particulières et assez variées. Tout en s'inspirant des caractérisations classiques basées sur le gradient, il convenait d'y ajouter des contraintes de nature plus logique que numérique. La mise en œuvre de ce filtre structurel par un automate s'est avérée très pratique dans sa conception et pour la mise au point des contraintes. Il devrait être possible de réemployer cette approche hybride dans d'autres contextes. La méthode demande simplement de connaître a priori la direction des axes recherchés, pour réaliser un balayage dans une direction très approximativement perpendiculaire. C'est l'automate qui décide ensuite quels profils d'extrema du gradient sont reconnus durant le balayage. Quelques contraintes numériques restent nécessaires, elles concernent le seuillage du gradient significatif et des limites sur la distance d'appariement, qui sont assimilables à un paramètre d'échelle des structures à caractériser.

La suite de la construction des axes s'inspire de techniques algorithmiques assez générales et elle suit les grands principes d'un processus de groupement perceptuel. On y a simplement privilégié la construction d'axes approximativement parallèles, dans les choix algorithmiques et dans le paramétrage des décisions de regroupement. Le nombre d'axes voulus n'est pas un paramètre de leur construction. Dans notre cas, il sert finalement de cible pour l'ajustement de l'échelle de détection, un paramètre essentiel qui résume un certain nombre d'aspects techniques de l'acquisition de l'image et qui détermine grandement la suite du traitement.

Dans son ensemble, la construction des axes s'inscrit dans le paradigme général classique *détection puis localisation*, en procédant au calcul des contours longitudinaux pour deux échelles adaptées en fonction des caractéristiques de l'image. La relative constance, pour une image donnée, des dimensions des formes nous permet une telle approche en deux phases et il n'est pas nécessaire d'envisager des approches multi-échelle beaucoup plus coûteuses. Le lien entre les deux échelles est assuré par la construction des axes à l'échelle de détection qui permettent ensuite la sélection des contours pertinents à l'échelle de détection. Le seul paramètre critique est la largeur moyenne d'un doigt et l'automatisation de sa première estimation impose quelques contraintes assez ordinaires sur le cadrage de l'image.

La segmentation le long d'un axe est une phase plus délicate car elle repose sur une contrainte de normalité par l'utilisation d'a priori sur les longueurs relatives des os. Un assouplissement de cette contrainte consiste à ne l'utiliser que pour déterminer un intervalle qui est supposé ne contenir qu'une et une seule articulation.

La reconnaissance d'une articulation utilise un filtre structurel, implanté encore une fois comme un automate, qui permet de gérer les quelques signatures caractéristiques des stades de développement des épiphyses et des brillances corticales. Cet automate prend en entrée les extrema des dérivées première et seconde de l'intensité le long de l'axe et doit être paramétré par les dimensions relatives des structures internes à une articulation. En pratique, il s'agit de règles structurelles différentes et donc de quelques automates différents, selon l'axe (cas particulier du pouce) et le rang sur l'axe, ce qui détermine le type de l'articulation (nombre d'épiphyses potentielles, proximales ou distales).

L'intérêt de cette approche est le fort contrôle structurel qui est permis par la définition d'un automate particulier et la conception d'un tel automate reste relativement simple. Le point plus délicat est le paramétrage numérique. En notant une certaine similitude avec des réseaux de Markov cachés, par le principe de fonctionnement, ou avec des réseaux de neurones, car nos contraintes numériques s'expriment par des inégalités linéaires, on pourrait envisager un apprentissage du paramétrage.

Nos résultats actuels fournissent déjà la localisation et une typologie des articulations qui permettent une mise en œuvre automatique d'un apprentissage. Mais le principe d'un traitement dirigé par un apprentissage serait surtout intéressant plus en aval, pour une véritable analyse de la forme bi-dimensionnelle des articulations qui trouve son utilité dans les applications comme l'âge osseux et qui permettrait aussi une validation *a posteriori* de la segmentation. Mais il faut voir que l'introduction d'un apprentissage pose un problème délicat qui est celui de la définition de la base d'apprentissage. Celle-ci doit alors décrire correctement un grand nombre d'anomalies anatomiques et cela demande une forte expertise médicale.

Le traitement du carpe est de nature beaucoup plus géométrique. Cette étape a demandé la définition d'une caractérisation très particulière des symétries centrales qui privilégie les contours à une distance typique connue. On l'utilise pour identifier des contours qui se rattachent à des points centraux qui sont estimés pas des constructions géométriques robustes à partir des métacarpiens. En exploitant ainsi l'information déjà acquise, on évite une mise en évidence directe qui serait coûteuse et ambiguë. La position de ces points est ensuite recalée à partir des contours obtenus, ce qui constitue un référentiel fiable pour des traitements ultérieurs. Il est à noter que ce principe reste effectif pour une tranche d'âge intermédiaire où les contours du carpe ne sont plus directement exploitables mais contribuent pourtant à déterminer un recalage pertinent. Nous en avons vérifié l'utilité dans une amélioration de la localisation des extrémités proximales des métacarpiens.

La méthode que nous avons développée permet donc une localisation fiable et précise malgré la grande variabilité technique et anatomique des images. Des mesures élémentaires sur un nombre d'images significatif nous ont permis de le vérifier expérimentalement. Les articulations

des phalanges et des métacarpiens sont bien traitées et d'une manière relativement indépendante de la morphologie examinée, au regard du paramétrage qui induit une certaine normalité a priori. Le traitement des extrémités distales est beaucoup plus fragile en raison du manque de contraste. Les extrémités proximales des métacarpiens sont finalement bien localisées quand on réutilise les résultats acquis sur ce carpe. Cela illustre parfaitement l'intérêt d'une résolution progressive où l'on alterne un filtrage par des apports de contraintes structurelles faibles et des phases de recalage sur l'image.

Nous avons fait l'inventaire des facteurs techniques d'acquisition qui peuvent contrarier un traitement correct des images. Les principales contraintes concernent le cadrage de la main et le positionnement des éléments étrangers dont la forme importe peu mais dont la présence dans la région d'intérêt peut poser des problèmes. Les critères de qualité que nous imposons sont finalement peu contraignants ; ils sont compatibles avec les protocoles radiographiques usuels et permettent l'utilisation d'images d'archives. Un cadrage un peu plus large qui serait simple à respecter lors de l'acquisition de nouvelles images, permettrait d'envisager aussi le traitement des épiphyses radiale et cubitale.

La précision atteinte dans la localisation des contours semble largement suffisante pour les besoins diagnostiques et ce, même avec des images de résolution 75 ou 150 dpi. La limite de précision qui est atteinte est un problème intrinsèque, lié au flou des formes arrondies et auquel s'ajoute certaines difficultés de définition de la mesure elle-même.

La dynamique des niveaux d'intensité est un critère plus important. Dans les zones de forte absorption, les blancs du négatif radiographique, il est techniquement relativement facile de contrôler la dynamique. Il faut noter que cela correspond à une faible exposition du capteur et le maintien d'une certaine dynamique invite alors à limiter la résolution spatiale. Dans les zones de faible absorption, le problème du contraste est plus délicat et une amélioration du contraste demanderait une augmentation de l'exposition.

Pour des images de l'ordre de 600×1000 pixels, les temps de traitement sont très faibles, de l'ordre de 1 à 2 secondes par image, ce qui permet de traiter facilement de grandes bases d'images.

Du point de vue médical, des développements particuliers et des expérimentations à plus grande échelle sont maintenant nécessaires pour déterminer les mesures les plus appropriées aux applications diagnostiques, par exemple la détermination de l'âge osseux. Nous disposons d'une localisation fiable qui permet des recalages géométriques, des mesures locales de longueur, densité, entre autres. Nous disposons également des informations de présence effective ou non de telle ou telle structure qui permettent de décliner plusieurs modèles si besoin.

Concernant une application à l'âge osseux, la localisation correcte des différentes structures permet d'envisager des techniques de mesure de forme pour reproduire une méthodologie de type TW2. Un traitement correct du carpe est important pour les premiers âges et notamment l'identification des os visibles aux âges critiques de 3 à 5 ans qui permet de gérer les nombreuses variations de chronologie. L'utilisation du carpe est possible, au moins jusqu'à 7 ou 8 ans, ce qui correspond aux limites de la méthode TW2 classique. La localisation des articulations entre phalanges et métacarpiens est suffisante pour les âges plus avancés. Notamment, les erreurs constatées sur les phalanges distales n'affectent pas la localisation de leur épiphyses, ce qui est important car cela constitue l'ultime source d'information pour les derniers stades de maturation.

En excluant les phalanges distales dont l'extrémité reste trop imprécise pour en mesurer correctement la longueur, on peut mettre en œuvre des diagnostics basé sur le profil métacarpo-phalangien.

Annexe A

Détails techniques

A.1 Rappels sur la détection des contours

Nous rappelons ici ce qui constitue l'état de l'art concernant la détection des contours. La théorie en la matière traite toujours de cas idéalisés, très éloignés de la réalité d'une image radiographique, autant par le "profil" des contours que par la nature du bruit. Nous pouvons pourtant dégager les idées directrices qui justifient la méthode qui est décrite en 2.2, 2.3 et 2.5.

A.1.1 Caractérisation des contours

Le niveau de gris dans l'image étant considéré comme une fonction des coordonnées, soit $I(x, y)$, suffisamment régulière, on note :

$$\mathbf{g} = \nabla I = \begin{pmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{pmatrix}, \text{ le gradient} \quad \text{et} \quad \mathbf{H} = \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial x \partial y} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix}, \text{ le hessien.}$$

Il existe essentiellement deux caractérisations des contours basées sur ces dérivées de l'image.

La caractérisation la plus "naturelle", au moins parce qu'elle traduit la perception visuelle d'un contour [Mar82], est celle des passages par zéro du laplacien $\Delta I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$.

En pratique, on détermine donc les pixels de contour comme ceux ayant un voisin dont le laplacien est du signe opposé [Har84]. On obtient ainsi par définition des contours fermés, ce qui est un avantage. Mais ces contours sont alors d'épaisseur 2, ce qui demande un amincissement un peu délicat si on veut préserver la symétrie et assurer une bonne localisation.

D'autre part, cette caractérisation demande un calcul explicite des dérivées secondes. C'est un inconvénient majeur car l'importance relative des hautes fréquences augmente avec le degré de dérivation, ce qu'on vérifie facilement par les transformées de Fourier. Or, ces hautes fréquences spatiales sont essentiellement caractéristiques du bruit et des textures, éléments généralement indésirables.

La seconde caractérisation détermine les contours là où le gradient de l'image est extrémal dans la direction du gradient, autrement dit, $\partial_{\mathbf{g}}(\mathbf{g}^2) = 0$, avec $\partial_{\mathbf{g}}$ qui note l'opérateur de dérivée partielle dans la direction du gradient¹. Et comme $\partial_{\mathbf{g}}(\mathbf{g}^2) = \frac{\mathbf{g}^T \nabla \mathbf{g}^2}{|\mathbf{g}|} = \frac{2 \mathbf{g}^T \mathbf{H} \mathbf{g}}{|\mathbf{g}|}$, cette caractérisation est équivalente à $\mathbf{g}^T \mathbf{H} \mathbf{g} = 0$. On notera que cette dernière expression fait encore intervenir des dérivées secondes.

Mais l'algorithme de suppression des non-maxima (NMS) de Canny [Can86] évite un usage explicite des dérivées secondes par une recherche directe des maxima du module du gradient

¹Plus généralement, $\partial_{\mathbf{u}}(F) = \frac{\mathbf{u}^T \nabla F}{|\mathbf{u}|}$, $\forall \mathbf{u}$ non nul.

dans la direction du gradient. C'est l'avantage principal de cette caractérisation qui est ainsi beaucoup plus robuste vis-à-vis du bruit et des textures.

L'algorithme de NMS (voir figure A.1) revient à considérer pour chaque pixel A , de coordonnées (i, j) , deux points P_1 et P_2 , voisins de A dans la direction du gradient \mathbf{g} en A . En ces deux points, on estime une valeur du module du gradient par interpolation entre les valeurs connues aux pixels adjacents, resp. B_1, C_1 et B_2, C_2 . Si le module du gradient, pour le pixel A considéré, est supérieur aux deux valeurs estimées en P_1 et P_2 , alors ce pixel A est retenu comme appartenant à un contour.

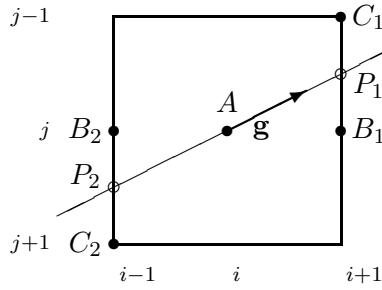


FIG. A.1 – Voisins pour l'algorithme de NMS

Il est important de noter que les deux caractérisations des contours que nous venons de considérer, ne sont pas équivalentes en général. En effet, en utilisant l'identité :

$$\Delta I = \text{trace}(\mathbf{H}) = \frac{\mathbf{u}^T \mathbf{H} \mathbf{u} + \mathbf{u}_\perp^T \mathbf{H} \mathbf{u}_\perp}{\mathbf{u}^2}, \quad \forall \mathbf{u} \text{ non nul}, |\mathbf{u}| = |\mathbf{u}_\perp| \text{ et } \mathbf{u} \times \mathbf{u}_\perp = 0$$

qui découle des propriétés fondamentales de \mathbf{H} , on obtient, pour \mathbf{g} non nul :

$$\Delta I = 0 \text{ et } \partial_{\mathbf{g}}(\mathbf{g}^2) = 0 \Rightarrow \mathbf{g}_\perp^T \mathbf{H} \mathbf{g}_\perp = 0.$$

Pour interpréter la contrainte $\mathbf{g}_\perp^T \mathbf{H} \mathbf{g}_\perp = 0$, lorsque \mathbf{g} est non nul, on peut s'intéresser aux variations de la direction du gradient. Soit, pour cela, $\theta_{\mathbf{g}} = \arctan(\frac{\partial I}{\partial y}/\frac{\partial I}{\partial x})$, supposé défini².

On obtient ainsi $\nabla \theta_{\mathbf{g}} = \frac{\mathbf{H} \mathbf{g}_\perp}{\mathbf{g}^2}$, d'où la dérivée de la direction $\theta_{\mathbf{g}}$ selon le vecteur \mathbf{g}_\perp :

$$\partial_{\mathbf{g}_\perp}(\theta_{\mathbf{g}}) = \frac{\mathbf{g}_\perp^T \mathbf{H} \mathbf{g}_\perp}{|\mathbf{g}|^3}.$$

Autrement dit, l'équivalence des deux caractérisations implique que $\partial_{\mathbf{g}_\perp}(\theta_{\mathbf{g}}) = 0$, c'est-à-dire que l'isophote (ligne de niveau, normale au gradient), doit être rectiligne, au premier ordre.

A contrario, là où l'isophote a une forte "courbure" (les "coins"), les caractérisations ne sont pas équivalentes et on observe effectivement des écarts relativement importants. La figure 2.23, page 67, montre un exemple pour une "marche à angle droit", avec des effets volontairement amplifiés par une très forte régularisation.

A.1.2 Calcul du gradient

Le calcul des dérivées d'un signal bruité est mal conditionné et, pour le régulariser, chaque composante du gradient est estimée par dérivation d'une image $\widehat{I}(x, y) = F_0(x, y) \star I(x, y)$, c'est-à-dire que I est lissée par convolution avec un filtre F_0 ad-hoc. Notamment, on préfère un filtre séparable et symétrique, soit $F_0(x, y) = f_0(x)f_0(y) = f_0(x) \star f_0(y)$, ce qui conduit à utiliser comme composantes du gradient les valeurs de $\frac{\partial \widehat{I}}{\partial x}$ et $\frac{\partial \widehat{I}}{\partial y}$, calculées par :

²Ou, sinon, $\theta_{\mathbf{g}} = \arccot(\frac{\partial I}{\partial x}/\frac{\partial I}{\partial y})$, l'une ou l'autre expression, au moins, étant localement définie.

$$\frac{\partial \widehat{f}}{\partial x} = f_1(x) * f_0(y) * I(x, y) \quad \text{et} \quad \frac{\partial \widehat{f}}{\partial y} = f_1(y) * f_0(x) * I(x, y)$$

où l'on note $f_1(u) = \frac{\partial f_0(u)}{\partial u}$.

Canny [Can83] a formalisé des critères de qualité pour de tels filtres, dans le cadre d'une détection de contours par les maxima du gradient. Soit, en mono-dimensionnel, à trouver les maxima de $s * f_1$, pour un signal particulier $s(x) = A.Y(x) + b(x)$, de type "marche" en $x = 0$, d'amplitude A (Y notant la fonction de Heavyside) et dégradé par un bruit b (additif), blanc, gaussien, de moyenne nulle et de variance σ_b^2 . Les critères formulés par Canny [Fau93, Can83] sont alors les suivants :

- un critère de détection qui est le rapport signal/bruit de $s * f_1$, en $x = 0$:

$$D = \frac{A}{\sigma_b} \frac{\int_{-\infty}^0 f_1(u) du}{\sqrt{\int_{-\infty}^{+\infty} f_1^2(u) du}} = \frac{A}{\sigma_b} \Sigma(f_1)$$

- un critère de localisation qui est l'inverse de l'écart-type de la position "du" maximum de $s * f_1$, soit :

$$L = \frac{A}{\sigma_b} \frac{|f'_1(0)|}{\sqrt{\int_{-\infty}^{+\infty} f'^2_1(u) du}} = \frac{A}{\sigma_b} \Lambda(f_1)$$

- un critère d'unicité qui traduit l'écart moyen entre deux maxima du bruit³ :

$$x_{\max}(f_1) = 2\pi \sqrt{\frac{\int_{-\infty}^{+\infty} f'^2_1(u) du}{\int_{-\infty}^{+\infty} f''^2_1(u) du}}$$

On remarquera que les critères D et L sont des invariants pour toute transformation linéaire de s , c'est-à-dire des intensités de l'image. Notons aussi que le rapport $\frac{D}{L}$ ne dépend que du filtre f_1 et pas de l'image, notamment pas du rapport $\frac{A}{\sigma_b}$. Enfin, on peut vérifier que le produit DL est invariant par un changement d'échelle (une reparamétrisation $x \rightarrow \lambda x$).

Les filtres de Deriche [Der87] sont des solutions au problème de maximiser le produit DL , avec contrainte sur $x_{\max}(f_1)$. Ces filtres, optimaux en ce sens particulier [Fau93], sont de la forme :

$$f_0(x) = c_0(1 + \alpha|x|)e^{-\alpha|x|} \quad \text{et} \quad f_1(x) = -c_0x\alpha^2e^{-\alpha|x|}$$

Ces filtres sont à réponse infinie, mais ils ont l'avantage d'être récursifs, avec une réalisation discrète très efficace et notamment meilleure que celle des filtres à support fini proposés initialement par Canny. Le facteur $c_0 = \frac{(1-e^{-\alpha})^2}{1+2e^{-\alpha}\alpha-e^{-2\alpha}}$ assure la normalisation de l'amplitude moyenne⁴.

On voit que ces filtres sont contrôlés par le seul paramètre α qui permet de régler facilement le compromis entre une bonne atténuation du bruit dans le champ gradient, avec α relativement faible, pour un critère $\Sigma(\alpha) = \sqrt{\frac{2}{\alpha}}$, et une bonne localisation des maxima du gradient, avec α relativement élevé, pour un critère $\Lambda(\alpha) = \sqrt{2\alpha}$. D'autre part, on a l'expression $x_{\max}(\alpha) = \frac{2\pi}{\alpha\sqrt{5}}$ et on retiendra notamment que la valeur limite $x_{\max} = 2$, soit un pixel intercalé, est obtenue pour $\alpha \approx 0.71$.

On peut facilement établir une correspondance avec les filtres gaussiens, en vérifiant que la "variance" du filtre f_0 est $\frac{4}{\alpha^2}$, soit l'équivalence $\sigma = \frac{2}{\alpha}$. Cela permet d'estimer la "portée" des opérateurs différentiels définis par les dérivées de f_0 . D'ailleurs, on remarque que, dans les

³On utilise alors seulement l'hypothèse que b est un bruit blanc pour appliquer un résultat formulé par Rice.

⁴L'amplitude moyenne est ainsi indépendante de α mais, très logiquement, les amplitudes maximales du signal filtré et de ses dérivées décroissent quand le lissage augmente, c'est-à-dire quand α diminue.

expressions de f_0 et f_1 , la variable x n'intervient que dans des produits αx (ou $\alpha|x|$) et il est donc ais  de mettre un lien de proportionnalit  entre ce param tre α et l' chelle de l'image.

Le principal inconveni nt des filtres de Deriche est une importante anisotropie⁵. Aussi, pour calculer le module du gradient, au lieu de l'expression classique $\sqrt{\frac{\partial I}{\partial x}^2 + \frac{\partial I}{\partial y}^2}$, nous utilisons l'expression $\frac{(\frac{\partial I}{\partial x} + \frac{\partial I}{\partial y})^3}{(\frac{\partial I}{\partial x} + \frac{\partial I}{\partial y})^2 + \frac{\partial I}{\partial x} \frac{\partial I}{\partial y}}$ qui permet notamment de minimiser les variations du module le long du contour d'un disque [ZF93, HM95].

A.1.3 Scale-space et analyse multi- chelle

La d finition du scale-space [Wit83, Koe84, Lin94] permet de g n raliser l'analyse d'une image $I(x, y)$   diff rentes  chelles en consid rant une fonction $\mathcal{I}(x, y, \varepsilon) = (I * L(\varepsilon))(x, y)$. Cela repose sur le choix d'un filtre de lissage $L(\varepsilon)$, "graduable" par un certain facteur d' chelle ε et tel que \mathcal{I} est suffisamment r guli re en x, y et ε .

L'une des principales propri t s de d finition du scale-space est celle de causalit  qui peut s'exprimer par : $\frac{\partial \mathcal{I}}{\partial \varepsilon} \Delta I > 0$. Cela se traduit en pratique par le fait qu'une singularit , notamment un contour, peut  tre "suivie" sans ambigu t  en affinant l' chelle. C'est aussi quivalent   dire qu'un accroissement du lissage ne cr e pas de nouvelle singularit . Si l'on impose un certain nombre d'autres propri t s comme une loi de semi-groupe⁶, l'isotropie et l'homog n it , on aboutit   la caract risation des filtres gaussiens, solutions de $\frac{\partial L}{\partial \varepsilon} = \Delta L$, soit en bi-dimensionnel : $L(\varepsilon)(x, y) = \frac{1}{4\pi\varepsilon} e^{-\frac{x^2+y^2}{4\varepsilon}}$, le lien avec la formulation classique d'une gaussienne  tant alors donn  par l'identit  $2\varepsilon = \sigma^2$, avec la pseudo-additivit  : $\sigma_{a \oplus b}^2 = \sigma_a^2 + \sigma_b^2$.

Consid rons maintenant une discr tisation du scale-space, avec les pas δ_x , δ_y et δ_ε . Pour permettre le "suivi" inter- chelle des contours, il faut respecter une certaine proximit  entre deux  chelles cons cutives. C'est une condition n cessaire pour pr senter notamment l'axiome de causalit . Avec le choix naturel $\delta_x = \delta_y = 1$ (pour un pixel), cela impose $\delta_\varepsilon \leq \frac{1}{4}$, soit encore $\delta_{\sigma^2} \leq \frac{1}{2}$. On retrouve ainsi la m me contrainte sur le pas de "temps" que celle qui est n cessaire   la stabilit  des sch mas de diffusion [WB97]. Enfin, on peut remarquer que la suite des filtres gaussiens d' chelles $\sigma_i = \sqrt{\frac{i+1}{2}}$ qui est induite par $\delta_{\sigma^2} = \frac{1}{2}$, est approxim e par la suite des filtres binomiaux pour lesquels on dispose d'un calcul par cascade d'additions relativement efficace. Mais cela repr sente pr s de 128  chelles pour remonter de $\sigma^2 = 64$ (soit $\sigma = 8$, $\alpha = 0.25$)   $\sigma^2 = 1$ (soit $\alpha = 2$) qui sont des valeurs courantes pour, respectivement, la d tection et la localisation.

En pratique, nous avons obtenus des r sultats convenables en utilisant une suite d' chelles ($\alpha_i = \sqrt[4]{2} \alpha_{i-1}$), nettement plus espac es, surtout aux  chelles grossi res,   condition de mettre en  uvre des r gles de coh rence inter- chelles assez sophistiqu es, tenant compte notamment de l'orientation du gradient. On notera aussi qu'une telle progression g om trique traduit bien la relation qui s' tablit entre une  chelle effective comme i , une fonction lin aire du logarithme de σ , et le degr  de d tails apparent [Lin93]. Mais cela demande encore de traiter 13  chelles. Avec moins d' chelles, le suivi inter- chelle devient trop ambigu. C'est le cas notamment avec l'ap roche dyadique qui supporte des optimisations particuli res [MZ92, WL98] et qui permettrait de n'utiliser que 4  chelles. En fait, la d composition dyadique est une repr sentation compacte, en ce sens comparable   la pyramide laplacienne [BA83], et qui demande  tre compl t e par des interpolations d' chelles [CM97].

⁵Les seuls filtres s parables et isotropes sont construits   base de filtres gaussiens, mais ils sont moins efficaces du point de vue des crit res de Canny [Fau93].

⁶ $L(\varepsilon_a) * L(\varepsilon_b) = L(\varepsilon_a + \varepsilon_b)$, $\forall \varepsilon_a, \varepsilon_b > 0$.

A.2 Précisions sur l'algorithme d'appariement des contours

Il s'agit de mettre en évidence la symétrie des os longs en construisant des paires de points de contours dont les milieux feront des points d'axes potentiels. De manière générale, si le long d'une ligne, on rencontre un pixel de contour "gauche" suivi d'un pixel de contour "droit", alors on produit le pixel médian. Mais il faut envisager des cas plus ambigus et nous avons inventorié 7 cas d'association de contours avec, pour chacun, des règles d'appariement spécifiques.

D'autre part, quelques contraintes de cohérence et des précautions opératoires sont requises, notamment pour écarter les contours trop proches de l'horizontale et pour assurer un espacement suffisant entre les points médians sur une même ligne. Enfin, on retiendra que cet algorithme peut produire quelques mauvais appariements et manquer certains autres qui seraient "bons". C'est la suite du traitement, en particulier le chaînage des pixels médians, qui permettra de corriger l'essentiel de ces erreurs.

A.2.1 Caractérisation des contours "gauches" et "droits"

On ne considère désormais que des points de contour "relativement verticaux", c'est-à-dire ceux où le gradient vérifie $\left| \frac{\partial I}{\partial x} / \frac{\partial I}{\partial y} \right| > s_d$, pour un certain seuil s_d fixé empiriquement.

Le choix de ce seuil n'est pas critique. Le but premier est d'éviter toute ambiguïté sur le signe de $\frac{\partial I}{\partial x}$ et le plus important est de choisir une valeur de seuil suffisamment faible pour prendre en compte l'essentiel des contours longitudinaux, même dans les cas de doigts assez nettement inclinés. Nous avons retenu la valeur 0.2, soit un angle limite de l'ordre de 10° entre le contour et l'horizontale, et le choix d'un seuil plus faible est sans conséquence notable sur le résultat final. Notons que ce critère n'interdit pas de retenir sur la même ligne plusieurs pixels consécutifs d'un même contour. Ce "petit" problème sera détecté et traité de manière appropriée durant le balayage (voir A.2.5).

Les pixels de contour qui vérifient $\left| \frac{\partial I}{\partial x} / \frac{\partial I}{\partial y} \right| > s_d$ et $\frac{\partial I}{\partial x} > 0$ seront les pixels de contour "gauches" (le gradient étant dirigé vers l'intérieur des doigts), que l'on notera '('. De même, on caractérise les pixels de contour "droits", ayant un gradient de direction opposée et notés ')'. Il va donc s'agir de former des paires '((), en associant un '(' et un ')'.

A.2.2 Contraintes d'appariement

Dans le cas où un seul contour d'une paire est effectivement détecté, il faut éviter de l'apparier avec un contour plus éloigné qui ne peut manifestement pas se rattacher à un même doigt. Pour cela, on va se fixer une distance maximale d'appariement d_{max} . Pour traiter convenablement le cas des doigts serrés, il faut une distance maximale d'appariement de l'ordre d'un cinquième de la largeur de l'image. Si l'on suppose que les marges de part et d'autre de la main ne sont pas trop importantes, on peut admettre que la largeur de l'image reflète assez bien l'inclinaison des doigts, ce qui ajuste correctement la distance maximale d'appariement.

Dans le cas de doigts bien écartés, le cinquième de l'image donne une valeur nettement surestimée pour la "largeur" d'un doigt. Cela reste une valeur acceptable comme mesure de l'espacement entre deux doigts, du moins en ce qui concerne leur partie distale. La contrainte de distance maximale est par contre totalement inefficace entre les métacarpiens voisins mais c'est sans conséquence notable, car cette contrainte est de toute façon peu utile pour les métacarpiens dont les contours sont généralement bien détectés.

Nous avons donc retenu une valeur de d_{max} égale à un cinquième de la largeur de l'image.

À l'inverse, si l'on considère une paire '((), manifestement trop étroite, au moins un des deux contours '(() ou ')' n'est certainement pas un "vrai". Il peut s'agir d'un bruit ou d'une texture interne à l'os, ou alors, on a détecté les contours du canal médullaire. Dans tous ces cas,

il faut renforcer la pertinence d'un appariement plus large et on introduit donc un paramètre de distance minimale d_{\min} dans l'analyse des séquences.

Nous avons fixé d_{\min} à un trentième de la largeur de l'image. C'est un compromis qui permet, en général, d'écartier les contours des canaux médullaires et néanmoins de valider les contours des extrémités des doigts lorsque leur écartement est important, ce qui conduit à une image relativement large.

Enfin, on va supposer que les structures anatomiques "vraies" ne peuvent pas donner plus de 4 contours par doigt et par ligne de balayage. La détection de la peau et du canal médullaire sont considérés comme exclusifs. On ne saurait d'ailleurs pas donner les règles générales qui permettraient d'interpréter correctement des configurations plus complexes qui contiennent certainement beaucoup de bruit.

Une troisième contrainte d'appariement est donc que, au maximum, seulement deux autres pixels de contours, '()' ou ')' quelconques, peuvent s'intercaler entre ceux qui forment la paire '()' valide.

A.2.3 Principe du balayage

La mise en œuvre des contraintes de distance maximale et d'au plus 4 contours par doigt consiste à examiner la séquence des pixels de contour au travers d'une fenêtre qui va se déplacer pendant le parcours d'une ligne. Chaque fois que l'on déplace la fenêtre vers la droite pour y entrer un nouveau pixel de contour, on va, si nécessaire, éliminer celui qui était le plus à gauche (pour la contrainte des 4), et aussi tout ceux qui sont trop éloignés du nouveau pixel entré (pour la contrainte par d_{\max}). On va aussi éliminer tous les ')', directement à gauche de la fenêtre (avant le premier '(') car ils ne peuvent plus être considérés pour de nouveaux appariements. La structure naturelle pour coder cette fenêtre glissante est une file de longueur maximale 4.

On peut vérifier que l'application de ces règles conduit à conserver la propriété :

la file est vide ou commence par un '()' .

Il y a alors 16 états possibles, c'est-à-dire la file vide, '()', '(()', ')()', puis 4 de longueur 3 et 8 de longueur 4. Parmi ces 16, il y en a 8 qui finissent par ')', y compris ')' tout seul. Ces cas ne seront pas analysés comme tels mais, pour ceux qui contiennent un ')' interne, celui-ci aura déjà été considéré dans l'analyse d'un cas précédent. Il reste donc exactement 7 cas à analyser qui correspondent aux mots de longueur maximale 4, commençant par '(' et finissant par ')'.

L'algorithme d'appariement s'exprime alors, pour chaque ligne de l'image, balayée de gauche à droite :

En début de ligne, la file est vide.

Balayer la ligne et, pour chaque nouveau pixel de contour **p** rencontré sur cette ligne :

- si la file est pleine, sortir le pixel le plus à gauche pour n'en conserver que 3,
- ajouter **p** à droite dans la file,
- sortir successivement, par la gauche de la file, tous les pixels
qui sont trop loin de **p** ou de type ')', éventuellement **p** lui-même,
- si la file n'est pas vide et que **p** est de type ')',
analyser le contenu de la file pour produire éventuellement un pixel médian,

Passer à la ligne suivante.

A.2.4 Analyse des configurations d'appariement

On va s'attacher à définir des règles de décision symétriques, c'est-à-dire que le traitement de l'image symétrique (par une symétrie verticale) devrait donner les points médians symétriques. En pratique, cette symétrie pourra présenter quelques défauts d'ordre numérique. Comme on travaille en coordonnées entières, les arrondis dans l'image symétrique peuvent produire des contours déplacés d'un pixel. Dans les cas limites, la validation par les distances minimale et maximale peut être alors affectée, ce qui se traduit par quelques points médians en plus ou en moins. Enfin, le calcul du pixel médian peut également donner lieu à un arrondi différent.

En règle générale, on va privilégier les appariements '()' les plus internes qui satisfont la contrainte de distance minimale. Cette règle est essentielle pour assurer un traitement cohérent "au vol" sans devoir reconsidérer les décisions antérieures. Cela implique aussi qu'un pixel de contour ne peut pas déterminer plus d'un pixel médian. Les règles de décision, cas par cas, sont alors les suivantes :

Cas '()' : si la contrainte de distance minimale est respectée, alors la paire est validée. Il est à noter que ce mot '()' peut rester comme préfixe des prochains cas à analyser.

Cas '(()' : si la paire formée avec le contour ')' interne respecte la contrainte de distance minimale, alors elle est validée. Sinon, on considère ensuite la distance entre le contour ')' externe et ')', pour valider ou non la paire correspondante.

Cas ')()' : on doit produire un résultat symétrique du précédent. On remarquera que si la paire avec le contour ')' interne satisfait la contrainte de distance minimale, alors elle aura été déjà validée lors du traitement du mot précédent qui avait ce ')' comme suffixe. Si cette paire ne satisfaisait pas la contrainte de distance minimale, alors seulement on considère le contour ')' externe, pour valider ou non la paire correspondante.

Cas '((())' : ce cas est comparable à ')()', avec un niveau de décision de plus. La paire avec le contour ')' externe n'est donc considérée que si aucun des deux contours ')' plus internes ne satisfait la contrainte de distance minimale.

Cas ')())' : c'est le cas symétrique du précédent et les contours ')' internes ont déjà été considérés, lors de l'analyse des mots qu'ils terminaient. La paire externe n'est considérée que si aucune des deux paires plus internes ne satisfait la contrainte de distance minimale.

Cas '()()' : ce cas a deux interprétations anatomiques possibles, soit il s'agit des contours d'un os avec son canal médullaire, soit il s'agit de deux doigts relativement rapprochés. Les autres cas correspondent à la présence de bruit. L'hypothèse du canal médullaire qui conduit à apparier les contours externes, n'est considérée que si aucune des deux paires avec un contour interne ne satisfait la contrainte de distance minimale. Sinon, on peut valider indépendamment une ou deux paires, en notant que la première l'aura déjà été lors du traitement du mot dont elle était le suffixe.

Cas '((())' : c'est le cas le plus ambigu, la seule interprétation anatomique est que l'on a détecté à la fois les contours de la peau et aussi ceux de l'os. Mais il peut y avoir une seule paire de "vrais" contours, avec un parasite intercalé et un parasite extérieur ou aussi deux parasites intérieurs. Si la paire interne satisfait la contrainte de distance minimale, elle a déjà été validée lors du traitement du mot précédent et on ne fait alors rien de plus.

Sinon, on doit envisager de valider indépendamment l'une ou les deux paires construites en utilisant un contour interne et un contour externe. On peut ainsi aboutir à la validation conjointe des deux paires "imbriquées". Alors l'une au moins est du bruit, mais on peut espérer que l'autre est valide. Il serait difficile de choisir a priori une des deux paires plutôt que l'autre et il est raisonnable de préserver la symétrie du traitement. En effet, la première de ces deux paires a été validée lors du traitement du mot précédent, avec élimination de l'appariement le plus interne qui est trop étroit, et on ne veut pas remettre en cause cette décision.

Finalement, c'est seulement si aucune des trois paires qu'il est possible de former avec un contour interne, ne satisfait la contrainte que l'on pourra considérer la paire des deux contours les plus externes et donc supposer que les deux contours intérieurs sont du bruit.

Cet exposé des cas privilégie la mise en évidence de la symétrie et de la progressivité de l'analyse. En pratique, pour ne pas avoir à reconsidérer inutilement les paires internes préfixes strictes, c'est-à-dire dans les cas '()' , '())' , '()()' et '((())' , on peut écarter les paires préfixes dès qu'elles sont validées. Ainsi, quand une telle paire reste dans la file à analyser, on sait qu'elle n'a pas été validée, donc qu'elle ne satisfait pas la contrainte de distance minimale. Cela revient à dire qu'il suffit d'examiner les appariements des contours '()' avec seulement le dernier contour ')'. .

La mise en œuvre de ce principe consiste à supprimer de la file un contour '()' dès qu'il est utilisé pour former une paire valide. On supprime donc aussi tout ce qui le précède et tous les contours ')' qui le suivent immédiatement. Le gain en efficacité est relativement modéré et l'algorithme reste linéaire en nombre de pixels de contour, mais cela explicite clairement qu'un pixel de contour '()' ou ')' ne détermine, au plus, qu'un seul pixel médian.

A.2.5 Traitement des redondances d'un contour

Il faut prévoir un traitement particulier pour les occurrences de points de contour consécutifs qui n'auraient pas été écartés par le test de caractérisation $\left| \frac{\partial I}{\partial x} / \frac{\partial I}{\partial y} \right| > s_d$. Cela correspond au cas d'un contour assez proche de l'horizontale. Par continuité de $\frac{\partial I}{\partial x}$, il s'agit alors de points de même orientation et ils doivent être traités comme une seule intersection de la ligne de balayage avec ce contour. Il faut supprimer cette redondance "au vol", pour conserver un mot pertinent de longueur maximale 4, sans alourdir l'algorithme par une détection "en avant" systématique.

Dans le cas d'une suite de pixels ')', consécutifs, le plus simple est de valider le premier rencontré qui est alors traité de manière ordinaire, et de ne pas considérer les suivants qui ne sont donc pas mis dans la file. Dans le cas d'une suite de pixels '()', pour des raisons de symétrie, c'est la dernière occurrence que l'on devra considérer. La première occurrence aura été mise dans la file, comme il se doit, sans déclencher d'analyse du mot puisqu'il s'agit d'un cas terminant par '()', mais elle sera remplacée par ses voisines, au fur et à mesure et cela sans toucher au reste de la file, sauf application de la contrainte de distance maximale.

Remarquons que cette manière de procéder a tendance à réduire la distance entre les pixels '()' et ')', dès que l'un des deux contours est nettement incliné. Cela a peu de conséquence lorsque les contours sont normalement espacés. Dans le cas contraire, cela favorise clairement l'élimination de la paire du fait de la contrainte de distance minimale. Cette élimination est souvent bénéfique, notamment aux extrémités des os où, par construction, les points médians ont tendance à dévier de l'axe de symétrie naturel.

Nous allons voir ci-dessous qu'il convient en fait de généraliser un peu ce mécanisme, pour écarter de la même manière des pixels de contour rencontrés sur une ligne de balayage et qui ne seraient pas directement consécutifs, mais pourtant situés à très faible distance les uns des autres. Il s'agit alors d'éviter une ambiguïté pour la construction des axes.

A.2.6 Distance minimale entre pixels médians

Pour assurer un chaînage non ambigu (voir 2.4), nous avons besoin de maintenir une certaine distance minimale Δ entre deux pixels médians de la même ligne. Selon les règles de connexité mises en œuvre, la distance minimale requise est $\Delta = 3$, soit deux pixels intercalés, ou $\Delta = 5$, voire $\Delta = 7$.

Un pixel médian se trouve nécessairement à une distance au moins égale à $\lfloor \frac{d_{\min}}{2} \rfloor$ de chacun de ses deux contours sources, d_{\min} étant la distance minimale requise pour valider cette paire. Comme, par construction, il y a au moins un pixel d'écart, c'est-à-dire une distance 2, entre deux contours ' et ' consécutifs, la distance minimale entre deux pixels médians est donc, en général, strictement supérieure à $2\lfloor \frac{d_{\min}}{2} \rfloor$, cette valeur étant de l'ordre d'un trentième de la largeur de l'image, ce qui reste largement supérieur au Δ voulu, pour les tailles d'images courantes. En fait, on constate que le nombre moyen de points médians par ligne reste très nettement inférieur à 30 et leur densité n'est donc pas un problème.

Il y a pourtant une exception dans le cas '()' lorsque l'on décide de valider conjointement les deux appariements entrelacés.

La distance entre les deux médians est alors égale à seulement $\lfloor \frac{\delta_i + \delta_j}{2} \rfloor$ où δ_i et δ_j notent respectivement les distances entre les deux ' et les deux '. Pour garantir l'écart minimal entre les deux pixels médians, il faut $\delta_i \geq \Delta$ et $\delta_j \geq \Delta$. Pour cela, on procède à la suppression des pixels de contour trop proches de manière similaire à ce qui est indiqué en A.2.5. Une occurrence du cas '()' qui aurait posé problème n'apparaît donc pas dans la file, avec la suppression d'au moins un des deux contours externes. Le mot qui est effectivement analysé est alors plus simple et il ne peut engendrer qu'un pixel médian au plus.

A.2.7 Influence du bruit

Lorsque le bruit augmente, il apparaît inévitablement des faux appariements. Ceux qui sont réalisés entre un "bon" et un "mauvais" pixels de contours ont pour conséquence d'interrompre localement la continuité des axes. Les étapes ultérieures pourront facilement corriger ces défauts qui, en général, n'affectent pas le même axe sur deux lignes consécutives.

Il peut aussi se produire des appariements entre deux faux contours, ce qui a pour résultat de générer des points médians supplémentaires. C'est ce qui est détecté lors de l'adaptation de l'échelle de calcul des contours (voir 2.3). Lorsqu'il y a trop de points médians par ligne, on essaye une échelle qui doit réduire le bruit.

Mais, s'il y a vraiment trop de bruit, les distances entre les pixels de contour sont nettement réduites et les contraintes d'appariement s'appliquent de manière chaotique. Il en résulte des variations non régulières du nombre de points médians par ligne qui peuvent mettre en échec l'heuristique d'adaptation de l'échelle. Nous n'avons effectivement constaté ce problème que sur quelques cas de très mauvaises images.

A.3 Mise en évidence des symétries

Une alternative à la construction des points médians par balayage serait le recours à une mise en évidence directe des axes symétries. Nous présentons ici la mise en œuvre d'une technique relativement générale [RWY95], avec les adaptations qui sont nécessaires pour renforcer le type de symétrie que nous recherchons. La caractérisation résultante est finalement assez similaire à celle qui est à la base de diverses variantes [MPL94, PEMF98] d'une technique de calcul du squelette d'une forme. Toutes ces méthodes ont en commun de procéder par accumulation et elles prennent en compte le gradient dans un voisinage du point qui est considéré comme candidat pour un centre de symétrie. Le coût des calculs, pour des résultats au mieux comparables, nous font écarter de telles techniques de recherche des symétries.

Il faut noter aussi que les caractérisations plus "locales" et relativement plus rapides, basées en particulier sur les extrema des valeurs propres du hessien [EGM⁺94], ne conviennent pas. D'une part, l'axe de symétrie d'un doigt ne constitue pas une ligne de "crête" et il faut conjuguer divers types d'extrema suivant le contexte. D'autre part, les échelles requises pour régulariser, causent des interactions entre les doigts et des déviations importantes des lignes d'extrema.

A.3.1 Principe

Nous noterons \mathbf{p} les coordonnées d'un pixel considéré en tant que point de référence d'une symétrie potentielle et, par abus, nous noterons aussi \mathbf{p} ce pixel lui-même. Il s'agit de quantifier la symétrie du gradient entre deux pixels, de positions symétriques par rapport à \mathbf{p} . Soit \mathbf{r} un vecteur, à coordonnées entières, qui sera un déplacement de part et d'autre de \mathbf{p} . On considère ainsi les deux pixels $\mathbf{p} + \mathbf{r}$ et $\mathbf{p} - \mathbf{r}$ et on note $\mathbf{g}_{\mathbf{p}+\mathbf{r}}$ et $\mathbf{g}_{\mathbf{p}-\mathbf{r}}$ les vecteurs gradients, calculés respectivement en ces deux pixels, et $\theta_{\mathbf{p}+\mathbf{r}}$ et $\theta_{\mathbf{p}-\mathbf{r}}$ les angles mesurés entre la direction de \mathbf{r} et les gradients respectifs en $\mathbf{p} + \mathbf{r}$ et $\mathbf{p} - \mathbf{r}$. Les valeurs de $\theta_{\mathbf{p}+\mathbf{r}}$, $\theta_{\mathbf{p}-\mathbf{r}}$, $|\mathbf{r}|$, $|\mathbf{g}_{\mathbf{p}+\mathbf{r}}|$ et $|\mathbf{g}_{\mathbf{p}-\mathbf{r}}|$ vont déterminer un critère :

$$C(\mathbf{p}, \mathbf{r}) = \Theta(\mathbf{p}, \mathbf{r}) D(|\mathbf{r}|) |\mathbf{g}_{\mathbf{p}+\mathbf{r}}| |\mathbf{g}_{\mathbf{p}-\mathbf{r}}|$$

pour la paire de pixels $\mathbf{p} + \mathbf{r}$ et $\mathbf{p} - \mathbf{r}$.

Cette expression combine ainsi de manière arbitraire, mais classique, un critère angulaire $\Theta(\mathbf{p}, \mathbf{r})$, une pondération $D(|\mathbf{r}|)$ et les modules du gradient. Dans une interprétation dense de l'algorithme, la qualité de la symétrie par rapport à \mathbf{p} s'exprime alors comme la somme du critère pondéré $C(\mathbf{p}, \mathbf{r})$ pour tous les \mathbf{r} définissant un domaine Γ autour de \mathbf{p} . Pour assurer une certaine localité, le déplacement \mathbf{r} est limité à un domaine fini qui, dans une approche a priori, sans information particulière sur l'orientation de l'image, doit être isotrope et constant. Autrement dit, on pose $0 < |\mathbf{r}| \leq r_{\max}$, pour un certain r_{\max} qui sera un paramètre, en relation avec la largeur attendue d'un os.

Dans [RWY95], le module du gradient $|\mathbf{g}|$ intervient dans une pondération en $\log(1 + |\mathbf{g}|)$. Dans notre cas, une pondération plus marquée est souhaitable car ce sont essentiellement les zones de fort gradient, c'est-à-dire les contours et leur voisinage, qui contribuent à caractériser la symétrie axiale. Les zones de plus faible gradient, les textures notamment, ne respectent pas cette symétrie et leur contribution ne peut que brouiller la mise en évidence.

Idéalement, pour une symétrie selon un axe perpendiculaire à \mathbf{r} , on a $\theta_{\mathbf{p}+\mathbf{r}} + \theta_{\mathbf{p}-\mathbf{r}} = \pi$, d'où le critère de symétrie angulaire $\cos(\pi - \theta_{\mathbf{p}+\mathbf{r}} - \theta_{\mathbf{p}-\mathbf{r}}) = \sin(\theta_{\mathbf{p}+\mathbf{r}}) \sin(\theta_{\mathbf{p}-\mathbf{r}}) - \cos(\theta_{\mathbf{p}+\mathbf{r}}) \cos(\theta_{\mathbf{p}-\mathbf{r}})$ qui est utilisé dans [RWY95].

Dans notre cas, une expression réduite à $-\cos(\theta_{\mathbf{p}+\mathbf{r}}) \cos(\theta_{\mathbf{p}-\mathbf{r}})$ donne de meilleurs résultats pour caractériser les axes des os longs. Notamment, dans le cas de deux gradient antiparallèles, soit deux contours parallèles en vis-à-vis, on évite ainsi une contribution négative du produit des sinus. On obtient donc une meilleure accumulation quand \mathbf{r} balaye un secteur angulaire de part et d'autre de la normale à l'axe. Dans [RWY95], une pondération supplémentaire vise spécifiquement à écarter les cas proches de $\theta_{\mathbf{p}+\mathbf{r}} = \theta_{\mathbf{p}-\mathbf{r}} = \pm \frac{\pi}{2}$ qui correspondent le plus souvent au cas d'un point \mathbf{p} sur un contour quasi rectiligne, avec \mathbf{r} tangent à ce contour. Ce problème ne se pose pas avec l'expression réduite $-\cos(\theta_{\mathbf{p}+\mathbf{r}}) \cos(\theta_{\mathbf{p}-\mathbf{r}})$ que nous utilisons.

On ajoute par contre une contrainte de signe sur chaque cosinus, pour exprimer que le gradient doit être dirigé vers l'axe. Cette contrainte permet de pénaliser les configurations du type ')'(' entre deux os, car celles-ci contribuent à révéler les symétries interdigitales qui sont aussi évidentes que celles que l'on recherche. D'où le critère angulaire :

$$\begin{aligned} \Theta(\mathbf{p}, \mathbf{r}) &= -\cos(\theta_{\mathbf{p}+\mathbf{r}}) \cos(\theta_{\mathbf{p}-\mathbf{r}}) && \text{si } \cos(\theta_{\mathbf{p}+\mathbf{r}}) < 0 \text{ et } \cos(\theta_{\mathbf{p}-\mathbf{r}}) > 0, \\ &= 0 && \text{sinon.} \end{aligned}$$

Avec un **ou** à la place du **et**, la contribution $\Theta(\mathbf{p}, \mathbf{r})$ serait strictement négative dans le cas de deux cosinus de même signe, ce qui aurait l'effet de mieux pénaliser ce couple (\mathbf{p}, \mathbf{r}) . En fait, cela influe très peu sur la mise en évidence des symétries. Dans [MPL94], l'appariement n'est pas explicité et les contributions des deux pixels $\mathbf{p} + \mathbf{r}$ et $\mathbf{p} - \mathbf{r}$ sont calculées séparément avant d'être sommées, ce qui correspond à un critère angulaire de la forme $\cos(\theta_{\mathbf{p}+\mathbf{r}}) - \cos(\theta_{\mathbf{p}-\mathbf{r}})$.

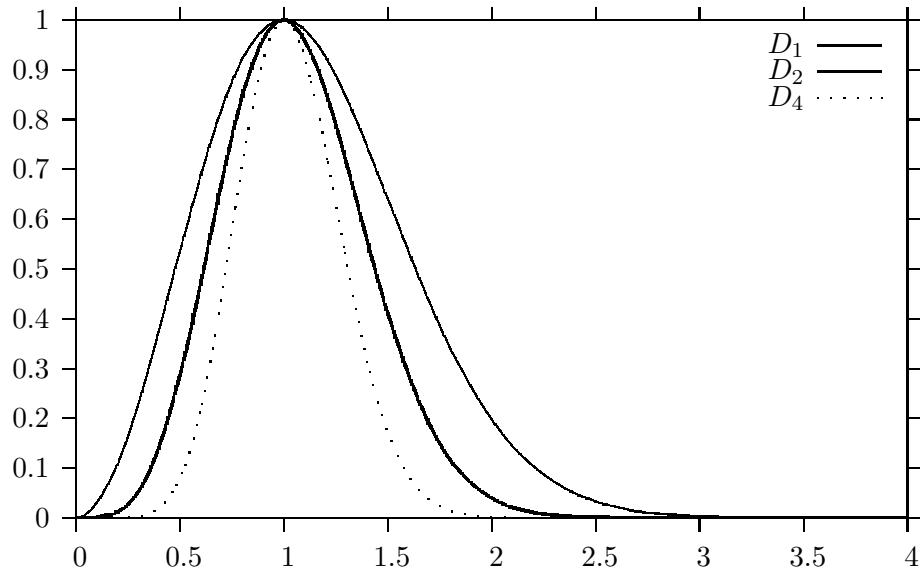


FIG. A.2 – Fonctions D_k de pondération selon la distance au point de référence

Dans [RWY95], la distance $|\mathbf{r}|$ intervient dans une pondération de la forme $e^{-\rho}$, avec $\rho = \frac{|\mathbf{r}|}{r_\gamma}$, pour un certain rayon r_γ , caractéristique de l'échelle des symétries. Dans notre cas, la pondération par la distance est particulièrement critique pour le traitement des configurations du type '()''. Une pondération décroissante a pour effet de renforcer la mise en évidence de chaque paire '()' interne, par rapport à la paire '()' externe. C'est bien ce qu'il faut pour éviter une symétrie interdigitale mais, dans le cas du canal médullaire, on voudrait obtenir l'effet inverse. En fait, on doit spécifier que l'information la plus pertinente se trouve à une certaine distance r_γ qui est à adapter en fonction de la largeur d'un os. Ainsi, dans [MPL94], le domaine Γ est un cercle de rayon $|\mathbf{r}| = r_\gamma$ qui est adapté à l'échelle de calcul du gradient. Les valeurs de $|\mathbf{r}|$ sont ensuite explorées lors d'une analyse multi-échelle. Dans notre cas, le réglage de r_γ est beaucoup plus délicat et ce paramètre doit être clairement séparé de l'échelle de calcul des contours.

Pour ce qui concerne les axes des os long, nous avons obtenu des résultats convenables avec une pondération de la forme $\rho^2 e^{-\rho^2}$. Plus généralement, on peut considérer la famille des fonctions $D_k(\rho) = (\rho^2 e^{1-\rho^2})^k$ avec un paramètre k qui, comme le montre la figure A.2, permet de contrôler la sélectivité autour de $\rho = 1$ correspondant à la distance privilégiée r_γ . On voit une certaine ressemblance avec les fonctions qui sont utilisées en 4.1 pour la caractérisation des contours du carpe, voir figure 4.1, page 121, mais les présentes sont plus pénalisantes sur les faibles distances. Quand k tend vers l'infini, on s'approche de la caractérisation "en Dirac" qui est utilisée dans [MPL94]. Il faut noter que le critère de distance garde encore une certaine tolérance, du fait de la "largeur" d'un contour, pour une part anatomique et pour une autre inhérente à son calcul.

En fait, il y a une certaine redondance entre l'échelle de calcul du gradient et notre paramètre k . Mais le choix de l'échelle va plutôt arbitrer de manière classique le compromis entre détection et localisation, en autorisant des variations assez larges, alors que la pondération par la distance détermine principalement la validation des symétries intéressantes. Pour la localisation, il s'avère intéressant d'utiliser une échelle relativement plus fine qui préserve les orientations du gradient, et de compter sur un complément de régularisation des amplitudes par la fonction de pondération D_k , avec k petit. Pour la détection, cet aspect est négligeable.

Pour le choix de r_γ , on peut alors trouver une valeur de compromis entre les effets recherchés notamment sur les deux cas du type '()''. Une détermination automatique est envisageable. Par exemple, selon un principe similaire à celui vu en 2.3, on peut procéder au comptage des points de symétrie mis en évidence, pour quantifier la plus ou moins bonne détection des 5 axes dans un processus itératif de recherche d'une valeur acceptable de r_γ .

Pour $k = 1$, on obtient ainsi le critère :

$$C(\mathbf{p}, \mathbf{r}) = \Theta(\mathbf{p}, \mathbf{r}) \rho^2 e^{-\rho^2} |\mathbf{g}_{\mathbf{p}+\mathbf{r}}| |\mathbf{g}_{\mathbf{p}-\mathbf{r}}|$$

En regroupant la contribution de $\mathbf{p} + \mathbf{r}$ dans le terme $\mathbf{r} \cdot \mathbf{g}_{\mathbf{p}+\mathbf{r}}$ qui est aussi celui utilisé dans [MPL94] par exemple, on aboutit à une expression relativement simple :

$$\begin{aligned} C(\mathbf{p}, \mathbf{r}) &= -\frac{1}{r_\gamma^2} (\mathbf{r} \cdot \mathbf{g}_{\mathbf{p}+\mathbf{r}}) (\mathbf{r} \cdot \mathbf{g}_{\mathbf{p}-\mathbf{r}}) e^{-\rho^2} && \text{si } \mathbf{r} \cdot \mathbf{g}_{\mathbf{p}+\mathbf{r}} < 0 \text{ et } \mathbf{r} \cdot \mathbf{g}_{\mathbf{p}-\mathbf{r}} > 0, \\ &= 0 && \text{sinon.} \end{aligned}$$

Si on procède à une caractérisation dense, les calculs sont très lourds car, pour chaque pixel \mathbf{p} considéré, on doit accumuler la réponse $C(\mathbf{p}, \mathbf{r})$ pour tous les \mathbf{r} d'un domaine Γ qui peut être relativement large. Il convient en effet de prendre r_{\max} de l'ordre de $2r_\gamma$, soit environ $2\pi r_\gamma^2$ paires, avec r_γ qui peut valoir plusieurs dizaines. Si on considère que r_γ et les dimensions de l'image sont liés, cela fait une loi en puissance 4 pour le traitement complet.

La simplification introduite dans [RWY95] consiste à ne considérer que les cas où $\mathbf{p} + \mathbf{r}$ et $\mathbf{p} - \mathbf{r}$ sont des pixels de contours. Pour chaque paire $(\mathbf{p}_1, \mathbf{p}_2)$ de pixels de contours, on accumule alors $C(\mathbf{p}, \mathbf{r})$ pour le "pixel"⁷ $\mathbf{p} = \frac{\mathbf{p}_1 + \mathbf{p}_2}{2}$ et $\mathbf{r} = \frac{\mathbf{p}_1 - \mathbf{p}_2}{2} \in \Gamma$. Une énumération quadratique de l'ensemble des paires $(\mathbf{p}_1, \mathbf{p}_2)$ est d'un coût raisonnable et, dans notre cas, des techniques de balayage qui exploitent la contrainte $|\mathbf{r}| \leq r_{\max}$ permettent d'approcher un coût linéaire.

A.3.2 Résultats

La figure A.3 montre les lignes de crête du critère de symétrie, obtenues pour l'image de référence, avec un seuillage à hystérosis, pour la caractérisation dense et pour quelques valeurs de r_γ . À titre indicatif, ces valeurs sont exprimées par rapport à L_{ref} , la largeur moyenne des appariements horizontaux que l'on a obtenus indépendamment. Pour cet exemple, la valeur $r_\gamma = \frac{1}{2}L_{\text{ref}}$ correspond à environ un vingtième de la largeur de l'image, ce qui est à peu près la limite pour un cadrage étroit et des doigts serrés.

On a utilisé ici des contours calculés à l'échelle de détection. On voit que la caractérisation dense est relativement tolérante dans le choix de r_γ , ce qui permet une estimation a priori assez facile et compatible avec les variations de la largeur effective le long d'un axe. La figure A.4 montre que les résultats sont comparables à ceux obtenus par notre algorithme d'appariement par balayage horizontal. Dans le cas d'axes notablement plus inclinés ou, à plus forte raison, divergents, une caractérisation isotrope éviterait les défauts qui sont inhérents au choix a priori d'une direction de balayage. On pourrait ainsi espérer une meilleure détection des axes. Mais le choix a priori de r_γ devient plus délicat et on voit figure A.3 qu'une surestimation d'un facteur 2 est déjà fortement nuisible. La figure A.5 montre un autre exemple de détection des axes.

À une échelle plus fine, propre à la localisation des contours, on observe des caractérisations multiples, avec notamment les cas triples du type '()' (2 doigts ou canal médullaire) ou du type '()' ('os" et "peau"). On peut même observer des multiplicités supérieures sur la configuration '()' (os avec canal médullaire et peau). La figure A.6 illustre bien ces problèmes et on voit que la caractérisation peut être très ambiguë, d'autant plus que les symétries (orientation, distance, module) sont imparfaites et que le critère $C(\mathbf{p}, \mathbf{r})$ pondère de manière

⁷Les coordonnées de \mathbf{p} (et \mathbf{r}) sont alors définies au demi-entier près.

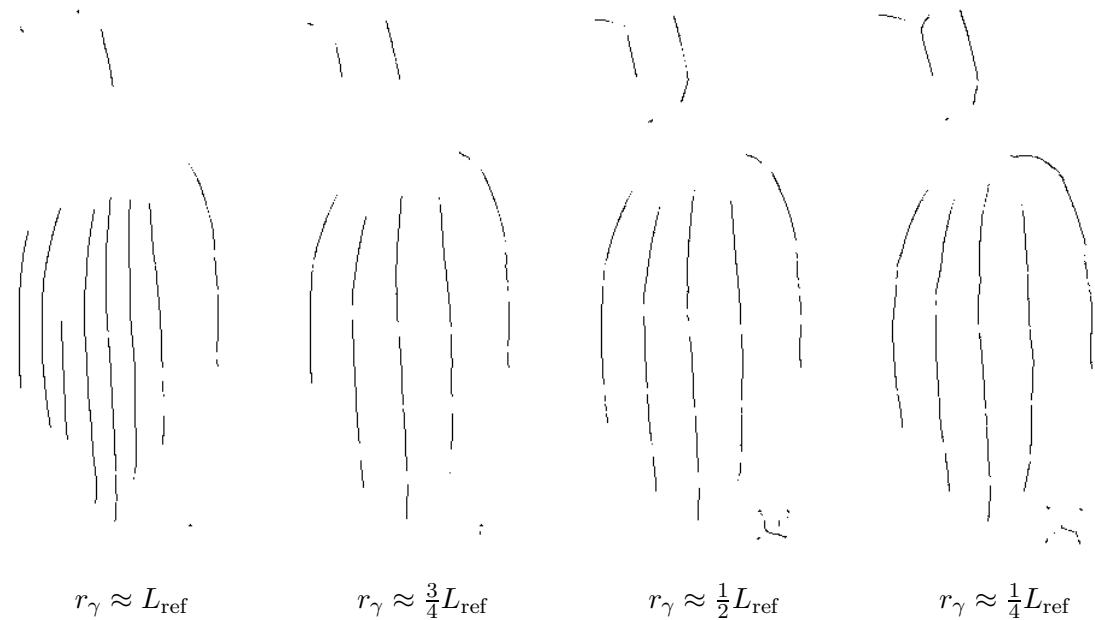
FIG. A.3 – Axes de symétrie, par une caractérisation dense, en fonction du rayon r_γ 

FIG. A.4 – Comparaison des trois techniques de caractérisation des points d'axes

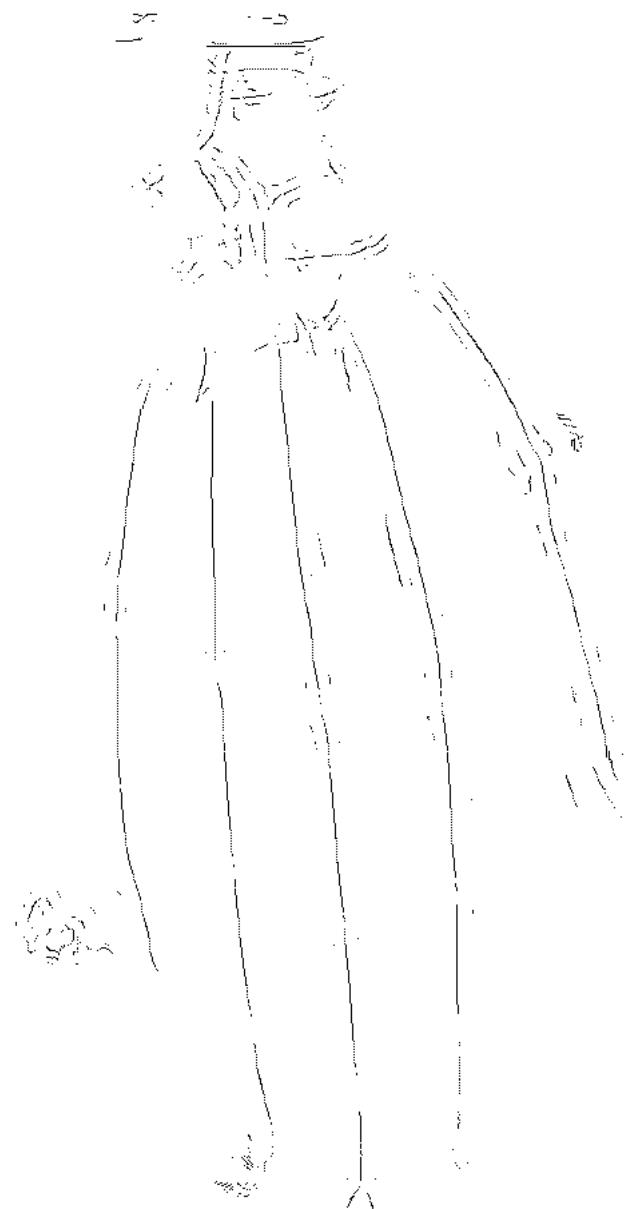


FIG. A.5 – Symétries à une échelle de détection



FIG. A.6 – Symétries à une échelle de localisation

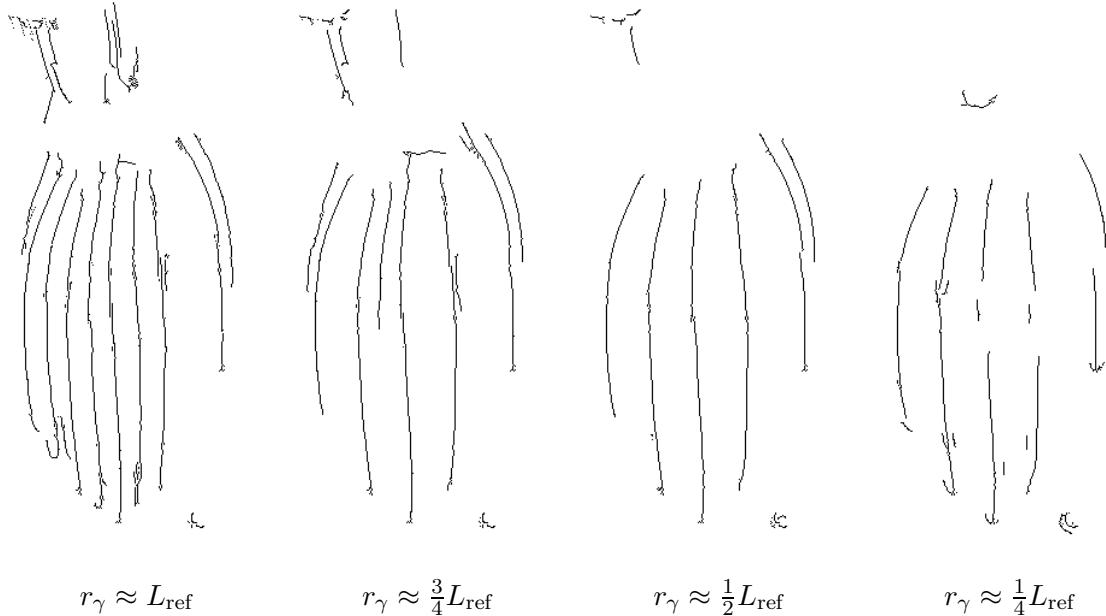


FIG. A.7 – Axes de symétrie des contours seuls en fonction du rayon r_γ

arbitraire les trois contributions. La méthode, essentiellement locale, ne permet pas d'identifier de telles configurations et de déterminer un choix adapté. Le choix de r_γ est alors plus sensible et cela rend impossible une sélection par seuillage.

En fait, le problème majeur est le coût du calcul. Pour une image de plutôt petite taille, comme celle de la figure A.3, avec L_{ref} de l'ordre de 20 pixels, la méthode dense requiert le calcul de $C(\mathbf{p}, \mathbf{r})$ pour une centaine paires, pour chaque pixel \mathbf{p} . Pour l'image de la figure A.5, r_γ est plutôt de l'ordre la trentaine de pixels. Un sous-échantillonnage d'un facteur 2 ou 4 est envisageable, mais cela reste excessif.

La figure A.7 montre les résultats obtenus, dans des conditions similaires, mais en se limitant aux pixels de contour, comme dans [RWY95]. Le calcul est ainsi nettement plus rapide, mais on perd un effet de régularisation qui résulte de l'accumulation du critère $C(\mathbf{p}, \mathbf{r})$ pour tous les \mathbf{r} du domaine Γ . Du fait des aléas d'arrondi, les contributions de plusieurs paires de pixels de contours pour un même pixel théorique \mathbf{p} peuvent se trouver plus ou moins dispersées sur les pixels voisins de \mathbf{p} . Cet étalement irrégulier complique encore la détection des lignes de maxima et leur seuillage, avec finalement des ambiguïtés de construction des axes. Enfin, en l'absence de cette régularisation supplémentaire, la qualité intrinsèque du gradient et des contours influent nettement sur le résultat. Dans les cas où l'échelle a priori ne convient pas, cette méthode ne fournit pas de mécanisme direct pour estimer une correction, sachant qu'un processus itératif, comme celui vu en 2.3, doit d'abord déterminer un choix relativement pointu de r_γ .

Aucune de ces deux méthodes de mise en évidence des symétries ne constitue donc une alternative intéressante pour la détection des axes des os longs. La méthode dense est trop lente, pour des résultats comparables à ceux obtenus par balayage, et la méthode rapide s'avère trop délicate à paramétriser. Pour les mêmes raisons, il est difficile d'utiliser ces méthodes pour obtenir, plus simplement, une estimation de l'orientation globale des axes de l'image. Si cette approche semble un peu plus générale que notre heuristique d'orientation vue en 2.1, des contraintes de cadrage s'imposent de toute façon pour assurer l'estimation a priori d'une échelle de calcul des contours et d'une largeur caractéristique r_γ .

Par ailleurs, les maxima locaux du critère peuvent mettre en évidence des centres de symétrie. Une telle approche serait envisageable pour les os du carpe. En fait, on aboutit aux mêmes

remarques sur le coût des calculs et la qualité, la méthode dense étant la seule à donner un résultat utilisable. Pour le carpe, l'identification d'un centre de symétrie comme centre d'un os est un problème compliqué, du fait de l'existence de nombreuses symétries entre os et aussi parce que la présence ou non d'un os est une inconnue a priori. Cette technique n'est donc pas appropriée, contrairement aux solutions particulières que nous avons décrites au chapitre 4.

A.4 Précisions sur l'algorithme de construction des axes

A.4.1 Calcul de la droite de régression linéaire d'une chaîne

Il s'agit d'exprimer la droite de régression associée à une chaîne \mathcal{A} par une fonction linéaire de la forme $x = X_{\mathcal{A}}(y) = \bar{x}_{\mathcal{A}} + \lambda_{\mathcal{A}} \cdot (y - \bar{y}_{\mathcal{A}})$ qui minimise l'erreur quadratique $\sum_{p \in \mathcal{A}} (x_p - X_{\mathcal{A}}(y_p))^2$. Soient $n_{\mathcal{A}}$ le nombre de pixels p d'une chaîne \mathcal{A} , puis $\bar{x}_{\mathcal{A}} = \frac{1}{n_{\mathcal{A}}} \sum_{p \in \mathcal{A}} x_p$, $\bar{y}_{\mathcal{A}} = \frac{1}{n_{\mathcal{A}}} \sum_{p \in \mathcal{A}} y_p$, $\bar{xy}_{\mathcal{A}} = \frac{1}{n_{\mathcal{A}}} \sum_{p \in \mathcal{A}} x_p y_p$ et $\bar{yy}_{\mathcal{A}} = \frac{1}{n_{\mathcal{A}}} \sum_{p \in \mathcal{A}} y_p^2$.

La résolution aux moindres carrés nous donne $\lambda_{\mathcal{A}} = \frac{\bar{xy}_{\mathcal{A}} - \bar{x}_{\mathcal{A}} \bar{y}_{\mathcal{A}}}{\bar{yy}_{\mathcal{A}} - (\bar{y}_{\mathcal{A}})^2}$.

A.4.2 Définition plus rigoureuse de la “distance” entre deux chaînes

Dans le cas de chaînes \mathcal{A} et \mathcal{B} nettement inclinées, le décalage horizontal $\delta_{x\mathcal{AB}}$ et l'espacement vertical $\delta_{y\mathcal{AB}}$ (voir 2.4 et figure 2.14, page 52, en bas à gauche) ne déterminent pas une mesure fidèle de la “distance” entre les deux chaînes. Nous allons donc définir ici une mesure plus rigoureuse mais voir aussi qu'elle manque de pertinence en pratique.

Soit $d_{\mathbf{d}_{\mathcal{A}}}^{\mathcal{B}}$ (voir figure 2.14, en bas à droite) la distance algébrique entre, d'une part, le point “distal” de la droite $x = X_{\mathcal{A}}(y)$, de coordonnées $(X_{\mathcal{A}}(y_{\mathbf{d}_{\mathcal{A}}}), y_{\mathbf{d}_{\mathcal{A}}})$ et, d'autre part, la droite $x = X_{\mathcal{B}}(y)$. Ces droites sont obtenues par régression linéaire sur les pixels des chaînes \mathcal{A} et \mathcal{B} et on les exprime respectivement par $X_{\mathcal{A}}(y) = \bar{x}_{\mathcal{A}} + \lambda_{\mathcal{A}} \cdot (y - \bar{y}_{\mathcal{A}})$ et $X_{\mathcal{B}}(y) = \bar{x}_{\mathcal{B}} + \lambda_{\mathcal{B}} \cdot (y - \bar{y}_{\mathcal{B}})$. On obtient ainsi l'expression $d_{\mathbf{d}_{\mathcal{A}}}^{\mathcal{B}} = \frac{X_{\mathcal{A}\mathcal{B}}(y_{\mathbf{d}_{\mathcal{A}}})}{\sqrt{1+\lambda_{\mathcal{B}}^2}}$, avec $X_{\mathcal{A}\mathcal{B}}(y) = X_{\mathcal{B}}(y) - X_{\mathcal{A}}(y)$.

Pour symétriser, il faut faire de même en échangeant les rôles de \mathcal{A} et \mathcal{B} , c'est-à-dire en considérant la projection sur la droite $x = X_{\mathcal{A}}(y)$ du point “proximal” $(X_{\mathcal{B}}(y_{\mathbf{p}_{\mathcal{B}}}), y_{\mathbf{p}_{\mathcal{B}}})$ de la droite $x = X_{\mathcal{B}}(y)$.

Cela consiste donc à calculer $d_{\mathbf{p}_{\mathcal{B}}}^{\mathcal{A}} = \frac{X_{\mathcal{A}\mathcal{B}}(y_{\mathbf{p}_{\mathcal{B}}})}{\sqrt{1+\lambda_{\mathcal{A}}^2}}$.

En prenant la moyenne, $\frac{d_{\mathbf{d}_{\mathcal{A}}}^{\mathcal{B}} + d_{\mathbf{p}_{\mathcal{B}}}^{\mathcal{A}}}{2}$, on obtient une mesure que l'on peut qualifier de décalage transversal. Ce calcul est, de toute évidence, nettement plus lourd que celui de $\delta_{x\mathcal{AB}}$ et, en pratique, on constate que cela n'apporte pas d'amélioration notable dans la construction des axes. Dans le cas le plus courant de chaînes plutôt verticales, $\lambda_{\mathcal{A}}$ et $\lambda_{\mathcal{B}}$ petits, les différences sont en effet négligeables, d'autant plus qu'en général il y a un couplage $(\mathcal{A}, \mathcal{B})$ réalisant une distance nettement moindre à celles pour les couplages concurrents.

Dans le cas de chaînes sensiblement inclinées et plutôt parallèles, $\lambda_{\mathcal{A}}$ et $\lambda_{\mathcal{B}}$ de même signe et relativement grands en valeur absolue, le décalage horizontal $\delta_{x\mathcal{AB}}$ est supérieur au décalage transversal, jusqu'à un facteur 2 pour $\lambda_{\mathcal{A}} = \lambda_{\mathcal{B}} = \pm 1$. Heureusement, l'ensemble des axes a souvent une direction approximativement commune, ce qui fait qu'à peu près le même facteur s'applique à tous les couplages envisagés, ou alors les axes sont divergents, ce qui implique une bonne séparation de toute façon.

Enfin, il reste le cas où $\lambda_{\mathcal{A}}$ et $\lambda_{\mathcal{B}}$ sont relativement grands en valeur absolue mais de signes opposés, c'est-à-dire une brisure très forte de l'axe potentiel. Cela ne se rencontre pas dans le chaînage des doigts, sauf cas du pouce replié pour lequel on n'a pas de problème de séparation. En

général, une des deux chaînes est un artefact périphérique et cela ne perturbe pas la construction des axes, même si cet artefact est finalement raccordé à l'extrémité d'un axe.

De même, on doit considérer $e_{\mathbf{d}_A}^B$, la distance entre, d'une part, la projection orthogonale de $(X_A(y_{\mathbf{d}_A}), y_{\mathbf{d}_A})$ sur $x = X_B(y)$ et, d'autre part, le point $(X_B(y_{\mathbf{p}_B}), y_{\mathbf{p}_B})$. Cette distance a pour expression $e_{\mathbf{d}_A}^B = \lambda_B \cdot d_{\mathbf{d}_A}^B + \sqrt{1 + \lambda_B^2} \cdot (y_{\mathbf{p}_B} - y_{\mathbf{d}_A})$, en remarquant que son premier terme peut être positif ou négatif alors que le second est toujours positif.

Pour symétriser, on calcule aussi $e_{\mathbf{p}_B}^A = \lambda_A \cdot d_{\mathbf{p}_B}^A + \sqrt{1 + \lambda_A^2} \cdot (y_{\mathbf{p}_B} - y_{\mathbf{d}_A})$ et un espacement longitudinal moyen $\frac{e_{\mathbf{d}_A}^B + e_{\mathbf{p}_B}^A}{2}$ dont l'analyse comparative avec δ_{yAB} n'est pas aisée. On peut seulement remarquer que son influence est nettement minorée dans le calcul de $D(\mathcal{A}, \mathcal{B})$.

On peut aussi noter que le rapport $d_{\mathbf{d}_A}^B / e_{\mathbf{d}_A}^B$ est la tangente de l'angle entre le segment qui joint les points $(X_A(y_{\mathbf{d}_A}), y_{\mathbf{d}_A})$ et $(X_B(y_{\mathbf{p}_B}), y_{\mathbf{p}_B})$, d'une part et la droite $x = X_B(y)$, d'autre part. Cette tangente intervient dans la mesure [Sau93] de la colinéarité du raccordement en \mathbf{p}_B et on a aussi une construction symétrique pour \mathbf{d}_A . Les deux angles sont de même signe pour un raccordement en "zig-zag" et leur somme est donc relativement pénalisante. Au contraire, les deux angles sont de signes opposés quand les axes convergent pour un raccordement en "chanfrein", voir figure 2.14, page 52. Néanmoins, on constate que ces angles vont diminuer si on augmente $y_{\mathbf{p}_B} - y_{\mathbf{d}_A}$, ce qui n'en fait pas des mesures très pertinentes.

D'autre part, pour caractériser " \mathcal{B} en dessous de \mathcal{A} ", on devrait logiquement considérer un espacement longitudinal positif, plutôt que $\delta_{yAB} > 0$. C'est a priori nettement différent et un espacement longitudinal positif est compatible avec $\delta_{yAB} < 0$. Mais les règles de distance minimale appliquées lors de l'appariement horizontal interdisent l'existence d'une telle configuration pour deux chaînes pouvant se rattacher à un même axe. On utilisera donc le critère $\delta_{yAB} > 0$, beaucoup plus simple.

Bibliographie

- [AB94] R. Adams and L. Bischof, “Seeded region growing,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641–647, 1994.
- [AFdLGMFAL04] S. Aja-Fernández, R. de Luis-García, M. Martín-Fernández, and C. Alberola-López, “A computational TW3 classifier for skeletal maturity assessment. a computing with words approach,” *J. of Biomedical Informatics*, vol. 37, no. 2, pp. 99–107, 2004.
- [ÅH96] K. Åström and A. Heyden, “Stochastic modelling and analysis of sub-pixel edge detection,” in *Proc. of the 13th International Conference on Pattern Recognition*, vol. 2, 1996, pp. 86–90.
- [AK96] Y. Amit and A. Kong, “Graphical templates for model registration,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 3, pp. 225–236, 1996.
- [BA83] P. J. Burt and E. H. Adelson, “The Laplacian pyramid as a compact image code,” *IEEE Trans. on Communications*, vol. 31, no. 4, pp. 532–540, 1983.
- [Bal81] D. H. Ballard, “Generalizing the Hough transform to detect arbitrary shapes,” *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, 1981.
- [BC98] H. Brönnimann and P. Chassaignet, “Techniques multi-résolutions coopératives pour l’analyse structurelle des os de la main,” LIX, École Polytechnique, rapport de recherche LIX/RR/98/01, 1998.
- [Ber84] V. Berzins, “Accuracy of Laplacian edge detectors,” *Computer Vision, Graphics and Image Processing*, vol. 27, no. 2, pp. 195–210, 1984.
- [Ber87] F. Bergholm, “Edge focusing,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 9, no. 6, pp. 726–741, 1987.
- [BG87] J. Bigün and G. Granlund, “Optimal orientation detection of linear symmetry,” in *Proc. of the First International Conference on Computer Vision*, 1987, pp. 433–438.
- [Big88] J. Bigün, “Recognition of local symmetries in gray value images by harmonic functions,” in *Proc. of the 9th International Conference on Pattern Recognition*, 1988, pp. 345–347.
- [BM92] P. J. Besl and N. D. McKay, “A method for registration of 3-D shapes,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.
- [BS94] K. L. Boyer and S. Sarkar, “On the localization performance measure and optimal edge detection,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 16, no. 1, pp. 106–108, 1994.
- [BT00] A. D. Brett and C. J. Taylor, “A method of automated landmark generation for automated 3D PDM construction,” *Image and Vision Computing*, vol. 18, no. 9, pp. 739–748, 2000.

- [Can83] J. F. Canny, "Finding edges and lines in images," Artificial Intelligence Lab., MIT, Technical Report AI-TR-720, 1983.
- [Can86] ——, "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
- [CC93] L. D. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-D and 3-D images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1131–1147, 1993.
- [CCCD93] V. Caselles, F. Catte, T. Coll, and F. Dibos, "A geometric model for active contours," *Numerische Mathematik*, vol. 66, pp. 1–31, 1993.
- [CD95] G. W. Cook and E. J. Delp, "Multiresolution sequential edge linking," in *Proc. of the International Conference on Image Processing*, 1995, pp. 41–44.
- [CET98] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," in *Proc. of the 5th European Conference on Computer Vision*, ser. Lecture Notes in Computer Science, vol. 1407, no. 2, 1998, pp. 484–498.
- [CGE94] P. Chassaignet, B. Giai-Checa, and A. El Karimi, "Contribution à la détermination automatique de l'âge osseux à partir de radiographies de la main," LIX, École Polytechnique, rapport de recherche LIX/RR/94/02, 1994.
- [Cha96] P. Chassaignet, "Modélisation et lecture automatisée de la radiographie de la main de l'enfant," premier rapport d'étape, AP-HP, LIX, École Polytechnique, juillet 1996.
- [CKS97] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International J. of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.
- [CL94] Y. L. Chang and X. Li, "Adaptive image region-growing," *IEEE Trans. on Image Processing*, vol. 3, no. 6, pp. 868–872, 1994.
- [CLMC92] F. Catte, P.-L. Lions, J.-M. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM J. of Numerical Analysis*, vol. 29, no. 1, pp. 182–193, 1992.
- [CM97] G. Cong and S. D. Ma, "Dyadic scale space," *Pattern Recognition*, vol. 30, no. 2, pp. 339–352, 1997.
- [CNHS99] P. Chassaignet, T. Nitescu, M. Hassan, and R. Stanescu, "Hand radiograph analysis for fully automatic bone age assessment," in *Proc. of Medical Imaging '99*, ser. SPIE Conference Series, vol. 3661, 1999, pp. 1254–1262.
- [Coh91] L. D. Cohen, "On active contour models and balloons," *Computer Vision, Graphics and Image Processing*, vol. 53, no. 2, pp. 211–218, 1991.
- [Cov97] M. Covell, "Eigen-points : control-point location using principal component analysis," in *Proc. of the Second International Conference on Automatic Face and Gesture Recognition*, 1997, pp. 122–127.
- [CRZL04] H. Chui, A. Rangarajan, J. Zhang, and C. M. Leonard, "Unsupervised learning of an atlas from unlabeled point-sets," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 2, pp. 160–172, 2004.
- [CTCG95] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.
- [Del78] A. Delmas, *Anatomie humaine, descriptive, topographique et fonctionnelle*, onzième ed. Masson, 1978, vol. 3.

- [Der87] R. Deriche, “Using Canny’s criteria to derive a recursively implemented optimal edge detector,” *International J. of Computer Vision*, vol. 2, no. 1, pp. 167–187, 1987.
- [Dev95] F. Devernay, “A Non-Maxima Suppression method for edge detection with sub-pixel accuracy,” INRIA, rapport de recherche RR-2724, 1995.
- [DG93] R. Deriche and G. Giraudon, “A computational approach for corner and vertex detection,” *International J. of Computer Vision*, vol. 10, no. 2, pp. 101–124, 1993.
- [DH72] R. O. Duda and P. E. Hart, “Use of the Hough transform to detect lines and curves in pictures,” *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.
- [DL52] R. Debré and M. Lelong, *Pédiatrie*. Flammarion, 1952.
- [dLGAFAAL02] R. de Luis-García, J. I. Arribas, S. Aja-Fernández, and C. Alberola-López, “A neural architecture for bone age assessment,” in *Proc. of the IASTED International Conference on Signal Processing, Pattern Recognition and Applications*, 2002, pp. 161–166.
- [dLGMAAAL03] R. de Luis-García, M. Martín-Fernández, J. I. Arribas, and C. Alberola-López, “A fully automatic algorithm for contour detection of bones in hand radiographs using active contours,” in *Proc. of the International Conference on Image Processing*, vol. 3, 2003, pp. 421–424.
- [DLR77] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *J. of the Royal Statistical Society, Series B*, vol. 39, no. 1, pp. 1–38, 1977.
- [DM98] I. L. Dryden and K. V. Mardia, *Statistical Shape Analysis*. Wiley, 1998.
- [DR92] J. Dolan and E. Riseman, “Computing curvilinear structure by token-based grouping,” in *Proc. of the Conference on Computer Vision and Pattern Recognition*, 1992, pp. 264–270.
- [DTC⁺02] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, “A minimum description length approach to statistical shape modeling,” *IEEE Trans. on Medical Imaging*, vol. 21, no. 5, pp. 525–537, 2002.
- [DV91] P. F. Dijkstra and H. W. Venema, “Metacarpophalangeal pattern profiles : Q-scores for ages from birth to 7 years,” *American J. of Medical Genetics*, vol. 40, pp. 107–114, 1991.
- [DV92] ———, “Metacarpophalangeal pattern profiles : Q-scores for ages from 3 years to adult with epiphyses,” *American J. of Medical Genetics*, vol. 43, pp. 1041–1043, 1992.
- [ED85] P. H. Eichel and E. J. Delp, “Sequential edge detection in correlated random fields,” in *Proc. of the Conference on Computer Vision and Pattern Recognition*, 1985, pp. 14–21.
- [Eff94] N. D. Efford, “Knowledge-based segmentation and feature analysis of hand-wrist radiographs,” University of Leeds, research report 94.31, 1994.
- [EGM⁺94] D. Eberly, R. Gardner, B. S. Morse, S. M. Pizer, and C. Scharlach, “Ridges for image analysis,” *J. of Mathematical Imaging and Vision*, vol. 4, no. 4, pp. 353–373, 1994.
- [El 93] A. El Karimi, “Applications cliniques des techniques de traitement d’image,” rapport de stage d’option, École Polytechnique, juillet 1993.

- [ER67] O. Eklof and H. Ringertz, “A method for assessment of skeletal maturity,” *Annals of Radiology*, vol. 10, pp. 330–336, 1967.
- [EZ96] J. H. Elder and S. W. Zucker, “Scale space localization, blur and contour-based image coding,” in *Proc. of the Conference on Computer Vision and Pattern Recognition*, 1996, pp. 27–34.
- [EZ98] ——, “Local scale control for edge detection and blur estimation,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 7, pp. 699–716, 1998.
- [Fau93] O. Faugeras, *Three-dimensional computer vision : a geometric viewpoint*. MIT Press, 1993, ch. 4, pp. 69–123.
- [FF62] L. R. Ford and D. R. Fulkerson, *Flows in networks*. Princeton University Press, 1962.
- [Fuc94] F. Fuchs, “Analyse structurelle des os du carpe pour la détermination de l’âge osseux,” rapport de stage d’option, École Polytechnique, juillet 1994.
- [GBB95] G. W. Gross, J. M. Boone, and D. M. Bishop, “Pediatric skeletal age : determination with neural networks,” *J. of Radiology*, vol. 195, pp. 689–695, 1995.
- [GI89] D. Gusfield and R. W. Irving, *The stable marriage problem : structure and algorithms*, ser. Foundations of Computing Series. MIT Press, 1989.
- [Gir87] G. Giraudon, “Chaînage efficace de contour,” INRIA, rapport de recherche RR-0605, 1987.
- [GL96] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Johns Hopkins University Press, 1996.
- [GM95] M. Gondran and M. Minoux, *Graphes et Algorithmes*, troisième ed. Eyrolles, 1995.
- [Goo91] C. Goodall, “Procrustes methods in the statistical analysis of shape,” *J. of the Royal Statistical Society, Series B*, vol. 53, no. 2, pp. 285–339, 1991.
- [GP83] W. W. Greulich and S. I. Pyle, *Radiographic atlas of skeletal development of hand and wrist*, 2nd ed. Stanford University Press, 1983.
- [GP93a] J. M. Gauch and S. M. Pizer, “The intensity axis of symmetry and its application to image segmentation,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 8, pp. 753–770, 1993.
- [GP93b] ——, “Multiresolution analysis of ridges and valleys in grey-scale images,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 6, pp. 635–646, 1993.
- [GS62] D. Gale and L. S. Shapley, “College admissions and the stability of marriage,” *The American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, 1962.
- [GSv03] J.-M. Geusebroek, A. W. M. Smeulders, and J. van de Weijer, “Fast anisotropic Gauss filtering,” *IEEE Trans. on Image Processing*, vol. 12, no. 8, pp. 938–943, 2003.
- [Har84] R. Haralick, “Digital step edges from zero crossing of second directional derivatives,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 6, no. 1, pp. 58–68, 1984.
- [HJCT07] C.-W. Hsieh, T.-L. Jong, Y.-H. Chou, and C.-M. Tiu, “Computerized geometric features of carpal bone for bone age estimation,” *Chinese Medical J.*, vol. 120, no. 9, pp. 767–770, 2007.

- [HJT07] C.-W. Hsieh, T.-L. Jong, and C.-M. Tiu, “Bone age estimation based on phalanx information with fuzzy constrain of carpals,” *Medical and Biological Engineering and Computing*, vol. 45, no. 3, pp. 283–295, 2007.
- [HLP07] C.-C. Han, C.-H. Lee, and W.-L. Peng, “Hand radiograph image segmentation using a coarse-to-fine strategy,” *Pattern Recognition*, vol. 40, no. 11, pp. 2994–3004, 2007.
- [HM95] R. Horaud and O. Monga, *Vision par ordinateur – outils fondamentaux*, deuxième ed. Hermès, 1995.
- [HP74] S. L. Horowitz and T. Pavlidis, “Picture segmentation by a directed split-and-merge procedure,” in *Proc. of the 2nd International Conference on Pattern Recognition*, 1974, pp. 424–433.
- [HSN⁺97] M. Hassan, R. Stanescu, T. Nitescu, G. Sebag, C. Garel, and P. Chassagnet, “Biométrie automatisée des phalanges de la main : aide à la détermination de la maturation osseuse,” communication aux Journées Françaises de Radiologie, Paris, novembre 1997.
- [Hub81] P. Huber, *Robust Statistics*, ser. Wiley Series in Probability and Statistics. Wiley, 1981.
- [HZ83] R. A. Hummel and S. W. Zucker, “On the foundation of relaxation labelling processes,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 5, no. 3, pp. 267–287, 1983.
- [JKS95] R. Jain, R. Kasturi, and B. G. Schunk, *Machine vision*. McGraw-Hill, 1995.
- [JM93] D. S. Johnson and C. C. McGeoch, Eds., *Network flows and matching : first DIMACS implementation challenge*, ser. DIMACS Series in Discrete Mathematics and Theoretical Computer Science. American Mathematical Society, 1993, vol. 12.
- [Kan81] T. Kanade, “Recovery of the three-dimensional shape of an object from a single view,” *J. of Artificial Intelligence*, vol. 17, pp. 409–460, 1981.
- [KG94] J. Koplowitz and V. Greco, “On the edge location error for local maximum and zero crossing edge detectors,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 16, no. 12, pp. 1207–1212, 1994.
- [KHEK76] M. Kuwahara, K. Hachimura, S. Eiho, and M. Kinoshita, “Processing of RI-angio-cardiographic images,” in *Digital processing of biomedical images*. Plenum Press, 1976, pp. 187–203.
- [Kir84] D. R. Kirks, *Practical pediatric imaging. Diagnostic radiology of infants and children*, 1st ed. Little, Brown & Company, 1984, ch. 6, pp. 198–201.
- [KK07] H.-J. Kim and W.-Y. Kim, “Computerized bone age assessment using DCT and LDA,” in *Proc. of Mirage 2007, Computer Vision/Computer Graphics Collaboration Techniques*, ser. Lecture Notes in Computer Science, vol. 4418, 2007, pp. 440–448.
- [Koe84] J. J. Koenderink, “The structure of images,” *Biological Cybernetics*, vol. 50, no. 5, pp. 363–370, 1984.
- [Kru56] J. B. Kruskal, “On the shortest spanning subtree of a graph and the traveling salesman problem,” *Proc. of the American Mathematical Society*, vol. 7, no. 1, pp. 48–50, 1956.
- [KSM06] U. Köthe, P. Stelldinger, and H. Meine, “Provably correct edgel linking and subpixel boundary reconstruction,” in *Pattern Recognition, Proc. of 28th*

- DAGM Symposium*, ser. Lecture Notes in Computer Science, vol. 4174, 2006, pp. 81–90.
- [Kuh55] H. W. Kuhn, “The hungarian method for the assignment problem,” *Naval Research Logistic Quarterly*, vol. 2, pp. 83–97, 1955.
- [KWT87] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes, active contour models,” in *Proc. of the First International Conference on Computer Vision*, 1987, pp. 259–268.
- [Lee80] J. S. Lee, “Digital image enhancement and noise filtering by use of local statistics,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 2, no. 2, pp. 165–168, 1980.
- [Lee81] ———, “Refined filtering of image noise using local statistics,” *Computer Graphics and Image Processing*, vol. 15, no. 4, pp. 380–389, 1981.
- [LH89] T. S. Levitt and M. W. Hedgcock, “Model-based analysis of hand radiographs,” in *Proc. of Medical Imaging III : PACS System Design and Evaluation*, ser. SPIE Conference Series, vol. 1093, 1989, pp. 563–570.
- [Lin93] T. Lindeberg, “Effective scale : a natural unit for measuring scale-space lifetime,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1068–1074, 1993.
- [Lin94] ———, *Scale-space theory in computer vision*, ser. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers, 1994.
- [Lin98] ———, “Feature detection with automatic scale selection,” *International J. of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.
- [Low85] D. G. Lowe, *Perceptual organization and visual recognition*. Kluwer Academic Publishers, 1985.
- [Low87] ———, “Three-dimensional object recognition from single two-dimensional images,” *J. of Artificial Intelligence*, vol. 31, no. 3, pp. 355–395, 1987.
- [LW88] Y. Lamdan and H. J. Wolfson, “Geometric hashing : a general and efficient model-based recognition scheme,” in *Proc. of the Second International Conference on Computer Vision*, 1988, pp. 238–249.
- [LXYZ04] P. Lin, F. Zhang, Y. Yang, and C.-X. Zheng, “Carpal-bone feature extraction analysis in skeletal age assessment based on deformable model,” *J. of Computer Science and Technology*, vol. 4, no. 3, pp. 152–156, 2004.
- [Mar76] A. Martelli, “An application of heuristic search methods to edge and contour detection,” *Communications of the ACM*, vol. 19, no. 2, pp. 73–83, 1976.
- [Mar82] D. Marr, *Vision : a computational investigation into the human representation and processing of visual information*. Freeman, 1982.
- [MCRS93] G. Manos, A. Y. Cairns, I. W. Ricketts, and D. Sinclair, “Automatic segmentation of hand-wrist radiographs,” *Image and Vision Computing*, vol. 11, no. 2, pp. 100–111, 1993.
- [MCRS94] ———, “Segmenting radiographs of the hand and wrist,” *Computer Methods and Programs in Biomedicine*, vol. 43, no. 3-4, pp. 227–237, 1994.
- [MD97] P. Montesinos and S. Dattenny, “Sub-pixel accuracy using recursive filtering,” in *Proc. of the 10th Scandinavian Conference on Image Analysis*, 1997, pp. 523–530.

- [MDDS01] A. M. Marques Da Silva, S. Delgado Olabarriaga, C. A. Dietrich, and C. A. A. Schmitz, “On determining a signature for skeletal maturity,” in *Proc. of the First International Conference on Color in Graphics and Image Processing*, 2001, pp. 246–251.
- [MFMFAL03] M. A. Martín-Fernández, M. Martín-Fernández, and C. Alberola-López, “Automatic bone age assessment : a registration approach,” in *Proc. of Medical Imaging '03 : Image Processing*, ser. SPIE Conference Series, vol. 5032, 2003, pp. 1765–1776.
- [MFMMMFAL05] M. A. Martín-Fernández, E. Muñoz-Moreno, M. Martín-Fernández, and C. Alberola-López, “Articulated registration : elastic registration based on a wire-model,” in *Proc. of Medical Imaging '05 : Image Processing*, ser. SPIE Conference Series, vol. 5747, 2005, pp. 182–191.
- [MGPH92] M. F. McNitt-Gray, E. Pietka, and H. K. Huang, “Image preprocessing for Picture Archiving and Communication System,” *Investigative Radiology*, vol. 27, no. 7, pp. 529–535, 1992.
- [MH80] D. Marr and E. Hildreth, “Theory of edge detection,” *Proc. of the Royal Society London, Series B*, vol. 207, pp. 187–217, 1980.
- [MH92] S. Mallat and W. L. Hwang, “Singularity detection and processing with wavelets,” *IEEE Trans. on Information Theory*, vol. 38, no. 2, pp. 617–643, 1992.
- [MN89] D. J. Michael and A. C. Nelson, “HANDX : a model-based system for automatic segmentation of bones from digital hand radiographs,” *IEEE Trans. on Medical Imaging*, vol. 8, no. 1, pp. 64–69, 1989.
- [MN92] R. Mohan and R. Nevatia, “Perceptual organization for scene segmentation and description,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, no. 6, pp. 616–635, 1992.
- [Mon71] U. Montanari, “On the optimal detection of curves in noisy pictures,” *Communications of the ACM*, vol. 14, no. 5, pp. 335–345, 1971.
- [MPL94] B. S. Morse, S. M. Pizer, and A. Liu, “Multiscale medial analysis of medical images,” *Image and Vision Computing*, vol. 12, no. 6, pp. 327–338, 1994.
- [MSC⁺00] S. Mahmoodi, B. S. Sharif, E. G. Chester, J. P. Owen, and R. Lee, “Skeletal growth estimation using radiographic image processing and analysis,” *IEEE Trans. on Information Technology in Biomedicine*, vol. 4, no. 4, pp. 292–297, 2000.
- [MSV95] R. Malladi, J. A. Sethian, and B. C. Vemuri, “Shape modeling with front propagation : a level set approach,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158–175, 1995.
- [MT96] T. McInerney and D. Terzopoulos, “Deformable models in medical image analysis : a survey,” *Medical Image Analysis*, vol. 1, no. 2, pp. 91–108, 1996.
- [MV98] J. B. A. Maintz and M. A. Viergever, “A survey of medical image registration,” *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.
- [MZ92] S. Mallat and S. Zhong, “Characterization of signals from multiscale edges,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–732, 1992.
- [Niț96] T. Nițescu, “Analyse d’images radiographiques pour la détermination automatique de l’âge osseux,” rapport de stage d’option, École Polytechnique, juillet 1996.

- [NvM⁺03] M. Niemeijer, B. van Ginneken, C. A. Maas, F. J. A. Beek, and M. A. Viergever, “Assessing the skeletal age from a hand radiograph : automating the Tanner-Whitehouse method,” in *Proc. of Medical Imaging ’03 : Image Processing*, ser. SPIE Conference Series, vol. 5032, 2003, pp. 1197–1205.
- [PBC⁺94] S. M. Pizer, C. A. Burbeck, J. M. Coggins, D. S. Fritsch, and B. S. Morse, “Object shape before boundary shape : scale-space medial axes,” *J. of Mathematical Imaging and Vision*, vol. 4, no. 3, pp. 303–313, 1994.
- [PEMF98] S. M. Pizer, D. Eberly, B. S. Morse, and D. S. Fritsch, “Zoom-invariant vision of figural shape : the mathematics of cores,” *Computer Vision and Image Understanding*, vol. 69, no. 1, pp. 55–71, 1998.
- [PF98] P. Perona and W. T. Freeman, “Factorization approach to grouping,” in *Proc. of the 5th European Conference on Computer Vision*, ser. Lecture Notes in Computer Science, vol. 1407, 1998, pp. 655–670.
- [PG97] A. K. Poznanski and S. Gartmann, “A bibliography covering the use of metacarpophalangeal pattern profile analysis in bone dysplasias, congenital malformation syndromes, and other disorders,” *J. of Pediatric Radiology*, vol. 27, pp. 358–365, 1997.
- [PGNG72] A. K. Poznanski, S. M. Garn, J. M. Nagy, and J. C. Gall, “Metacarpophalangeal pattern profiles in the evaluation of skeletal malformations,” *J. of Radiology*, vol. 104, pp. 1–11, 1972.
- [PGP⁺01] E. Pietka, A. Gertych, S. Pospiech, F. Cao, H. K. Huang, and V. Gilsanz, “Computer-assisted bone age assessment : image preprocessing and epiphyseal/metaphyseal ROI extraction,” *IEEE Trans. on Medical Imaging*, vol. 20, no. 8, pp. 715–729, 2001.
- [PH95] E. Pietka and H. K. Huang, “Epiphyseal fusion assessment based on wavelets decomposition analysis,” *Computerized Medical Imaging and Graphics*, vol. 19, no. 6, pp. 465–472, 1995.
- [Pie95] E. Pietka, “Computer-assisted bone age assessment based on features automatically extracted from a hand radiograph,” *Computerized Medical Imaging and Graphics*, vol. 19, no. 3, pp. 251–259, 1995.
- [PKKH93] E. Pietka, L. Kaabi, M. L. Kuo, and H. K. Huang, “Feature extraction in carpal-bone analysis,” *IEEE Trans. on Medical Imaging*, vol. 12, no. 1, pp. 44–49, 1993.
- [PL90] T. Pavlidis and Y. Liow, “Integrating region growing and edge detection,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 3, pp. 225–233, 1990.
- [PM90] P. Perona and J. Malik, “Scale-space and edge detection using anisotropic diffusion,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [PMGHH92] E. Pietka, M. F. McNitt-Gray, T. Hall, and H. K. Huang, “Computerized bone analysis of hand radiographs,” in *Proc. of Medical Imaging VI : Image Processing*, ser. SPIE Conference Series, vol. 1652, 1992, pp. 522–528.
- [PMGKH91] E. Pietka, M. F. McNitt-Gray, M. L. Kuo, and H. K. Huang, “Computer-assisted phalangeal analysis in skeletal age assessment,” *IEEE Trans. on Medical Imaging*, vol. 10, no. 4, pp. 616–620, 1991.
- [Poz74] A. K. Poznanski, *The hand in radiologic diagnosis*. W. B. Saunders, 1974.

- [PPKGC03] E. Pietka, S. Pospiech-Kurkowska, A. Gertych, and F. Cao, “Integration of computer assisted bone age assessment with clinical PACS,” *Computerized Medical Imaging and Graphics*, vol. 27, no. 2-3, pp. 217–228, 2003.
- [PSL90] A. Pothen, H. D. Simon, and K.-P. Liou, “Partitioning sparse matrices with eigenvectors of graphs,” *SIAM J. of Matrix Analysis and Applications*, vol. 11, no. 3, pp. 430–452, 1990.
- [PTVF92a] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C : the art of scientific computing*, 2nd ed. Cambridge University Press, 1992, ch. 15.5, pp. 681–688.
- [PTVF92b] ———, *Numerical Recipes in C : the art of scientific computing*, 2nd ed. Cambridge University Press, 1992, ch. 15.4, pp. 671–681.
- [Rad93] P. I. Radeva, “A rule-based approach to hand x-ray image segmentation,” in *Proc. of Computer Analysis of Images and Patterns*, ser. Lecture Notes in Computer Science, vol. 719, 1993, pp. 641–648.
- [RCB97] A. Rangarajan, H. Chui, and F. L. Bookstein, “The softassign procrustes matching algorithm,” in *Proc. of the 15th International Conference on Information Processing in Medical Imaging*, ser. Lecture Notes in Computer Science, vol. 1230, 1997, pp. 29–42.
- [RCN⁺95] M. Rucci, G. Coppini, I. Nicoletti, D. Cheli, and G. Valli, “Automatic analysis of hand radiographs for the assessment of skeletal age : a subsymbolic approach,” *Computers and Biomedical Research*, vol. 28, pp. 239–256, 1995.
- [RCT88] A. F. Roche, W. C. Chumlea, and D. Thissen, *Assessing the skeletal maturity of the hand-wrist : FELS method*. C. C. Thomas, 1988.
- [Ris78] J. Rissanen, “Modeling by shortest data description,” *Automatica*, vol. 14, pp. 465–471, 1978.
- [Ris89] T. Risse, “Hough transform for line recognition : complexity of evidence accumulation and cluster detection,” *Computer Vision, Graphics and Image Processing*, vol. 46, no. 3, pp. 327–345, 1989.
- [RL87] P. J. Rousseeuw and A. M. Leroy, *Robust regression and oulier detection*, ser. Wiley Series in Probability and Statistics. Wiley, 1987.
- [Roh94] K. Rohr, “Localization properties of direct corner detectors,” *J. of Mathematical Imaging and Vision*, vol. 4, no. 2, pp. 139–150, 1994.
- [RWY90] D. Reisfeld, H. Wolfson, and Y. Yeshurun, “Detection of interest points using symmetry,” in *Proc. of the Third International Conference on Computer Vision*, 1990, pp. 62–65.
- [RWY95] ———, “Context free attentional operators : the generalized symmetry transform,” *International J. of Computer Vision*, vol. 14, no. 2, pp. 119–130, 1995.
- [Sai96] C. Saintillan, “Un modèle de croissance pour les os du carpe,” rapport de stage d’option, École Polytechnique, juillet 1996.
- [Sau93] E. Saund, “Identifying salient circular arcs on curves,” *Computer Vision and Image Understanding*, vol. 58, no. 3, pp. 327–337, 1993.
- [SB93] S. Sarkar and K. L. Boyer, “Perceptual organization in computer vision : a review and a proposal for a classificatory structure,” *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 23, no. 2, pp. 382–399, 1993.
- [SB98] ———, “Quantitative measures of change based on feature organization : eigenvalues and eigenvectors,” *Computer Vision and Image Understanding*, vol. 71, no. 1, pp. 110–136, 1998.

- [SD92] L. H. Staib and J. S. Duncan, “Boundary finding with parametrically deformable models,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, no. 11, pp. 1061–1075, 1992.
- [Sez90] M. I. Sezan, “A peak detection algorithm and its application to histogram-based image data reduction,” *Computer Vision, Graphics and Image Processing*, vol. 49, no. 1, pp. 36–51, 1990.
- [SKBG96] G. Székely, A. Kelemen, C. Brechbühler, and G. Gerig, “Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible fourier contour and surface models,” *Medical Image Analysis*, vol. 1, no. 1, pp. 19–34, 1996.
- [SM97] J. Shi and J. Malik, “Normalized cuts and image segmentation,” in *Proc. of the Conference on Computer Vision and Pattern Recognition*, 1997, pp. 731–737.
- [SMB00] C. Schmid, R. Mohr, and C. Bauckhage, “Evaluation of interest point detectors,” *International J. of Computer Vision*, vol. 37, no. 2, pp. 151–172, 2000.
- [Spr] P. Sprawls, “The physical principles of medical imaging,” online textbook. [Online]. Available : <http://www.sprawls.org/ppmi2>
- [SS04] M. Sezgin and B. Sankur, “Survey over image thresholding techniques and quantitative performance evaluation,” *J. of Electronic Imaging*, vol. 13, no. 1, pp. 146–165, 2004.
- [Ste98] C. Steger, “Evaluation of subpixel line and edge detection precision and accuracy,” in *Proc. of the ISPRS Commission III Symposium*, ser. International Archives of Photogrammetry and Remote Sensing, vol. 32, no. 3/1, 1998, pp. 256–264.
- [Ste00] R.-C. Stefanescu, “Utilisation des contours actifs dans le traitement automatique des radiographies de la main de l’enfant,” rapport de stage d’option, École Polytechnique, juillet 2000.
- [Sto87] G. Stockman, “Object recognition and localization via pose clustering,” *Computer Vision, Graphics and Image Processing*, vol. 40, no. 3, pp. 361–387, 1987.
- [SU88] A. Sha’ashua and S. Ullman, “Structural saliency : the detection of globally salient structures using a locally connected network,” in *Proc. of the Second International Conference on Computer Vision*, 1988, pp. 321–327.
- [SVV90] J. Serrat, J. M. Vitrià, and J. J. Villanueva, “Analysis of x-ray images for bone age assessment,” in *Proc. of Visual Communications and Image Processing ’90*, ser. SPIE Conference Series, vol. 1360, 1990, pp. 716–723.
- [TBA05] C. Tauber, H. Batatia, and A. Ayache, “A general quasi-automatic initialization for snakes : application to ultrasound images,” in *Proc. of the International Conference on Image Processing*, vol. 2, 2005, pp. 806–809.
- [Td90] H. D. Tagare and R. J. P. deFigueiredo, “On the localization performance measure and optimal edge detection,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 12, pp. 1186–1190, 1990.
- [THGC01] J. M. Tanner, M. J. R. Healy, H. Goldstein, and N. Cameron, *Assessment of skeletal maturity and prediction of adult height (TW3 method)*, 3rd ed. W. B. Saunders, 2001.

- [TP91] M. Turk and A. Pentland, “Eigenfaces for recognition,” *J. of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.
- [Tud99] L. Tuduri, “Analyse structurelle des os de la main,” rapport de stage, ENSPM, novembre 1999.
- [TWC⁺83] J. M. Tanner, R. H. Whitehouse, N. Cameron, W. A. Marshall, M. J. R. Healy, and H. Goldstein, *Assessment of skeletal maturity and prediction of adult height (TW2 method)*, 2nd ed. Academic Press, 1983.
- [Vap95] V. Vapnik, *The nature of statistical learning theory*. Springer-Verlag, 1995.
- [Vv94] P. W. Verbeek and L. J. van Vliet, “On the location error of curved edges in low-pass filtered 2-D and 3-D images,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 16, no. 7, pp. 726–733, 1994.
- [WB97] J. Weickert and B. Benhamouda, “Why the Perona-Malik filter works,” Department of Computer Science, University of Copenhagen, Technical Report DIKU-TR-97/22, 1997.
- [Wei97] J. Weickert, “A review of nonlinear diffusion filtering,” in *Proc. of the First International Conference on Scale-Space Theory in Computer Vision*, ser. Lecture Notes in Computer Science, vol. 1252, 1997, pp. 3–28.
- [WH03] S.-Y. Wan and W. E. Higgins, “Symmetric region growing,” *IEEE Trans. on Image Processing*, vol. 12, no. 9, pp. 1007–1015, 2003.
- [Wit83] A. P. Witkin, “Scale space filtering,” in *Proc. of the 8th International Joint Conference on Artificial Intelligence*, vol. 2, 1983, pp. 1019–1022.
- [WL98] Y.-P. Wang and S. L. Lee, “Scale-space derived from B-splines,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 10, pp. 1040–1055, 1998.
- [XP98a] C. Xu and J. L. Prince, “Generalized gradient vector flow external forces for active contours,” *Signal Processing*, vol. 71, no. 2, pp. 131–139, 1998.
- [XP98b] ———, “Snakes, shapes, and gradient vector flow,” *IEEE Trans. on Image Processing*, vol. 7, no. 3, pp. 359–369, 1998.
- [Zad99] L. A. Zadeh, “From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions,” *IEEE Trans. on Circuits and Systems*, vol. 46, no. 1, pp. 105–119, 1999.
- [ZF93] D. Ziou and J.-P. Fabre, “Rotation invariance in edge detection,” in *Proc. of Computer Analysis of Images and Patterns*, ser. Lecture Notes in Computer Science, vol. 719, 1993, pp. 141–148.
- [ZGL07] A. Zhang, A. Gertych, and B. J. Liu, “Automatic bone age assessment for young children from newborn to 7-year-old using carpal bones,” *Computerized Medical Imaging and Graphics*, vol. 31, no. 4-5, pp. 299–310, 2007.

Table des figures

1	Les principales parties de la main et du poignet	7
2	Désignation des parties d'un os long	8
3	Exemples d'ossification des métacarpiens et phalanges proximales en fonction de l'âge	8
4	Exemples d'ossification du carpe en fonction de l'âge	9
5	Désignation des os du carpe (vue schématique)	10
6	Exemples de variation chronologique, pour des âges osseux de l'ordre de 3 à 4 ans	10
7	Principe projectif de la radiographie	13
1.1	Les "vrais" contours ne peuvent pas être caractérisés seulement par un seuillage	18
2.1	Orientation de l'image	34
2.2	Mise en évidence d'une direction principale des contours	36
2.3	Maxima du gradient en fonction du paramètre d'échelle α	38
2.4	Petites variations du paramètre d'échelle α	40
2.5	Histogramme des maxima du gradient et seuils de détection, pour $\alpha = 0.25$	41
2.6	Histogramme des maxima du gradient en fonction du paramètre d'échelle α	42
2.7	Nombre de maxima du gradient en fonction du paramètre d'échelle α	43
2.8	Seuillage et appariement des contours (échelle $\alpha = 0.25$)	44
2.9	Exemples de construction des points médians	45
2.10	Nombre de points médians en fonction de l'échelle α	47
2.11	Exemples de chaînage par connexité	49
2.12	Priorités de chaînage pour un pixel médian	50
2.13	Raccordement naturel des chaînes d'axes	51
2.14	Continuité de deux chaînes d'axe	52
2.15	Exemple de raccordement difficile	58
2.16	Découpage des chaînes parasites	59
2.17	Résultats obtenus avec une image insuffisamment contrastée	62
2.18	Exemples de profils des niveaux de gris dans une image	63
2.19	Seuils de détection des contours, pour l'image de la fig. 2.17	63
2.20	Nombre de points médians en fonction de l'échelle	63
2.21	Histogrammes des valeurs du module du gradient	65
2.22	Profil théorique de l'épaisseur d'absorption d'un os	66
2.23	Définitions dans le contour d'un "coin" selon la caractérisation employée	67
2.24	Mesure de la largeur d'une métaphyse	68
2.25	Maxima du gradient en fonction du paramètre d'échelle	69
2.26	Sélection des contours à partir des axes	71
2.27	Contours longitudinaux à échelle fine ($\alpha \approx 1.6$) et nouveaux axes	73
2.28	Variations de la dérivée le long de l'axe et segmentation résultante	75
2.29	Longueurs de la première phalange et du métacarpien, pour un axe $j > 1$	76
2.30	Exemples de signatures plus complexes	77

3.1 Histogramme des scores, tous os confondus, pour les images de “bonne” qualité	91
3.2 Visualisation des écarts relatifs pour l’image d’exemple	92
3.3 Exemple de variations dans la segmentation, en fonction des paramètres de normalité	95
3.4 Histogramme des différences relatives, entre les jeux de paramètres 0 % et 5 %, sur les métacarpiens	98
3.5 Histogramme des différences relatives, entre les jeux de paramètres 0 % et 5 %, sur les phalanges distales	98
3.6 Exemple d’extrémités concaves sur une phalange proximale	104
3.7 Scores obtenus selon l’âge	107
3.8 Relation entre scores effectifs et scores a priori	107
3.9 Ossification des articulations en fonction de l’âge osseux	109
3.10 Médianes de l’ossification des articulations en fonction de l’âge osseux	109
3.11 Profil métacarpo-phalangien pour les images 188.1 et 424.1	111
3.12 Trois premiers modes de covariance des axes, avec un recalage standard	112
3.13 Premiers modes, avec recalage sur les 3 phalanges proximales centrales	113
3.14 Quelques modes particuliers, avec recalage sur les 3 métacarpiens centraux	114
3.15 Exemple de difficulté de mise en œuvre d’un a priori statistique	116
4.1 Fonctions f_k de pondération selon la distance au point de référence	121
4.2 Cas de points de référence en dehors de l’os	121
4.3 Cas d’une chaîne de maxima partagée par deux contours	122
4.4 Cas du chevauchement des régions osseuses	123
4.5 Variations de E_K^2 pour des dipôles	126
4.6 Comparaison des diverses méthodes de recalage d’un point de référence	127
4.7 Exemple de placement des points de référence	133
4.8 Exemples de correction des extrémités proximales des métacarpiens	134
A.1 Voisins pour l’algorithme de NMS	142
A.2 Fonctions D_k de pondération selon la distance au point de référence	151
A.3 Axes de symétrie, par une caractérisation dense, en fonction du rayon r_γ	153
A.4 Comparaison des trois techniques de caractérisation des points d’axes	153
A.5 Symétries à une échelle de détection	154
A.6 Symétries à une échelle de localisation	155
A.7 Axes de symétrie des contours seuls en fonction du rayon r_γ	156
B.1 Détection des axes	179
B.2 Identification des axes, nouveaux appariements et médians	179
B.3 Segmentation résultante	180
B.4 Segmentation obtenue pour les images 142.1 et 144.1	181
B.5 Segmentation obtenue pour les images 155.1 et 156.1	181
B.6 Segmentation obtenue pour les images 161.1 et 164.1	182
B.7 Segmentation obtenue pour les images 170.1 et 173.1	182
B.8 Segmentation obtenue pour les images 188.1 et 203.1	183
B.9 Segmentation obtenue pour les images 221.1 et 227.1	183
B.10 Segmentation obtenue pour les images 230.1 et 232.1	184
B.11 Segmentation obtenue pour les images 258.1 et 292.1	184
B.12 Segmentation obtenue pour les images 424.1 et 434.1	185
B.13 Segmentation obtenue pour les images 533.1 et 666.2	185
B.14 Segmentation obtenue pour les images 671.1 et 727.1	186
B.15 Segmentation obtenue pour les images 731.1 et 748.1	186
B.16 Segmentation obtenue pour les images 749.1 et 780.1	187

B.17 Segmentation obtenue pour les images 810.1 et 819.1 187

Liste des tableaux

1	Âges normaux d'ossification des épiphyses, selon [Del78]	8
2	Âges normaux de début d'ossification du carpe, selon [Del78] et [DL52]	9
1.1	Répartition par qualité d'image et de dossier	31
2.1	Valeurs des coefficients “normaux” de longueur relative des os et des axes	76
3.1	Aboutissement du traitement des axes des doigts	85
3.2	Nombre d’itérations lors de la localisation des axes	86
3.3	Nombre de mesures aberrantes par image	88
3.4	Nombre de mesures aberrantes, os par os, pour les images de “bonne” qualité	88
3.5	Moyenne (en gras) et écart-type du score de segmentation \tilde{q}_{ij} , os par os	89
3.6	Moyenne (en gras) et écart-type du score de segmentation, par groupe d’os	90
3.7	Quantiles des scores, par groupe d’os, pour les images de “bonne” qualité	90
3.8	Scores de segmentation, os par os, pour l’image d’exemple	90
3.9	Nombre d’images exploitables selon les jeux de paramètres à comparer	94
3.10	Variation des mesures résultantes (en pixels), os par os, pour l’image d’exemple	95
3.11	Estimation de la qualité de la segmentation selon les jeux de paramètres	96
3.12	Nombre de mesures différentes, os par os, entre les jeux de paramètres 0 % et 5 %	97
3.13	Nombre de mesures différentes, os par os, entre les jeux de paramètres 0 % et 10 %	97
3.14	Nombre de mesures différentes, os par os, entre les jeux de paramètres 5 % et 10 %	97
3.15	Nombre de mesures différentes, os par os, entre les jeux de paramètres 0 % et -5 %	97
3.16	Nombre de mesures différentes par image, entre les jeux de paramètres 0 % et 5 %	99
3.17	Nombre de mesures différentes par image, entre les jeux de paramètres 0 % et 10 %	99
3.18	Nombre de mesures différentes par image, entre les jeux de paramètres 5 % et 10 %	99
3.19	Nombre de mesures différentes par image, entre les jeux de paramètres 0 % et -5 %	99
3.20	Corrélation entre deux mesures pour les jeux de paramètres 0 % et 5 %	100
3.21	Corrélation entre deux mesures pour les jeux de paramètres 0 % et 10 %	100
3.22	Corrélation entre deux mesures pour les jeux de paramètres 5 % et 10 %	101
3.23	Corrélation entre deux mesures pour les jeux de paramètres 0 % et -5 %	101
3.24	Scores a priori sur un échantillon de 28 images	105
3.25	Scores effectifs sur l’échantillon	106

Annexe B

Illustrations en couleurs

Cette page est intentionnellement laissée vide.

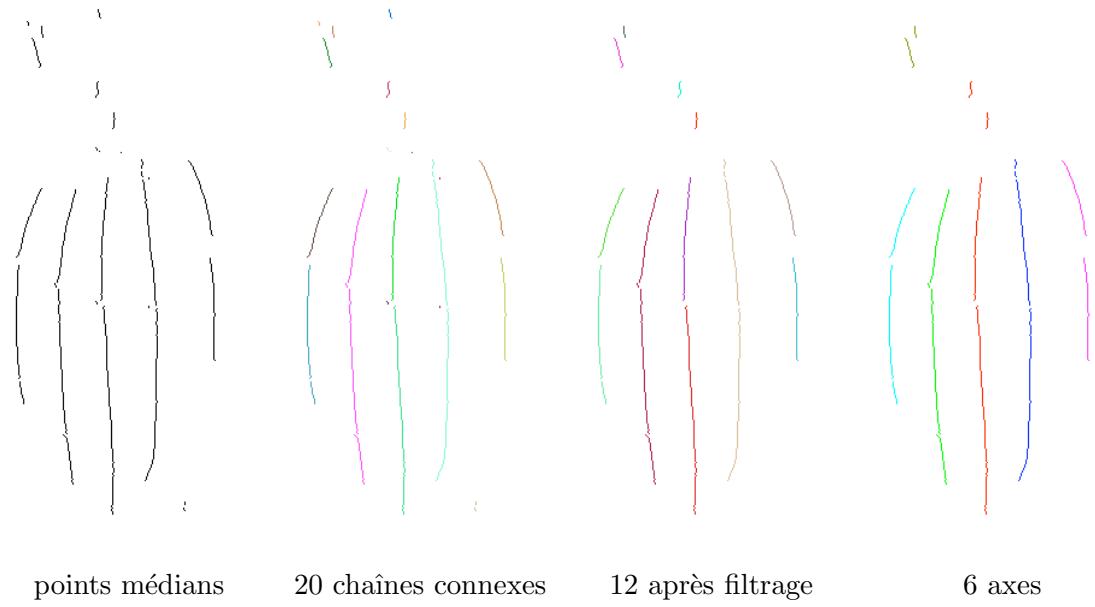


FIG. B.1 – Détection des axes

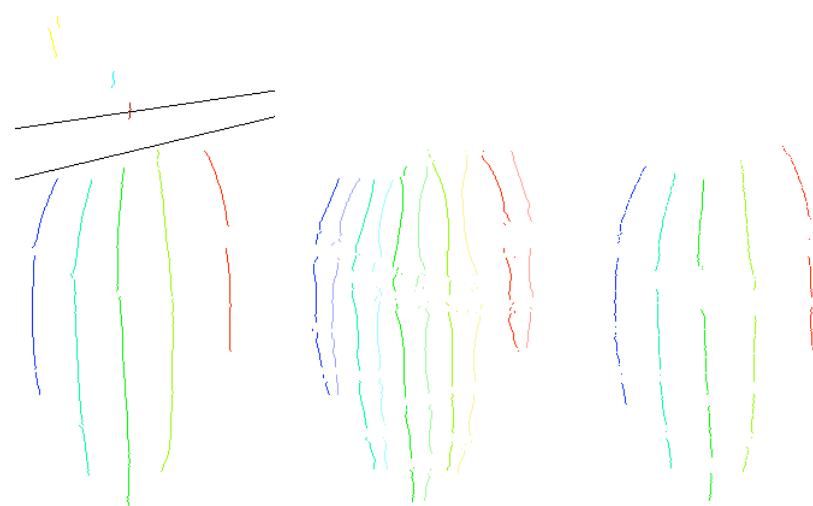


FIG. B.2 – Identification des axes, nouveaux appariements et médians

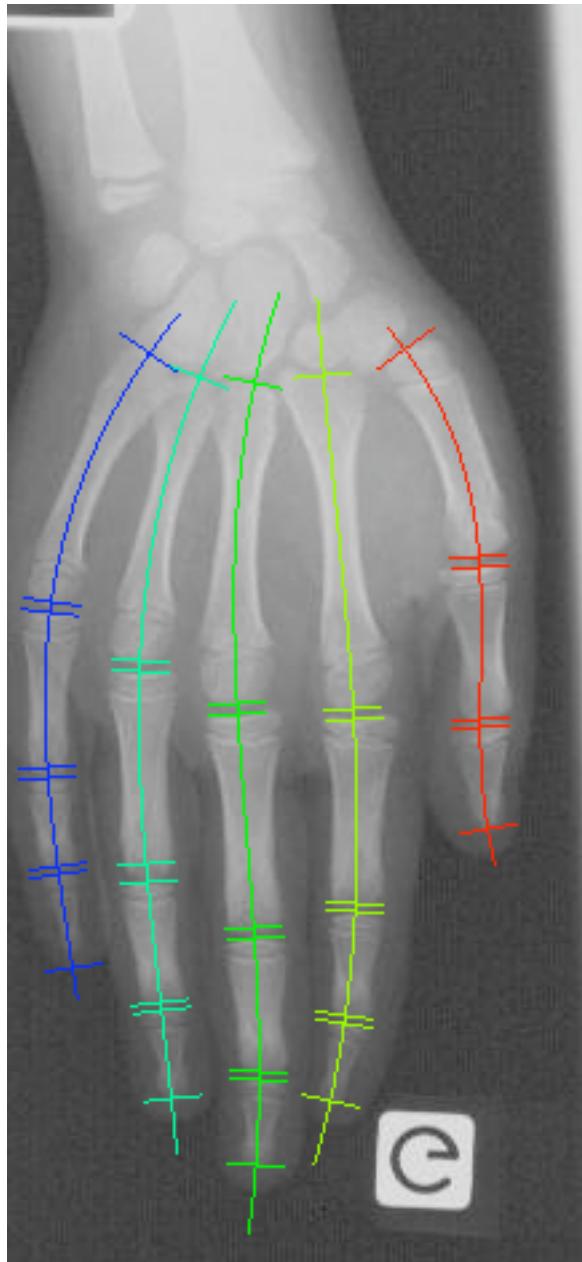


FIG. B.3 – Segmentation résultante

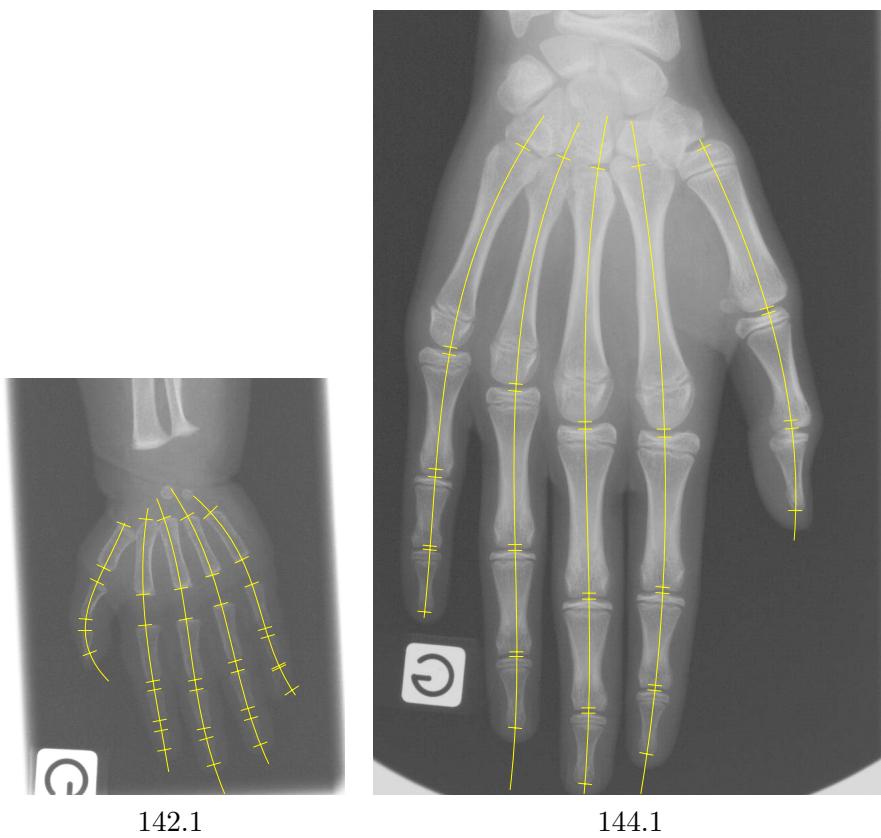


FIG. B.4 – Segmentation obtenue pour les images 142.1 et 144.1

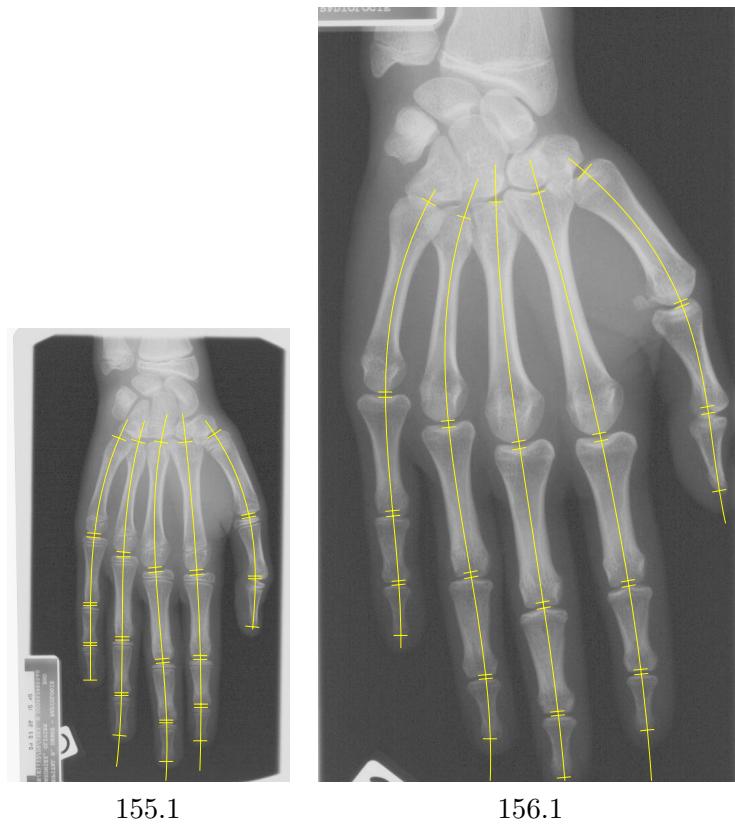
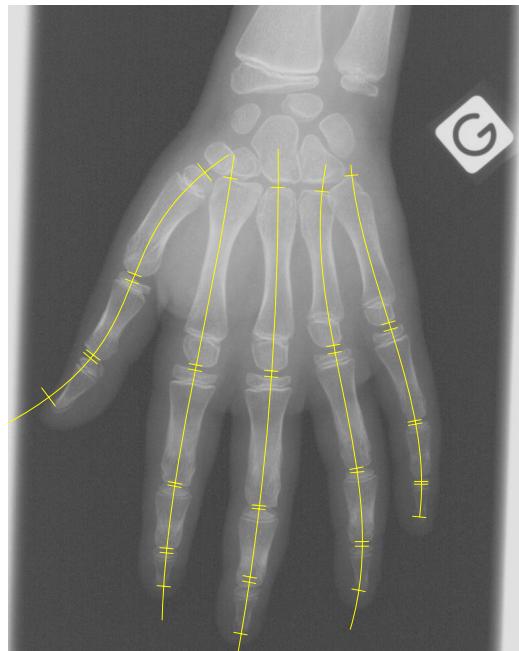
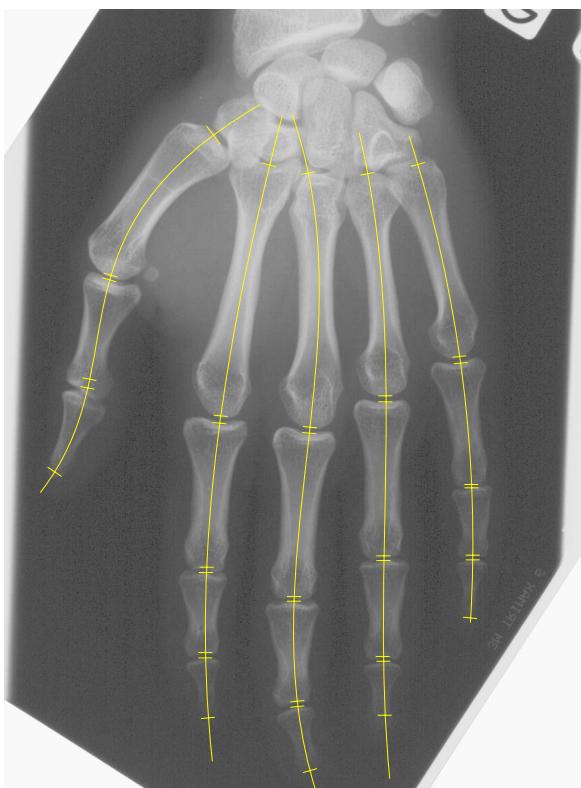


FIG. B.5 – Segmentation obtenue pour les images 155.1 et 156.1

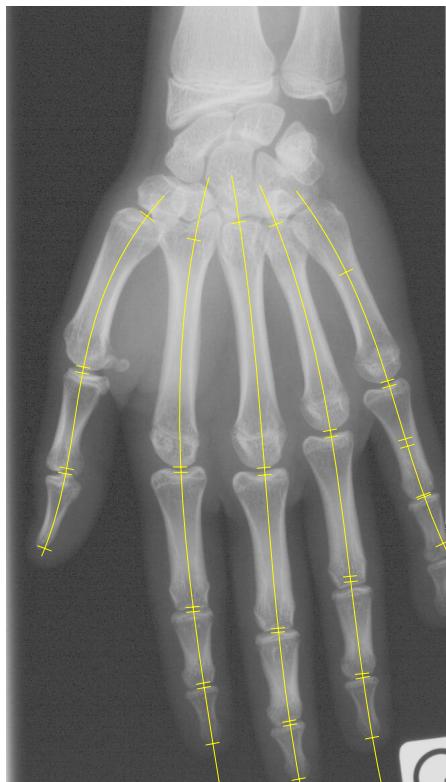


161.1

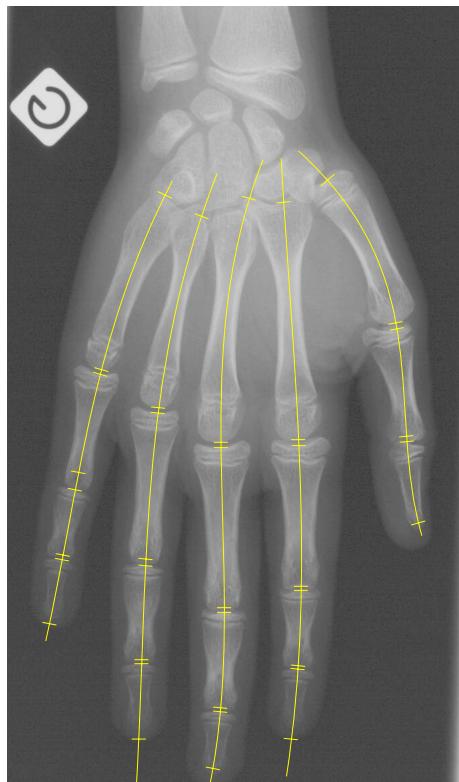


164.1

FIG. B.6 – Segmentation obtenue pour les images 161.1 et 164.1



170.1



173.1

FIG. B.7 – Segmentation obtenue pour les images 170.1 et 173.1

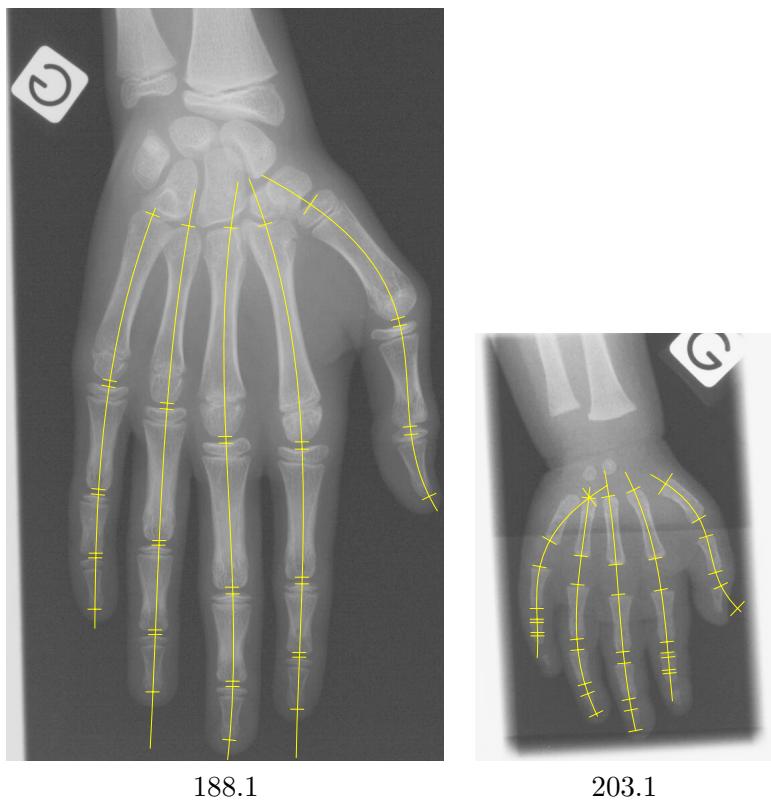


FIG. B.8 – Segmentation obtenue pour les images 188.1 et 203.1

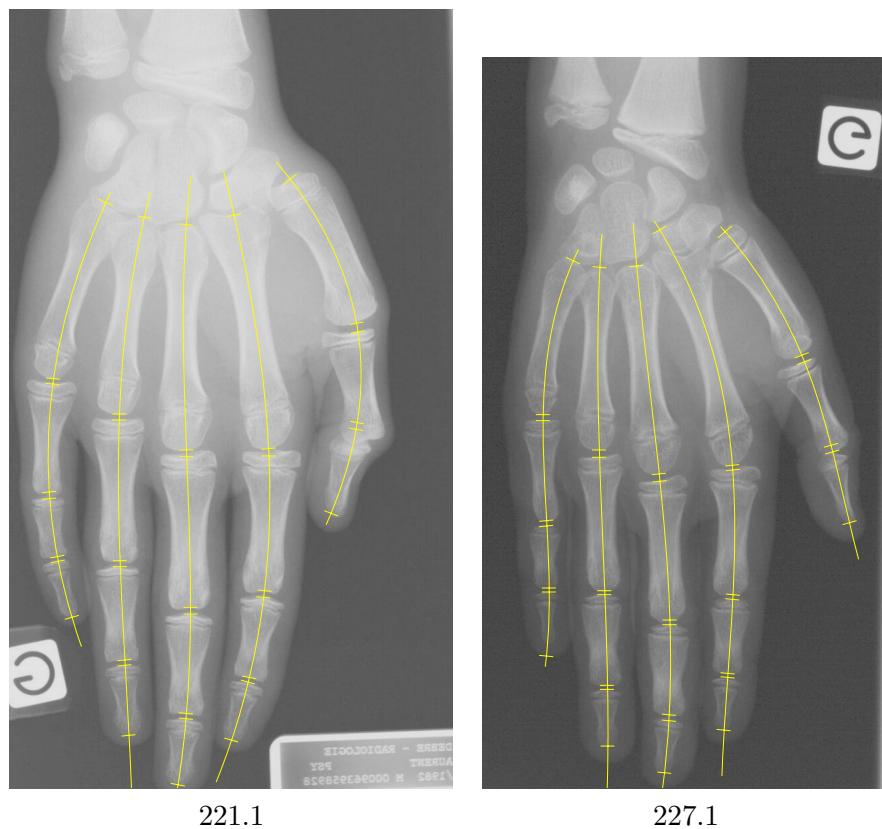


FIG. B.9 – Segmentation obtenue pour les images 221.1 et 227.1

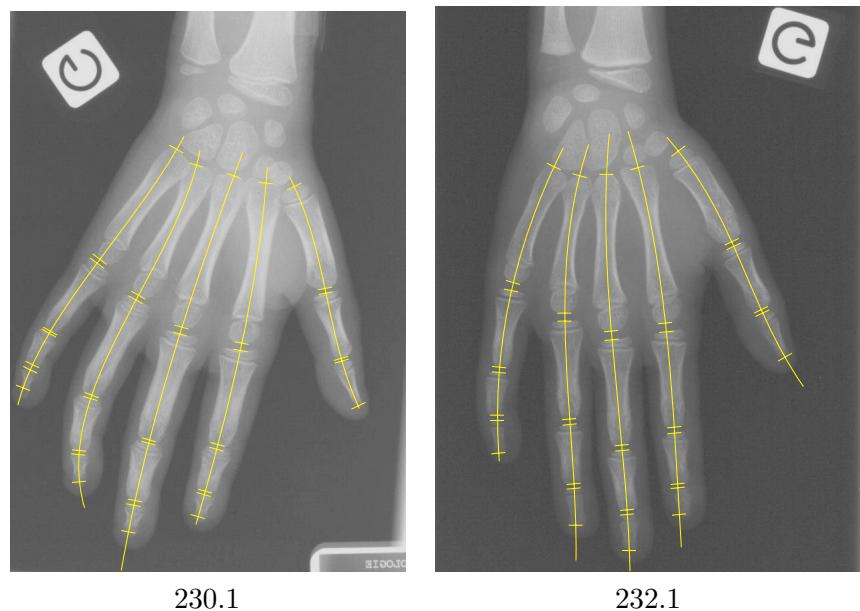


FIG. B.10 – Segmentation obtenue pour les images 230.1 et 232.1

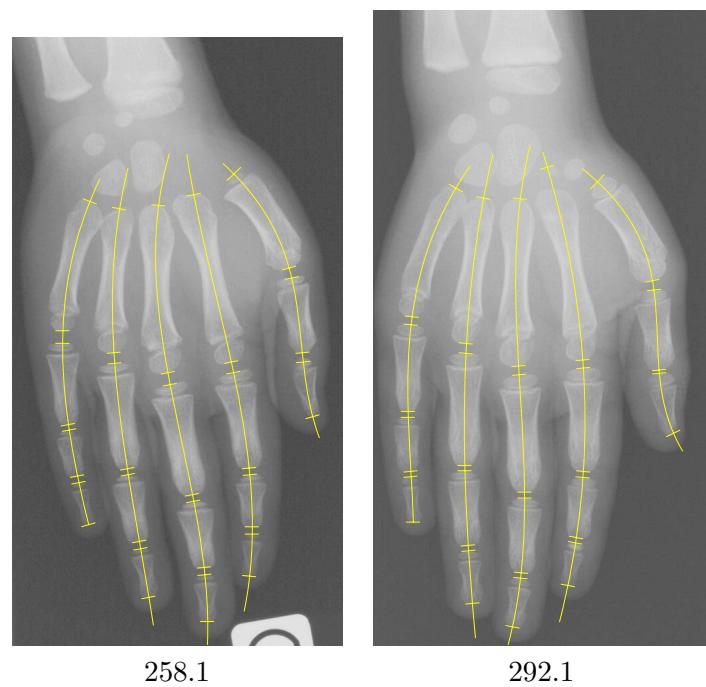


FIG. B.11 – Segmentation obtenue pour les images 258.1 et 292.1

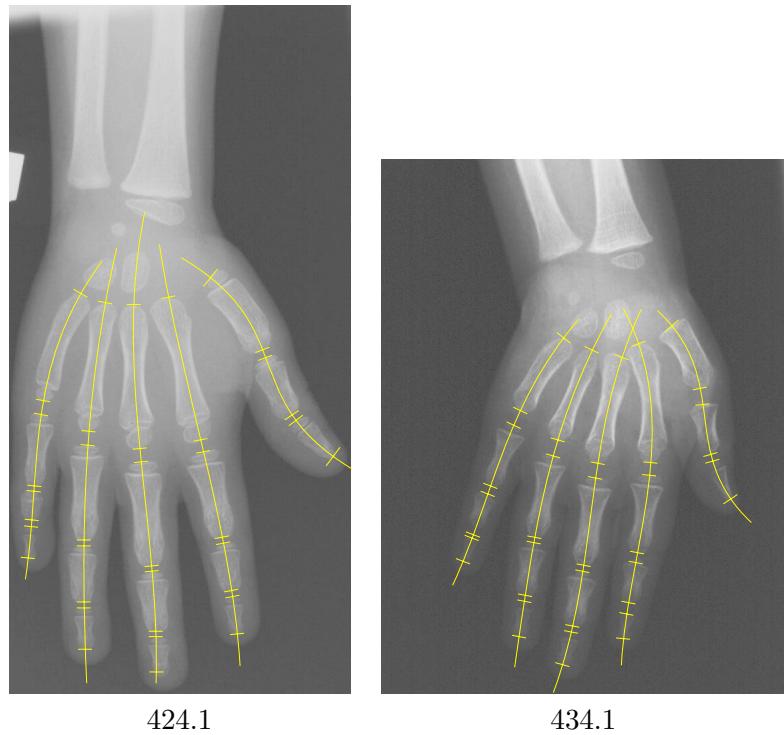


FIG. B.12 – Segmentation obtenue pour les images 424.1 et 434.1

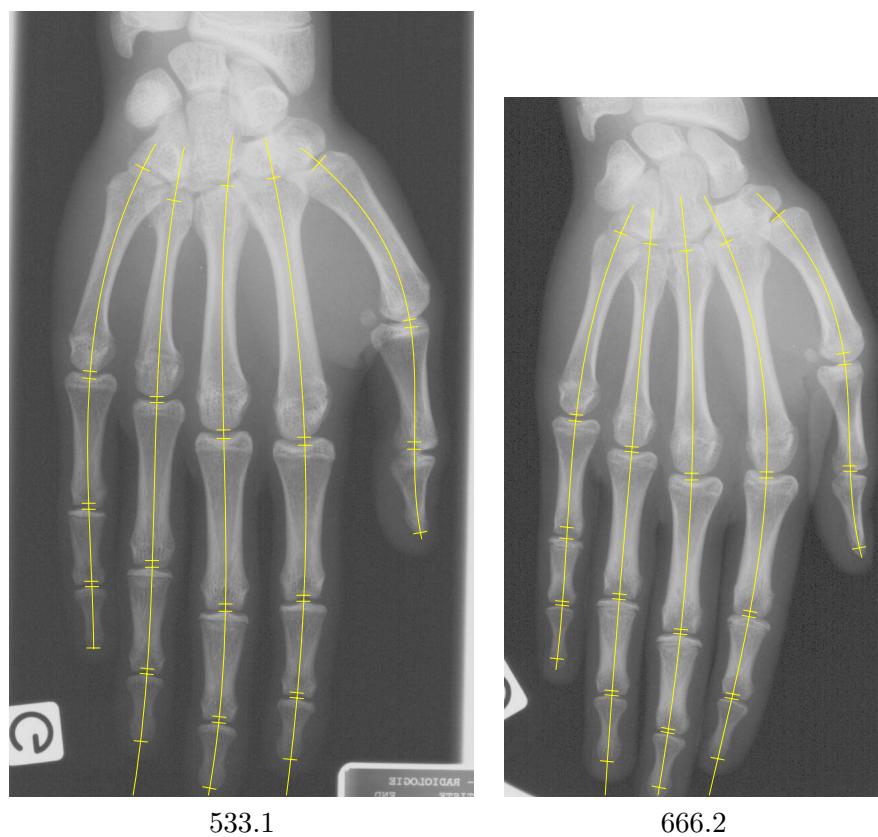


FIG. B.13 – Segmentation obtenue pour les images 533.1 et 666.2

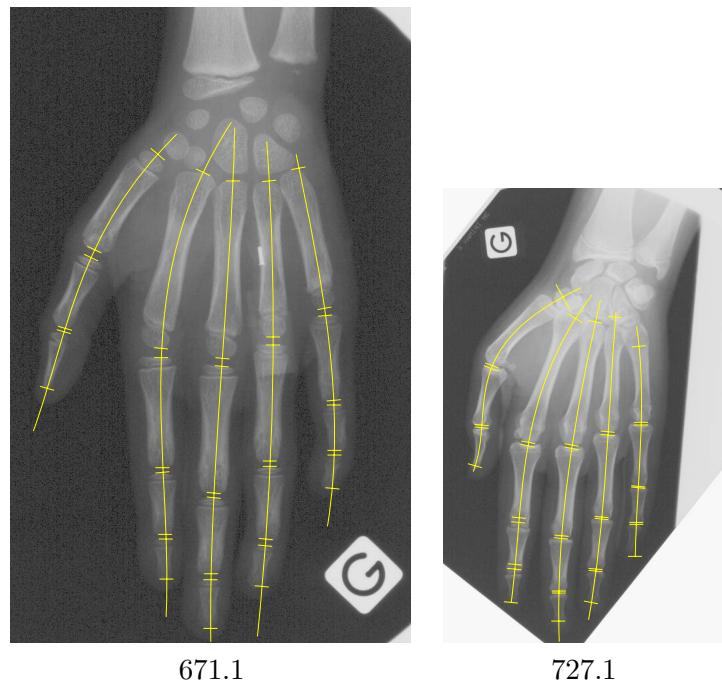


FIG. B.14 – Segmentation obtenue pour les images 671.1 et 727.1

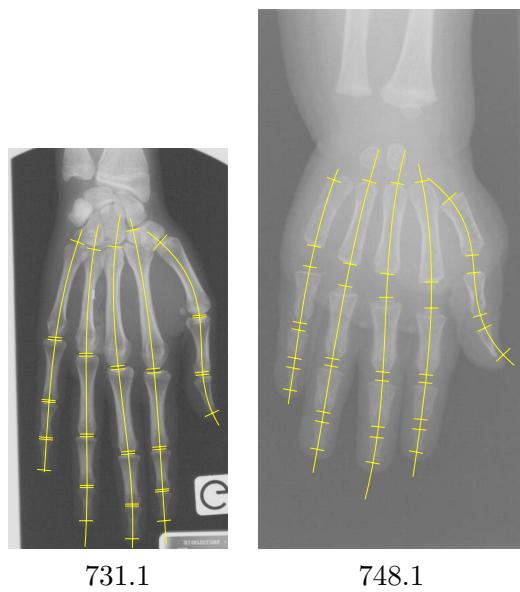


FIG. B.15 – Segmentation obtenue pour les images 731.1 et 748.1

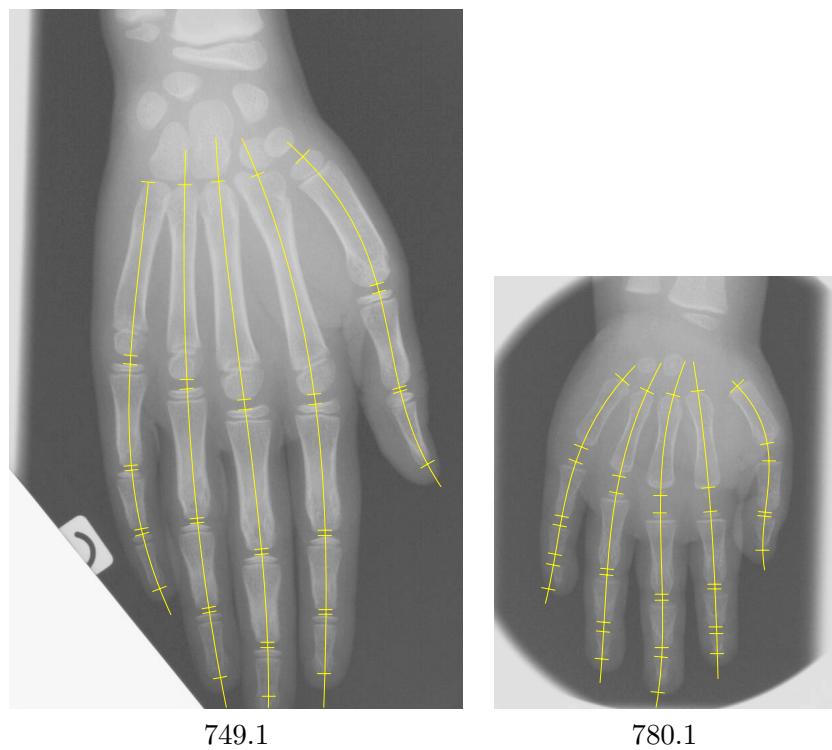


FIG. B.16 – Segmentation obtenue pour les images 749.1 et 780.1

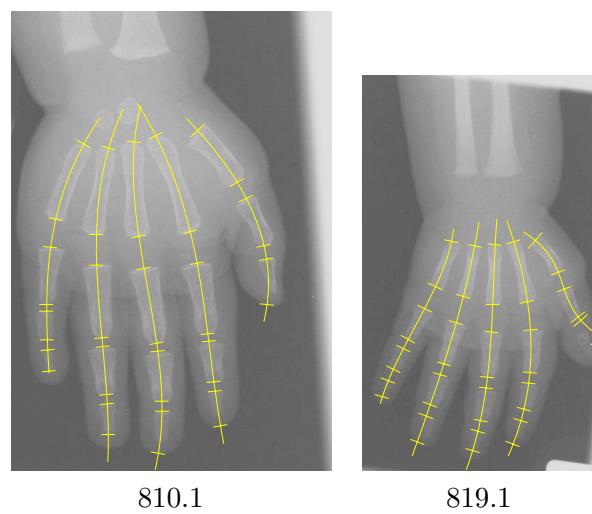


FIG. B.17 – Segmentation obtenue pour les images 810.1 et 819.1