ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2017-ENST-0066

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Signal et Images »

présentée et soutenue publiquement par

Timothée EVAIN

le 6 décembre 2017

Nouvelles méthodes de segmentation en imagerie tomographique volumique à faisceau conique dentaire

Directrice de thèse : **Pr. Isabelle BLOCH**
Directeur de thèse : **Pr. Jamal ATIF**

Jury

Mme Beatriz MARCOTEGUI, Rapportrice
Professeur, Centre de Morphologie Mathématique, MINES ParisTech, Fontainebleau, France

M. Hugues TALBOT, Rapporteur
Professeur, Équipe A3SI, LIGM, UMR 8049, Université Paris-Est Marne-la-Vallée, Paris, France

Mme Corinne LAGORRE, Examinatrice
Maître de Conférences HDR, LISSI, Université Paris-Est Créteil, Vitry-sur-Seine, France

M. Lionel PREVOST, Examinateur
Professeur, LDR, ESIEA, Paris, France

M. Thierry GÉRAUD, Examinateur
Professeur, LRDE, EPITA, Kremlin-Bicêtre, France

M. Xavier RIPOCHE, Examinateur
Chef d'équipe, Équipe Algorithmes, Carestream Dental, Croissy-Beaubourg, France

TELECOM ParisTech
école de l'Institut Mines-Télécom - membre de ParisTech
46 rue Barrault 75013 Paris - (+33) 1 45 81 77 77 - www.telecom-paristech.fr

THÈSE

« la victoire totale ne s'obtient que par des succès partiels. »

–Sénèque, De la colère, II-29.

Résumé

La démocratisation récente de la tomographie à faisceau conique (CBCT) ainsi que les études prouvant son utilité diagnostique ont conduit à un nombre important de cabinets dentaires s'équipant de scanners, et cette modalité est en passe de devenir la référence dans le domaine grâce à l'accès à la dimension de profondeur qui n'existait pas jusqu'ici avec les radiographies traditionnelles. Cependant, du fait de la nouveauté de cette technique d'imagerie, le domaine du traitement des images qui en sont issues est peu développé, les praticiens ont à l'heure actuelle à disposition des outils rudimentaires pour traiter les images, et la majorité de l'analyse se fait encore par l'observation visuelle. Il existe donc une forte demande pour la création d'outils aptes à assister et à renforcer la prise en charge des patients. Les images possèdent plusieurs particularités par rapport à des images de scanners traditionnels. Premièrement, le bruit est accru par une dispersion plus importante et le contraste est plus faible. Deuxièmement, il n'existe pas de calibration de référence de ces scanners, impliquant qu'il n'est pas possible d'utiliser l'échelle de Hounsfield. Troisièmement, le contexte dentaire induit une grande variabilité dans le contenu de l'image, l'aire maxillo-faciale évoluant grandement au cours de la vie. S'ajoute à cela la présence attendue d'éléments métalliques, qui, par interaction physique avec les photons X, conduisent à des dégradations très importantes dans l'image (stries hyper-intenses, zones de perte de données), qui ne sont pas toujours contenues à un niveau local. Les structures d'intérêt recherchées sont principalement les dents et les arches dentaires qu'elles constituent, mais aussi toutes les structures attenantes qui sont impliquées dans la santé bucco-dentaire, comme par exemple les fosses nasales, la mandibule et ses canaux mandibulaires, les os maxillaires, ou encore le pharynx.

Nos travaux sont divisés en deux parties. Dans un premier temps, nous proposons une nouvelle méthode de segmentation semi-automatique des dents, adaptée à un usage par des praticiens. À partir d'une base restreinte de dents annotées, un modèle statistique de forme et d'intensité a été réalisé en suivant la méthodologie des modèles actifs de forme. En utilisant une information de position donnée par un utilisateur, la forme moyenne du modèle est approximativement adaptée aux données étudiées, puis les modes de déformation nous permettent de définir un espace potentiel de forme de la dent. Dans cet espace est alors réalisée une segmentation, dont le problème est formulé comme une énergie composée de termes a priori sur l'intensité, la position et la forme grâce au modèle. Cette énergie est minimisée par un algorithme de coupure de graphe, qui permet un temps de calcul raisonnable ainsi que des interactions de correction aisée. Une étape de post-traitement reposant sur de la morphologie mathématique vient corriger les erreurs résiduelles de la coupure de graphe. L'analyse des résultats montre une bonne qualité de segmentation avec un coefficient de Dice moyen de $0,958$ lorsque les images sont suffisamment résolues (identifiées comme ayant une taille de voxels inférieure à 200 microns), ou de $0,91$ autrement, le tout avec un coût temporel raisonnable (de l'ordre de 15 à 20 minutes pour 32 dents, soit une denture complète) et adapté à des configurations matérielles modestes. Cette méthode est intégrée dans une application avec une interface graphique adaptée aux praticiens, et dotée d'un outil de correction manuelle, qui permet des retouches et une validation rapide des segmentations. Un travail préliminaire sur l'extension de la méthode à une segmentation multi-objets montre qu'il est nécessaire d'obtenir des contraintes de formes plus strictes que celles que nous avons utilisées dans le cadre mono-objet, ainsi que de modifier l'algorithme d'optimisation, car le temps consommé devient incompatible avec la pratique clinique.

La deuxième partie est plus prospective et concerne la création d'un modèle structurel de l'aire maxillo-faciale, dans l'objectif d'obtenir une segmentation reposant sur un système à base de connaissances. Ce modèle est un graphe conceptuel où les structures et leurs interactions sont représentées comme des concepts. Un tel modèle permet de formaliser les connaissances a priori sur l'anatomie, qui

sont susceptibles d'améliorer la reconnaissance des structures en contraignant le problème. Ce modèle est enrichi par des informations sur l'intensité, la forme et les relations spatiales des structures. Nous avons notamment identifié et étendu au domaine 3D les relations d'alignement et "le long de", en utilisant la théorie des ensembles flous, qui permet une modélisation de la nature intrinsèquement imprécise de telles relations.

Abstract

The recent availability of Cone-Beam Computed Tomography (CBCT) and its usefulness for diagnosis proven by clinical studies have convinced many dental offices to acquire this equipment. CBCT is now the new standard dental imaging procedure thanks to the new depth dimension, which was lacking in traditional radiography. However, since this technique is still in its infancy, the CBCT image processing field is underdeveloped, practitionners have only access to basic tools to analyze the images, and most diagnoses are made by visual inspection. Thus there is a strong demand from dentists for new tools able to assist in the patients' healthcare. CBCT images exhibit some specificities compared to traditional CT scans. First, the noise is increased due to increased scattering, and the contrast is globally lower. Secondly, no reference calibration has yet been developed for CBCT scanners, thus it is currently impossible to use the Hounsfield scale. Finally, the dental context implies a great variabilty of image content because the maxillo-facial area can change dramatically during life. Moreover, metallic elements are expected to be found, which lead to important image alterations due to the physical interaction with X photons. These artefacts can be hyper-intense rays or blackened area (loss of data) which can impact the whole dataset. The clinically interesting structures are obviously the teeth and the arches, but also all the adjacent structures which are involved in oral health, like the nasal cavities, the mandibula with its mandibular canals, the maxillar bones or the pharynx.

This thesis is organized in two parts. First, we propose a new semi-automatic segmentation method for each tooth, usable by medical practitioners. From a small annotated teeth database, a statistical model including shape and intensity is created with an active shape model procedure. Using a positional information given by the user, the mean shape is approximately registered to the use case, before using deformation modes to generate a possible region where the tooth can be located. A segmentation is done in this space, formulated as an energy composed by a priori terms on intensity, position and shape thanks to the mean shape. This energy is minimized by a graph-cut algorithm, featuring short computation times and easy interaction for correction. Post-processing based on mathematical morphology allows us to correct any residual missegmentation from the graph-cut. Results show a good segmentation quality with a 0.958 mean Dice coefficient when the image quality is sufficient (i.e. when the voxel size is less than 200 microns), or 0.91 otherwise, for an acceptable time cost (from 15 to 20 minutes for 32 teeth, i.e. a full dentition) and executable on modest hardware configuration. This method is made available through a graphical user interface suitable for dentists together with manual correction and validation tools. A preliminary work extending the method to multi-objects segmentation shows that we need to get more stringent shape contraints than the mean model, as well as to improve the optimization algorithm since the computation time is too long for practical clinical use.

The second, more prospective, part is about the creation of a structural model of the maxillo-facial area, heading toward a global segmentation guided by a knowledge based system. This model is a conceptual graph where organs and their interaction are seen as concepts. Such a model formalizes the prior knowledge on the anatomy, which can improve the structure identification by constraining the segmentation problem. This model includes information on intensity, shape and the spatial relationships between organs. In particular, we identified and extended to 3D the "along" and alignment spatial relations, using fuzzy sets theory, which allows for a modelling of the intrinsically imprecise nature of such relations.

Remerciements

Voici donc les remerciements que j'ai pour tous les gens qui m'ont entouré, de près ou de loin, durant ces années de thèse. Vous avez été nombreux autour de moi, donc mes plus plates excuses si jamais je vous oublie dans les lignes qui suivent, vous n'en êtes pas moins dans mes pensées !

En premier lieu je voudrais remercier Beatriz, Corinne, Lionel, Hugues et Théo, pour avoir accepté d'être mon jury et d'avoir lu ma thèse. Un merci supplémentaire à Beatriz et à Hugues pour leurs rapports détaillés et clairs qui ont permis de rendre le manuscrit plus pertinent et d'ouvrir de nouvelles perspectives pour mes travaux, ce qui m'a été très utile !

Je pense qu'il est difficile d'exprimer à quel point un directeur de thèse est une personne importante dans le bon déroulé d'une thèse et donc dans la vie d'un doctorant, d'un point de vue académique, mais avant tout d'un point de vue humain. À cet égard, je ne saurais jamais assez te remercier Isabelle, d'abord d'avoir accepté de me prendre en thèse, mais aussi et surtout pour tout le soutien que tu m'as donné en me guidant au fur et à mesure sur les routes sinueuses de la recherche et en me rassurant lorsque j'ai profondément douté d'arriver au bout de mon parcours. Le tout toujours avec gentillesse, bienveillance et en souriant, même lorsqu'il a fallu me pousser. Merci pour ces (presque) quatre années qui m'ont fait évoluer et grandir, je mesure maintenant la chance que j'ai eu !

Merci évidemment à toi aussi Jamal pour avoir accepté d'être mon co-directeur. Tes remarques et indications ont été à l'origine de nombreuses progressions dans mes travaux et ma thèse ne serait pas celle qu'elle est aujourd'hui sans toi. Malgré quelques incompréhensions mutuelles sur la méthodologie à adopter, je suis heureux d'avoir fait cette thèse avec ton aide, ce qui m'a fait énormément progresser.

Xavier, merci mille fois d'avoir été là pour encadrer le côté industriel de ces travaux. Le cours des choses a fait que tu as dû assumer seul ma supervision chez Carestream alors que ça n'était pas prévu. De plus, je n'ai jamais ressenti la moindre pression de ta part lorsque les résultats tardaient à arriver alors même que je sais que tu en recevais de ta hiérarchie. Nos discussions m'ont permis à de nombreuses reprises de réorienter ou d'ajuster mes recherches pour correspondre au mieux aux attentes industrielles et pratiques. J'ai passé de très bons moments dans l'entreprise, et tu y es pour une bonne part. Alors merci.

Puisque nous parlons de Carestream, je veux remercier toutes les personnes qui m'ont accompagné tout au long de ces années. Bien évidemment en premier lieu l'équipe Algo avec Delphine, Karima, Leila, Émilie, Gaël R., Jean-Pascal, Luc, Ahmed, Elyes, Pascal, Mamadou, Sylvain et Mohammed, sans oublier Yannick. Je me suis vraiment senti intégré et j'ai pris plaisir à discuter et collaborer avec vous sur nos problématiques respectives.

Merci aussi à toutes les autres équipes que j'ai côtoyé au cours de mes longues pérégrinations de bureau en bureau : Sabrina, Axelle, Sandra, Hervé, Ronan, Gaël V., Samuel, Xavier, Walid, Yazid, Adbelslam, Christophe, Loïc, Philippe, Sébastien, Guillaume, Michel, Laurent, Bilel, Florian, Younesse, Patrick, Romain, Shinji, Guy et tous ceux que j'oublie. Merci pour tous ces excellents moments passés en votre compagnie et mention spéciale à l'équipe de volley !

Merci à toi Aude pour tes précieux conseils qui m'ont aidé à créer une interface plus pratique et plus ergonomique pour mon prototype.

Stéphane, j'ai passé d'excellentes journées lorsque nous partagions un bureau, et plus généralement merci pour ton avis expert et toutes les données que tu m'as fournies, morceaux de musique y compris : keep rocking ragazzo !

Une dédicace toute particulière à mon équipe de drôles de dames : Chloé, Anna, Colombe, Claire et Alice. Votre bureau est probablement le plus chaleureux et à l'ambiance bon enfant que j'ai eu le plaisir d'occuper. L'alternance des discussions aussi bien grivoises que raffinées autour d'une sucrerie ou d'un carré de chocolat était un réel plaisir à partager !

Non content d'avoir fait partie d'une entreprise, j'ai aussi eu le plaisir de faire partie du laboratoire TSI de Télécom. Ma première pensée va à toi, chère Hélène 2D : j'ai adoré partager notre C01, que nous y riions ou que nous y râlions ensemble. N'oublie pas de faire de beaux schémas, et surtout de mettre à la poubelle les boîtes de café vides ! Nos pauses café sur les marches de la cour vont me manquer sois en sûre.

Merci aussi à toi Ana, pour tous les bons moments et toutes les découvertes autour de ta Roumanie natale. J'espère que ton passage en France a été agréable et n'hésite pas à nous faire signe si tu reviens. Je te souhaite tout le meilleur et du succès lors de ton retour : mult noroc !

Le reste de l'équipe Images n'est pas en reste, alors merci Romane, Yongchao, la bande des italiens Alessio/Giovanni/Samy (une bonne dose de Caffe, c'est toujours bon pour les neurones, mais je préfère la pizza personnellement, même si je dois bouger des meubles pour ça), Nicolas, Ivan, Alessandro, et sans oublier Pietro (merci pour toutes tes propositions de relecture de thèse, et pour tout l'investissement pour faire évoluer l'équipe).

Je pense aussi à toute l'équipe 3D et ses joyeux drilles : Émilie, Hélène 3D (spécialiste en cabanes de guirlandes et lapins spatiaux), Malik le panda furieux, Gilles qui a non pas une araignée au plafond mais une équation (du rendu), Sylvain R./Victor/Thibault le triplet geek (plus que les autres en tous cas), Adrien, Jeremy, Elie, Christophe, Jean-Marc et Tamy. J'ai passé d'excellentes pauses et soirées avec vous, qui m'ont bien changé les idées.

Je ne manque pas de remercier aussi tous les autres qui ont compté (même ceux qui font du radar...), à savoir Sonia, Flora, Guillaume Tartav', Sylvain L., Yifan, Gang, Rémi, Abigail, Paul, Antoine, Alasdair, Clément, David, Cécilia, Clara, et Baptiste. Mention spéciale pour l'homme qui cherchait des paraboles et des ballons ovales, j'ai nommé Guillaume T., merci à toi pour toutes ces discussions joyeuses autour du rugby et tous les bons moments en ta compagnie.

Merci aussi à toutes les équipes administratives qui ont toujours été très à l'écoute et serviable, notamment Céline et Teresa, Salimatou et Florence. Merci de m'avoir autant facilité la vie !

Mes pensées les plus affectueuses à toi Thierry, qui de collègue est vite devenu un proche, merci pour les soirées jeux de société et pizzas issues de Marcel. Rendez-vous très vite pour la prochaine.

Enfin, je voudrais aussi remercier ma famille pour avoir toujours été là. Je leur dédie ce manuscrit.

À mes parents, Evelyne et Jean-Marie, pour leur présence réconfortante, discrète mais inconditionnelle.

À mes grands-parents, Marie-Thérèse et Jean, Nicole et Pierre, qui m'ont toujours poussé à aller plus haut.

À mon frère JB qui m'a inspiré et qui m'a montré que rien n'est hors d'atteinte avec de la volonté et du travail.

Et pour finir à ma femme, Élodie, qui m'a soutenu en permanence et durant les moments les plus éprouvants pour me permettre de terminer mes travaux. Sans elle, ce manuscrit n'existerait pas.

Je terminerai cette logorrhée en image, qui résume assez bien ce qu'est finalement un docteur :).

Sommaire

1 Introduction **1**
1.1 Contexte applicatif . . . 1
1.1.1 Les dents dans la pratique clinique . . . 1
1.1.2 Autres structures dans la prise en charge . . . 3
1.2 Les besoins des dentistes et la réponse industrielle . . . 4
1.3 Les enjeux scientifiques . . . 6

2 État de l'art **9**
2.1 Segmentation . . . 9
2.1.1 Sans apprentissage . . . 9
2.1.2 Avec apprentissage . . . 14
2.1.3 Approches globales guidées par des connaissances . . . 21
2.2 Relations spatiales . . . 22
2.3 Conclusion . . . 23

3 Segmentation mono-objet des dents **25**
3.1 Connaissances a priori sur la forme et l'apparence des dents en CBCT . . . 27
3.1.1 Caractéristiques dentaires . . . 27
3.1.2 Choix du modèle . . . 28
3.1.3 Création des modèles . . . 28
3.2 Le modèle statistique comme contrainte spatiale . . . 31
3.2.1 Pré-traitement . . . 32
3.2.2 Recalage du modèle d'apparence moyenne . . . 33
3.2.3 Génération d'informations a priori . . . 35
3.3 Segmentation contrainte de l'image . . . 37
3.3.1 Formulation du problème de segmentation d'une dent . . . 37
3.3.2 Méthode de résolution . . . 40
3.3.3 Post-traitement . . . 42
3.4 Expérimentation . . . 43
3.4.1 Protocole . . . 43
3.4.2 Influence des paramètres et des données . . . 45
3.4.3 Comparaison avec d'autres approches . . . 54
3.4.4 Analyse de la performance sur des volumes de résolution 0,15 mm . . . 60
3.5 Conclusion . . . 64

4 Vers une approche globale par un modèle structurel **67**
4.1 Extension multi-objets de la segmentation . . . 67
4.1.1 Problématique . . . 67
4.1.2 Nouvelle formulation de la segmentation . . . 67
4.1.3 Expérimentation . . . 70
4.1.4 Discussion . . . 71
4.2 Modèle structurel . . . 72
4.2.1 Motivation . . . 72

4.2.2 Formalisation graphique et attributs . . . 72
4.3 Extension de deux relations spatiales en 3D . . . 75
4.3.1 Relation “le long de” . . . 75
4.3.2 Relation d’alignement . . . 83
4.4 Conclusion . . . 87
4.4.1 Utilisation pratique du modèle structurel . . . 87
4.4.2 Amélioration de l’efficacité de la segmentation multi-objets . . . 87

5 Application : Outil de segmentation de l’arche dentaire à destination des praticiens 89
5.1 Mise en œuvre . . . 89
5.2 Interface graphique et outils de correction manuelle . . . 91
5.3 Applications cliniques . . . 97
5.4 Perspectives . . . 99

6 Conclusion et perspectives 103
6.1 Apport de la thèse . . . 103
6.1.1 1er axe : une méthode de segmentation dentaire . . . 103
6.1.2 2e axe : un modèle structurel maxillo-facial . . . 104
6.1.3 3e axe : un logiciel pour praticiens . . . 104
6.2 Perspectives . . . 105
6.2.1 La segmentation en CBCT maxillo-facial . . . 105
6.2.2 Vers une conception globale du problème d’analyse d’images CBCT dentaires . . . 106
6.2.3 Développement des solutions applicatives . . . 107

A Approche morphologique de la segmentation dentaire 109
A.1 Arbres de composantes . . . 109
A.2 Conception initiale du problème . . . 109
A.3 Révision de l’approche . . . 111
A.4 Vers une approche par forme . . . 112

B Exemples de résultats de segmentation 117
B.1 Segmentation mono-objet . . . 117
B.2 Segmentation multi-objets . . . 126

C Caractérisation du canal mandibulaire 129
C.1 Analyse de l’intensité . . . 130
C.2 Analyse de la texture par matrice de co-occurrences . . . 132
C.3 Analyse de la décomposition en ondelettes . . . 135
C.4 Conclusion . . . 139

D Description du modèle structurel 141

E Représentation graphique du modèle structurel 149

F Algorithme de localisation d’arche dentaire 155

Liste des figures

1.1 Présentation du contexte anatomique dentaire. 1
1.2 Système de notation ISO. Denture définitive (entre 11 et 48) et denture temporaire (entre 51 et 85). Cette notation prend le point de vue du dentiste : haut=maxillaire, bas=mandibule, gauche=coté droit du patient et droite=coté gauche du patient. . . . 3
1.3 Anatomie d'une dent (molaire mandibulaire). Crédit : infovisual.info. 3
1.4 Exemple de cliché panoramique ou orthopantomogramme. 4
1.5 Principe de fonctionnement du CBCT comparé au scanner traditionnel (CT). On remarque que l'acquisition d'un plan entier en CBCT conduit à capturer plus de photons diffusés (flèches rouges) que dans l'acquisition coupe par coupe du CT, ce qui affaiblit le contraste et bruite l'image. Les reconstructions finales sont issues des projections de base (CBCT) plutôt que des reconstructions de coupes (CT). Schéma traduit de [108]. 5
1.6 Exemples de différentes acquisitions. Les grands champs de vue possèdent une résolution isotrope de 0,3mm alors que les petits champs permettent une résolution de 0,075mm. 6
1.7 Illustrations des problématiques rencontrées dans le cadre de l'utilisation de la tomographie volumique à faisceau conique pour l'imagerie dentaire. 8

2.1 Approche proposée initialement dans [58]. L'association d'un ensemble de niveaux par dent est proposé dans [57] et la différenciation des dents en contact durant l'occlusion dans [125]. Illustration traduite de [125]. 11
2.2 Champ harmonique généré à partir du canal. Pour chaque série est présentée la zone d'intérêt, le champ harmonique et l'harmonique retenue. (a) Vue axiale, (b) vue sagittale et (c) vue coronale. Illustration de [130]. 13
2.3 Temps de calcul (en secondes) utilisé par les méthodes d'optimisation sur graphe en fonction du nombre de voxels. Notez la performance du *powerwatershed* (en rouge) face à la marche aléatoire (en vert) et à la coupure de graphe (en bleu). Dans le cadre multi-objets, ce gain est multiplié par le nombre d'étiquettes. Illustration de [40]. . . . 14
2.4 Méthode de segmentation par ligne de partage des eaux. (a) Im, (b) MIP, (c) seuillage de MIP, (d) Im_m, (e) Im_c, (f) Im_s, (g) ligne de partage des eaux superposée à Im, (h) résultat final après fusion des bassins de faible aire. Illustration de [110]. 15
2.5 Schéma de l'utilisation des maillages de formes comme contrainte de forme dans une segmentation par marche aléatoire. Illustration de [102]. 16
2.6 Architecture utilisée pour la segmentation des structures cérébrales, à partir du réseau pré-entraîné VGG16. Chaque couche finale synthétise l'information d'un niveau du réseau. Illustration de [127]. 20

3.1 Schéma de la méthode de segmentation dentaire. La partie hors ligne s'effectue en amont afin de générer les modèles nécessaires à l'application des contraintes de la partie en ligne. 26
3.2 Apparence générale des dents dans un volume CBCT. (a) Notons le léger liseré autour de la dent, le ligament parodontal, qui permet de distinguer de l'os cortical. (b) Nous voyons distinctement l'émail (forte intensité), la dentine (intensité modérée) ainsi que la chambre pulpaire prolongée par des canaux radiculaires (faible intensité). (c,d) La présence de métal conduit à des pertes de données sous la forme de raies hyper-intenses et de zones hypo-intenses. 27

3.3 Modèle statistique d'incisive mandibulaire. (a) Forme moyenne ($\bar{\mathbf{M}}$) avec l'isoligne 0 en rouge. (b) Variation associée au premier mode (premier vecteur propre mis sous la forme d'image). L'ajout ou la soustraction de cette variation à $\bar{\mathbf{M}}$ permet d'explorer l'espace des formes selon cet axe. (c,d) Deux exemples de formes binaires issues d'une variation selon le premier mode de déformation. Nous notons ici que ce dernier encode la variabilité de la taille des incisives (dents plus ou moins grandes). (e) Dent d'apparence moyenne $\bar{\mathbf{A}}$. 29

3.4 Distribution d'intensité de la catégorie des incisives mandibulaires (bleu) superposée à celle du fond (rouge). Les 256 groupes d'intensités représentent une gamme réelle de [0,5120] (chaque groupe contient donc 20 niveaux de gris). Cette gamme correspond aux intensités représentatives des tissus rencontrés. Des valeurs supérieures peuvent exister mais elles proviennent d'éléments métalliques et ces dernières ont été limités à 5120. Le pic final de la courbe s'explique donc par la présence de métal dans certaines incisives segmentées. 31

3.5 Densités de probabilités des intensités dentaires pour nos données face à l'intensité du fond. Le profil global est représenté par la courbe cyan, et les cas extrêmes des patients 13 (courbe verte) et 15 (courbe violette) sont superposés. 32

3.6 Cas du patient 13 : apparence des dents et segmentation manuelle de référence. Nous soulignons les dégradations engendrées par la présence des bagues dentaires métalliques. 33

3.7 Cas du patient 15 : l'augmentation de l'intensité moyenne dentaire s'explique par une forte proportion d'éléments métalliques dans les dents. 33

3.8 Pré-traitement. (a) Points sélectionnés par l'utilisateur (zoom sur rendu 3D) qui indique également le numéro dentaire (41 ici), pour procéder à l'extraction de la zone d'intérêt (vue frontale (b), sagittale (c), et axiale (d)). 34

3.9 Recalage du modèle d'apparence moyenne. Pour chaque vue (a,b,c) sont présentés l'image, l'initialisation puis le recalage final. Notons que parfois, comme présenté ici, l'optimisation ne semble pas procurer une amélioration notable, voir même dégrade le résultat. Ici cela provient du fait que la dent d'apparence moyenne est constituée de fortes intensités que la corrélation pénalise lors du débordement sur le fond (voir vue frontale (a)). 34

3.10 Illustration des cartes de contraintes pour la segmentation. (a) Carte des étiquettes. La couleur permet de distinguer la provenance de la contrainte : rouge pour la zone où la déformation ne peut pas entraîner la dent (marqueur du fond), bleue pour la zone dentaire qui existe quelle que soit la déformation (marqueur de la dent) et verte pour les graines sur le plus court chemin entre les points donnés par l'utilisateur (marqueur de la dent). (b) Carte de distance ϕ. Les couleurs marquent une distance infinie. . . . 36

3.11 Représentation graphique du comportement des termes de l'équation 3.5. (a) $R(\lambda_{v_i})$ en fonction de la probabilité. (b,c) $D(0)$ et $D(1)$ en fonction de la distance. (d,e) B pour un σ de 50 et de 100 respectivement, en fonction de la différence d'intensité. 39

3.12 Représentation du problème de coupure de graphe. Schéma traduit de [21]. 41

3.13 Illustration du processus de post-traitement de la segmentation. (a) État initial de la segmentation. (b) Carte de distance interne inversée. (c) Résultat de la ligne de partage des eaux sur (b). (d) Sélection du bassin central. (e) Résultat du traitement après application de la régularisation. 43

3.14 Impact des données de patients en occlusion. 45

3.15 Aspect des différents types de données au niveau dentaire. Les grands champs de vue possèdent une résolution isotrope de 0,3mm alors que les petits champs permettent une résolution de 0,075mm. 46

3.16 Impact du placement des points donnés par l'utilisateur sur la segmentation finale (score de Dice). Le placement correct est noté "Référence". Les autres cas correspondent à une erreur de placement de 3 mm et sont : point coronal trop éloigné de la surface (CourExt) ou trop dans la couronne (CourInt), points d'apex trop éloigné (ApexExt) ou trop dans la racine (ApexInt), points coronal ou d'apex décalés dans le sens vestibulaire, lingual, distal et mésial respectivement (CourVest, CourLing, CourDist, CourMesi ApexVest, ApexLing, ApexDist et ApexMesi). Les paramètres du recalage sont : 100 itérations, $0,001$ de pas minimal, $0,001$ de tolérance au gradient. Les paramètres de la coupure de graphe sont : $\alpha = 1$, $\beta = 5$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$. 47
3.17 Illustration de l'influence de la position des points donnés par l'utilisateur. Dans l'ordre des lignes : sélection correcte, points d'apex en dehors des racines, points d'apex trop haut dans les racines, point de la couronne trop éloigné de la surface, point de la couronne au niveau de la chambre pulpaire. Chaque ligne présente successivement les résultats du recalage (vignettes 2 à 4) et de la coupure de graphe (vignettes 5 à 7). . . 48
3.18 Illustration de l'influence de la sensibilité au gradient dans la précision du recalage. Le nombre d'itérations requis pour obtenir ces résultats est respectivement 0 (l'initialisation est suffisamment correcte), 4 et 13. L'amélioration peut se noter au niveau de l'alignement de la forme dans la vue axiale. 48
3.19 Impact des paramètres de recalage sur la segmentation finale. Les valeurs fixes par défaut sont 100 itérations, tolérance au gradient et pas minimal de 0,001, un seul niveau de recalage, $\alpha = 1$, $\beta = 5$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$. (a) Score de Dice moyen en fonction du nombre d'itérations. (b) Score de Dice moyen en fonction de la tolérance au gradient. (c) Score de Dice moyen en fonction du nombre de niveaux d'échelle. (d) Score de Dice moyen en fonction du pas minimal accepté. 49
3.20 Illustration de l'influence du paramètre σ dans le terme de régularisation reposant sur le gradient. Les autres paramètres sont $\alpha = 1$, $\beta = 5$, $\gamma = 15$, et $\xi = 1$. 50
3.21 Illustration de l'influence du terme B de coupes aux gradients lorsque son poids γ varie. Les autres paramètres sont $\alpha = 1$, $\beta = 5$, $\xi = 1$ et $\sigma = 100$. 50
3.22 Illustration de l'influence du terme R d'a priori sur l'intensité lorsque son poids α varie. Les autres paramètres sont $\beta = 5$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$. 51
3.23 Illustration de l'influence du terme D d'a priori sur la distance lorsque son poids β varie. Les autres paramètres sont $\alpha = 1$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$. 52
3.24 Illustration de l'influence du terme T lorsque son poids ξ varie. Les autres paramètres sont $\alpha = 1$, $\beta = 5$, $\gamma = 15$ et $\sigma = 100$. 52
3.25 Influence des paramètres de la segmentation par coupure de graphe sur la segmentation finale. Les valeurs fixes par défaut sont 100 itérations, tolérance au gradient et pas minimal de 0,001, un seul niveau de recalage, $\alpha = 1$, $\beta = 5$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$. (a) Score de Dice moyen en fonction de σ. (b) Score de Dice moyen en fonction de γ. (c) Score de Dice moyen en fonction de α. (d) Score de Dice moyen en fonction de β. (e) Score de Dice moyen en fonction de ξ. 53
3.26 Comparaison des résultats des approches sur notre base de test : forêt aléatoire avec descripteurs RISC [59] (violet), modèle déformable avec a priori de forme [85] (bleu), *powerwatershed* [40] (vert) et notre méthode [49] (rouge). Pour chaque numéro dentaire est indiqué le score de Dice associé. La première moitié des cas correspond à des échantillons issus de volumes avec une taille de voxels de 300 microns, contre 150 microns pour la seconde moitié. La perte de qualité des résultats à 300 microns est notable. . . 55
3.27 Exemple de résultat produit par l'approche par forêt aléatoire avec descripteurs RISC. On note la prise en compte de l'os cortical dans la segmentation, et une absence du canal radiculaire. 55
3.28 Exemple de résultat produit par l'approche par modèle déformable avec a priori de forme. La vue axiale permet de bien voir que la contrainte de forme (en bleu) se positionne sur des gradients importants qui ne correspondent pas aux contours dentaires. . 56
3.29 Cas particulier de la dent numéro 33 segmentée par *powerwatershed*. 56

3.30 Segmentation d'une image avec le réseau de neurones : (a) image d'entrée du réseau, (b) zoom sur une partie de l'image d'entrée, où l'on peut noter des zones de teinte bleue, qui correspondent à une information existante dans la coupe supérieure, et des zones de teinte rouge, correspondant à l'information provenant de la coupe inférieure, (c) image des annotations superposées à l'image d'entrée (les couleurs ne reflètent pas la numérotation dentaire). 57
3.31 Exemples de comparaison de la segmentation produite par le réseau superposée à la segmentation de référence. (a) Cas issu de la base de validation, (b) cas de la mâchoire édentée, (c) cas d'un volume en grand champ. 58
3.32 Résultats du score de Dice obtenus par le réseau de neurones : (a) pour chaque numéro dentaire sur l'ensemble de validation, (b) sur une arche mandibulaire partiellement édentée, (c) sur un ensemble de volume en grand champ. 59
3.33 Comparaison de la formulation proposée avec l'algorithme CV utilisé en interne. (a-b) Résultats sur un volume d'endodontie (taille de voxel isotrope de 75 microns) pour notre formulation puis pour le CV. (c-d) Résultats sur un volume de champ moyen (taille de voxel isotrope de 150 microns) pour notre formulation puis pour le CV. (e) Résultats sur un volume de grand champ (taille de voxel isotrope de 300 microns) : ce dernier cas met en échec la recherche des canaux dans l'algorithme CV, d'où l'absence de résultats. Chaque série de vignette présente successivement : la vue sagittale de l'image et de la segmentation, puis la vue coronale et enfin la vue axiale. 61
3.34 Résultats du score de Dice pour chacun des 64 échantillons. La différence du ton de couleur indique un patient différent. 62
3.35 Résultats du score de Dice moyen selon le type dentaire. 63
3.36 Résultats de spécificité et de sensibilité pour chaque échantillon. Les couleurs différencient l'étape de segmentation : rouge pour la coupure de graphe et bleue pour le post-traitement. 63
3.37 Image correspondant à l'échantillon numéro 1 de la figure 3.36 et sa segmentation (sortie de la coupure de graphe). Remarquons la très forte dégradation induite par la présence de métal, notamment la présence de faux contours le long des raies hyper-intenses, qui met en échec la segmentation. 64
3.38 Cas correspondant à l'échantillon numéro 2 de la figure 3.36 et sa segmentation (sortie de la coupure de graphe). Il est possible de voir que la segmentation référence (b) prend en compte toute une partie périphérique non retenue par la segmentation automatique (c). 64
3.39 Image correspondant à l'échantillon numéro 3 de la figure 3.36 et sa segmentation (sortie de la coupure de graphe). Remarquons le gradient très important induit par la proximité des sinus avec les racines de la molaire maxillaire qui conduit la segmentation à intégrer tout l'espace inter-radiculaire. 65

4.1 Résultats de la segmentations multi-objets en termes de sensibilité et de spécificité. Le nuage des résultats multi-objets (bleu) est nettement en retrait comparativement au nuage des résultats mono-objet (rouge). 71
4.2 Résultats du score de Dice pour chaque échantillon, en rouge pour le cas mono-objet et en bleu pour le cas multi-objets. La différence de ton d'une même couleur indique un patient différent. Il est possible de constater une régression de la qualité de segmentation pour la quasi-totalité des échantillons lors du passage à l'algorithme multi-objets. . . . 71
4.3 Illustration des cas erronés produits par la segmentation multi-objets. (a) Lorsque les racines des dents sont peu distantes, l'espace interdentaire est attribué à une des dents adjacentes par l'optimisation (cas de faible spécificité). (b) La solution sous-optimale trouvée conduit à ne pas segmenter toute la partie dentaire de la dent centrale malgré les marqueurs (3^{e} vignette). Le marqueur de la molaire en rouge est erroné : le plus court chemin du point de la racine mésiale est passé par la racine distale. Ce cas conduit à une chute de la sensibilité. 72

4.4 Représentation graphique du modèle d'une dent. Les structures sont représentées par les formes rectangulaires tandis que les relations sont de forme ovoïde. Capture d'écran depuis l'application CoGui. 73
4.5 Cas où la définition e_{3D} ne convient pas. La région β est allongée alors que ni A ni B ne le sont : ils ne sont pas le long l'un de l'autre. 76
4.6 Cas où l'aire des surfaces de contact n'est pas discriminante. La forme de cette interface doit être prise en compte. 77
4.7 Cas où la distance apporte une information supplémentaire sur la configuration des objets. 78
4.8 Objets manuellement segmentés de la région maxillo-faciale. Les canaux mandibulaire sont en vert, les planchers mandibulaires en rouge et la dent 36 est en bleu. 80
4.9 Résultats obtenus sur les objets nets. La couleur différencie les patients. 81
4.10 Résultats obtenus pour les objets flous. 82
4.11 Moyenne des scores obtenus pour les degrés sélectionnés (α_2 et $F\alpha_2$) sur 6 patients. Le premier ensemble de résultats correspond à α_2 (bleu) et le second à $F\alpha_2$ (rouge). L'écart-type est montré sous la forme de barres d'erreur. 83
4.12 Direction dans un espace 3D. 84
4.13 Forme de l'élément structurant utilisé pour la dilatation des histogrammes d'orientation. 85
4.14 Agencement spatial des dents utilisées pour tester l'alignement. 86

5.1 Flux de travail utilisé dans le prototype logiciel. Les numéros dentaires et les ensembles de points sont les éléments nécessaires à la segmentation, donnés par l'utilisateur. . . . 90
5.2 Interface graphique correspondant à la première étape de traitement. 92
5.3 Interface graphique correspondant à la seconde étape de traitement : état initial. . . . 93
5.4 Interface graphique correspondant à la seconde étape de traitement : sélection d'une première dent, l'utilisateur coche le numéro dentaire, puis indique les points. 93
5.5 Interface graphique correspondant à la seconde étape de traitement : sélection d'une autre dent, de la même manière. 94
5.6 Interface graphique correspondant à la troisième étape : segmentation initiale de la dent 41. 94
5.7 Interface graphique correspondant à la troisième étape : segmentation initiale de la dent 42. 95
5.8 Aspect des outils de correction manuelle : l'ajout de marqueurs internes est signalé par une couleur verte, tandis que ceux externes sont oranges (rouge pour l'outil). 96
5.9 Interface graphique correspondant à la troisième étape : correction du résultat de la dent 42. 96
5.10 Interface graphique correspondant à la troisième étape : segmentation finale de la dent 42. 97
5.11 Interface graphique correspondant à la quatrième étape : nouvelle analyse ou sauvegarde des données. 97
5.12 Organisation des résultats enregistrés. 98
5.13 Fenêtre de contrôle de l'outil pour développeur. 98
5.14 Exemple de : (a) modèle d'étude dentaire, ou "plâtre", (b) scan intra-oral, (c) guide chirurgical, qui s'adapte à la denture afin de faciliter la pose d'implant. 100
5.15 Exemple de segmentation dentaire contrainte par le maillage recalé dans le cadre du projet PDIP (*Patient Driven Implant Planning*). Toutes les dents sont correctement segmentées grâce à la connaissance de l'interface entre les couronnes provenant du maillage. 101
5.16 Illustration du projet PDIP. (a-c) Vue 3D des structures dures (os et dents), avec recalage du maillage du scan intra-oral puis ajout d'un implant. (d) Vue pseudo-panoramique, où est représentée en vert la limite du maillage qui épouse correctement la forme des dents, ainsi que l'implant en jaune-orangé. (e) Vue axiale, reprenant les codes de couleurs de la vue pseudo-panoramique. 102

A.1 Exemple d'arbres de composantes : (a) image synthétique, (b) arbre maximum, (c) arbre minimum et (d) arbre des formes. 110
A.2 Représentation simplifiée de l'aspect d'une dent (molaire) dans nos images. Si la dent est correctement séparée de l'os cortical adjacent par le ligament parodontal, l'identification de la composante connexe correspondante à l'émail devrait correspondre à une segmentation de la dent. 110
A.3 Représentation des composantes connexes de nœuds extraits d'un arbre de forme pour une incisive maxillaire, dans un parcours de la racine aux feuilles. Notez qu'au fur et à mesure, les parties appartenant aux structures environnantes disparaissent, mais la dent se dégrade. 111
A.4 Illustration du filtrage des composantes de dents par une composante dégradée. La fermeture effectuée sur (a) donne un marqueur suffisamment précis pour pour récupérer les voxels de la dent de (b). (c) est le résultat de l'intersection. 112
A.5 Processus de segmentation d'une dent. 112
A.6 Illustration de la reconstruction effectuée sur les composantes connexes de l'arbre pour éliminer un maximum de structures indésirables. (a) Image binaire d'une composante connexe. (b) Résultat final de la reconstruction. (c) Coupe initiale et son traitement (sélection de l'élément central, dilatation et remplissage) donnant le masque de comparaison avec la coupe suivante. (d) Coupe adjacente, intersection avec le masque et masque pour itération suivante. (e) Exemple de résultat dans une coupe plus proche des racines. (f) Illustration de l'échec du filtrage lorsque l'élément indésirable est directement adjacent. 114
A.7 Optimisation finale du modèle statistique (a) et approximation finale de la segmentation (b) par masquage de la composante. 115
A.8 Scores de Dice obtenus pour chaque dent par notre approche utilisant sur le modèle statistique. Le score moyen est de $0,781 \pm 0,065$. 116

B.1 Résultats obtenus par dent individuelle avec la méthode présentée dans le chapitre 3. Pour chaque série est présentée la coupe axiale puis la coupe coronale et enfin la coupe sagittale, chacune avec leur résultat. 121
B.2 Superposition d'une segmentation complète d'un volume par la méthode présentée dans le chapitre 3. 123
B.3 Rendu 3D des résultats de segmentation d'un volume complet avec la méthode présentée dans le chapitre 3. 125
B.4 Exemples de résultats avec la méthode proposée dans le chapitre 4. Notons que la prise en compte des dents adjacentes permet de mieux identifier les contours de chaque dent, notamment au niveau des couronnes, grâce à la compétition entre chaque étiquette. Nous présentons des coupes axiales ainsi que des coupes sagittales ou coronales selon l'axe permettant de mieux distinguer les interfaces entre les dents. 127

C.1 Identification du canal sur les données CBCT. 129
C.2 Étude des histogrammes du canal mandibulaire (courbe bleue) et de son environnement. Les trois épaisseurs considérées sont 0,5 mm (courbe violette), 1mm (courbe verte) et 1,5mm (courbe rouge). Les moyennes et écarts-types du canal et de l'épaisseur 1,5mm sont indiquées dans la table C.1. 131
C.3 Étude des histogrammes des différentes strates du canal mandibulaire. Le squelette (courbe bleue) est généré à partir de la segmentation manuelle, puis dilaté successivement par un élément structurant sphérique de 2 voxels de rayon jusqu'à retrouver la segmentation initiale (courbe orange). L'ordre des couleurs est : bleue, rouge, verte, violette, turquoise puis orange. 132
C.4 Aspect comparé des matrices de co-occurrences, avec celle du canal (gauche) et de l'environnement (droite). Pour chaque patient, le modèle de répartition est similaire. La différence de taille correspond à la plus grande diversité de niveaux de gris dans la partie environnementale que dans le canal. 133

C.5 Valeurs caractéristiques calculées pour les matrices des zones d'intérêt (en bleu pour le canal, en rouge pour son environnement). Seule la différence d'énergie est statistiquement significative (test de Student à 1%). (a) Énergie, (b) entropie, (c) inertie, (d) homogénéité. 134
C.6 Ondelette mère utilisée (CDF-5/3). 135
C.5 Énergie comparée des coefficients d'ondelettes (en bleu pour le canal, en rouge pour l'environnement) pour chaque patient. Chaque ensemble est numéroté par $(\pm n, \pm n, \pm n)$, où n est le niveau de décomposition, et le signe indique le filtre utilisé dans cette direction (correspondant successivement à X, Y et Z) : + pour le passe bas (donc l'approximation) et − pour le passe-haut (donc les détails). 138

D.1 Représentation 3D de la forme des fosses nasales. (a) Vue de face. (b) Vue de profil. Images provenant de www.fauquierent.net, extraites d'une vidéo de chirurgie des sinus. 142
D.2 Schéma des os du crâne humain en façade (a) et des voies aéro-digestives (b). Illustrations extraites de Wikipédia. 143
D.3 Anatomie de la mandibule. Illustrations de Wikipédia. 144
D.4 Illustration des formes rencontrées pour la chambre pulpaire. (a) Incisives et canine (à droite). (b) Prémolaire mandibulaire. (c) Prémolaire maxillaire. (d) Molaires. 147

E.1 Modèle structurel de la zone maxillo-faciale. 154

F.1 Étapes de la sélection des points. 156
F.2 Résultats de l'approximation de l'arche. Sont représentés les points saillants (vert), l'enveloppe convexe initiale ouverte (rouge), l'enveloppe convexe suréchantillonnée (jaune), l'approximation parabolique externe (bleu) et le modèle final (violet). 156

Liste des tables

1.1 Avantages et inconvénients du CBCT dans la région maxillo-faciale, traduit de [43]. . . 6

2.1 Table récapitulative des approches de segmentation envisagées. 21

3.1 Répartition des dents segmentées manuellement en classes pour la construction d'un modèle statistique de forme. 30
3.2 Intensité moyenne dentaire de nos données. Les appellations CS8100 3D et CS9300 correspondent aux noms commerciaux de différents scanners Carestream. Le CS9300 est un scanner permettant l'acquisition de grands champs grâce à un grand capteur, mais d'une génération plus ancienne que le CD8100 3D, qui lui ne permet d'acquérir que des champs moyens (de taille maximale d'environ 10 cm de côté). 32
3.3 Répartition des poids sur les arêtes du graphe utilisé dans notre formulation de coupure de graphe. 42
3.4 Résultats moyens sur l'ensemble de la base pour la sensibilité, la spécificité, la précision et le coefficient de Dice 62

4.1 Résultats moyens sur l'ensemble de la base pour la sensibilité, la spécificité, la précision et le coefficient de Dice (voir la section 3.4.1 pour la discussion sur la pertinence de la spécificité et de la précision). 70
4.2 Table de correspondance entre les classes qualitatives d'intensité et leur valeur dans nos données. Pour rappel, les données issues des scanners Carestream sont encodées sur 16 bits, soit des valeurs d'intensités comprises entre 0 et 65535. L'impact de la présence de métal dans la dernière classe entraîne une variabilité très importante. 74
4.3 Résumé des degrés d'élongation proposés. 79

C.1 Table récapitulative des moyennes d'intensités du canal et de son environnement à 1,5 mm dans nos cas d'étude. 130

Chapitre 1

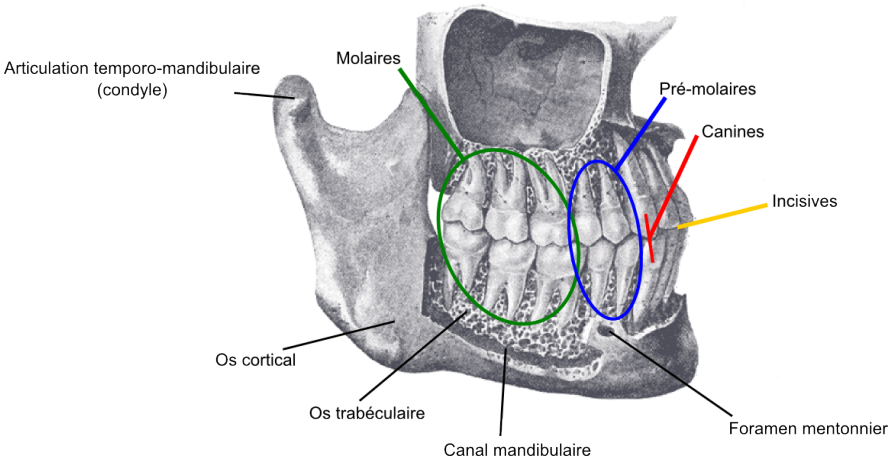
Introduction

1.1 Contexte applicatif

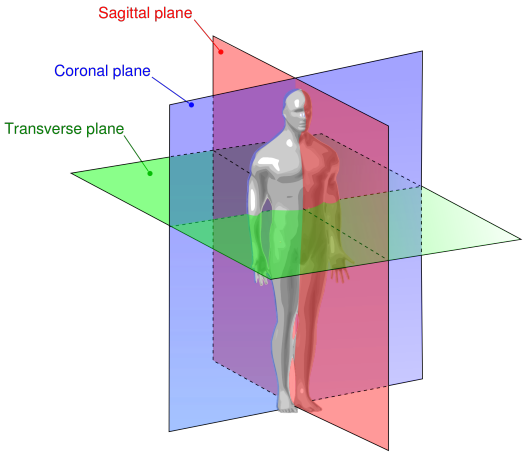
Cette thèse à pour visée le développement d'outils algorithmiques de vision par ordinateur appliqués au domaine du soin bucco-dentaire. L'espace anatomique (figure 1.1) exploré se concentre donc sur la bouche mais aussi sur les voies aériennes supérieures (pharynx et sinus) et les articulations temporo-mandibulaires. Notre travail s'adresse donc aux dentistes, généralistes ou spécialistes, et aux radiologues, par l'assistance du traitement d'image. Si les outils d'imagerie de ce domaine sont nombreux, nous nous focaliserons sur une modalité : la tomographie volumique à faisceau conique (*Cone-Beam Computed Tomography*, CBCT), appartenant à la catégorie des modalités utilisant les rayons X.

1.1.1 Les dents dans la pratique clinique

Les dents sont les structures principales de la zone maxillo-faciale. Elles constituent la première étape du processus d'alimentation qui nous permet de subvenir à nos besoins. Selon leur type, elles remplissent une tâche bien précise qui permet de constituer la première forme du bol alimentaire. Les dents antérieures (incisives et canines) servent à séparer la nourriture en morceaux grossiers qui seront réduits en pâte par les dents postérieures (prémolaires et molaires). Cette transformation physique, associée à une action chimique et enzymatique de la salive, est la première phase de la digestion qui permet d'extraire les nutriments des aliments. L'importance de la mastication pour une bonne



(a) Structures de la zone maxillo-faciale (dessin de W. Spalteholz ; theodora.com/anatomy).



(b) Nomenclature des plans anatomiques (le plan transverse est aussi nommé plan axial). Illustration de Y. Mrabet (Wikipédia).

FIGURE 1.1 – Présentation du contexte anatomique dentaire.

digestion a déjà été mise en lumière et est régulièrement rappelée par les professionnels de santé [104]. Les dents, étant dans la partie externe du corps (comme tout le tube digestif), subissent donc de multiples contraintes, d'ordre mécanique (mastication, fractures, poussée dentaire), d'ordre physico-chimique (pH variable dû aux aliments, à la salive et aux boissons, action enzymatique) mais aussi d'ordre micro-biologique lorsque surviennent des caries ou des abcès. De plus, ces organes présentent une évolution particulière au regard du reste du corps en trois phases : d'abord la mise en place de la denture temporaire ("dents de lait", apparition entre 6 mois et 3 ans), renouvellement complet par la denture définitive (de 6 à 12 ans) puis éruption des troisièmes molaires ("dents de sagesse", autour de 20 ans).
L'environnement dentaire est donc constamment en évolution, ce qui explique l'attention particulière portée à ces organes et que la majorité des pratiques cliniques dans cette zone y soit liée. Il existe plusieurs nomenclatures pour identifier les dents :

— le système ISO 3950,

— le système universel de numérotation,

— la notation de Palmer.

La plus répandue, et celle recommandée par l'organisation mondiale de la santé, étant le système ISO, nous utiliserons cette dernière comme référence dans ces travaux (figure 1.2). La dent est composée de multiples parties, et selon la spécialité du praticien, elle sera abordée dans son intégralité ou en détail. Un schéma explicatif de cette anatomie est présenté en figure 1.3. Voici les différents niveaux d'intervention sur cet organe selon la spécialité dentaire :

— Odontologie conservatrice : principalement des travaux de surface comme le détartrage et le blanchiment ainsi que le traitement des caries.

— Prosthétique : restauration totale ou partielle des couronnes (comblement avec amalgame ou composite, couronnes complètes, bridges).

— Orthodontie : vérification de la disposition générale et de l'occlusion, traitement des alignements et correction de position.

— Implantologie : section ou extraction des dents, forage, pose d'implants totaux ou partiels.

— Endodontie : forage, limage et obturation de l'intérieur de la dent.

— Chirurgie : extraction des dents (incluses ou externalisées).

Voici ici un petit récapitulaif des termes utilisés lorsqu'on se réfère aux dents :

— Vestibulaire : concerne la partie externe de l'arcade dentaire, espace entre les dents et les joues ou l'ouverture buccale (ex : la face vestibulaire d'une incisive).

— Lingual : par opposition à vestibulaire, concerne la partie interne de l'arcade et désigne l'espace entre les dents et la langue.

— Mésial : concerne la partie orientée vers le centre de l'arcade, vers les dent antérieures.

— Distal : par opposition à mésial, concerne la partie orientée vers le fond de l'arcade.

— Occlusal : concerne les parties en opposition lorsque la mâchoire est serrée ou en mastication.

— Cuspide : éminence de la couronne de forme convexe présente sur la face occlusale des dents.

— Sillon : fossé présent sur la couronne entre les cuspides.

— Cervical : concerne le collet, on parle par exemple de limite cervicale pour la jonction entre la dent et la gencive.

— Radiculaire : concerne les racines d'une dent.

— Coronal : concerne la couronne dentaire (à ne pas confondre avec le plan anatomique).

Du fait de cette place prépondérante dans le quotidien clinique, l'identification et la segmentation des dents revêt un intérêt majeur pour être en mesure de faciliter le diagnostic, la prise en charge et le traitement des patients. Dans un premier temps cela permet de faire ressortir l'information contenue dans les volumes 3D pour faciliter l'analyse et dans un second temps de développer des aides au traitement comme par exemple une simulation d'évolution orthodontique, un protocole d'intervention optimisé, ou la compilation de mesures cliniques automatiquement.

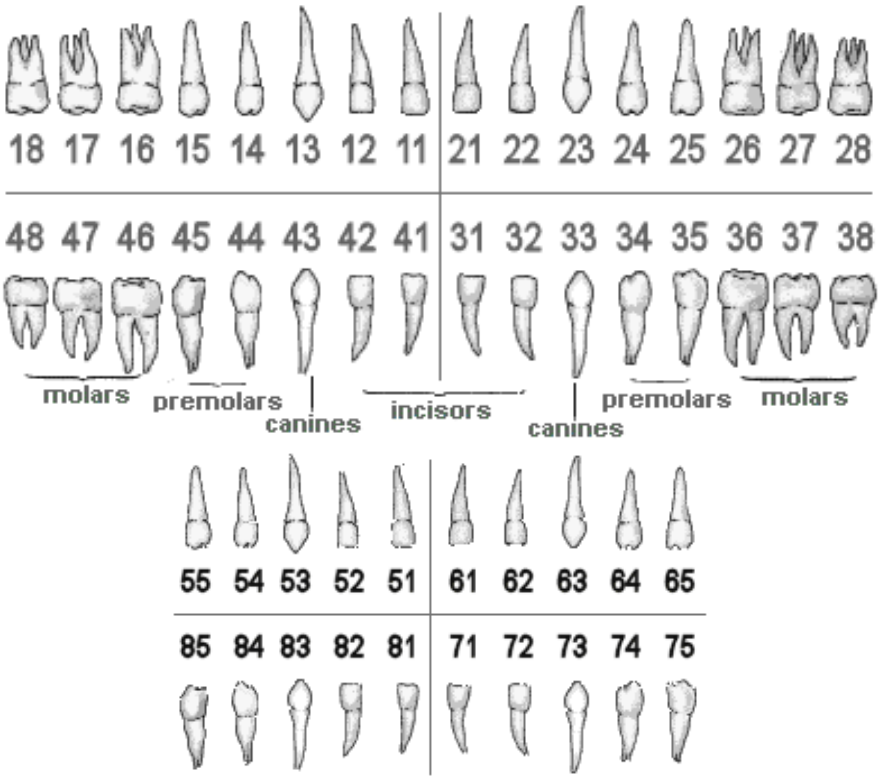


FIGURE 1.2 – Système de notation ISO. Denture définitive (entre 11 et 48) et denture temporaire (entre 51 et 85). Cette notation prend le point de vue du dentiste : haut=maxillaire, bas=mandibule, gauche=coté droit du patient et droite=coté gauche du patient.

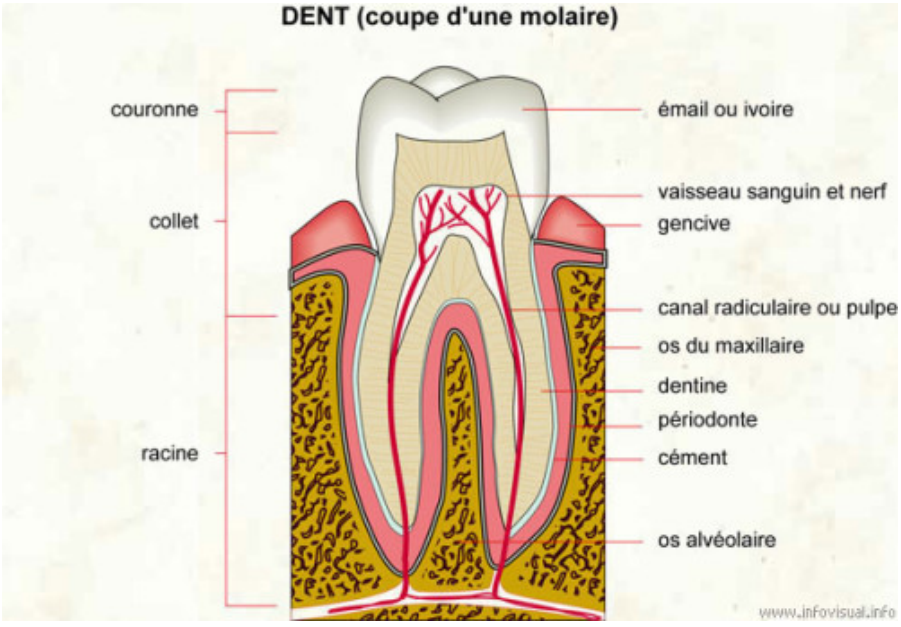


FIGURE 1.3 – Anatomie d'une dent (molaire mandibulaire). Crédit : infovisual.info.

1.1.2 Autres structures dans la prise en charge

En dehors des dents, d'autres structures présentent un intérêt dans leur identification et/ou délimitation à des fins de soins.

Les os mandibulaire et maxillaires

Les identifier permet d'analyser leur constitution, leur morphologie, le rapport entre la partie corticale et la partie trabéculaire et leur santé respective. Si pour le moment il n'est pas possible d'estimer la densité osseuse, le développement tout récent des appareils de CBCT avec compteur de photons à double énergie permettra peut être d'ouvrir le champ de l'ostéodensitométrie.

Les voies aériennes

Les sinus étant positionnés directement au-dessus de l'arche maxillaire, il arrive parfois que les racines des dents postérieures supérieures pénètrent directement dans le sinus. Toute intervention doit alors se faire de manière précautionneuse pour ne pas entraîner de complications comme des sinusites. Du côté du pharynx, il existe un lien entre la volumétrie aérienne et le développement de troubles orthodontiques comme des malocclusions en raison d'une mauvaise respiration, notamment chez les enfants.

Les canaux mandibulaires

Ce sont des structures tubulaires internes duales (gauche et droite) de la mandibule, elles commencent au niveau de l'épine de Spix (ou lingula mandibulaire), qui se situe sur la surface interne de la branche montante, pour ressortir au niveau des foramens mentonniers, petites ouvertures situées sous les prémolaires. Ces canaux sont d'une grande importance car ils renferment les nerfs et les artères alvéolaires et mentonniers qui irriguent la mandibule ainsi que la partie basse du visage. Il arrive fréquemment que ces canaux soient enchevêtrés ou à proximité immédiate des racines dentaires. Toute intervention doit alors être précédée d'une analyse de leur position pour éviter des traumatismes pouvant conduire à des hémorragies importantes ou des anesthésies définitives.

1.2 Les besoins des dentistes et la réponse industrielle

Dans la pratique, l'imagerie intervient de multiples façons. Notamment, l'imagerie à rayons X permet l'exploration en profondeur des structures. Il est de pratique courante de réaliser des clichés dits "panoramiques" (ou orthopantomogrammes, figure 1.4) qui représentent la mise à plat de l'arche dentaire et permettent d'analyser les structures internes invisibles à l'œil nu. Les praticiens s'en servent afin d'observer l'environnement dentaire, pour voir le développement des organes, leur interaction, et repérer les éléments intervenant dans le processus de soin (les structures sensibles, les pathologies, les positionnements problématiques, etc.). Cependant ce cliché représente de manière plane une information volumique, rendant difficile l'interprétation spatiale dans le sens de la projection.

FIGURE 1.4 – Exemple de cliché panoramique ou orthopantomogramme.

L'arrivée des scanners à faisceau conique et leur démocratisation a permis d'obtenir une information pleinement volumique comme le scanner hélicoïdal, avec toutefois des spécificités non négligeables, comme présenté dans la table 1.1. Cette technique, dont le principe est rappelé dans la figure 1.5, est

relativement récente dans son application grand-public (une centaine de cabinets équipés en 2007 selon un rapport de la Haute Autorité de Santé de 2009 [106], principalement en ville) ce qui fait que la littérature concernant le traitement de telles images est encore peu développée. Les praticiens ont donc maintenant à disposition une représentation équivalente à la réalité spatiale du patient, mais qu'ils doivent analyser manuellement afin d'extraire les informations recherchées, ce qui est coûteux en temps et demande une expertise pour obtenir une bonne reproductibilité.
Carestream Dental (partenaire industriel de la thèse), en tant que fournisseur de scanners CBCT et de solutions logicielles associées pour les cabinets dentaires, cherche à proposer une assistance dans l'analyse de ces données aux utilisateurs. La diversité des acquisitions issues des scanners Carestream est présentée en figure 1.6. Nous avons des volumes dits "grands champs" qui contiennent presque l'ensemble du crâne, et possèdent une résolution isotrope de 0,3mm, des volumes "moyens champs" qui ciblent la région buccale et possèdent une résolution entre 0,18 et 0,15mm et enfin des "petits champs" qui permettent l'exploration d'une fraction de denture avec une résolution de 0,075mm. Ces différents volumes répondent à des usages différents, les plus grands apportant une information plus générale de posture, d'alignement et de comportement en occlusion intéressant principalement les orthodontistes alors que les plus petits apportent des informations ciblées sur une partie précise, ce qu'utilisent par exemple les endodontistes. Les outils à la disposition des praticiens, pour le moment, sont de l'ordre de la mesure de distance, de la présentation en coupes et de la manipulation du rendu 3D. Mais il existe une vraie demande pour des traitements plus complexes des images qui permettraient d'extraire des informations de plus haut niveau pour le renfort de la prise en charge : ces outils devront avoir un apport clinique et diagnostique suffisant, tout en gardant un temps de calcul et d'utilisation raisonnable, pour permettre une adoption par les praticiens.

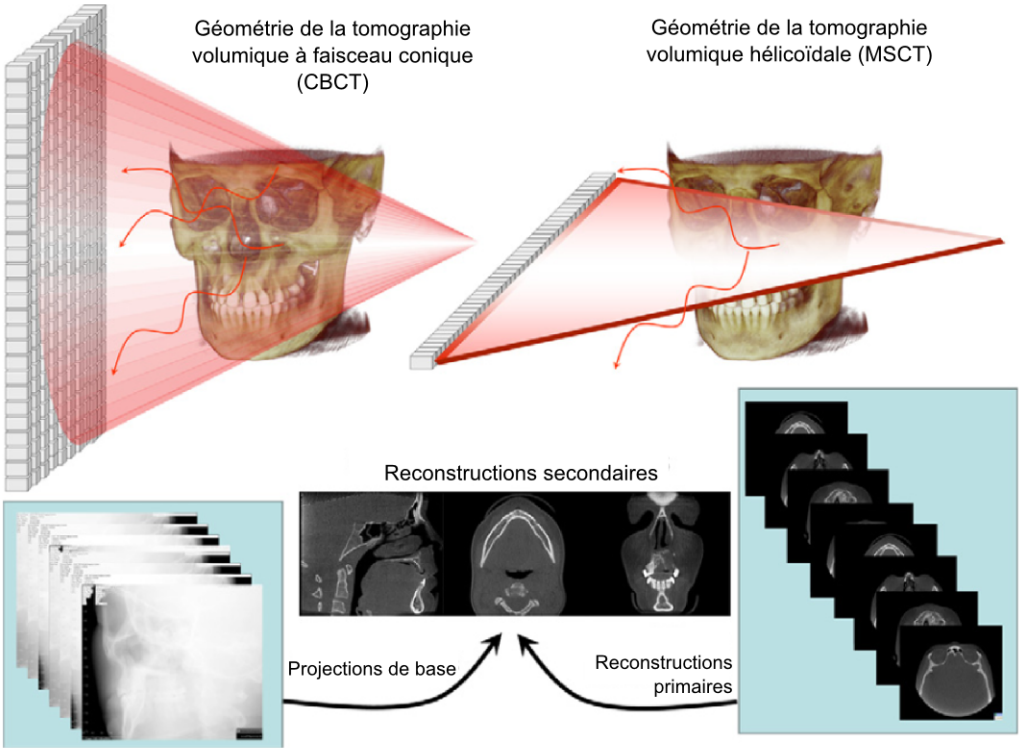
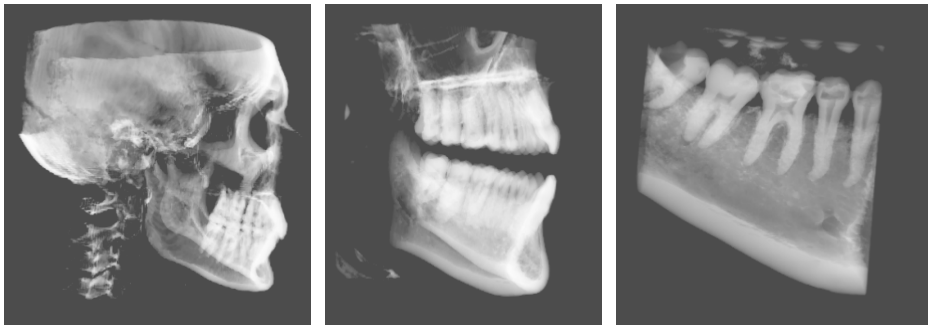


FIGURE 1.5 – Principe de fonctionnement du CBCT comparé au scanner traditionnel (CT). On remarque que l'acquisition d'un plan entier en CBCT conduit à capturer plus de photons diffusés (flèches rouges) que dans l'acquisition coupe par coupe du CT, ce qui affaiblit le contraste et bruite l'image. Les reconstructions finales sont issues des projections de base (CBCT) plutôt que des reconstructions de coupes (CT). Schéma traduit de [108].

Avantages
- données 3D
- données à taille réelle
- potentiel pour générer toutes les images 2D (par exemple orthopantomogrammes, céphalogrammes latéraux, images de l'articulation temporo-mandibulaire)
- potentiel pour des acquisitions en position assise naturelle
- voxels isotropes
- haute résolution (par exemple trabéculation osseuse, ligament parodontal (PDL), racines dentaires)
- dose de radiation plus faible que le CT hélicoïdal
- impact réduit des artéfacts métalliques
- coût plus faible que le CT hélicoïdal
- facilité d'accès
- imagerie de cabinet
- prise en main aisée
- compact
- compatible avec le format DICOM
- logiciels de visualisation et de post-traitement intuitifs
- économe en énergie comparé au CT hélicoïdal

Inconvénients
- faible contraste (dépendant du type de détecteur à rayons X)
- taille limitée du capteur induisant une taille de champ de vue et de volume limitée
- renseignements limités sur le tissu mou interne
- bruit accru du rayonnement diffusé résultant en une perte de résolution de contraste
- artéfacts de mouvement affectant tout le volume
- artéfacts de coupure (causés par le fait que les projections acquises avec la sélection d'une région d'intérêt ne contiennent pas tout l'objet)
- ne peut pas être utilisé pour l'estimation des unités Hounsfield (HU)

TABLE 1.1 – Avantages et inconvénients du CBCT dans la région maxillo-faciale, traduit de [43].



(a) (b) (c)

FIGURE 1.6 – Exemples de différentes acquisitions. Les grands champs de vue possèdent une résolution isotrope de 0,3mm alors que les petits champs permettent une résolution de 0,075mm.

1.3 Les enjeux scientifiques

Quel peut être l'apport de travaux de recherche dans ce cadre ? Le développement de l'analyse d'images dans les autres domaines médicaux a montré le gain potentiel réalisé en termes de connaissance et de compréhension pour l'amélioration du parcours médical (prise en charge plus rapide,

meilleur diagnostic, réduction des incidents opératoires, traitements personnalisés). Nous souhaitons donc amener dans les cabinets dentaires des outils et des méthodes de segmentation d'image qui correspondent aux besoins cliniques, afin de combler le manque actuel d'automatisation de l'analyse des données des patients.
Quelles sont les questions soulevées ? Premièrement il faut identifier quels éléments des images sont les plus intéressants pour les praticiens. Le service rendu sera important si nous offrons des informations utiles à la pratique mais aussi si ces informations sont peu aisées à extraire visuellement. Deuxièmement, comment être capable d'extraire toutes les informations désirées compte tenu des particularités inhérentes à la modalité et au contexte clinique ? En effet s'il existe un schéma global sur lequel s'appuyer (l'anatomie), la diversité des variations individuelles et des pathologies présentée par les patients tend à en affaiblir la pertinence. S'il est possible d'évaluer et de prendre en compte l'impact de la variabilité physiologique, il est beaucoup plus difficile de faire de même avec la variabilité pathologique. D'un autre côté, l'imagerie CBCT produit des images avec un faible contraste, ce qui complique fortement les tentatives d'identification et de délimitation entre structures en dehors du schéma tissu mou/tissu dur, d'autant qu'il n'est pas possible d'utiliser les unités Hounsfield (échelle de correspondance entre intensité apparente et type de structure), les différents constructeurs de scanners CBCT ne s'étant pas entendu sur une calibration de référence. S'ajoute à cela que la présence de métal est fréquente dans la région maxillo-faciale (couronnes, bagues, dentiers), ce qui provoque de très fortes dégradations de l'image résultant d'interactions physiques lors de l'acquisition, qui perturbent la reconstruction des volumes et ce malgré l'utilisation du CBCT (moins sensible aux artéfacts métalliques que le CT traditionnel, voir table 1.1). Ces problématiques sont illustrées en figure 1.7. Troisièmement du point de vue industriel de cette thèse, quels sont les critères techniques qui s'imposent dès lors que nous souhaitons réaliser des outils prêts pour les dentistes ? En dehors de la méthodologie, qu'attend l'utilisateur en termes de précision, de résultat, d'ergonomie ? Qu'est ce qui va faire que notre méthode sera adoptée lorsque celle du concurrent sera délaissée ? Enfin, comment évaluer les performances de nos propositions ? Il est important de noter que contrairement à d'autres champs de l'imagerie médicale, et du fait de la relative nouveauté et du peu de mobilisation globale des acteurs de santés sur la thématique, il n'existe pas de bases d'images libres d'accès qui permettraient une analyse robuste et un moyen direct de comparaison. Les données de travail vont donc être spécifiques aux scanners Carestream et la disponibilité de données segmentées et annotées manuellement réduite. Pour résumer, notre approche doit tenir de compte de :

— la pertinence des structures considérées,

— la faible disponibilité de données annotées,

— l'hétérogénéité des images cliniques due aux différents scanners Carestream, champs de vue et variabilité anatomique,

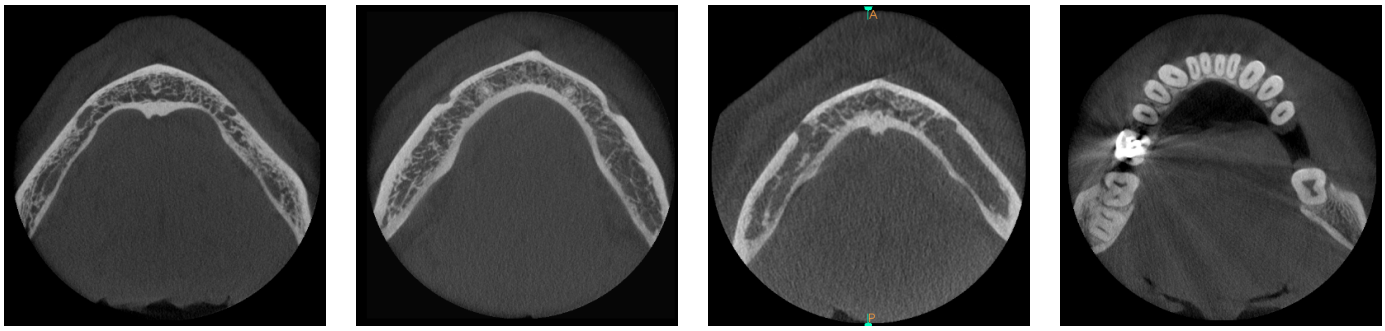
— la spécificité de la modalité,

— la nécessité d'obtenir des méthodes aptes à fonctionner sur des ordinateurs de configuration courante,

— la demande des praticiens d'avoir un contrôle sur le résultat, ce qui passe par une possibilité d'interaction avec un temps raisonnable.

Les débouchés possibles sont nombreux : cela permet d'envisager plus facilement des procédures comme les poses d'implants, l'orthodontie avec simulation temporelle, la création automatique de couronnes, ou encore la construction de rapports cliniques mettant en valeur les spécificités du patient.

Nous avons choisi de focaliser notre travail sur la segmentation de la denture, c'est-à-dire des dents, de manière individuelle. Ce choix s'explique par le fait que ces structures sont l'enjeu principal des traitements qui s'effectuent dans les cabinets médicaux, et que l'information clinique qui découle de leur identification intéresse toutes les spécialités du soin dentaire. De plus, l'analyse visuelle d'une dent est une tâche peu aisée puisque cet organe présente une forme complexe, difficile à conceptualiser par la simple inspection des coupes d'un volume. Lorsqu'on s'attache à analyser les relations entre les dents, cela devient encore plus complexe, ce qui en fait un sujet de choix pour notre étude. Pour parer aux problématiques liées à la qualité de l'image, nous avons choisi de contraindre notre identification



(a) Forme d'arche en "U". (b) Forme d'arche en "V". (c) Les tissus mous sont peu distincts. (d) La présence de métal dégrade fortement le signal.

FIGURE 1.7 – Illustrations des problématiques rencontrées dans le cadre de l'utilisation de la tomographie volumique à faisceau conique pour l'imagerie dentaire.

par l'ajout d'informations statistiques a priori, extraites de volumes où les dents ont été annotées manuellement par un expert.

Ces informations sont utilisées dans un processus semi-automatique de segmentation en deux temps, où un modèle statistique pré-établi de dent est comparé à l'image inconnue avant de contraindre la segmentation pour laquelle nous employons des méthodes de résolution de type coupure de graphe qui permettent à la fois rapidité d'exécution, faible coût en ressources de calcul (CPU d'entrée de gamme, peu de mémoire) et facilité d'édition. L'utilisateur est intégré au processus puisqu'il définit quelle structure il souhaite segmenter et qu'il a la possibilité de modifier le résultat. Les résultats sont comparés aux méthodes existantes dans le domaine dentaire ainsi qu'aux méthodes de l'état de l'art. Au moment de l'écriture, la validation clinique n'a pas encore pu être faite.

Le chapitre 2 présentera une revue non exhaustive de la littérature autour du sujet. Le chapitre 3 décrira notre méthode de segmentation dentaire mono-objet avant que le chapitre 4 ne présente une extension multi-objets. Il couvrira aussi les travaux effectués pour tenter de synthétiser les connaissances sur la région buccale et ouvrira la porte à une approche par modèle structurel. La mise en application de notre méthode sera faite par la création d'une application fonctionnelle discutée au chapitre 5. La conclusion et les perspectives ouvertes par ces travaux seront discutés dans le chapitre 6

Nous résumons ici les contributions des travaux :

— Création d'une méthode de segmentation dentaire dans les volumes de tomographie à faisceau conique (chapitre 3).

— Réalisation d'un modèle structurel pour le raisonnement sur la zone maxillo-faciale avec extension des relations spatiales "Alignés" et "Le long de" dans le domaine 3D (chapitre 4).

— Réalisation d'une application avec interface graphique prête à l'emploi (chapitre 5).

Chapitre 2

État de l'art

Ce chapitre comprend deux parties faisant une revue non exhaustive des méthodes existantes sur les grands axes abordés dans la thèse, et ce que nous cherchons à obtenir. La première partie traite des méthodes de segmentation. Nous y discutons aussi rapidement des propositions qui existent pour la définition et l'utilisation de systèmes à base de connaissances pour guider l'identification du contenu des images en seconde partie. La deuxième partie aborde plus précisément le domaine des relations spatiales, et notamment celles qui ont servi de base à nos travaux d'extension au domaine 3D.

2.1 Segmentation

Nous présentons ici les méthodes de la littérature concernant la segmentation d'images dans l'objectif d'une application à la délimitation de dents. Les critères recherchés dans les approches sont, en plus de la précision (voir chapitre 1) :

- un temps de calcul et un coût en ressources (informatiques) faibles, pour permettre l'adoption de la méthode dans un cadre clinique,
- une facilité d'intégration de contraintes a priori, pour renforcer l'identification des dents, mais aussi dans l'objectif d'étendre la méthode à d'autres structures de la zone maxillo-faciale,
- une ouverture à une interaction avec un utilisateur, pour donner la possibilité au praticien de corriger les résultats,
- une possibilité d'extension à la segmentation simultanée de plusieurs objets.

Les deux volets de cette partie sont différenciés selon leur emploi de l'apprentissage.

2.1.1 Sans apprentissage

Méthodes de minimisation d'énergie

Ces méthodes formulent le problème de segmentation sous la forme d'une énergie qui doit être la plus faible possible, cette dernière traduisant un écart à un modèle idéal de segmentation. L'énergie est exprimée comme un compromis entre plusieurs contraintes, qui proviennent de sources différentes. Chaque terme peut représenter une attache aux données, qui relie la segmentation à des informations extérieures, ou une régularisation, qui dépend de l'état du système à un instant donné. Pour minimiser cette fonction, deux types d'approches sont principalement utilisées. Premièrement il y a les modèles déformables représentés soit par des contours paramétriques [30, 31, 79], soit par des ensembles de niveaux (*level-sets*) [100, 111], qui sont implicites.
À partir d'une position initiale, un modèle déformable évolue selon une équation différentielle fonction du temps pour minimiser l'énergie, pendant un nombre d'itérations fixe ou jusqu'à ce qu'un critère comme une erreur minimale soit atteint. La formulation continue donne lieu à des schémas de résolution qui sont discrétisés, par exemple par la méthode des différences finies ou par des schémas plus optimisés [1].
Le deuxième type d'approche concerne les méthodes sur graphes de type champ de Markov. La formulation est directement discrète, et pour minimiser l'énergie sur le graphe, les trois méthodes les plus

courantes sont la coupure de graphe [22, 23, 24, 51, 68], qui cherche la coupe de moindre coût séparant le graphe en deux, la marche aléatoire [66] qui résout le problème de Dirichlet (trouver une fonction harmonique, minimisant ici l'énergie) par résolution d'un système linéaire, et enfin les méthodes de type recuit simulé reposant sur la modélisation du problème comme un système thermodynamique, mais nous ne les détaillerons pas, ces méthodes souffrant d'un temps de calcul habituellement élevé comparativement aux deux précédentes. Dans ce qui suit, nous donnons des exemples des deux premières classes d'approches appliquées à la segmentation des dents. Soulignons que la formulation du problème de segmentation sous la forme d'une énergie se prête bien à l'intégration d'une contrainte donnée par un utilisateur, c'est-à-dire qu'une correction peut être prise en compte directement au niveau de la résolution.
Dans le contexte de la segmentation en imagerie médicale, les énergies prennent en compte le plus souvent deux éléments : l'apparence des voxels, qui fonde les approches par régions (voir par exemple [31]), et le gradient, qui permet de définir les approches par contours (voir par exemple [30]). Ces deux caractéristiques n'étant pas toujours suffisantes pour modéliser le problème de segmentation, notamment dans le cadre de l'imagerie dentaire où leur pertinence est remise en cause par la présence de métal, il est possible d'utiliser des contraintes supplémentaires, comme par exemple la forme.

Dans [58], les dents sont segmentées individuellement pour chaque coupe axiale grâce à des informations a priori d'intensité et de forme. L'énergie est formulée comme :

$$E_S(\Psi) = \frac{\mathcal{A}}{2}\int_\Omega (|\nabla\Psi(x)|-1)^2 \ dx + \mathcal{B}\int_\Omega (g(x)+\mathcal{P}\phi(x)^2)\delta(\Psi(x))|\nabla\Psi(x)| \ dx + \mathcal{C}\int_\Omega \log\left(\frac{p_B(\mathrm{Im}(x))}{p_A(\mathrm{Im}(x))}\right) H(-\Psi(x)) \ dx, \quad (2.1)$$

où Ψ est la fonction de distance du contour d'un ensemble de niveaux, x un point du domaine Ω de l'image, g une fonction de potentiel de gradient (par exemple $\frac{1}{1+|\nabla \mathrm{Im}(x)|}$), δ la fonction Dirac, ϕ une carte de distance issue d'une segmentation précédente, p_A et p_B des probabilités d'appartenir à l'objet et au fond respectivement, $\mathrm{Im}(x)$ est l'intensité au point x et H la fonction de Heaviside ($\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ et $\mathcal{P}$ sont des poids). Le premier terme est une force interne qui force Ψ à rester une fonction de distance signée (pour assurer une évolution du contour cohérente sans ré-initialiser la carte de distance, voir [86] pour plus de précision), le second est une force externe attirant le contour vers les forts gradients et le contraignant à rester proche d'une précédente segmentation. Enfin le troisième terme est une vraisemblance d'intensité. Pour la partie des couronnes, ces dernières apparaissant le plus souvent fusionnées sans délimitation claire, deux ensembles de niveaux sont employés simultanément pour utiliser la compétition des régions et la formulation devient :

$$E_C(\Psi_1,\Psi_2) = E_S(\Psi_1) + E_S(\Psi_2) + \mathcal{D}\int_\Omega H(-\Psi_1(x))H(-\Psi_2(x)) \ dx,$$

où E_S est défini par l'équation 2.1 et $\mathcal{D}$ est un poids. Le troisième terme sert à pénaliser le recouvrement entre les régions délimitées par les courbes. Les informations de vraisemblance et de forme sont obtenues à partir d'une segmentation initiale donnée par l'utilisateur sur une coupe optimale (c'est-à-dire où les dents sont disjointes). Le coefficient moyen de Dice rapporté sur 3 dents est de 0,978, et le temps de calcul pour un volume entier (124 coupes de résolution isotrope de 300 microns) est d'environ 5 minutes. La segmentation multi-objets est effectuée de manière implicite en tirant parti de la capacité de l'ensemble de niveaux à représenter n'importe quelle topologie (mais le numéro dentaire n'est donc pas automatiquement connu), en initialisant manuellement tous les contours en même temps sur la coupe initiale. Cette initialisation par tracé du contour de chaque dent est simple pour un utilisateur, mais tend à devenir fastidieuse pour une dentition complète. Les principaux défauts de cette approche sont que d'une part il n'est pas toujours possible de trouver une coupe où toutes les dents sont séparées pour initialiser le contour, et d'autre part la contrainte de distance ne repose que sur la segmentation précédente, ce qui tend à propager les erreurs et permet des formes qui ne sont pas celles d'une dent. Enfin, si le traitement coupe par coupe permet un gain de temps de calcul, l'initialisation de chaque coupe étant hypothétiquement très proche de la solution (permettant d'effectuer peu d'itérations), aucune cohérence spatiale dans la direction orthogonale aux coupes n'est

imposée. Notons qu'ici l'interaction de correction est peu aisée car le modèle déformable représente plusieurs dents en même temps, ce qui rend difficile les contraintes locales (c'est aussi le cas pour des contraintes a priori). Deux extensions de cette approche ont été proposées : dans [57] est proposé d'associer un ensemble de niveaux différent à chaque dent pour les identifier, ce qui permet de corriger le problème d'interaction au coût d'un doublement du temps de calcul, et dans [125] est proposée une méthode pour traiter les volumes en occlusion (voir figure 2.1). Notez qu'il est possible d'utiliser les relations spatiales dans une telle approche, voir [5, 36].

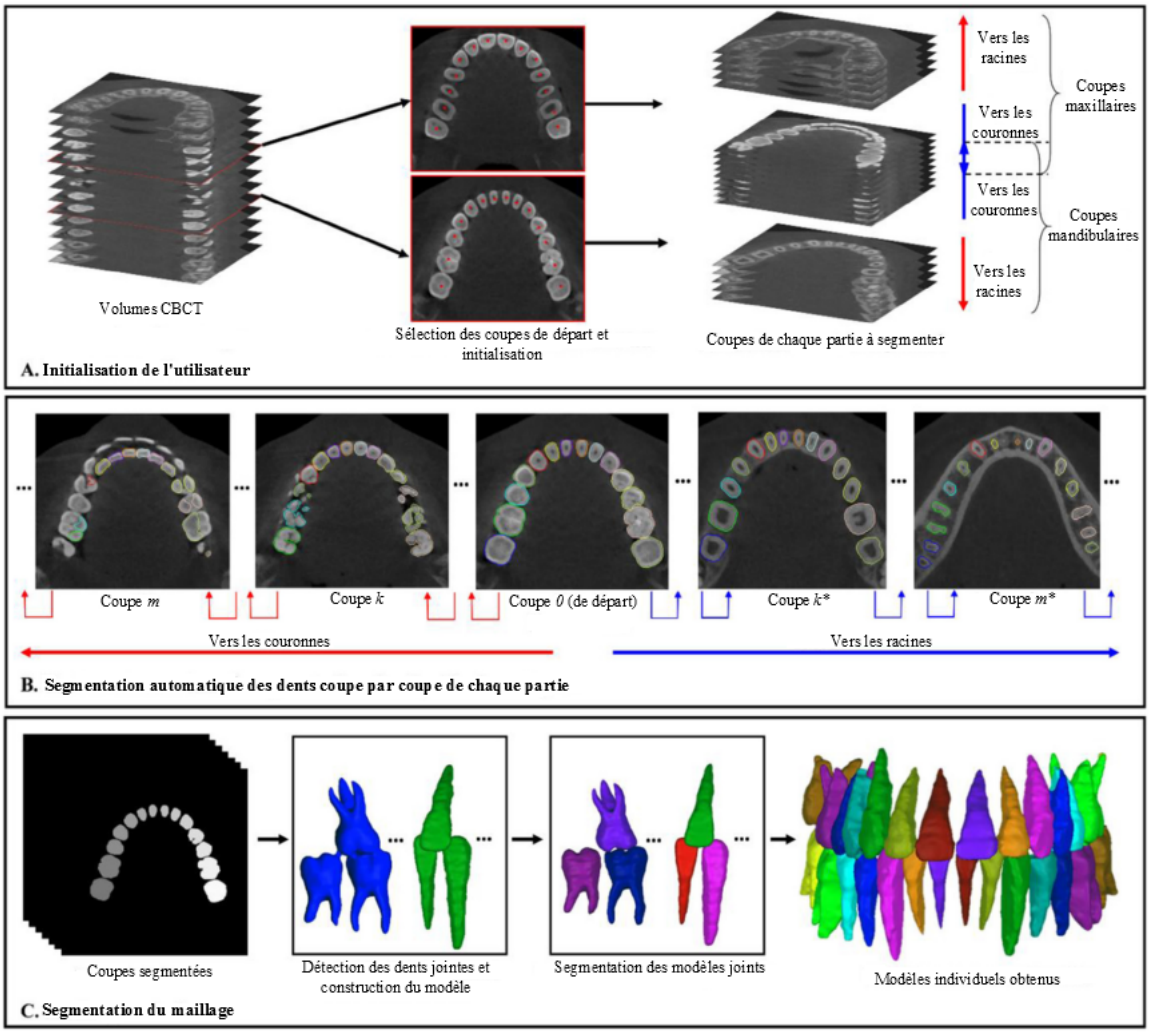


FIGURE 2.1 – Approche proposée initialement dans [58]. L'association d'un ensemble de niveaux par dent est proposé dans [57] et la différenciation des dents en contact durant l'occlusion dans [125]. Illustration traduite de [125].

Dans [73] est présentée une approche par coupure de graphe. À partir de marqueurs indiqués par un utilisateur, le problème de segmentation d'une dent est formulé comme :

$$E(\mathcal{L}) = \sum_{v_i \in \mathcal{V}} -\ln p(\mathrm{Im}(v_i)|\lambda_{v_i}) + \sum_{(v_i,v_j)\in\mathcal{E}} \exp\left(-\frac{(\mathrm{Im}(v_i)-\mathrm{Im}(v_j))^2}{2\sigma^2}\right)\left(\frac{1}{\mathrm{dist}(v_i,v_j)}\right)\cdot|\lambda_{v_i}-\lambda_{v_j}|, \quad (2.2)$$

où $\mathcal{L}$ représente une solution (partition de l'image, ou étiquetage), v_i, v_j sont des nœuds du graphe ($\mathcal{V}$ et $\mathcal{E}$ les ensembles des nœuds et des arêtes respectivement), et $\lambda_{v_i}, \lambda_{v_j}$ sont les étiquettes associées. Le premier terme traduit une vraisemblance sur l'intensité (dont les densité de probabilités ont été obtenues à partir d'une binarisation de l'image d'après la méthode d'Otsu) alors que le second favorise la coupe entre les voxels de forte différence d'intensité. Le score de Dice moyen est de 0,89 pour un temps de calcul de 21 secondes par dent en moyenne sur 25 dents (8 incisives, 6 canines, 7 prémolaires

et 4 molaires, issues de volumes d'une résolution isotrope de 300 microns). Cette étude s'attachant à démontrer les capacités de la coupure de graphe pour la segmentation dentaire, les marqueurs du fond et de l'objet sont obtenus par dilatation et érosion respectivement de la vérité terrain (segmentation manuelle), et chaque dent est segmentée séparément, mais cette formulation est extensible à plusieurs objets, et la coupure de graphe permet d'intégrer facilement une contrainte a priori ou donnée par un utilisateur.

En formulant la segmentation comme un problème de Dirichlet, les auteurs de [130] utilisent le canal radiculaire (obtenu par le plus court chemin de faible intensité entre deux points donnés par un utilisateur) comme marqueur pour identifier les dents. La matrice laplacienne est définie comme :

$$L_{v_i v_j} = \begin{cases} \sum_{v_k \neq v_i} W_{v_i v_k} & \text{si } v_i = v_j, \\ -W_{v_i v_j} & \text{si } v_i, v_j \text{ sont voisins}, \\ 0 & \text{sinon}, \end{cases} \tag{2.3}$$

avec

$$W_{v_i v_j} = \mathcal{A} \cdot \exp\left(-\frac{|\mathrm{Im}(v_i) - \mathrm{Im}(v_j)|^2}{\mathrm{dist}(v_i, v_j)^2}\right),$$

où $\mathcal{A}$ est un poids dépendant du gradient, *dist* la distance euclidienne et v_k appartient à un voisinage défini en 26-connexité. Le champ harmonique φ (voir figure 2.2) est obtenu en résolvant $L \cdot \varphi = 0$ par la méthode des moindres carrés. La segmentation finale est obtenue en sélectionnant l'harmonique h (c'est-à-dire l'iso-surface) qui obtient la meilleure réponse au score Sc :

$$Sc = \frac{1}{|h|}\left(\mathcal{B}\sum_{x\in h}|\nabla \mathrm{Im}(x)| + \mathcal{C}\sum_{x\in h}\left(\frac{1}{|\mathcal{N}_x|}\sum_{y\in\mathcal{N}_x}(\overrightarrow{n_x}\cdot\overrightarrow{n_y})\right)\right),$$

où x, y sont des voxels, $\overrightarrow{n_x}, \overrightarrow{n_y}$ les vecteurs normaux à h en ces points, $|\nabla \mathrm{Im}(x)|$ l'amplitude du gradient et $\mathcal{N}_x$ l'ensemble des voxels appartenant à h dans un voisinage de deux voxels de x ($\mathcal{B}$ et $\mathcal{C}$ sont des poids). Le premier terme favorise les harmoniques positionnées sur les forts gradients tandis que le second permet d'éliminer les courbes très bruitées en favorisant les faible courbures (en partant du principe que le contour de l'objet varie régulièrement). Le coefficient de Dice moyen sur les 16 dents mandibulaires de 8 volumes de patients (de résolution isotrope de 250 microns) est de $0,98 \pm 0,01$ pour un temps de calcul proche du quart d'heure. Les auteurs indiquent que la méthode n'est pas apte à traiter les cas en occlusion, et présente des erreurs lorsque plusieurs canaux sont présents dans une même racine. L'approche ici employée est quasiment similaire à une méthode de marche aléatoire, et peut donc être étendue à une segmentation multi-objets ainsi qu'à l'utilisation de contraintes a priori ou de l'utilisateur.

Nous présentons maintenant des méthodes qui n'ont pas été appliquées pour la segmentation dentaire, mais qui présentent à nos yeux des éléments intéressants pour satisfaire les critères que nous recherchons.

Dans [40] est proposée une méthode d'optimisation énergétique qui provient de la formulation d'une ligne de partage des eaux sur un graphe, le *powerwatershed*, et qui démontre de bonnes performances en temps de calcul, grâce à l'association d'un algorithme de forêt couvrante maximale et de la marche aléatoire, qui procède au calcul de la probabilité des nœuds d'appartenir aux différentes étiquettes uniquement lorsqu'ils constituent un "plateau", c'est-à-dire une zone du graphe reliée aux marqueurs (donnés par l'utilisateur) par des arêtes de poids équivalents. Notez qu'ici λ_{v_i} est une variable continue à valeurs dans $[0, 1]$ et que les étiquettes correspondent à une valeur particulière de cette variable. L'énergie est :

$$E(\mathcal{L}) = \sum_{v_i\in\mathcal{V}} \mathcal{W}\cdot|\lambda_{v_i} - M_{v_i}|^2 + \sum_{(v_i,v_j)\in\mathcal{E}}\left(\max_{x\in\Omega}|\nabla \mathrm{Im}(x)| - |\mathrm{Im}(v_i) - \mathrm{Im}(v_j)|\right)\cdot|\lambda_{v_i} - \lambda_{v_j}|^2,$$

avec

$$\mathcal{W} = \begin{cases} +\infty & \text{si} \quad M_{v_i} \in \{0, 1\}, \\ 0 & \text{sinon}, \end{cases}$$

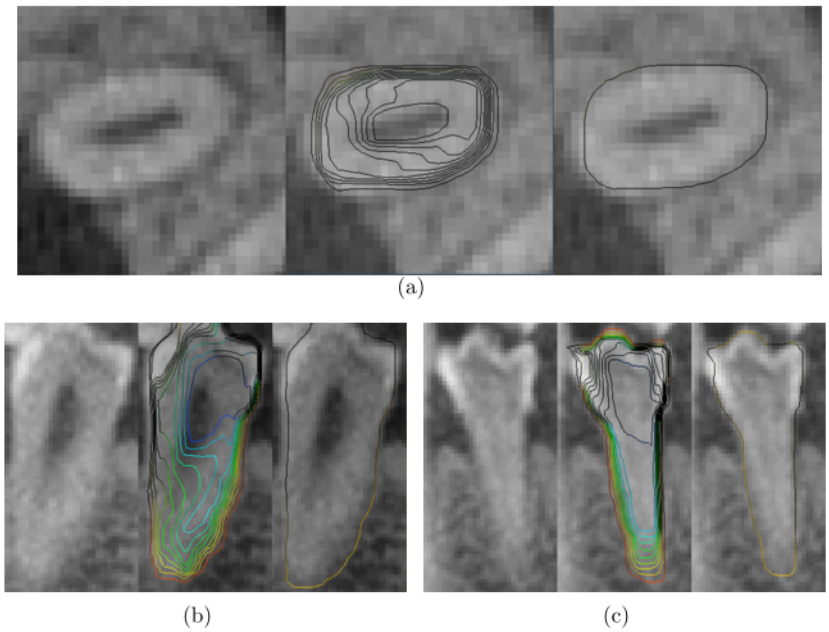


FIGURE 2.2 – Champ harmonique généré à partir du canal. Pour chaque série est présentée la zone d'intérêt, le champ harmonique et l'harmonique retenue. (a) Vue axiale, (b) vue sagittale et (c) vue coronale. Illustration de [130].

où M_{v_i} désigne la valeur du sommet v_i dans une image de marqueurs (où 0 et 1 désignent respectivement le fond et l'objet), et $\nabla \mathrm{Im}$ est le gradient de l'image. Le premier terme permet de contraindre la segmentation à respecter les marqueurs donnés par un utilisateur, tandis que le second favorise la séparation des partitions sur les forts gradients. La minimisation de l'énergie produit une image de probabilité à valeurs dans $[0, 1]$, et la partition définitive est alors obtenue comme :

$$\forall v_i,\ \bar{\mathcal{L}}(v_i) = \begin{cases} 1 & \text{si } \lambda_{v_i} \geqslant 0,5, \\ 0 & \text{si } \lambda_{v_i} < 0,5. \end{cases}$$

Les résultats quantitatifs montrent un score de précision d'environ $0,96$ sur la base d'images GrabCut, avec une robustesse au placement des marqueurs initiaux accrue par rapport à la coupure de graphe et à la marche aléatoire. Cette approche n'a pas été appliquée pour de la segmentation dentaire, mais ses performances en temps de calcul (voir figure 2.3) en font une bonne candidate pour remplacer l'approche par coupure de graphe dans le cadre multi-objets (voir chapitre 3). Au niveau de l'utilisation, cette approche possède les mêmes facilités d'ouverture aux contraintes que la coupure de graphe et la marche aléatoire.

Méthodes par morphologie mathématique

Deux paradigmes différents sont utilisés en morphologie mathématique pour la segmentation. Le premier est de considérer les variations d'intensité comme un relief composé de bassins (identifiés par les minima locaux), où les crêtes forment alors les contours des objets : c'est l'approche par ligne de partage des eaux [9]. La sursegmentation est usuellement limitée par l'utilisation de marqueurs. Notez que le *powerwatershed* est l'équivalent de cette approche lorsqu'elle est formulée comme un problème de type champ de Markov. Le second paradigme repose sur des filtres créant des zones plates. Les approches qui exploitent cette vision reposent sur des filtres connectés, qui peuvent être mis en oeuvre de manière efficace sur des représentations hiérarchiques comme les arbres de composantes [97, 107] ou de formes [29, 62, 126]. Ces filtres agissent sur les nœuds de l'arbre selon un attribut calculé préalablement comme le volume ou le contraste. Nous ne détaillerons pas plus ici cette approche, mais des exemples d'applications dans le monde médical peuvent être trouvés dans [27, 96, 119] (voir aussi l'annexe A, où nous avons tenté cette approche).

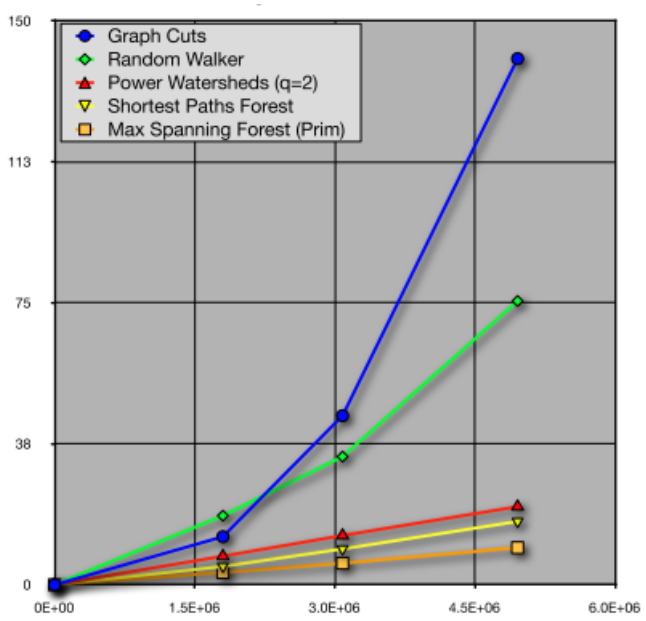


FIGURE 2.3 – Temps de calcul (en secondes) utilisé par les méthodes d'optimisation sur graphe en fonction du nombre de voxels. Notez la performance du *powerwatershed* (en rouge) face à la marche aléatoire (en vert) et à la coupure de graphe (en bleu). Dans le cadre multi-objets, ce gain est multiplié par le nombre d'étiquettes. Illustration de [40].

Une méthode pour segmenter puis compter les dents reposant sur la segmentation par ligne de partage des eaux est présentée dans [110]. À partir d'une coupe axiale Im (figure 2.4 (a)) où les dents sont bien disjointes et d'une projection maximale d'intensité MIP (image axiale où la valeur du pixel (x, y) est celle la plus élevée parmi tous les voxels $(x, y, 0...Z)$, voir figure 2.4 (b)) la segmentation est réalisée selon le processus suivant :

1. Im_m (figure 2.4 (d)) : masquage de Im par une image seuillée de MIP (figure 2.4 (c)),
2. Im_c (figure 2.4 (e)) : "bouchage de trous" (*hole filling*) pour chaque niveau de gris de Im_m,
3. Im_f (figure 2.4 (f)) : obtention d'une image simplifiée telle que $((\text{Im}_c \circ SE) \bullet SE)$ où $\circ$ et $\bullet$ sont respectivement une ouverture et une fermeture par un élément structurant circulaire SE,
4. segmentation de Im_f par ligne de partage des eaux (figure 2.4 (h)),
5. les composantes connexes résultantes ayant une aire trop faible pour être une dent sont ensuite fusionnées entre elles pour réduire la sur-segmentation, voir figure 2.4 (h).

Les résultats montrent un coefficient de Dice moyen de $0,97$ pour un temps de calcul d'environ 2 secondes pour chacune des 1350 coupes axiales (de résolution non précisée) de la base de test. Un avantage de cette approche est que la segmentation est multi-objets par nature. La ligne de partage des eaux avec marqueurs est de plus particulièrement adaptées à l'interaction avec un utilisateur. Il est plus difficile en revanche d'intégrer des contraintes a priori. La méthode présentée ici a pour hypothèse que toutes les dents sont alignées suivant l'axe axial pour obtenir une projection d'intensité maximale représentative de l'arche, or cette hypothèse n'est pas toujours vraie, notamment pour les dents antérieures. De plus les résultats présentés ne possèdent pas d'artéfacts métalliques. Il paraît donc difficile d'étendre cette méthode pour la segmentation de toutes les coupes axiales.

2.1.2 Avec apprentissage

Méthodes de minimisation d'énergie

La souplesse de la formulation par énergie permet d'inclure des contraintes préalablement apprises sur des données annotées dans les énergies présentées dans la section 2.1.1.

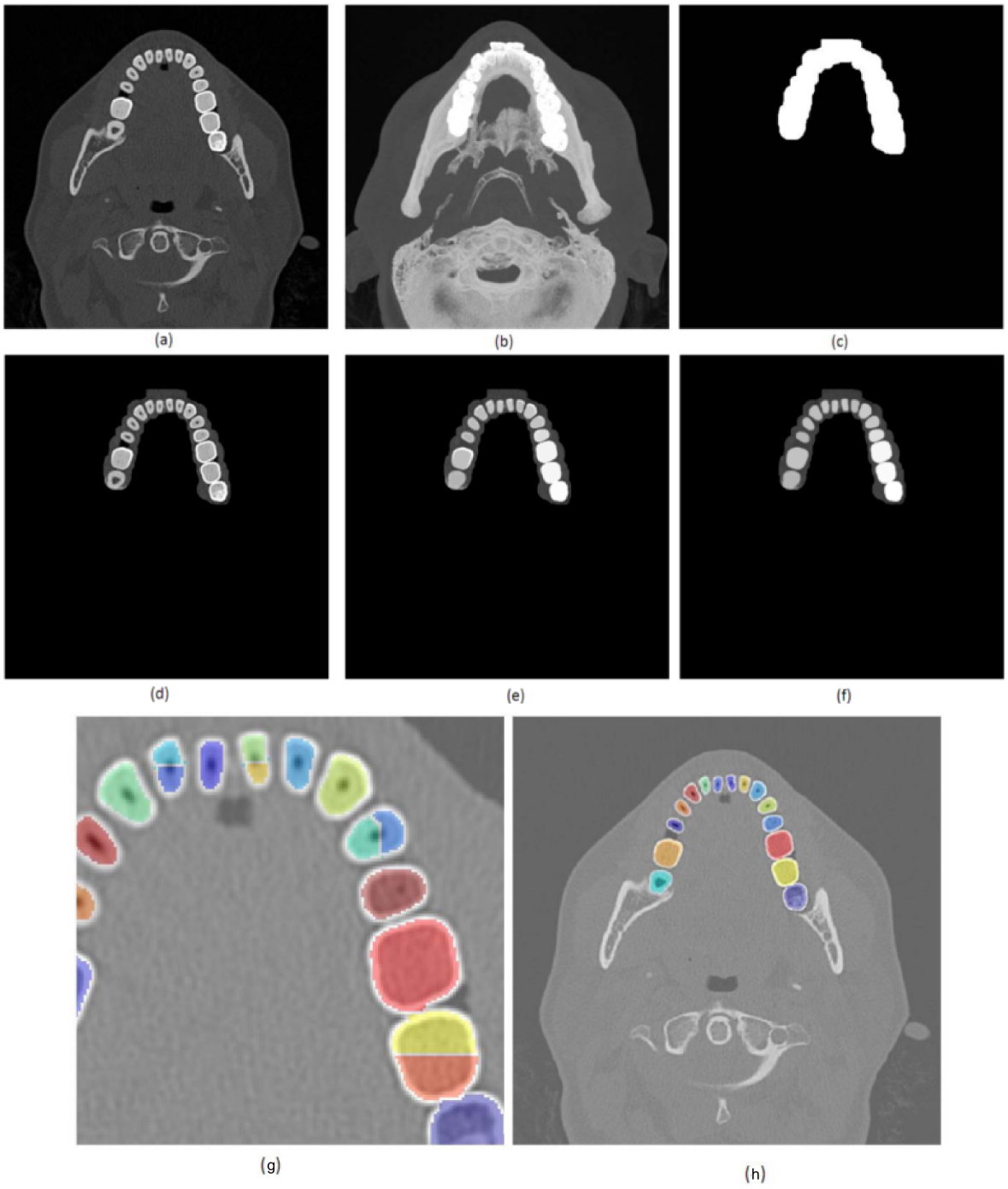


FIGURE 2.4 – Méthode de segmentation par ligne de partage des eaux. (a) Im, (b) MIP, (c) seuillage de MIP, (d) Im_m, (e) Im_c, (f) Im_s, (g) ligne de partage des eaux superposée à Im, (h) résultat final après fusion des bassins de faible aire. Illustration de [110].

Dans [102], l'apprentissage d'une forme repose sur des spécimens de dents qui sont numérisés sous la forme de maillage pour servir de modèles de référence. Une contrainte d'écart au maillage est ajoutée à l'énergie de la segmentation. À partir d'une initialisation donnée par un utilisateur, les dents sont segmentées grâce au processus suivant (voir figure 2.5) :

1. segmentation initiale par marche aléatoire suivant des critères d'intensités (voir ci-après),
2. recalage non rigide du maillage sur la segmentation,
3. segmentation par marche aléatoire contrainte par la position du maillage,
4. répétition des étapes 2 et 3 jusqu'à ce que le maillage du modèle n'évolue plus (ou en-deçà d'un

seuil prédéfini),

5. segmentation finale par une itération finale de l'étape 3.

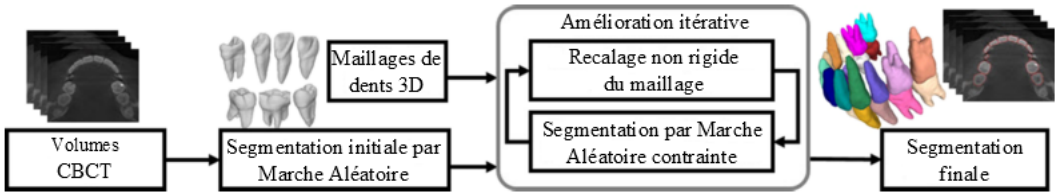


FIGURE 2.5 – Schéma de l'utilisation des maillages de formes comme contrainte de forme dans une segmentation par marche aléatoire. Illustration de [102].

Le problème de segmentation initial est formulé par l'énergie telle que

$$E(X) = H||X - \tilde{X}||^2 + \mathcal{A}X^\intercal LX + \mathcal{B}||X||^2$$

et résolu par le système linéaire suivant :

$$(H + \mathcal{A}L + \mathcal{B}I)X = H\tilde{X}$$

avec I la matrice identité, X la matrice de probabilité recherchée et $\tilde{X}$ la matrice des marqueurs donnés par l'utilisateur ($\mathcal{A}$ et $\mathcal{B}$ sont des poids). La matrice H est définie comme

$$H_{v_i v_j} = \begin{cases} 1 & \text{si } v_i = v_j \text{ et } v_i \text{ est un marqueur,} \\ 0 & \text{sinon,} \end{cases}$$

tandis que la matrice L (laplacienne) est identique à celle de l'équation 2.3 avec

$$W_{v_i v_j} = \begin{cases} \exp\left(-\left((\mathrm{Im}(v_i) - \mathrm{Im}(v_j))^2 + \mathcal{C}||f(v_i) - f(v_j)||^2 \right) \right) & \text{si } v_i, v_j \text{ sont voisins,} \\ 0 & \text{sinon,} \end{cases}$$

où $\mathcal{C}$ est un poids et $f(v_i)$ un descripteur contextuel, ici défini comme la différence d'intensité cumulée entre un voisinage centré sur v_i et un nombre prédéfini de voisinages décalés aléatoirement dans un rayon fixé. Ce descripteur permet d'atténuer l'impact du bruit sur l'intensité en prenant en compte le contexte spatial.
Le recalage non rigide est effectué en minimisant la distance entre les sommets du maillage et les voxels du contour de la segmentation par résolution d'un système linéaire (voir [102] pour plus de précisions). Le modèle déformé permet alors de définir une probabilité d'appartenance à la dent :

$$\widehat{X}(v_i) = \frac{c_{app}(v_i) + c_{pos}(r(v_i), \theta(v_i))}{2},$$

où c_{app} est la vraisemblance selon l'intensité et c_{pos} une vraisemblance selon la position du sommet par rapport au maillage. Sous forme d'énergie la contrainte s'écrit alors :

$$E_{example} = ||X - \widehat{X}||^2,$$

et le problème de segmentation devient alors

$$E(X) = \mathcal{D} \cdot \left(H||X - \tilde{X}||^2 + \mathcal{A}X^\intercal LX + \mathcal{B}||X||^2 \right) + \mathcal{F} \cdot \left(||X - \widehat{X}||^2 \right),$$

où $\mathcal{D}$ et $\mathcal{F}$ sont des poids, ce qui sous forme linéaire se traduit comme

$$(\mathcal{D}H + \mathcal{D}\mathcal{A}L + \mathcal{D}\mathcal{B}I + \mathcal{F}I)X = \mathcal{D}H\tilde{X} + \mathcal{F}\widehat{X}.$$

Le recalage et la segmentation contrainte sont répétés jusqu'à ce que la forme du maillage n'évolue plus. Le temps de calcul pour une dent est de $1,2$ minutes en moyenne, pour un coefficient de Dice moyen de $0,965$, obtenus sur 20 volumes (de résolution isotrope de 300 microns). Une limitation de la méthode repose sur la dépendance du recalage du modèle de forme sur la segmentation. Si la délimitation initiale est peu pertinente, ou si elle dérive durant le processus itératif, alors la forme va se déformer pour donner une contrainte ne représentant plus la dent, amplifiant le phénomène. Les auteurs indiquent que c'est notamment le cas en présence d'artéfacts métalliques dégradants.

En dehors des applications dentaires, plusieurs propositions d'utilisation de contraintes a priori ont retenu notre attention. Une autre manière d'ajouter une contrainte de forme est présentée dans [54]. Les auteurs apprennent une forme modèle à partir d'une segmentation manuelle, et utilisent une carte de distances ϕ pour la représenter. Ils définissent alors le terme :

$$E_{shape} = \sum_{(v_i,v_j)\in\mathcal{E}} \phi\left(\frac{v_i+v_j}{2}\right),$$

qui est ajouté à l'équation 2.2 pour une optimisation par coupure de graphe.
Toujours dans le but d'ajouter une contrainte de forme, les auteurs de [32] optimisent un modèle actif de forme [39] appris sur des images annotées, pour ajouter à l'énergie :

$$E_{model} = \sum_{v_i,v_j\in\mathcal{V}} 1 - \exp\left(-\frac{\text{dist}(v_i,\mathcal{S})}{r}\right),$$

où $\text{dist}(v_i,\mathcal{S})$ est la distance euclidienne du sommet à la forme optimisée $\mathcal{S}$, et r est le rayon du cercle circonscrit à $\mathcal{S}$. Ce terme est ajouté à l'équation 2.2 pour réaliser une segmentation finale par coupure de graphe.

L'idée d'utiliser un modèle statistique est aussi utilisée dans [85], où la segmentation est réalisée par une optimisation conjointe du modèle (appris comme dans [39] sur des segmentations manuelles) et d'un modèle déformable représenté par ensemble de niveaux. L'énergie de la courbe Ψ est celle du contour actif géodésique [30] :

$$E(\Psi) = \int_\Psi g(x)|\Psi'(x)|\ dx,$$

où g est une fonction de potentiel de gradient (par exemple $\frac{1}{1+|\nabla \text{Im}(x)|}$). L'équation d'évolution dérivée de cette énergie est :

$$\frac{\partial\Psi}{\partial t} = g(c+\kappa)|\nabla\Psi| + \nabla\Psi\cdot\nabla g,$$

avec c une force ballon et κ la courbure. L'équation est transformée directement en :

$$\frac{\partial\Psi}{\partial t} = \mathcal{A}\left(g(c+\kappa)|\nabla\Psi| + \nabla\Psi\cdot\nabla g\right) + \mathcal{B}\left(\Psi^* - \Psi\right), \tag{2.4}$$

où Ψ^* est une courbe finale estimée à partir du modèle statistique par un maximum a posteriori :

$$\Psi^* = \arg\max_{s,p} P(s,p \mid \Psi, \nabla\text{Im}),$$

avec s,p les paramètres de forme et de pose respectivement, et

$$P(s,p|\Psi,\nabla\text{Im}) = \frac{P(\Psi|s,p)P(\nabla\text{Im}|s,p,\Psi)P(s)P(p)}{P(\Psi,\nabla\text{Im})},$$

$$P(\nabla\text{Im}|s,p,\Psi) = \exp(-|\mathcal{G}(\Psi^*(s,p)) - |\nabla\text{Im}|\ |^2), \quad P(p) = \mathcal{K},$$

$$P(s) = \frac{1}{\sqrt{(2\pi)^k|M_k|}}\exp\left(-\frac{1}{2}s^\intercal M_k^{-1}s\right), \qquad P(\Psi|s,p) = \exp(-V_{out}),$$

où $\mathcal{G}$ désigne une fonction gaussienne qui modélise la relation entre la position de la courbe finale et le gradient suivant la direction normale, k est le nombre de modes de déformation du modèle statistique, avec M_k la matrice diagonale des k valeurs singulières associées aux modes de déformation du modèle,

et V_{out} le volume de la région circonscrite par la courbe Ψ qui n'est pas aussi incluse dans Ψ^*. Le terme $P(\Psi|s,p)$ (probabilité de voir la courbe actuelle sachant les paramètres du modèle) pénalise le contour qui dépasse le modèle estimé, alors que $P(\nabla \mathrm{Im}|s,p,\Psi)$ permet d'attirer le modèle estimé sur les forts gradients de l'image. Enfin, la probabilité des paramètres de forme du modèle suit une distribution gaussienne des formes autour du modèle moyen. Aucune hypothèse n'est faite sur la position que doit prendre le modèle, d'où une distribution uniforme (de constante $\mathcal{K}$) de la probabilité des paramètres de pose $P(p)$. Ainsi le modèle de forme permet de contraindre le modèle déformable à le respecter grâce à la minimisation de la différence entre les deux (second terme de l'équation 2.4).

Les auteurs rapportent des erreurs de distance maximale d'environ 2 mm pour la segmentation du corps calleux sur des coupes cérébrales, et de 2,7 à 4,4 mm pour la segmentation de vertèbres en 3D (pour un temps de calcul de 6 minutes approximativement). L'optimisation jointe du modèle et de la segmentation nous semble pertinente pour adapter efficacement la contrainte de forme au cas étudié, contrairement aux méthodes où la contrainte est fixée pour l'étape de segmentation. Dans le cadre dentaire, cela pourrait permettre une meilleure adaptation à la variabilité de la forme des dents. À l'instar des méthodes par ensembles de niveaux citées ci-avant, la formulation est ouverte à l'usage de contraintes a priori supplémentaires, ainsi qu'à celles d'un utilisateur pourvu que chaque objet soit segmenté par un modèle déformable différent dans le cas multi-objets (ce qui tend à augmenter le temps de calcul).

Un modèle de forme peut aussi être utilisé pour définir une contrainte géométrique. Dans [69], le modèle statistique est utilisé pour créer des échantillons déformés selon les modes autorisés. Les échantillons sont ensuite comparés pour identifier le domaine spatial susceptible d'être atteint par la déformation du modèle. Une segmentation par coupure de graphe utilisant la formulation de [54] est ensuite effectuée uniquement dans cet espace. Nous utilisons cette approche géométrique dans la formulation présentée dans les chapitres suivants, et plus d'informations peuvent être trouvées dans la partie 3.2.3.

Méthodes par forêts aléatoires

Les forêts aléatoires [2, 74] sont l'extension ensembliste des arbres de décision [95]. L'idée est d'utiliser un ensemble de classificateurs "faibles" (*weak classifiers*) pour réduire le sur-apprentissage. Une forêt aléatoire est composée de multiples arbres de décision raisonnant sur des ensembles de caractéristiques où de l'aléatoire est introduit dans le processus de décision durant l'apprentissage pour décorréler chaque arbre des autres. À chaque nœud de décision, seul un sous-ensemble de caractéristiques, choisi aléatoirement, est considéré. Des extensions pour introduire encore plus d'aléatoire ont été proposées dans [26, 63]. Chaque arbre participe alors à la réponse finale du classificateur. Cette approche à été notamment utilisée dans les images de scanners CT pour la détection et la localisation multi-organes (voir [41, 42, 60]). Pour la segmentation, elle est le plus souvent couplée avec des modèles déformables ou des atlas en temps que source de contraintes [64, 76], mais rarement directement pour la classification. Pour segmenter une image avec des forêts aléatoires, un ensemble de caractéristiques est calculé pour chaque voxel, et les arbres sont construits en utilisant un critère permettant de maximiser la séparation des échantillons de classes différentes à chaque nœud. La classification s'effectue en obtenant la réponse de chaque arbre à l'échantillon testé et en attribuant la classe ayant le plus de votes. Un des problèmes de la classification des voxels est le manque de prise en compte de la cohérence spatiale dû à une décision au niveau du voxel. Des propositions pour la prendre en compte existent, par exemple avec les forêts enchevêtrées [94] qui utilisent les déductions déjà effectuées pour renforcer les suivantes.

Dans [59] est présentée une approche intéressante pour segmenter les côtes dans des images de CT, en les différenciant des autres os grâce à des descripteurs pertinents intégrant le contexte, et ce directement par classification. Dans le contexte dentaire, cela pourrait permettre d'identifier et de segmenter les dents en les différenciant de l'os cortical environnant. De plus la méthode est naturellement compatible avec l'identification multi-objets. La forêt aléatoire, de taille $\mathcal{N}_F$, est construite pour identifier les côtes dans des volumes couvrant tout le thorax. Partant d'un ensemble FV d'échantillons d'apprentissage (vecteurs de caractéristiques calculés sur des voxels annotés), un arbre de décision binaire est construit en considérant uniquement un sous-ensemble $T \subset FV$ choisi aléatoirement, de

taille $\mathcal{N}_T$. Pour chaque nœud de l'arbre est assigné un ensemble de paires "variable-seuil de séparation" : le sous-ensemble de caractéristiques, de taille $\mathcal{N}_V$, est choisi aléatoirement, ainsi que les seuils de séparation (dans la plage de valeurs de la caractéristique concernée). Chaque paire produit ainsi une séparation binaire de T en T_A et T_B, qui est évaluée par son gain G avec le coefficient de Gini $\mathcal{G}$ (coefficient de dispersion) :

$$G = \mathcal{G}(T) - \mathcal{G}(T_A) \cdot \frac{|T_A|}{|T|} - \mathcal{G}(T_B) \cdot \frac{|T_B|}{|T|},$$

$$\mathcal{G} = 2\frac{\mathcal{N}_0 \cdot \mathcal{N}_1}{|T|^2},$$

où $\mathcal{N}_0$ et $\mathcal{N}_1$ représentent les nombres d'échantillons appartenant à la classe 0 et 1 respectivement dans T. La paire "variable-seuil de séparation" qui maximise le gain est retenue pour ce nœud, et les ensembles T_A et T_B correspondants sont utilisés pour générer les nœuds fils. Cette construction est répétée jusqu'à ce que l'arbre ait atteint une profondeur d_{max} prédéfinie ou que $\mathcal{G}(T) = 0$. Les feuilles de l'arbre votent pour la classe la plus représentée dans leurs ensembles d'échantillons. Lors de la phase de classification, le vecteur de caractéristiques représentant un voxel de classe inconnue est donné en entrée de la forêt. Chaque arbre va alors faire correspondre ce vecteur à une de ses feuilles en fonction des seuils appris et voter pour la classe correspondante. La classe emportant le plus de votes est assignée au voxel. L'intégration de relations spatiales ou des corrections de l'utilisateur est peu aisée dans les forêts aléatoires.
Dans cette étude, trois expériences pour différencier les côtes des autres os ont été menées avec trois vecteurs descripteurs différents :

— l'intensité moyenne dans 60 voisinages aléatoires autour du voxel considéré,

— à partir des matrices de covariance des voxels osseux (un voxel osseux en scanner est un voxel d'une valeur $> 400HU$ (unités Hounsfield)) compris dans 21 voisinages sphériques croissants, la valeur de la plus grande valeur singulière et le ratio de cette dernière avec la plus petite sont utilisées pour générer un vecteur de 42 valeurs représentant les axes d'inerties pris par les structures osseuses du voisinage,

— à partir de 56 bandes obtenues en générant aléatoirement des sphères concentriques sur le point considéré, un vecteur de 56 valeurs où chacune représente le nombre de voxels osseux dans une bande donnée est utilisé. Ce descripteur représente le volume d'os présent dans l'environnement du voxel.

Ces descripteurs tirent parti du fait qu'il est possible en image CT de définir précisément la classe des voxels osseux d'après leur intensité grâce aux unités Hounsfield, ce qui n'est pas directement utilisable en CBCT. Cependant, compte tenu de la stabilité des distributions d'intensités des données dentaires que nous utilisons (voir section 3.1.3), il est possible d'évaluer ces descripteurs sur les dents (voir section 3.4.3).

Méthodes par réseaux de neurones

Les récents développements [82] autour des réseaux de neurones convolutifs ont permis des adaptations de ces approches à la segmentation. Le principe général de ces réseaux est de prendre en entrée une image, et de la convoluer successivement par un nombre important de noyaux, dont les valeurs sont les paramètres à optimiser. L'optimisation est effectuée par des méthodes de type descente de gradient, grâce à la rétropropagation de l'erreur. Ces méthodes sont donc des problèmes à très haute dimension (dépassant souvent 10^8 paramètres), qui nécessitent des capacités de calculs importantes pour l'apprentissage. En pratique, l'apprentissage est effectué par calcul distribué sur des GPU pour obtenir des temps d'entraînement satisfaisants. Plusieurs architectures ont été proposées pour la segmentation, en ajoutant des couches entièrement connectées en fin de réseau pour obtenir une décision par voxel [89, 113], en inversant l'architecture pour déconvoluer la réponse [99], ou encore en obtenant une réponse moyenne pour un champ réceptif donné [78]. Un des problèmes importants encore ouverts de l'utilisation de telles méthodes est l'obtention d'une base d'images annotées suffisamment

importante pour pourvoir obtenir de bonnes performances de reconnaissance. Bien qu'il soit possible de recourir à des techniques d'augmentation de données, elles ne peuvent suffire pour répondre aux problématiques de la variabilité morphologique et surtout pathologique. Or l'obtention de données médicales est particulièrement difficile. Dans [103, 117], les auteurs montrent qu'il est possible de s'affranchir partiellement de cette dépendance en utilisant l'apprentissage par transfert : premièrement un réseau pré-entraîné sur des images naturelles peut généraliser la reconnaissance, et deuxièmement il est possible de le faire évoluer pour traiter des données médicales plutôt que de faire un apprentissage entier.

Partant de ce constat, dans [127] est proposée une méthode pour segmenter les différentes parties cérébrales dans des volumes IRM 3D chez le nouveau-né et l'adulte. Le réseau initial est celui à 16 couches décrit dans [113] (VGG16), qui à été entraîné sur la base ImageNet. Les couches entièrement connectées finales sont enlevées, et remplacées par des couches spécialisées qui exploitent l'information de chaque niveau du réseau (voir figure 2.6). Le réseau étant initialement prévu pour des images 2D naturelles, une partie de l'information tri-dimensionnelle est introduite en constituant des images 2D RVB où chaque canal de couleur est en fait une coupe axiale de volume. Chaque image représente donc l'information de trois coupes successives. De cette manière, la qualité de segmentation est renforcée puisqu'il existe une cohérence spatiale entre les images 2D. Cette méthode est particulièrement intéressante car elle permet de prendre en compte l'information 3D tout en ayant des temps d'inférence faibles (jusqu'à trois fois plus rapide qu'une analyse en 3D complète pour un même volume). La plupart des autres réseaux utilisant l'information tridimensionnelle (par exemple [46, 78]) affichent des temps d'inférence longs, de l'ordre de plusieurs minutes. De plus l'architecture permet l'utilisation d'échantillons de tailles variables, ce qui n'est pas le cas des réseaux habituels utilisant des couches finales entièrement connectées.

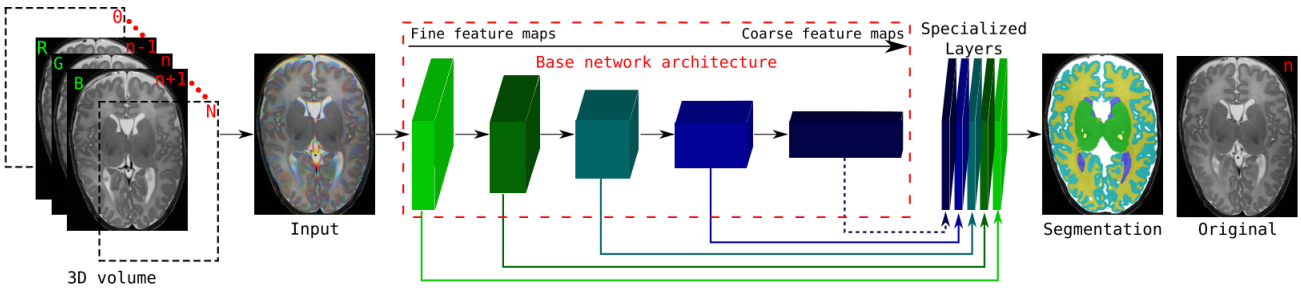


FIGURE 2.6 – Architecture utilisée pour la segmentation des structures cérébrales, à partir du réseau pré-entraîné VGG16. Chaque couche finale synthétise l'information d'un niveau du réseau. Illustration de [127].

Les scores de Dice moyens rapportés sur les structures cérébrales sont compris entre $0,79$ et $0,94$ selon la structure et l'âge du patient (le moins bon score étant obtenu pour la segmentation de la matière grise corticale chez le nouveau-né). Le temps d'apprentissage est d'environ 3 heures, et le temps d'inférence de l'ordre de la seconde. Appliquée au contexte dentaire, cette méthode permettrait de réaliser une segmentation multi-objets de manière immédiate et automatique, ce qui serait un grand avantage car économisant du temps à l'utilisateur. D'un point de vue robustesse, ce type d'approche nous semble pouvoir traiter les dents surnuméraires ou manquantes plus aisément que les autres, grâce à l'analyse implicite de l'ensemble du champ de vue qui apporte des informations supplémentaires de position et de relation entre les dents, comparativement aux autres approches (qui peuvent les intégrer mais elles doivent être explicitement définies). De plus, le réseau pourrait être entraîné à différencier directement les numéros dentaires, apportant un gain supplémentaire comparativement aux méthodes avec énergie. Le désavantage principal des réseaux de neurones est la difficulté de contraindre l'apprentissage, et par extension d'ajouter des informations a priori ou des corrections, même si des stratégies tentent d'y répondre (l'apprentissage par renforcement par exemple).

Approche	Avantages	Inconvénients
Minimisation d'énergie par modèle déformable	-Formulation simple et extensible -Facilement interprétable -Interaction aisée -Peu coûteux en ressources informatiques	-Formulation multi-objets coûteuse en temps de calcul -Nécessite une initialisation -Généralisation dépendant de la formulation
Minimisation d'énergie par méthode sur graphe	-Formulation simple et extensible -Facilement interprétable -Interaction aisée -Peu coûteux en ressources informatiques	-Temps de calcul très dépendant de l'algorithme -Nécessite une initialisation -Généralisation dépendant de la formulation
Morphologie mathématique	-Formulation simple -Facilement interprétable -Prise en compte inhérente des formes -Peu coûteux en ressources informatiques -Extension multi-objets immédiate	-L'objet recherché doit correspondre à un bassin versant ou une zone plate -Peu adaptée en présence de métal causant de faux contours -Formalisation des contraintes d'intensité peu évidente -Pas de prise en compte directe des contraintes a priori (hors marqueurs d'un utilisateur)
Forêts aléatoires	-Méthode automatique -Extension multi-objets immédiate -Bonne généralisation si la base d'apprentissage est adaptée -Peu coûteux en ressources informatiques	-Interaction difficile -Choix de descripteurs pertinents -Cohérence spatiale dépendante des descripteurs -Coûteux en temps de calcul en 3D -Performances dépendant de la base d'apprentissage
Réseaux de neurones	-Méthode automatique -Extension multi-objets immédiate -Adaptable à l'information 3D -Information spatiale et relationnelle implicite sur l'ensemble du champ de vue	-Interaction difficile -Difficilement interprétable et extensible (caractéristiques implicites) -Coûteux en ressources informatiques -Coûteux en temps de calcul en 3D -Performances dépendant de la base d'apprentissage

TABLE 2.1 – Table récapitulative des approches de segmentation envisagées.

2.1.3 Approches globales guidées par des connaissances

Dans le but de parvenir à reconnaître et segmenter le plus grand nombre possible de structures dans les images, une des possibilités est de guider la segmentation par un système à base de connaissance [6, 83]. Un tel système permet de modéliser et d'organiser les informations anatomiques et radiomé-

triques disponibles de manière structurée. Il comprend les connaissances a priori sur un problème (connaissance symbolique), la traduction formelle des ces connaissances, c'est-à-dire exprimées sous une forme utilisable numériquement, et un schéma de réflexion pour la résolution en fonction des deux éléments précédents. La représentation du système peut prendre de multiples formes : règles, logiques de descriptions, réseaux sémantiques ou ontologies, qui différent dans leur manière de raisonner. Les informations numériques sont quant à elles dépendantes du sujet et dans le cas des images peuvent être des distributions d'intensités, des relations spatiales, des modèles de forme, etc. L'exploitation du modèle se traduit alors par une mise en correspondance des informations du modèle avec un cas étudié en considérant le modèle comme une contrainte que la segmentation doit respecter. Utiliser un système à base de connaissances soulève trois problèmes principaux :

— puisque les connaissances sont explicites, le système sera alors limité à un contexte particulier,

— il est nécessaire de trouver comment faire correspondre les connaissances symboliques aux informations numériques, problème plus connu sous le nom de "fossé sémantique" [114],

— les connaissance étant imparfaites, il n'est pas certain que le modèle corresponde à la réalité des observations, et cette erreur doit être prise en compte.

Le modèle que nous présentons dans la partie 4.2 a vocation à servir de base à la création d'un tel système. Des exemples d'inspiration peuvent être trouvés dans [52, 98, 120], qui présentent des modèles reposant sur des graphes structurés par des relations spatiales, permettant la reconnaissance d'images cérébrales et satellites (voir ces références pour plus de détails sur l'exploitation des modèles). Le problème de segmentation global est décomposé en plusieurs segmentations individuelles, qui sont réalisées par une minimisation d'énergie sous contraintes issues du modèle, de manière similaire à ce que nous présentons au chapitre 3. Soulignons que dans le cadre médical, une attention particulière doit être portée à la question de la représentation et de la prise en compte des pathologies, compte tenu de leur forte probabilité d'apparition dans les données [3, 4]. Ces approches nous semblent particulièrement bien adaptées à notre objectif car la représentation par graphe est aisément manipulable et extensible, donc adaptable selon les besoins des praticiens, et du fait de la compatibilité entre la formulation de la segmentation et l'ajout de contraintes a priori.

2.2 Relations spatiales

Les relations spatiales sont d'une grande importance dans la perception et la reconnaissance visuelle [10, 34, 72], et ont prouvé être une information pertinente, notamment dans l'identification des structures cérébrales dans des systèmes à base de connaissances [35, 36, 52, 98], dans des méthodes de segmentation par modèles déformables [5, 17] et dans l'imagerie satellite [120, 122]. Elles sont le plus souvent exprimées en langage naturel, et leur conceptualisation est par nature souvent imprécise [10, 80]. Prendre en compte cette nature lors de la traduction des relations dans un langage formel [55, 56] afin de les utiliser dans les problématiques de vision est un élément important pour conserver toute la pertinence de l'information, mais la plupart des propositions dépendent des objets impliqués (entre une région et une ligne [112], entre des lignes [25] ou encore entre des sphères [84]), ce qui réduit leur utilité à des cas très spécifiques. Un des outils qui participe à cette prise en compte est la théorie des ensembles flous [105, 128, 129], qui considère qu'une région spatiale est définie sur l'ensemble de l'espace, à des niveaux d'appartenance variables. Les relations spatiales sont donc tout particulièrement indiquées pour être formulées comme floues, et bon nombre de relations ont déjà été étudiées dans la littérature, comme les relations directionnelles [14, 19] (devant, à droite, sous, etc.), les relations de distance (proche, loin, "à x centimètres", etc.) et l'adjacence [16], la relation "entre" [18], l'encerclement [121], l'alignement [45, 122] ou encore la relation "le long de" [112, 118] (une revue est présentée dans [15]). À notre connaissance, il n'existe pas d'étude sur les relations spatiales entre les dents dans les images. Les relations les plus simples peuvent être classées selon trois types, les relations directionnelles, les distances ou les relations topologiques (inclus, adjacent...). Les relations plus complexes (autour, entre...) peuvent reposer sur une composition de relations simples ou avoir une définition propre. Les relations telles que les relations directionnelles ou de distance ont été définies dans des espaces de dimension quelconque, mais il existe des relations plus complexes non décrites en

3D telles que l'alignement ou "le long de". Nous proposons des extensions des définitions de ces deux relations dans la section 4.3.

2.3 Conclusion

En résumé, la littérature de la segmentation des dents comprend trois types de méthodes : les optimisations par modèles déformables et par graphe, ainsi qu'une par morphologie mathématique. À notre connaissance, il n'existe pas d'applications utilisant des approches par classification (forêt aléatoire ou réseau de neurones) pour cette tâche, mais des exemples peuvent être trouvés dans l'identification du type dentaire [91], ou dans la différenciation maxillaire/mandibulaire [123]. Pour les applications de segmentation, toutes les méthodes partagent les mêmes limitations. En premier lieu, la différenciation entre la dentine et l'os cortical est peu évidente car reposant sur des structures de faible taille, dont l'information est facilement perdue si la qualité de l'image est diminuée. Deuxièmement, les cas de patients en occlusion sont des cas complexes par l'absence d'interfaces visibles fiables au niveau des couronnes. Enfin, la présence de métal conduit à des altérations non prévisibles et souvent dégradant les contours, qui perturbent la reconnaissance des dents. Toutes les études qui démontrent des résultats du niveau de l'état de l'art s'appuient soit sur des contraintes anatomiques soit sur l'identification de structures très significatives comme les canaux radiculaires. Les méthodes qui formulent le problème de segmentation comme une optimisation d'énergie donnent les résultats les plus précis.

Pour notre application, les données peuvent contenir des artéfacts métalliques et il peut être difficile d'identifier les structures internes des dents. Nous souhaitons aussi que la méthode de segmentation soit ouverte à une manipulation aisée et rapide par le praticien. Les formulations efficaces par ensembles de niveaux utilisent des optimisations ("bande étroite" ou représentation creuse) pour réduire la recherche autour du contour précédent, et gagner en vitesse de calcul. Cela nous semble problématique pour faire évoluer rapidement la segmentation lorsque les marqueurs initiaux sont éloignés du contour attendu ou que l'interaction de modification est effectuée loin de la segmentation. Les méthodes qui reposent sur la classification (forêts aléatoires ou autres) pour effectuer la segmentation nécessitent de trouver des descripteurs pertinents, qui prennent en compte le contexte et la cohérence spatiale. Or plus les descripteurs sont complexes plus ces méthodes sont coûteuses en temps de calcul. De plus elles ne se prêtent pas facilement à une interaction extérieure. Le cas des réseaux de neurones est intéressant, car nous avons évoqué qu'il était possible de les utiliser, malgré une pénurie de données annotées, grâce au transfert d'apprentissage. Cependant, deux éléments sont problématiques. Premièrement il est difficile d'interagir avec le réseau pour la prise en compte d'informations données par l'utilisateur, même si des propositions en ce sens commencent à apparaître, notamment à travers l'apprentissage par renforcement [33, 92]. Enfin, l'interprétation du fonctionnement de la segmentation est peu évidente. Des outils ont été créés [20] pour analyser les couches de convolution, mais le résultat n'est pas toujours compréhensible pour un humain. Il est de ce fait compliqué de comprendre sur quoi se repose le réseau pour segmenter, et donc de l'améliorer. C'est particulièrement problématique dans le contexte dentaire, où la variabilité de l'aspect des images est très importante du fait de la présence d'artéfacts métalliques, de positions mandibulaires différentes, et de l'évolution des arcades dentaires au cours du temps.

C'est pourquoi les approches qui utilisent une formulation sur un graphe (coupure de graphe ou marche aléatoire) nous semblent bien adaptées au problème puisqu'elles permettent une interaction et une correction faciles par l'utilisateur, pour des temps de calcul raisonnables, tout en ayant une formulation intelligible.

Ce choix de formulation est aussi en adéquation avec l'objectif d'utiliser des relations spatiales, qui sont facilement adaptables dans ce cadre. De manière plus prospective, nous avons vu que ces approches sont aussi utilisées dans les approches de segmentation guidées par des systèmes à base de connaissances, ce qui permettra de faire le lien entre notre modèle structurel et la méthode de segmentation.

Chapitre 3

Segmentation mono-objet des dents

Dans ce chapitre nous proposons une méthode de segmentation d'une dent. La méthode comporte trois parties. Dans un premier temps, des modèles statistiques des dents sont établis à partir de données annotées (partie dite hors-ligne, section 3.1). Ces modèles servent ensuite à contraindre la segmentation dans une nouvelle image (partie dite en-ligne). Lors de l'analyse de nouvelles données, le modèle correspondant au numéro dentaire indiqué par l'utilisateur est utilisé de manière à déduire l'espace où la dent peut se trouver (section 3.2) avant de procéder à la segmentation sous contraintes de celle-ci (section 3.3). Un schéma d'ensemble récapitulatif est présenté en figure 3.1.

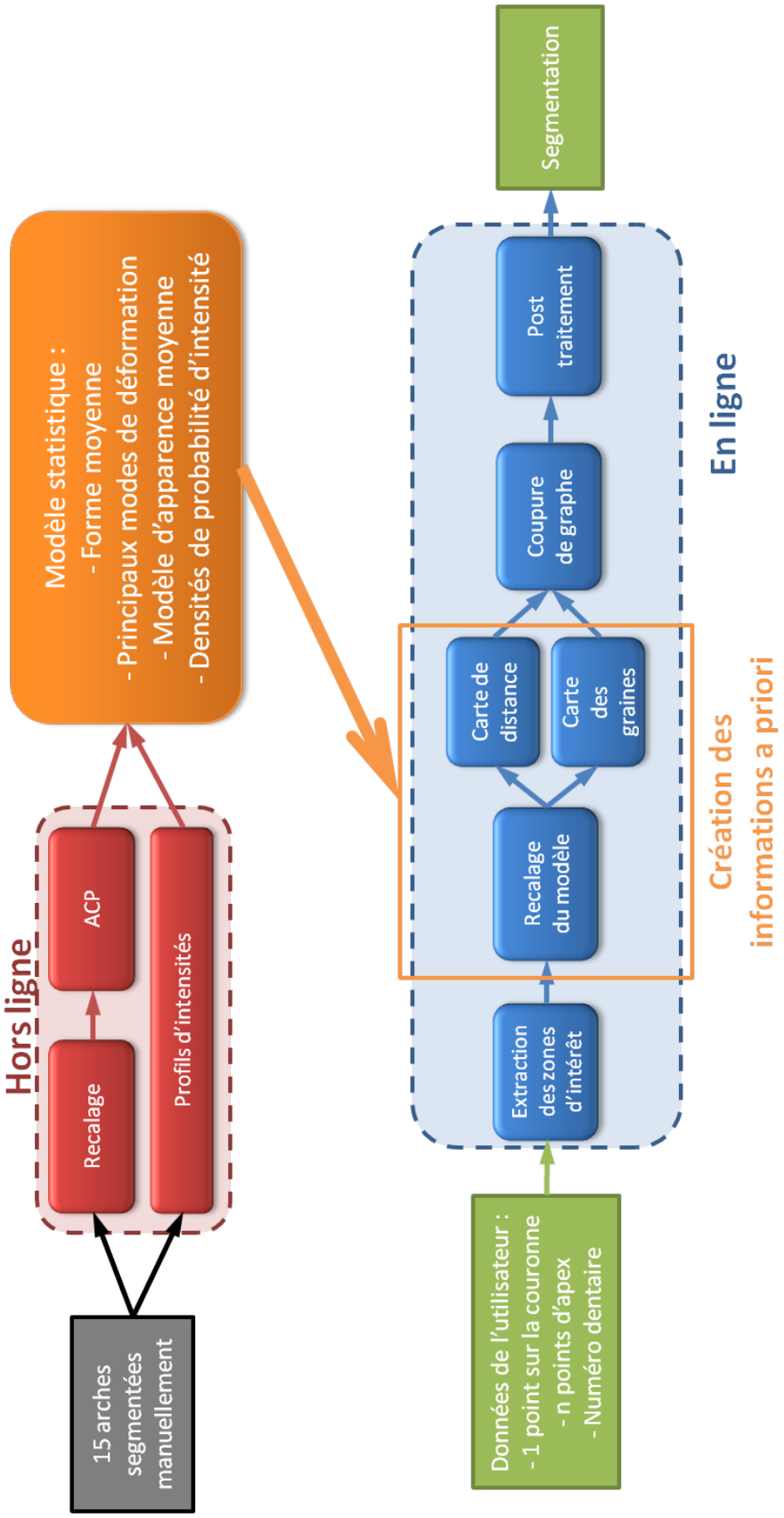
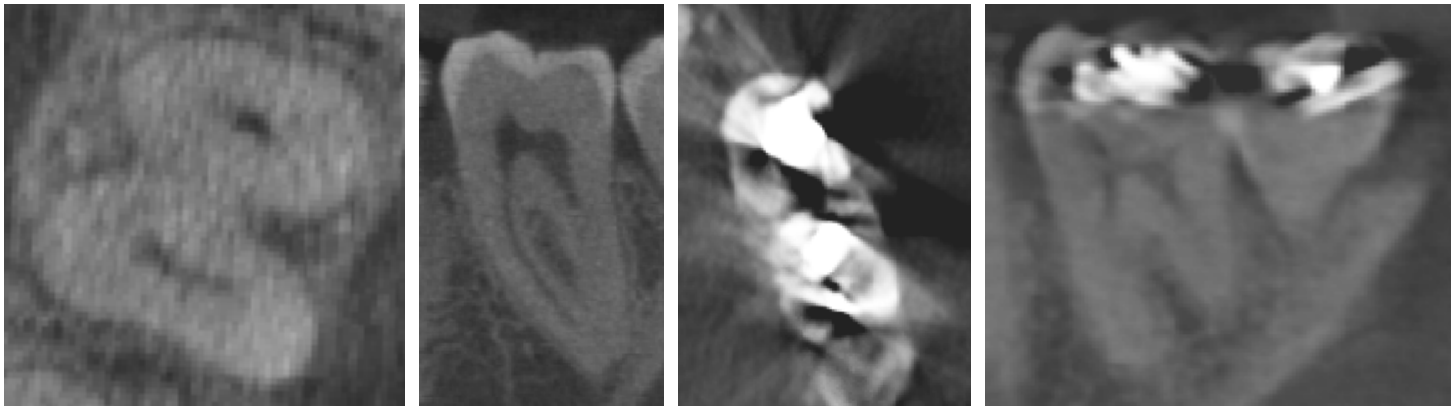


FIGURE 3.1 – Schéma de la méthode de segmentation dentaire. La partie hors ligne s'effectue en amont afin de générer les modèles nécessaires à l'application des contraintes de la partie en ligne.

3.1 Connaissances a priori sur la forme et l'apparence des dents en CBCT

3.1.1 Caractéristiques dentaires

Comment se caractérise une dent dans un volume CBCT ? Il s'agit d'une structure qui présente une forte intensité dans sa partie coronale car l'émail est très radio-opaque (il absorbe beaucoup de photons), et une intensité moyenne pour sa partie interne et inférieure (opacité due à la dentine). Lorsque la résolution et la qualité d'image le permettent, nous observons une structure peu intense en son centre avec des prolongations jusqu'aux apex (extrémités des racines), c'est la chambre pulpaire prolongée par les canaux radiculaires. Dans l'ensemble la dent se différencie assez bien de son environnement sauf avec les structures dures, telles que l'os cortical qu'elle traverse, qui, de par sa composition proche de la dentine, apparaît similaire en intensité. La distinction peut parfois être faite lorsque le tissu conjonctif qui entoure la partie incluse de la dent (ligament parodontal) est visible, car il est peu intense et marque la séparation (voir figure 3.2). L'aspect de la dent ne présente pas de texture spécifique. La forme est en revanche relativement particulière, une partie ovoïde (couronne et collet) avec des prolongements (racines) dont le nombre et la forme dépendent du type de dent. Lorsqu'une partie de la dent est métallique, par une couronne, un comblement ou un implant, nous retrouvons la très forte intensité ainsi que les artéfacts inhérents au métal qui conduisent à une perte des données d'importance variable (en fonction des paramètres d'acquisition et de la méthode de reconstruction volumique). Il est intéressant de noter que l'intensité moyenne des dents s'est révélée stable dans nos données quel que soit le type de scanner et de champ de vue considérés (exemples en figure 3.15) : la table 3.4 et la figure 3.5 illustrent que les valeurs des structures osseuses sont sans équivoque par rapport à celles des autres types de tissus. Ce dernier point est dû au jeu de données utilisé, qui provient d'un seul fournisseur de scanner et d'une seule méthodologie de reconstruction. Il est probable que des données issues d'autres constructeurs auraient un aspect différent. Pour rappel (voir chapitre 1), la méthode CBCT ne permet pas l'estimation des unités Hounsfield et par conséquent ne possède pas de référentiel commun entre les différents scanners.



(a) Coupe axiale d'une molaire.

(b) Coupe sagittale d'une molaire.

(c) Comblements métalliques dans des couronnes de molaires (vue axiale).

(d) Comblements métalliques dans des couronnes de molaires (vue sagittale).

FIGURE 3.2 – Apparence générale des dents dans un volume CBCT. (a) Notons le léger liseré autour de la dent, le ligament parodontal, qui permet de distinguer de l'os cortical. (b) Nous voyons distinctement l'émail (forte intensité), la dentine (intensité modérée) ainsi que la chambre pulpaire prolongée par des canaux radiculaires (faible intensité). (c,d) La présence de métal conduit à des pertes de données sous la forme de raies hyper-intenses et de zones hypo-intenses.

Quelles sont donc les problématiques associées à la segmentation dentaire ? Le premier obstacle majeur est l'absence de délimitation fiable pour séparer chaque dent des tissus durs environnants. Au niveau externe, les couronnes sont adjacentes entre elles et les scanners ne possèdent pas une

résolution suffisante pour les individualiser, ce qui se traduit par un chapelet d'émail hyper-intense dans les images. C'est aussi le cas lorsque les patients sont en occlusion (mâchoire serrée), ce qui brouille la délimitation des faces occlusales. Au niveau interne, la similarité entre l'os cortical et la dentine est si importante que même un œil expert peine à les séparer dès lors qu'ils sont en contact. Ce phénomène est amplifié dès que la résolution décroît, ce qui tend à faire disparaître le ligament parodontal qui est la seule délimitation visible. Le deuxième obstacle est dû au contexte dentaire lui même : la présence de métal de manière récurrente et usuelle. Comme présenté précédemment, cela détruit fortement l'information contenue dans les images, en masquant des zones entières ou en introduisant de faux contours. Enfin, si la morphologie générale est connue, la variabilité physiologique des formes dentaires reste importante. Soulignons en revanche que les pathologies impactent peu notre problématique. Les fractures conduisent soit à une absence partielle des couronnes, ce qui n'est pas pénalisant, soit à une fine fissure dans les racines, ce qui ne modifie pas réellement notre contexte. Les caries vont entraîner des zones hypo-intenses à la surface coronale. Les perte osseuses conduisent à une rétraction de la limite cervicale, ce qui ne modifie pas la morphologie de la dent. Les abcès et lésions péri-apicales quant à eux tendent à apparaître comme des masses sombres qui peuvent modifier la position des dents, mais peu leur forme intrinsèque.

3.1.2 Choix du modèle

Afin de rassembler des informations a priori pour guider la segmentation, des modèles statistiques (figure 3.3) ont été créés à partir d'une base de 15 volumes de patients où les dents ont été segmentées par un expert radiologue. L'estimation de ce modèle repose sur l'analyse d'un ensemble d'échantillons, ici des données segmentées. Une fois le modèle établi, le problème de reconnaissance et de segmentation peut être contraint par l'hypothèse que la nouvelle image est aussi issue de ce modèle, et ne doit donc pas en dévier. Notre choix d'utiliser une telle approche repose sur le fait que les dents, comme en général tout organe, possèdent une variabilité importante qu'il est difficile de prédire, mais que cette variabilité se répartit autour d'une structure commune : une couronne composée d'émail et de dentine, se prolongeant par un certain nombre de racines elles aussi composées de dentine, le tout incluant une chambre pulpaire et des canaux radiculaires. Ce modèle structurel, que nous concevons ici comme un ensemble unique (nous ne distinguons pas les composants dentaires), se traduit par une forme et une apparence spécifiques (précisées dans la partie suivante) qui, modélisées, nous permettent d'obtenir de l'information supplémentaire pour mieux segmenter les images. La première ligne de la figure 3.1 résume le procédé qui permet, à partir de données annotées, d'estimer des modèles dentaires.

3.1.3 Création des modèles

Les dents ont été réparties en 10 classes différentes de manière à représenter des groupes de morphologie cohérente (table 3.1). Pour chaque catégorie, le modèle statistique est composé des éléments suivants :

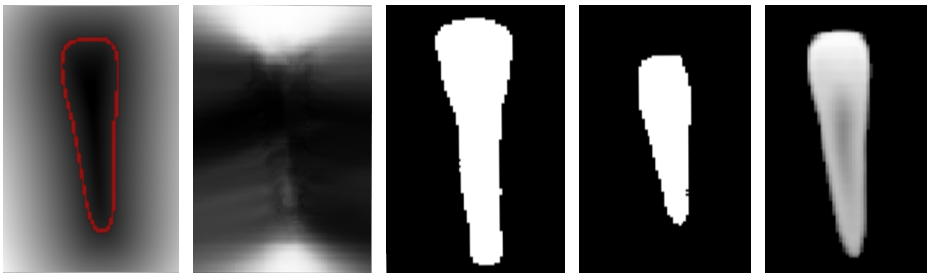
— une forme moyenne $\bar{\mathbf{M}}$ (sous la forme d'une carte de distances au contour),

— les 5 principaux modes de variations de forme $\{\mathbf{V}_1, ..., \mathbf{V}_5\}$ (nous verrons que cinq modes sont suffisamment représentatifs des variations possibles),

— une dent d'apparence moyenne $\bar{\mathbf{A}}$,

— une densité de probabilité des intensités.

Ces éléments ont été obtenus en suivant le processus décrit dans [38, 47] pour réaliser un modèle d'apparence actif et qui est rappelé ici. Dans un premier temps, il est nécessaire de recaler les N échantillons entre eux pour maximiser le recouvrement. Chaque échantillon est une image de dent segmentée où le fond a été masqué avec une valeur d'intensité nulle (semblable à l'image (e) de la figure 3.3). Les échantillons sont ensuite binarisés puis recalés par une transformation rigide (rotation et translation uniquement), afin d'éliminer les variations de pose pour ne retenir que les variations intrinsèques, sur un échantillon de référence choisi arbitrairement (pour la première itération) ou sur la forme moyenne binaire (pour les suivantes) [39]. Ce processus dit "de recalage des formes" est itéré plusieurs fois en mettant à jour la forme moyenne binaire jusqu'à convergence, c'est-à-dire



(a) (b) (c) (d) (e)

FIGURE 3.3 – Modèle statistique d'incisive mandibulaire. (a) Forme moyenne ($\bar{\mathbf{M}}$) avec l'isoligne 0 en rouge. (b) Variation associée au premier mode (premier vecteur propre mis sous la forme d'image). L'ajout ou la soustraction de cette variation à $\bar{\mathbf{M}}$ permet d'explorer l'espace des formes selon cet axe. (c,d) Deux exemples de formes binaires issues d'une variation selon le premier mode de déformation. Nous notons ici que ce dernier encode la variabilité de la taille des incisives (dents plus ou moins grandes). (e) Dent d'apparence moyenne $\bar{\mathbf{A}}$.

jusqu'à ce qu'elle soit stable. Chaque échantillon binaire recalé est alors transformé en représentation implicite sous forme de carte de distance au contour, notée $\mathbf{m}_i$, les valeurs positives et négatives étant respectivement à l'extérieur et à l'intérieur de l'objet. En considérant le modèle comme une image, cela permet de s'affranchir de la contrainte des modèles de distribution de points qui nécessitent de trouver une correspondance entre les points du modèle et de l'image. La forme moyenne $\bar{\mathbf{M}}$ est obtenue par :

$$\bar{\mathbf{M}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{m}_i.$$

Notons que $\mathbf{m}_i$ est un vecteur qui, dans la formulation originale, correspond à un ensemble de points de repère (modèle à distribution de points), mais qu'ici il contient l'ensemble des valeurs de distance de tous les voxels, ce qui implique que tous les échantillons aient le même nombre de voxels p. Une analyse en composantes principales (ACP) est ensuite réalisée sur ces données afin d'obtenir les principaux modes de variation. L'ACP est réalisée selon la méthodologie décrite dans [65] et rappelée ici. Soit $\mathbf{D}$ la matrice $p \times N$ telle que $\mathbf{D} = [\mathbf{m}_1, \mathbf{m}_2, ..., \mathbf{m}_N]$:

1. Calcul de la matrice de Gram $\mathbf{I} = \mathbf{D}^{\mathrm{T}}\mathbf{D}$.
2. Calcul de la matrice de dispersion $\mathbf{C_I} = \frac{1}{N}\left[(\mathbf{I} - \mu_{\mathbf{I}}e)(\mathbf{I} - \mu_{\mathbf{I}}e)^{\mathrm{T}}\right]$ où $\mu_{\mathbf{I}}$ est le vecteur moyen des vecteurs colonnes de $\mathbf{I}$ et $e = [1, 1, 1, ..., 1]$ (de dimension $1 \times N$).
3. Calcul des vecteur propres $(\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_N)$ de $\mathbf{I}^{-1}\mathbf{C_I}$.
4. Obtention des principaux modes de déformations $(\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_N)$ avec $\mathbf{V}_i = \mathbf{D}\mathbf{u}_i$.

L'avantage de cette méthode sur l'habituelle matrice de covariance est que la décomposition en vecteurs propres (coûteuse en calcul) s'effectue sur une matrice $N \times N$ plutôt que sur une matrice $p \times p$ et que $N << p$ (le nombre d'échantillons est nettement inférieur au nombre de voxels en 3D). N'importe quelle forme comprise dans le modèle peut être approchée par :

$$\mathbf{m} = \bar{\mathbf{M}} + \sum_{i=1}^{N} w_i \mathbf{V}_i, \tag{3.1}$$

où w_i est le poids donné au mode. Dans notre cas, nous avons conservé les cinq premiers modes, leur variance cumulée représentant plus de 90% de la variance totale pour toutes les catégories. La dent d'apparence moyenne $\bar{\mathbf{A}}$ est quant à elle obtenue de la manière suivante :

1. La transformation rigide obtenue lors du recalage des formes est appliquée aux échantillons originaux, avant binarisation.
2. Un recalage non rigide par B-splines est appliqué de manière à faire correspondre au mieux les échantillons à la forme moyenne binaire grâce à un critère d'information mutuelle [90].

Catégorie	Échantillons	Morphologie générale habituelle
Incisives mandibulaires	56	Forme elliptique avec une seule racine et l'extrémité coronale aplatie formant une arête. Chambre pulpaire de diamètre similaire au canal.
Canines mandibulaires	27	Forme elliptique particulièrement longue avec une seule racine. Partie coronale en pointe (mono-cuspide). Chambre pulpaire de diamètre supérieur au canal.
Prémolaires mandibulaires	52	Forme d'amande avec une seule racine possédant un ou deux canaux, partie coronale avec deux cuspides. Chambre pulpaire bien visible.
Molaires mandibulaires	64	Partie coronale multicuspidée nettement distinguable de la partie racine (au nombre de deux). Chambre pulpaire importante prolongée par deux canaux dans la racine mésiale, et un ou deux canaux dans la racine distale.
Incisives maxillaires	51	Similaires aux mandibulaires mais de taille plus importante.
Canines maxillaires	26	Similaires aux mandibulaires.
Premières prémolaire maxillaires	25	Anatomie très variable, de forme globale en amande. Le plus souvent deux racines avec deux canaux, mais parfois une seule racine ou un seul canal.
Secondes prémolaires maxillaires	26	Anatomie très variable, de forme globale en amande. Le plus souvent une racine avec un canal, mais parfois deux racines ou deux canaux.
Molaires maxillaires gauches	31	Partie coronale multicuspidée nettement distinguable de la partie racine (au nombre de trois), plus imposante que les mandibulaires. Les racines sont disposées en triangle avec deux s'alignant du côté vestibulaire quand la dernière, usuellement plus longue et divergente, part vers le palais.
Molaires maxillaires droites	31	Similaires aux maxillaires gauches en miroir.
Total	389	

TABLE 3.1 – Répartition des dents segmentées manuellement en classes pour la construction d'un modèle statistique de forme.

3. La valeur finale en chaque point est la moyenne des intensités après recalage non rigide.

Enfin, pour modéliser l'aspect de chaque catégorie, la densité de probabilité des intensités est calculée à partir de l'ensemble des intensités des voxels appartenant aux dents segmentées (cela est pertinent car les données utilisées ont des profils d'intensité stables, voir ci-après), comme l'histogramme normalisé. Nous procédons de même avec les voxels non segmentés pour constituer la fonction associée au fond des images (figure 3.4).

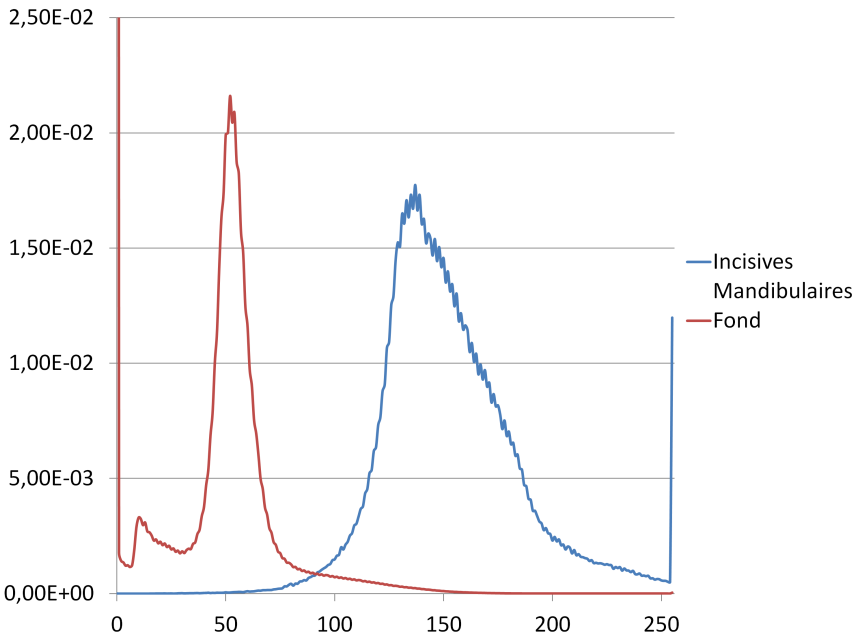


FIGURE 3.4 – Distribution d'intensité de la catégorie des incisives mandibulaires (bleu) superposée à celle du fond (rouge). Les 256 groupes d'intensités représentent une gamme réelle de [0,5120] (chaque groupe contient donc 20 niveaux de gris). Cette gamme correspond aux intensités représentatives des tissus rencontrés. Des valeurs supérieures peuvent exister mais elles proviennent d'éléments métalliques et ces dernières ont été limités à 5120. Le pic final de la courbe s'explique donc par la présence de métal dans certaines incisives segmentées.

Compte tenu de la stabilité des intensités dans nos données, il ne nous a pas paru critique de modéliser la variabilité de l'apparence (voir table 3.2). Les deux cas extrêmes sont superposés sur la figure 3.5. Nous pouvons voir que la variation est peu importante et ne vient pas se confondre avec le fond. De plus, une part importante de cette variabilité ne relève pas d'une modification intrinsèque d'apparence des dents, mais de la présence de métal. Par exemple dans le cas 13 (figure 3.6), le patient présente un appareil orthodontique constitué de bagues et de fils en métal (ou "arcs"), qui sont adjacents aux dents, entraînant une dégradation forte des données dentaires, avec une chute des intensités et un signal imprécis dans la périphérie de cet appareil, expliquant l'aspect de la densité de probabilité. De même, le cas du patient 15 (figure 3.7) montre l'augmentation de l'intensité moyenne par la présence de métal à l'intérieur des dents. Le caractère complexe de l'impact des interactions physiques sur l'image finale rend difficile la modélisation en amont d'une telle variabilité. C'est pourquoi nous avons considéré l'apparence dentaire comme stable, sans en modéliser les variations.

Néanmoins, la généralité du modèle pourrait être renforcée en prenant en compte ces variations, notamment si l'on souhaitait intégrer des images de sources différentes. Nous souhaitons attirer l'attention sur le fait que si la présentation des densités de probabilités semble indiquer que les dents sont facilement séparables du fond, c'est parce que la quantité de voxels de basse intensité est nettement supérieure à celle des voxels d'os cortical, d'où la faible probabilité d'une intensité similaire à celle de la dent dans le fond. L'intensité seule n'est donc pas suffisante pour distinguer la dent de l'os cortical (que nous considérons comme fond), ses valeurs étant comprises dans une plage approximative de $[1500, 3500]$ (soit les groupes de 75 à 175 sur les figures 3.4 et 3.5).

3.2 Le modèle statistique comme contrainte spatiale

Une fois les modèles établis, l'utilisateur peut procéder à l'analyse de nouvelles données. Dans la première étape (figure 3.1), son interaction est utilisée pour générer des contraintes utilisées dans la seconde étape.

Patient (Scanner ; Type de champ)	Moyenne de l'intensité dentaire
1 (CS8100 3D ; Moyen)	$2737,5 \pm 601,5$
2 (CS8100 3D ; Moyen)	$2562,2 \pm 572,2$
3 (CS8100 3D ; Moyen)	$2782,3 \pm 671,3$
4 (CS8100 3D ; Moyen)	$3026,8 \pm 769,5$
5 (CS9300 ; Grand)	$2312,8 \pm 441,2$
6 (CS9300 ; Grand)	$2218,5 \pm 442,8$
7 (CS9300 ; Grand)	$2319,5 \pm 498,3$
8 (CS9300 ; Grand)	$2428,9 \pm 488,4$
9 (CS9300 ; Grand)	$2191,6 \pm 489,4$
10 (CS9300 ; Grand)	$2346,5 \pm 558,6$
11 (CS9300 ; Grand)	$2439,4 \pm 565,6$
12 (CS9300 ; Grand)	$2498,1 \pm 588,8$
13 (CS9300 ; Grand)	$2154,4 \pm 458,9$
14 (CS9300 ; Grand)	$2580,5 \pm 646,2$
15 (CS8100 3D ; Moyen)	$3361,1 \pm 621,8$
Moyenne globale	$2530,7 \pm 320,9$

TABLE 3.2 – Intensité moyenne dentaire de nos données. Les appellations CS8100 3D et CS9300 correspondent aux noms commerciaux de différents scanners Carestream. Le CS9300 est un scanner permettant l'acquisition de grands champs grâce à un grand capteur, mais d'une génération plus ancienne que le CD8100 3D, qui lui ne permet d'acquérir que des champs moyens (de taille maximale d'environ 10 cm de côté).

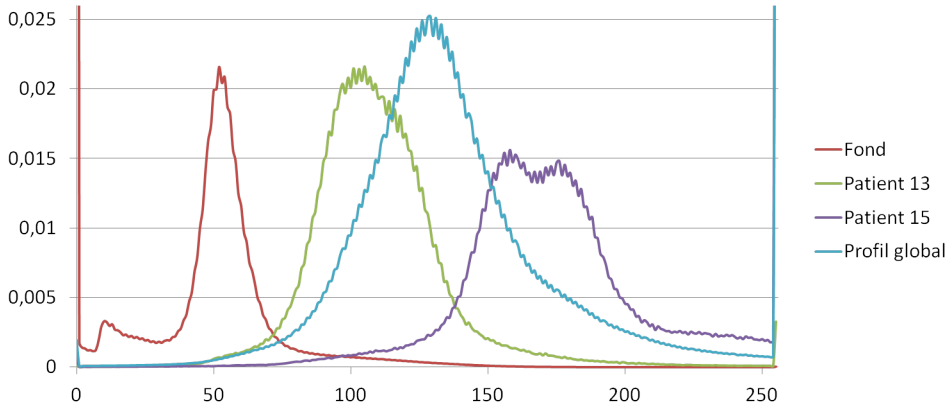


FIGURE 3.5 – Densités de probabilités des intensités dentaires pour nos données face à l'intensité du fond. Le profil global est représenté par la courbe cyan, et les cas extrêmes des patients 13 (courbe verte) et 15 (courbe violette) sont superposés.

3.2.1 Pré-traitement

L'utilisateur doit, pour chaque dent qu'il souhaite segmenter, donner son numéro dentaire ainsi qu'un ensemble de points. Il indique un point "couronne" qui se situe à la surface au centre de l'arête dentaire ou de la face occlusale (approximativement au centre des sillons ou entre les cuspides) ainsi qu'un point à chaque apex (extrémité des racines). Ces points servent à calculer l'axe principal de la dent (entre le point couronne et le barycentre des points radiculaires). Le volume est alors transformé pour aligner l'axe dentaire avec la direction verticale, avant d'extraire une zone d'intérêt autour de la dent, de manière à réduire la taille de l'image et à éliminer un maximum d'information parasite pour le recalage (figure 3.8). La taille de la zone d'extraction est fixée à l'avance en fonction du type de dent indiqué. Elle est choisie pour être l'espace minimal pouvant contenir toute la variabilité de taille

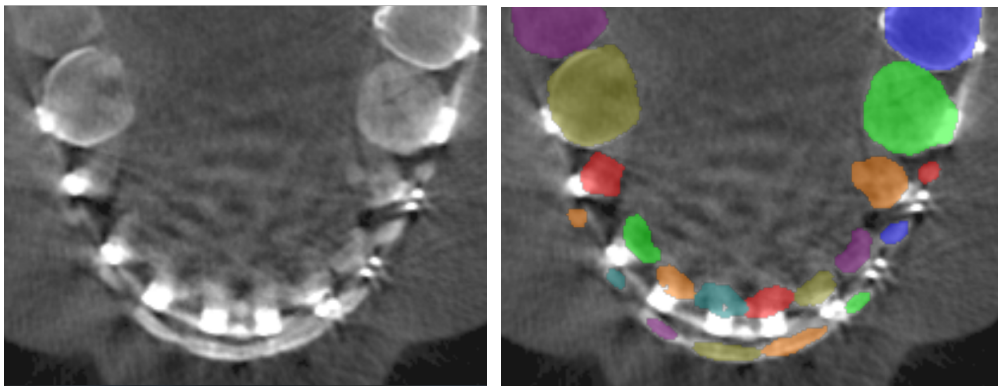


FIGURE 3.6 – Cas du patient 13 : apparence des dents et segmentation manuelle de référence. Nous soulignons les dégradations engendrées par la présence des bagues dentaires métalliques.

FIGURE 3.7 – Cas du patient 15 : l'augmentation de l'intensité moyenne dentaire s'explique par une forte proportion d'éléments métalliques dans les dents.

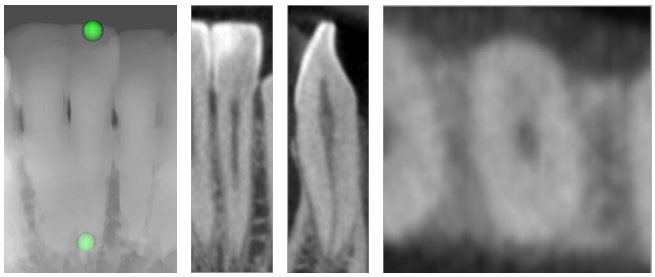
du type dentaire concerné, d'après les mesures anatomiques [11] :

- incisives maxillaires : 10mm × 10mm × (z+4)mm,
- incisives mandibulaires : 8mm × 8mm × (z+4)mm,
- canines : 10mm × 10mm × (z+4)mm,
- prémolaires : 12mm × 12mm × (z+4)mm,
- molaires maxillaires : 18mm × 18mm × (z+4)mm,
- molaires mandibulaires : 15mm × 15mm × (z+4)mm,

où z représente la longueur du segment entre le point couronne et le barycentre des points d'apex.

3.2.2 Recalage du modèle d'apparence moyenne

Nous effectuons ensuite à un recalage du modèle d'apparence moyenne (figure 3.3 (e)) sur la zone extraite. Pour initialiser le processus, le modèle est positionné de manière à faire correspondre son axe central avec l'axe vertical ainsi que son point couronne avec celui de la dent. Lorsqu'il y a plusieurs points d'apex, la direction principale dans le plan axial du nuage de points est calculée pour pouvoir aligner celle du modèle dessus par rotation autour de l'axe vertical.
L'optimisation de la pose est ensuite effectuée de manière rigide (translation et rotation uniquement)



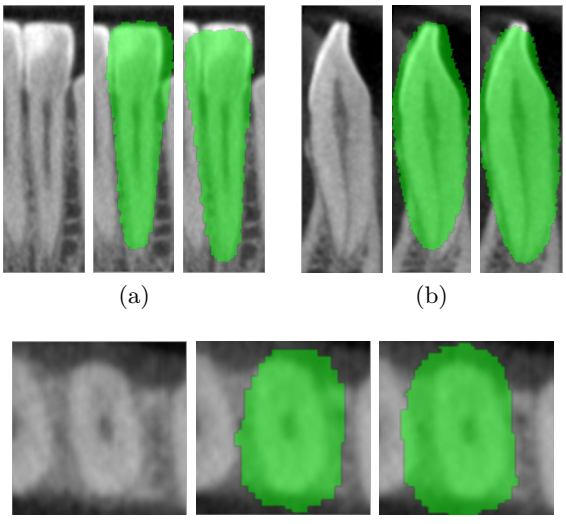
(a) (b) (c) (d)

FIGURE 3.8 – Pré-traitement. (a) Points sélectionnés par l'utilisateur (zoom sur rendu 3D) qui indique également le numéro dentaire (41 ici), pour procéder à l'extraction de la zone d'intérêt (vue frontale (b), sagittale (c), et axiale (d)).

par une descente de gradient. Le critère utilisé est la corrélation de Pearson :

$$Corr(\mathrm{Im}, \mathcal{T}(\bar{\mathbf{A}})) = \frac{\sum_{x\in\Omega}(\mathrm{Im}(x)\cdot\mathcal{T}(\bar{\mathbf{A}})(x))}{\sqrt{\sum_{x\in\Omega}\mathrm{Im}(x)^2\sum_{x\in\Omega}\mathcal{T}(\bar{\mathbf{A}})(x)^2}},$$

où Im et $\bar{\mathbf{A}}$ sont l'image et la dent d'apparence moyenne respectivement et $\mathcal{T}$ est la transformation de recalage (Ω est le support de ces images). Le choix de cette mesure est dû au fait que nous avons constaté qu'elle produisait des résultats plus pertinents que l'information mutuelle (utilisée dans la création des modèles) sur nos images, avec notamment une meilleure superposition au niveau des couronnes et un alignement de l'axe d'inertie plus proche. Afin d'atténuer l'impact des éléments métalliques sur la corrélation, les voxels ayant une intensité correspondant à du métal (valeur supérieure à un seuil empirique de 5000 dans notre cas, seuil maximal constaté pour des structures anatomiques dans notre base) ne sont pas considérés dans le calcul. Un exemple est donné en figure 3.9.



(a) (b)

(c)

FIGURE 3.9 – Recalage du modèle d'apparence moyenne. Pour chaque vue (a,b,c) sont présentés l'image, l'initialisation puis le recalage final. Notons que parfois, comme présenté ici, l'optimisation ne semble pas procurer une amélioration notable, voir même dégrade le résultat. Ici cela provient du fait que la dent d'apparence moyenne est constituée de fortes intensités que la corrélation pénalise lors du débordement sur le fond (voir vue frontale (a)).

3.2.3 Génération d'informations a priori

Maintenant que nous connaissons la correspondance entre le modèle et l'image, nous procédons à la construction de contraintes qui serviront de base à la segmentation. Notamment, l'espace des formes existantes permet de constituer une zone spatiale probable pour la dent. En nous inspirant de [69], nous générons des échantillons déformés issus de notre modèle (équation 3.1). Plus précisément, nous utilisons les formes binaires de ces échantillons :

$$S_i^{\pm} = \mathcal{H}(-(\bar{\mathbf{M}} \pm \mathrm{C}\sqrt{\lambda_i}\mathbf{V}_i)) \tag{3.2}$$

où $\bar{\mathbf{M}}$ est la forme moyenne, $\mathbf{V}_i$ le i^{e} mode de déformation associé à sa valeur propre λ_i et $\mathcal{H}$ la fonction de Heaviside, conventionnellement définie comme $\mathcal{H}(x) = 1$ si $x \geqslant 0$, $\mathcal{H}(x) = 0$ si $x < 0$ (ce qui explique l'inversion de signe de la carte de distance, pour avoir la valeur 1 à l'intérieur de la forme). La constante C (strictement positive) contrôle l'amplitude de la déformation, et donc l'étendue de l'espace des formes. En comparaison avec l'équation 3.1, notez qu'ici $w_i = \mathrm{C}\sqrt{\lambda_i}$. Pour chaque mode sont générés deux échantillons (selon le sens de parcours du mode de déformation), ce qui nous donne dans notre cas 10 échantillons ($S_1^+, S_2^+, S_3^+, S_4^+, S_5^+, S_1^-, S_2^-, S_3^-, S_4^-, S_5^-$). Ces derniers sont comparés à la forme moyenne binarisée afin d'établir un masque de l'espace où les déformations sont susceptibles de positionner la dent :

$$\mathrm{Masque}(x) = \begin{cases} 1 & \text{si} \quad \exists\, i : S_i^{\pm}(x) \neq \mathcal{H}(-\bar{\mathbf{M}}(x)), \\ 0 & \text{sinon.} \end{cases} \tag{3.3}$$

La valeur 1 sera donnée à tout voxel du masque s'il existe au moins un échantillon pour lequel le voxel de même position a une valeur différente du voxel de la forme moyenne binaire. Le masque est utilisé pour générer la première information a priori, une carte de marqueurs :

$$\mathrm{Marqueurs}(x) = \begin{cases} -1 & \text{si} \quad \mathrm{Masque}(x) = 1, \\ 0 & \text{si} \quad \mathrm{Masque}(x) = 0 \text{ et } \mathcal{H}(-\bar{\mathbf{M}}(x)) = 0, \\ 1 & \text{si} \quad \mathrm{Masque}(x) = 0 \text{ et } \mathcal{H}(-\bar{\mathbf{M}}(x)) = 1. \end{cases} \tag{3.4}$$

Cette carte permet de forcer la segmentation à respecter les zones associées aux étiquettes (0 pour le fond, 1 pour la dent) et à optimiser la zone inconnue (-1 ici). Enfin, nous nous servons des points donnés par l'utilisateur pour ajouter des graines supplémentaires. À partir du point couronne (sur la surface de la dent ou légèrement en dehors selon la précision du praticien), nous sélectionnons un point que nous savons inclus dans la couronne (en pratique, l'axe de la dent étant aligné avec l'axe vertical, il suffit de prendre un point de coordonnée axiale modifiée d'une valeur empirique selon que nous traitons une dent maxillaire ou mandibulaire) afin d'établir le plus court chemin entre ce point et les points radiculaires. Le critère de distance repose sur l'intensité : plus l'intensité du voxel est grande, plus le chemin tend à l'emprunter. Cela permet de s'assurer que le chemin reste dans la dent pour atteindre les apex. Cependant notons que lorsqu'ils sont visibles, les canaux radiculaires sont moins intenses que la dentine, or ils fournissent un chemin central idéal pour accéder aux apex. Pour cette raison, avant d'effectuer le calcul du chemin, nous procédons à une modification de l'intensité de ces canaux. L'intensité est seuillée par la méthode d'Otsu [101] pour une division en deux classes ainsi que pour une division en trois classes, donnant trois seuils, t_2, $\mathrm{t}_3^{inférieur}$ et $\mathrm{t}_3^{supérieur}$, où $\mathrm{t}_3^{inférieur} \leqslant \mathrm{t}_2 \leqslant \mathrm{t}_3^{supérieur}$. Habituellement le seuillage à t_2 permet d'obtenir une dent bien formée mais dont les canaux sont remplis (inclus dans le seuillage) alors que le seuil $\mathrm{t}_3^{supérieur}$ conduit à une forme de dent dégradée et non fermée. Nous échantillonnons donc des valeurs de seuils dans une gamme $[\mathrm{t}_2, \mathrm{t}_3^{supérieur}]$ et le seuil qui permet de segmenter le plus de canaux en termes de volume est sélectionné (les canaux sont obtenus comme les trous des composantes connexes de l'image seuillée). Nous attribuons alors une valeur maximale aux voxels des canaux ainsi identifiés pour le calcul du chemin. Cela correspond à l'algorithme suivant :

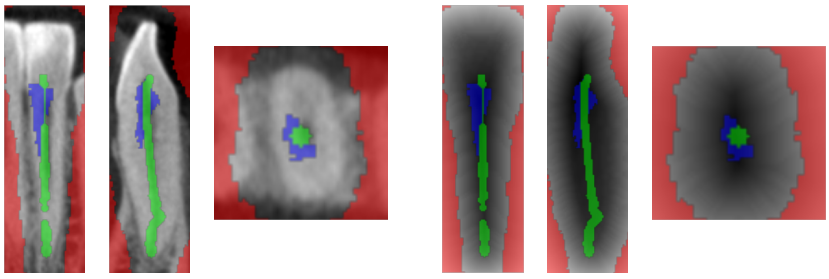
```
[t2] = Otsu_2Classes(Im)
[t3_inf,t3_sup] = Otsu_3Classes(Im)
// t3_inf < t2 < t3_sup
MeilleurSeuil = 0
```

```
MeilleurVolume = 0

Pour i de 0 à 5:
    Seuil = t2 + ( (t3_sup - t2) * (i / 5) )
    Im_Seuillée = Seuillage(Im, Seuil)
    Im_Remplie = ComblementTopologique(ImSeuillée)
    Im_Canaux = Im_Remplie - Im_Seuillée
    Si Volume(Im_Canaux)>MeilleurVolume alors:
        MeilleurVolume = Volume(Im_Canaux)
        MeilleurSeuil = Seuil

ImSeuilléeFinale = Seuillage(Im, MeilleurSeuil)
Pour chaque voxel x de Im:
    Si Im_SeuilléeFinale(x)!=0 alors:
        ImChemin(x) = 65535 // ou 255 si image 8 bits

Chemin = PlusCourtChemin(ImChemin)
```

Lorsque le chemin est construit, nous utilisons une dilatation morphologique avec un élément sphérique de 0,75mm de rayon afin de le rendre plus épais avant de l'ajouter aux marqueurs en tant que dent. Cet élément est illustré dans la figure 3.10 (a).
Le masque est aussi appliqué à la forme moyenne pour former une carte de distance "creuse" ϕ :

$$\phi(x) = \begin{cases} \bar{\mathbf{M}}(x) & \text{si} \quad \text{Masque}(x) = 1, \\ +\infty & \text{si} \quad \text{Masque}(x) = 0 \text{ et } \mathcal{H}(-\bar{\mathbf{M}}(x)) = 0, \\ -\infty & \text{si} \quad \text{Masque}(x) = 0 \text{ et } \mathcal{H}(-\bar{\mathbf{M}}(x)) = 1. \end{cases}$$

Dans l'espace inconnu de segmentation est conservée la valeur de distance à la forme moyenne, en revanche, lorsque nous sommes dans les marqueurs, la valeur de distance est infinie, ce qui agit comme une contrainte dure lors de la segmentation pour maintenir la cohérence des marqueurs (le signe dépend du fait d'être à l'intérieur ($-\infty$) ou à l'extérieur ($+\infty$) de la forme moyenne binaire). Enfin, nous appliquons à ϕ la transformation $\mathcal{T}$ issue du recalage pour la positionner dans l'espace des nouvelles données. Cette carte constitue la seconde information a priori utilisée dans la segmentation (figure 3.10 (b)). Nous soulignons ici que, quelle que soit la dent traitée, nous utilisons le modèle moyen pour générer nos contraintes, ce qui constitue une limitation car la moyenne n'est pas toujours représentative des contours réels.



(a) (b)

FIGURE 3.10 – Illustration des cartes de contraintes pour la segmentation. (a) Carte des étiquettes. La couleur permet de distinguer la provenance de la contrainte : rouge pour la zone où la déformation ne peut pas entraîner la dent (marqueur du fond), bleue pour la zone dentaire qui existe quelle que soit la déformation (marqueur de la dent) et verte pour les graines sur le plus court chemin entre les points donnés par l'utilisateur (marqueur de la dent). (b) Carte de distance ϕ. Les couleurs marquent une distance infinie.

3.3 Segmentation contrainte de l'image

Lorsque les informations a priori ont été établies, la seconde partie du processus consiste en une segmentation contrainte, formulée comme la minimisation d'une énergie, par une approche par coupure de graphe (voir figure 3.1).

3.3.1 Formulation du problème de segmentation d'une dent

Les premiers essais (voir annexe A) ont démontré qu'il était difficile de se reposer uniquement sur l'information d'intensité des images pour établir une segmentation satisfaisante. L'ajout d'informations issues des modèles précédemment établis permet de renforcer la spécificité de notre processus. Nous avons choisi de formuler notre problème sous la forme d'une minimisation énergétique car cette formulation permet naturellement l'incorporation de contraintes et de connaissances a priori. Segmenter revient alors à trouver les paramètres de l'équation qui minimisent l'énergie. Dans notre cas, cette formulation est représentée sur un graphe. Soit le graphe non orienté $G = \{\mathcal{V}, \mathcal{E}\}$ avec $\mathcal{V}$ l'ensemble des sommets et $\mathcal{E}$ l'ensemble des arêtes, $\mathcal{L}$ l'ensemble de variables aléatoires associé à $\mathcal{V}$, à valeurs dans $\Lambda = \{0, 1\}$ (ensemble des étiquettes, ici le fond et l'objet, notées de manière générale λ) et $\mathcal{C}$ un sous-ensemble de $\mathcal{V}$. On dit que $\mathcal{C}$ est une clique si et seulement si pour n'importe quelle paire de sommets de $\mathcal{C}$, les sommets sont adjacents entre eux (c'est-à-dire liés par une arête). Grâce à la théorie des champs aléatoires, il est possible de résoudre l'équation énergétique comme la somme des potentiels sur les cliques. Une réalisation des variables aléatoires constitue une partition entre les étiquettes, donc une segmentation. Pour la suite, par souci de simplicité, nous utiliserons de manière interchangeable la notion de voxels et de sommets s'agissant de la représentation graphique, notés v_i. Comment traduire notre connaissance sous forme énergétique ?
En premier lieu, nous savons que l'intensité de la dent est différente de celle de la majorité de son environnement. Nous intégrons donc cette information dans le terme de vraisemblance classique suivant :

$$R(\lambda) = \begin{cases} -\ln \mathrm{P}(\mathrm{Im}(v_i) \mid \mathrm{O}) & \text{si} \quad \lambda_{v_i} = 1 \quad \text{(objet)}, \\ -\ln \mathrm{P}(\mathrm{Im}(v_i) \mid \mathrm{F}) & \text{si} \quad \lambda_{v_i} = 0 \quad \text{(fond)}, \end{cases}$$

où $\mathrm{P}(\mathrm{Im}(v_i) \mid \mathrm{O})$ est la probabilité du niveau d'intensité au sommet v_i issue de la densité de probabilité de l'objet (ou du fond pour $\mathrm{P}(\mathrm{Im}(v_i) \mid \mathrm{F})$), comme présenté en figure 3.4, et λ_{v_i} l'étiquette du sommet v_i. Plus la probabilité est élevée, plus le terme est faible (figure 3.11 (a)). Cela favorise donc l'étiquette de la structure qui correspond le mieux en apparence au sommet. C'est donc un terme de vraisemblance.
La seconde information dont nous disposons est l'emplacement approximatif de la dent grâce au recalage. Il semble logique que les sommets proches ou inclus dans cette forme appartiennent à la dent [32]. Nous utilisons donc ϕ (voir figure 3.10 (a)) pour estimer la distance des sommets :

$$D(\lambda) = \begin{cases} \frac{1}{1+\exp(\phi(v_i))} & \text{si} \quad \lambda_{v_i} = 0 \quad \text{(fond)}, \\ 1 - \frac{1}{1+\exp(\phi(v_i))} & \text{si} \quad \lambda_{v_i} = 1 \quad \text{(objet)}. \end{cases}$$

Si la distance du sommet v_i à la forme est importante, alors la valeur de la fraction tendra vers 0. À l'inverse si la distance est faible voir négative (à l'intérieur de la forme) alors le rapport tendra vers 1 (figure 3.11 (b,c)). Ainsi, plus le sommet considéré sera éloigné de la forme, plus l'étiquette du fond sera favorisée face à l'étiquette de l'objet, et inversement, constituant un terme d'a priori de forme.
Cependant, avec ces termes, d'une part aucune dépendance spatiale n'est incluse, ce qui conduit à l'apparition de "bruit" dans la segmentation avec des zones hétérogènes, et d'autre part il ne nous est pas possible de différencier l'os cortical de la dent, puisqu'il possède une intensité similaire tout en étant proche, adjacent même, de cette dernière. Nous savons que bien que l'apparence soit semblable, il existe parfois une délimitation entre ces deux structures. Nous souhaitons donc que la partition s'effectue au niveau des contours de l'image, qui représentent les transitions entre structures. Cela nous conduit à utiliser le terme déjà introduit dans [22] :

$$B(v_i, v_j) = \exp\left(-\frac{(\mathrm{Im}(v_i) - \mathrm{Im}(v_j))^2}{2\sigma^2}\right)\left(\frac{1}{\mathrm{dist}(v_i, v_j)}\right),$$

où $\mathrm{Im}(v_i)$ et $\mathrm{Im}(v_j)$ sont les valeurs d'intensités aux sommets voisins v_i et v_j (en 26-connexité), σ un paramètre de tolérance et $\mathrm{dist}(v_i, v_j)$ la distance euclidienne entre v_i et v_j. Plus l'intensité entre les sommets est similaire, plus l'exponentielle tend vers 1. Inversement, si les sommets sont très dissemblables, alors le terme tend vers 0 (figure 3.11 (d,e)). La seconde fraction sert de normalisation par rapport à la distance entre les éléments considérés. Ce terme permet donc de favoriser la séparation au niveau des gradients importants (donc hypothétiquement au niveau des contours des structures de l'image) et de renforcer la cohérence de la séparation, agissant donc comme une régularisation. Le paramètre σ contrôle la dynamique de la fonction et est choisi en fonction de la gamme d'intensité de l'image et de l'écart d'intensité toléré.
Nous avons cependant vu que cette délimitation d'intensité n'était pas toujours existante dans les images. Dans ce cas, nous nous reposons sur la forme recalée pour contraindre la séparation :

$$T(v_i, v_j) = |\phi\left(\frac{v_i + v_j}{2}\right)|,$$

où ϕ est la carte de distance déjà utilisée dans $D(l_i)$ (terme inspiré par [54]). Ici, il s'agit de prendre la valeur absolue du point à mi-chemin entre v_i et v_j. L'idée est de favoriser la séparation au niveau du contour de la forme moyenne. Ce terme sera faible si v_i et v_j se trouvent de part et d'autre de ce contour et élevé si les sommets sont éloignés de la frontière.
Enfin pour minimiser la longueur de la séparation, nous utilisons une variation du delta de Kronecker ($\delta_{\lambda_{v_i} \neq \lambda_{v_j}} = 1 - \delta_{\lambda_{v_i}\lambda_{v_j}}$) dans la régularisation :

$$\delta_{\lambda_{v_i} \neq \lambda_{v_j}} = \begin{cases} 1 & \text{si} \quad \lambda_{v_i} \neq \lambda_{v_j}, \\ 0 & \text{si} \quad \lambda_{v_i} = \lambda_{v_j}. \end{cases}$$

Voici donc la forme de l'énergie que nous cherchons à minimiser :

$$E(\mathcal{L}) = \sum_{v_i \in \mathcal{V}} (\alpha R(\lambda_{v_i}) + \beta D(\lambda_{v_i})) + \sum_{(v_i, v_j) \in \mathcal{E}} (\gamma B(v_i, v_j) + \xi T(v_i, v_j)) \cdot \delta_{\lambda_{v_i} \neq \lambda_{v_j}}, \tag{3.5}$$

où les paramètres α, β, γ et ξ sont les paramètres de pondération positifs de chaque terme, dont la valeur est fixée manuellement et dépend de l'application (voir section 3.4). Le problème de segmentation est alors de trouver la réalisation $\mathcal{L}^*$ dont les λ_{v_i} minimisent E soit :

$$\mathcal{L}^* = \underset{\mathcal{L}=\{\lambda_{v_i} \in \Lambda,\ v_i \in \mathcal{V}\}}{\arg\min} E(\mathcal{L}). \tag{3.6}$$

Pour qu'une énergie de la forme $E(\mathcal{L}) = \sum_{v_i \in \mathcal{V}} E(\lambda_{v_i}) + \sum_{(v_i, v_j) \in \mathcal{E}} E(\lambda_{v_i}, \lambda_{v_j})$ puisse être représentée sur un graphe, il faut qu'elle soit régulière (condition de sous-modularité, [81]), c'est-à-dire que son terme d'interaction binaire $E(\lambda_{v_i}, \lambda_{v_j})$ satisfasse :

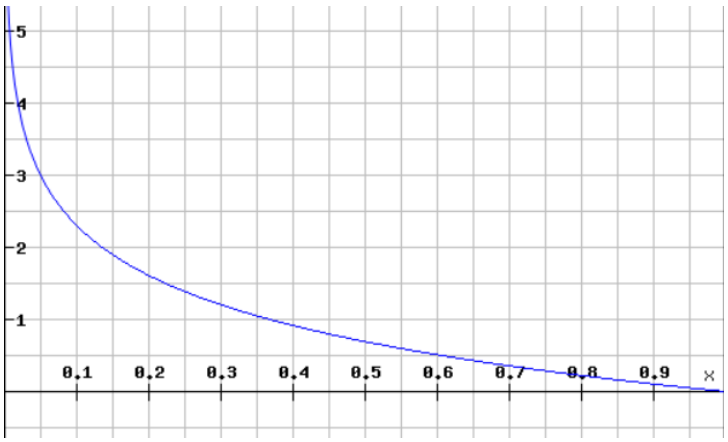
$$E(0,0) + E(1,1) \leqslant E(1,0) + E(0,1).$$

Dans notre cas, nous avons :

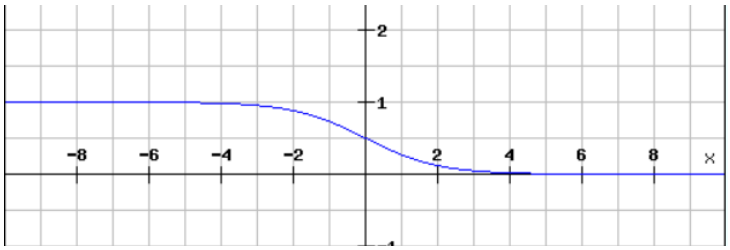
$$E(\lambda_{v_i}, \lambda_{v_j}) = (\gamma B(v_i, v_j) + \xi T(v_i, v_j)) \cdot \delta_{\lambda_{v_i} \neq \lambda_{v_j}}.$$

Lorsque $\lambda_{v_i} = \lambda_{v_j}$ alors $E(\lambda_{v_i}, \lambda_{v_j}) = 0$ car $\delta_{\lambda_{v_i} \neq \lambda_{v_j}} = 0$. Donc $E(0,0) + E(1,1) = 0$.
Lorsque $\lambda_{v_i} \neq \lambda_{v_j}$ alors $\delta_{\lambda_{v_i} \neq \lambda_{v_j}} = 1$. Le terme $B(v_i, v_j)$ est un produit d'une fonction exponentielle et d'un rapport strictement positif (car $\mathrm{dist}(v_i, v_j) > 0$) et est donc strictement positif. Les pondérations étant positives, $\gamma B(v_i, v_j) \geqslant 0$. Le terme $T(v_i, v_j)$ est positif par la fonction de valeur absolue. La somme $\gamma B(v_i, v_j) + \xi T(v_i, v_j)$ est donc strictement positive, impliquant que $E(\lambda_{v_i}, \lambda_{v_j}) > 0$ lorsque $\lambda_{v_i} \neq \lambda_{v_j}$, d'où $E(1,0) + E(0,1) > 0$, et donc

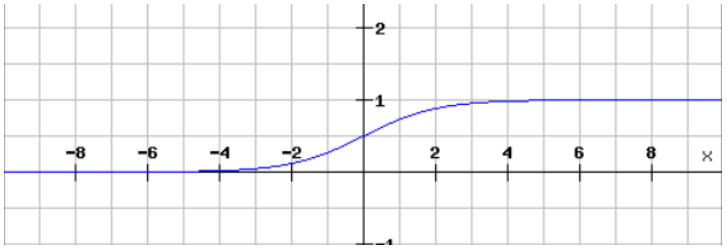
$$E(0,0) + E(1,1) < E(1,0) + E(0,1).$$



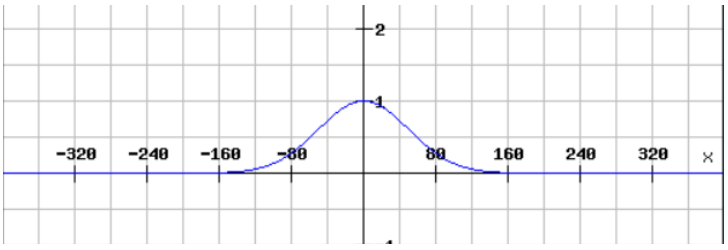
(a)



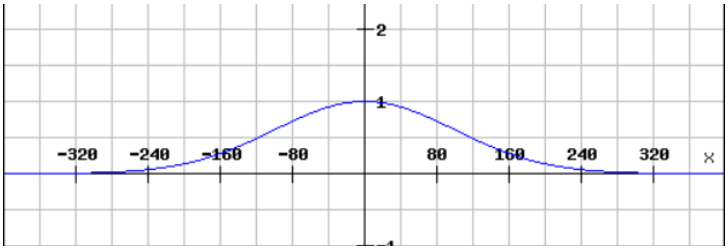
(b)



(c)



(d)



(e)

FIGURE 3.11 – Représentation graphique du comportement des termes de l'équation 3.5. (a) $R(\lambda_{v_i})$ en fonction de la probabilité. (b,c) $D(0)$ et $D(1)$ en fonction de la distance. (d,e) B pour un σ de 50 et de 100 respectivement, en fonction de la différence d'intensité.

3.3.2 Méthode de résolution

Si de manière générale, l'optimisation énergétique discrète est un problème NP-difficile [87], de multiples solutions ont été proposées afin d'optimiser ce type de problème. Les premières propositions reposant sur des schémas itératifs comme l'algorithme de recuit simulé [61], ou l'algorithme des modes conditionnels itérés (*ICM* [8]) ont pour défaut d'être coûteuses en temps de calcul et de ne fournir qu'une approximation de la solution (globale pour le recuit, locale pour l'ICM). En 1989, Greig et al. [68] montrent qu'il est possible de trouver la solution exacte au maximum a posteriori pour un champ aléatoire de Markov (*MAP-MRF*) de premier ordre (c'est-à-dire avec des potentiels fonction de deux sommets au maximum) avec variables binaires, or cette probabilité est liée à l'énergie (et donc à notre problème) par :

$$P(\mathcal{L}) = \frac{1}{Z}\exp(-E(\mathcal{L})),$$

où Z est la fonction de partition. Partant du graphe des voxels augmenté de deux sommets auxiliaires appelés "terminaux" (et différenciés en "source" et "puits") qui relient tous les autres sommets, une "coupe" est l'ensemble des arêtes qui partitionne en deux le graphe et sépare les deux terminaux. Plus formellement, une coupe comprend pour chaque v_i exactement une arête de terminal (celle auquelle l'étiquette ne correspond pas) ainsi que toutes les arêtes entre sommets pour lesquelles $\lambda_{v_i} \neq \lambda_{v_j}$. Le coût d'une telle séparation est appelé la capacité de la coupe et s'écrit comme la somme des potentiels (énergies) des arêtes composant cette coupe (voir ci-après). La coupe de plus faible capacité sur le graphe est dite coupe minimum (ou minimale), et est solution du problème de la minimisation énergétique. Une illustration du principe est présentée en figure 3.12 sur une image bi-dimensionnelle. Dans notre cas, le graphe prend la forme d'un pavé droit où chaque sommet est lié à ses 26 voisins ainsi qu'aux deux terminaux. Notre énergie est donc décomposée sur l'ensemble des cliques afin de correspondre à la formulation de coupure de graphe (*graph-cut*), et la valeur des arêtes de notre graphe est explicitée dans la table 3.3. Soulignons que dans ce cadre, les termes unaires de l'énergie (R et D) sont simulés par les termes binaires entre les v_i et les sommets terminaux (où 0 et 1 correspondent au puits et à la source respectivement), ce qui donne :

$$\begin{aligned} Capacité = \sum_{(v_i,Source)\in Coupe} (\alpha R(0) + \beta D(0)) + \sum_{(v_i,Puits)\in Coupe} (\alpha R(1) + \beta D(1)) + \\ \sum_{(v_i,v_j)\in Coupe} (\gamma B(v_i,v_j) + \xi T(v_i,v_j)). \quad (3.7) \end{aligned}$$

Notez que puisque la coupe contient uniquement les arêtes entre les nœuds d'étiquettes différentes, $\delta_{\lambda_{v_i}\neq\lambda_{v_j}} = 1$. Nous pondérons les arêtes entre les sommets et le puits par le coût d'étiquetage à l'objet (puisqu'en coupant ce lien, le sommet reste attaché à la source, donc à l'objet), et les arêtes entre les sommets et la source par le coût d'étiquetage au fond (de manière similaire, le sommet restant attaché au puits). Le graphe ainsi construit, la capacité est strictement équivalente à l'équation 3.5, car chaque v_i est nécessairement séparé d'un seul nœud terminal, c'est-à-dire que pour tout v_i, on a $(v_i, Source) \in Coupe \Leftrightarrow (v_i, Puits) \notin Coupe$ et inversement (il y a donc exactement $|\mathcal{V}|$ arêtes reliant des sommets aux terminaux dans la coupe).

Ford et Fulkerson [51] ont montré que le problème dual de la coupe minimale était celui du flot maximal dans un réseau et ont décrit un algorithme de coupure de graphe en temps polynomial pour sa résolution. Dans notre cas, nous utilisons l'algorithme de Boykov et Kolmogorov [23] qui est une version optimisée pour les images de l'algorithme par augmentation de flot. Pour rappel, c'est un algorithme itératif qui utilise un graphe résiduel encodant l'état du flot et fonctionne en trois phases :

1. Croissance : deux arbres de recherche parcourant les nœuds non explorés sont construits en partant des terminaux en suivant les arêtes non saturées jusqu'à ce que les arbres se rencontrent et forment un chemin entre les terminaux.
2. Augmentation : le flot est augmenté de manière à saturer la capacité minimale rencontrée le long du chemin.

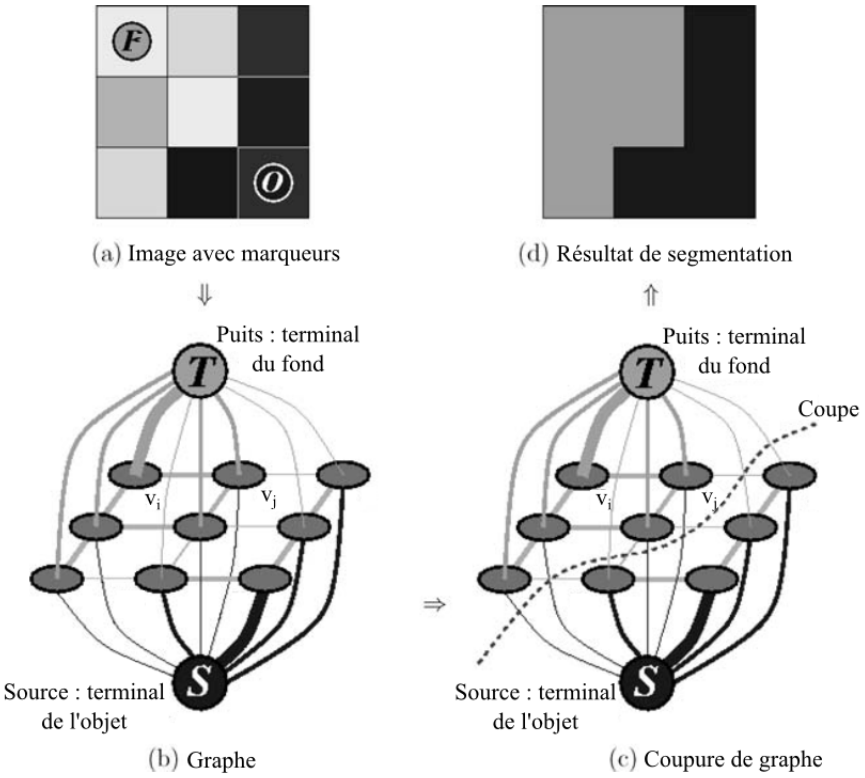


FIGURE 3.12 – Représentation du problème de coupure de graphe. Schéma traduit de [21].

3. Reconstitution : l'augmentation du flot peut conduire à déconnecter des branches dans chaque arbre. Chaque partie déconnectée est rattachée s'il existe un autre chemin non saturé avec l'arbre originel, sinon elles sont réinitialisées comme étant non explorées.

L'algorithme s'arrête lorsqu'il n'est plus possible de faire croître les arbres et qu'ils sont séparés par des arêtes saturées. Le flot maximal est atteint et la coupe minimale correspond à l'ensemble des arêtes saturées. Les détails d'implémentation peuvent être trouvés dans l'article sus-cité.

Arête	Poids	Situation
(v_i, v_j)	$\gamma B(v_i, v_j) + \xi T(v_i, v_j)$	$\forall (v_i, v_j) \in \mathcal{V}$
	$\alpha R(0) + \beta D(0)$	Marqueurs$(v_i) = -1$
$(v_i, Source)$	$+\infty$	Marqueurs$(v_i) = 1$
	0	Marqueurs$(v_i) = 0$
	$\alpha R(1) + \beta D(1)$	Marqueurs$(v_i) = -1$
$(v_i, Puits)$	$+\infty$	Marqueurs$(v_i) = 0$
	0	Marqueurs$(v_i) = 1$
(v_i, v_j)	$\gamma(\exp\left(-\frac{(\mathrm{Im}(v_i)-\mathrm{Im}(v_j))^2}{2\sigma^2}\right)\left(\frac{1}{\mathrm{dist}(v_i,v_j)}\right)) + \xi\lvert\phi\left(\frac{v_i+v_j}{2}\right)\rvert$	$\forall \{v_i, v_j\} \in \mathcal{V}$
	$\alpha(-\ln \mathrm{P}(\mathrm{Im}(v_i) \mid \mathrm{F})) + \beta(\frac{1}{1+\exp(\phi(v_i))})$	Marqueurs$(v_i) = -1$
$(v_i, Source)$	$+\infty$	Marqueurs$(v_i) = 1$
	0	Marqueurs$(v_i) = 0$
	$\alpha(-\ln \mathrm{P}(\mathrm{Im}(v_i) \mid \mathrm{O})) + \beta(1 - \frac{1}{1+\exp(\phi(v_i))})$	Marqueurs$(v_i) = -1$
$(v_i, Puits)$	$+\infty$	Marqueurs$(v_i) = 0$
	0	Marqueurs$(v_i) = 1$

TABLE 3.3 – Répartition des poids sur les arêtes du graphe utilisé dans notre formulation de coupure de graphe.

3.3.3 Post-traitement

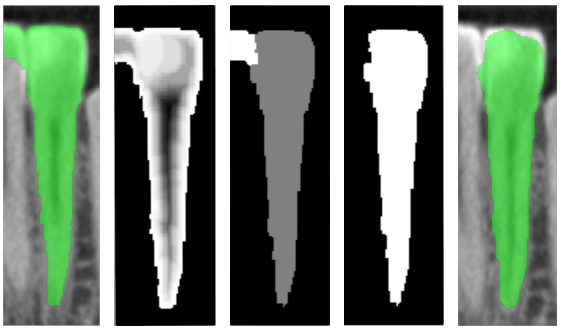
À la suite des tests préliminaires, il s'est avéré que le résultat de la coupure de graphe était sensible à la présence des couronnes adjacentes dans l'image, notamment lorsqu'elles apparaissaient jointes (c'est-à-dire sans gradient inter-couronnes visible). Un traitement postérieur est appliqué afin d'éliminer les éventuelles fuites de segmentation. Lorsque de telles erreurs surviennent, l'aspect général est celui de deux objets convexes accolés. Nous utilisons donc dans un premier temps une stratégie de séparation d'objets classique en morphologie mathématique :

1. Calcul de la carte de distance interne inversée de la segmentation. Cette carte peut être vue comme un relief, les contours de l'objet formant les crêtes et le centre un bassin versant.
2. Calcul de la ligne de partage des eaux, séparant tous les bassins versants (c'est-à-dire tous les minima locaux). Pour éviter une fragmentation trop importante due à des bassins peu significatifs, nous fusionnons tous les bassins ayant une profondeur inférieure à un seuil choisi empiriquement.
3. À ce stade, les parties de dents différentes doivent reposer dans des bassins différents. Nous sélectionnons donc le bassin central de l'image, qui correspond à la dent initialement ciblée.

```
CarteDistances = TransforméeDistanceEuclidienne(Image_Segmentation)
Pour chaque voxel x de CarteDistances:
    Si CarteDistances(x) > 0:
        CarteDistances(x) = |CarteDistances(x) - max(CarteDistances)|

Bassins = PartageDesEaux(CarteDistances, ProfondeurMinimale)
```

```
7
  DistanceMinimale = +∞
9 Pour chaque bassin b de Bassins:
      Si DistanceEuclidienne(Barycentre(b),CentreGeométriqueImage) < DistanceMinimale:
11        BassinSelectionné = b
          DistanceMinimale = DistanceEuclidienne(Barycentre(b),CentreGeométriqueImage)
13
  Pour chaque voxel x de Bassins:
15    Si Bassins(x) n'est pas dans BassinSelectionné:
          Bassins(x) = 0
17
  SegmentationFinale = OuvertureAdaptative(Bassins)
```

Pour finir, une régularisation est appliquée sous la forme de l'application d'une ouverture morphologique avec un élément structurant sphérique de taille décroissante selon la partie de dent considérée. Le volume est découpé en trois sections de hauteur entre les points donnés par l'utilisateur qui correspondent à la couronne, à la partie centrale et à la partie des apex. La partie coronale est traitée avec l'élément de plus grande taille tandis que la partie terminale est ouverte par le plus petit (le rayon des éléments structurants est respectivement 1mm, 0,6mm et 0,3mm). Un exemple est donné en figure 3.13.



(a) (b) (c) (d) (e)

FIGURE 3.13 – Illustration du processus de post-traitement de la segmentation. (a) État initial de la segmentation. (b) Carte de distance interne inversée. (c) Résultat de la ligne de partage des eaux sur (b). (d) Sélection du bassin central. (e) Résultat du traitement après application de la régularisation.

3.4 Expérimentation

3.4.1 Protocole

Expériences et jeux de données

Afin d'étudier la performance de notre méthode, nous avons procédé en trois étapes. En premier lieu nous présenterons comment nos résultats évoluent sous l'influence des différents paramètres qui composent notre approche, entre les données, l'interaction effectuée par l'utilisateur, l'étape de recalage et la segmentation. La seconde étape est une étude de la performance comparée à celle de quatre autres approches : un modèle déformable par ensemble de niveaux avec a priori de forme, un *powerwatershed*, une forêt aléatoire et enfin un réseau de neurones convolutifs. La base d'images pour ces deux parties est constituée de 32 échantillons (zones d'intérêt contenant une dent), qui représentent l'ensemble des types dentaires, et provenant de quatre patients différents : les volumes des patients 1 et 2 (champs moyens), de taille de voxels isotrope de 0,15mm, et ceux des patients 5 et 7 (grands champs), de taille de voxels isotrope de 0,3mm (les numéros de patient se réfèrent à la table 3.2). L'approche par réseau de neurones possède sa propre base d'images compte tenu de son architecture pour l'apprentissage faisant appel à des images 2D, qui sera expliquée dans le paragraphe dédié.

La dernière partie sera une analyse des résultats obtenus par l'algorithme en terme de qualité de segmentation et de ressources de calcul, sur une base d'images constituée de 64 échantillons, issues de deux dentures adultes complètes, celles des patients 1 et 2. Les zones d'intérêt sont constituées de 450 000 à 1 500 000 voxels selon le type de dent concerné. Cette base diffère principalement par le choix de se concentrer sur des volumes de résolution inférieure à 200 microns, pour étudier les performances en éliminant les volumes dont nous savons que la qualité d'images est insuffisante pour faire apparaître correctement les structures délimitant une dent (parodonte).

Critères d'évaluation

La mesure de référence prise pour l'ensemble de nos expériences est le coefficient de Dice, ou indice de recouvrement, qui mesure le volume partagé par deux segmentations au regard des volumes individuels. C'est une mesure couramment utilisée en imagerie médicale pour la validation de segmentations d'organes. Des mesures plus pertinentes existent pour des structures aux formes particulière (tubulaires par exemple) telles que l'erreur de distance moyenne, ou la distance de Hausdorff. Le coefficient de Dice est calculé par :

$$Dice = \frac{2|S^1_{ref} \cap S^1_{test}|}{|S^1_{ref}| + |S^1_{test}|},$$

où S^1_{ref} correspond aux voxels de la vérité terrain et S^1_{test} est son pendant dans la segmentation testée. Il peut également être obtenu à partir des valeurs de la matrice de confusion, qui définit quatre catégories :

— Vrais positifs (VP) : voxels étiquetés comme objet lors de la segmentation, et qui appartiennent bien à l'objet selon la référence.

— Faux positifs (FP) : voxels étiquetés comme objet alors qu'ils font partie de la classe fond en réalité.

— Faux négatifs (FN) : voxels étiquetés comme fond alors qu'ils appartiennent à l'objet d'après la référence.

— Vrais négatifs (VN) : voxels étiquetés comme fond et qui le sont effectivement (notez qu'ils sont ici peu pertinents car le fond peut être étendu artificiellement).

Le coefficient de Dice s'obtient alors par :

$$Dice = \frac{2VP}{2VP + FP + FN}.$$

Pour la dernière expérience, nous utilisons cette matrice de confusion pour calculer en plus les mesures suivantes : la sensibilité, la spécificité, et la précision. Ces mesures permettront d'analyser plus largement le comportement de l'algorithme et de comparer avec l'analyse multi-objets du chapitre 4. La sensibilité, aussi nommée "rappel" ou taux de vrais positifs, mesure la capacité du système à identifier et récupérer correctement les objets ciblés et est calculée par :

$$Sensibilité = \frac{VP}{VP + FN}.$$

La spécificité, ou taux de vrais négatifs, mesure la capacité du système à reconnaître le fond et se formule comme :

$$Spécificité = \frac{VN}{VN + FP}.$$

La précision (*accuracy*, à ne pas confondre avec la valeur prédictive positive, nommée *precision* en anglais) mesure la capacité globale du système à classer correctement quelle que soit la catégorie :

$$Précision = \frac{VP + VN}{VP + FP + VN + FN}.$$

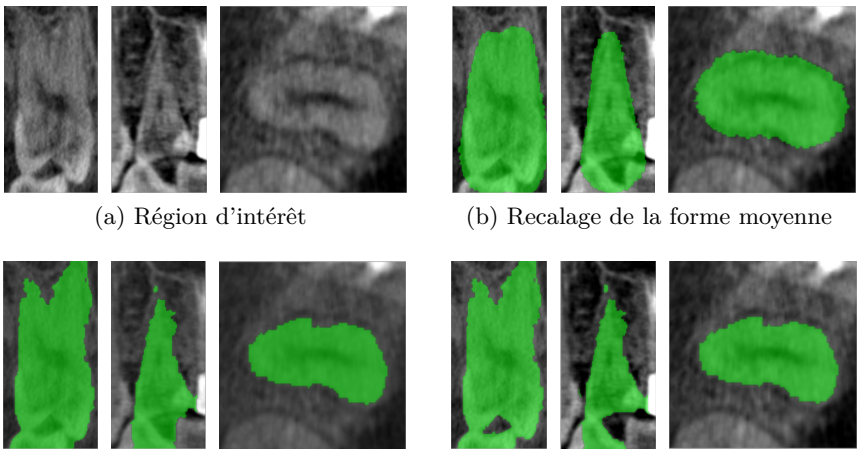
Ces deux dernières mesures utilisant les vrais négatifs, elles ne sont pertinentes que lorsque la taille de l'image est cohérente avec la taille de l'organe étudié, pour ne pas donner des résultats biaisés par un

nombre arbitrairement grand de voxels du fond. Dans notre cas, les zones d'intérêt étant similaires à des boîtes englobantes, la proportion du fond et de l'objet sont proches, et ces mesures nous semblant suffisamment pertinentes, elles seront utilisées pour étudier plus précisément le comportement de l'algorithme dans la section 3.4.4, ainsi que pour établir une comparaison avec l'extension multi-objets de la partie 4.1.3. Pour plus d'informations sur les mesures de comparaison en imagerie médicale, le lecteur peut se référer à la revue de Taha et Hanbury [116].

3.4.2 Influence des paramètres et des données

Type de données

Parmi nos données, deux facteurs vont avoir un impact important sur le comportement de notre méthode. En premier lieu, il faut distinguer les patients dont les images sont acquises en occlusion, c'est-à-dire mâchoire fermée et contractée, avec les dents supérieures et inférieures en contact, et les patients acquis avec un dispositif en plastique à mordre qui permet une acquisition avec les deux arches séparées. En occlusion, la limite entre les couronnes des dents opposées est difficilement visible, notamment aux plus basses résolutions. En figure 3.14 est présenté un cas d'occlusion sur une prémolaire maxillaire. Nous pouvons voir que, dès le recalage, le modèle a dérivé vers la dent inférieure, ce qui entraîne une segmentation de sa couronne. De plus, l'espace inter-dentaire est aussi compris dans la segmentation, puisque considéré à l'intérieur du modèle (il est éliminé si seules les contraintes d'intensité et de gradient sont utilisées). Le succès de la méthode dans un cas d'occlusion dépendra fortement du résultat du recalage. Si celui-ci dérive, la segmentation sera erronée, alors que s'il est réussi, les contraintes de formes permettront d'empêcher la segmentation de la dent opposée.
Le deuxième facteur impactant des données est la résolution des images. L'information contenue dans un même volume de l'espace physique sera bien plus faible dans les grands champs, et les structures bien moins définies (figure 3.15) que dans les moyens ou petits champs (le nombre total de voxels étant similaire). Notamment au niveau de la séparation avec l'os cortical, ces volumes en grand champ seront plus facilement sujets à des erreurs de segmentation selon la morphologie du patient, car la largeur moyenne du ligament parodontal (délimitation entre os et dent) est de 0,25mm, soit moins que la taille du voxel. Cela nous conduit à considérer les volumes ayant une résolution $> 0,2$mm comme étant peu adaptés à la segmentation dentaire. C'est néanmoins avec ce type de données que travaille la majorité des orthodontistes, ce qui constitue un problème qui reste ouvert.



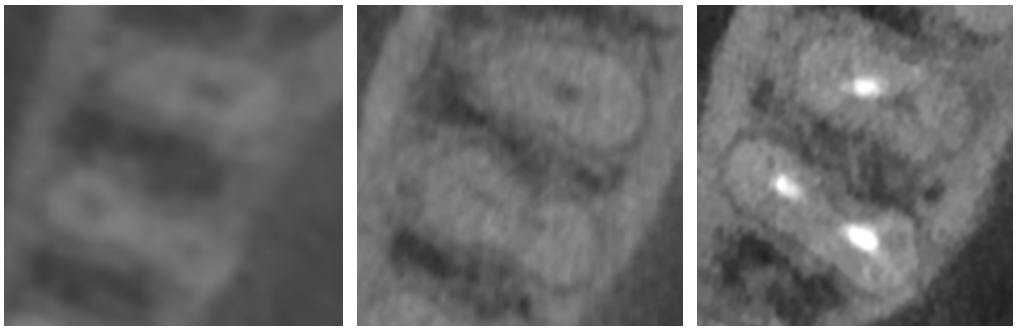
(a) Région d'intérêt

(b) Recalage de la forme moyenne

(c) Segmentation avec contraintes de forme ($\alpha = 1$, $\beta = 1$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$)

(d) Segmentation sans contraintes de forme ($\alpha = 1$, $\beta = 0$, $\gamma = 15$, $\sigma = 100$ et $\xi = 0$)

FIGURE 3.14 – Impact des données de patients en occlusion.



(a) Résolution de 0,3mm (b) Résolution de 0,15mm (c) Résolution de 0,075mm

FIGURE 3.15 – Aspect des différents types de données au niveau dentaire. Les grands champs de vue possèdent une résolution isotrope de 0,3mm alors que les petits champs permettent une résolution de 0,075mm.

Points donnés par l'utilisateur

La première étape est principalement conditionnée par les points donnés par l'utilisateur. Dans la figure 3.17, cinq situations sont illustrées. Dans le premier cas est représentée une sélection correcte de points telle que décrite dans la section 3.2.1 et le résultat des étapes suivantes (recalage et coupure de graphe). Les quatre autres cas présentent une erreur de sélection. En figure 3.17 (b) les points d'apex ont été sélectionnés trop bas en dehors des racines. La zone d'intérêt est plus grande que la normale, mais contient bien l'ensemble de la dent. Le recalage est correct et ne présente pas de différence avec le cas initial. La segmentation présente une différence au niveau des racines, car le marqueur de l'objet issu du calcul entre les points émerge alors de la dent par les racines. Dans le cas (c) où les points d'apex ont été positionnés trop hauts, nous voyons que la zone extraite est trop petite et la section terminale des racines est perdue. En dehors de cela, le recalage est correct et la segmentation semblable à celle obtenue à partir d'une sélection correcte pour la partie visible. Lorsque le point de la couronne est positionné trop haut (cas (d)), l'étendue spatiale est trop importante, mais cela n'affecte pas le recalage (notons qu'ici la mâchoire n'est pas en occlusion, ce qui n'entraîne pas la segmentation de la couronne opposée) et peu la segmentation. Enfin lorsque le point de la couronne est sélectionné dans la dent, nous retrouvons le problème de dent tronquée comme dans (c), mais cela n'affecte pas le recalage ni la segmentation des parties visibles. Pour résumer, la position du point de la couronne a peu d'influence pourvu qu'il ne soit pas inclus trop profondément dans la dent, ce qui est un comportement voulu, puisqu'il est parfois difficile de cibler correctement la surface notamment pour les molaires qui ont des formes coronales complexes. Une sélection approximative entre les cuspides est alors satisfaisante. Les positions aux apex doivent en revanche être plus précises sous peine d'observer des débordements dus à la construction des marqueurs de la dent. Une étude de l'impact sur le score de Dice de la segmentation finale est présentée dans la figure 3.16, où une erreur de 3 mm (cette valeur étant fixée empiriquement, correspondant à une limite visuelle au delà de laquelle le mauvais placement est sans équivoque) est introduite dans le placement d'un point selon les axes de la dent. Si la sélection de référence donne les meilleurs résultats avec un coefficient de Dice de $0,931$, nous notons que la dégradation entraînée par une erreur de placement est plutôt contenue, avec un coefficient de Dice de $0,872$ dans le pire des cas (ApexDist). Plus le placement est précis, meilleurs seront les résultats, mais cette sélection peut être faite de manière approximative tout en permettant d'obtenir des résultats qui restent cohérents en ciblant la dent.

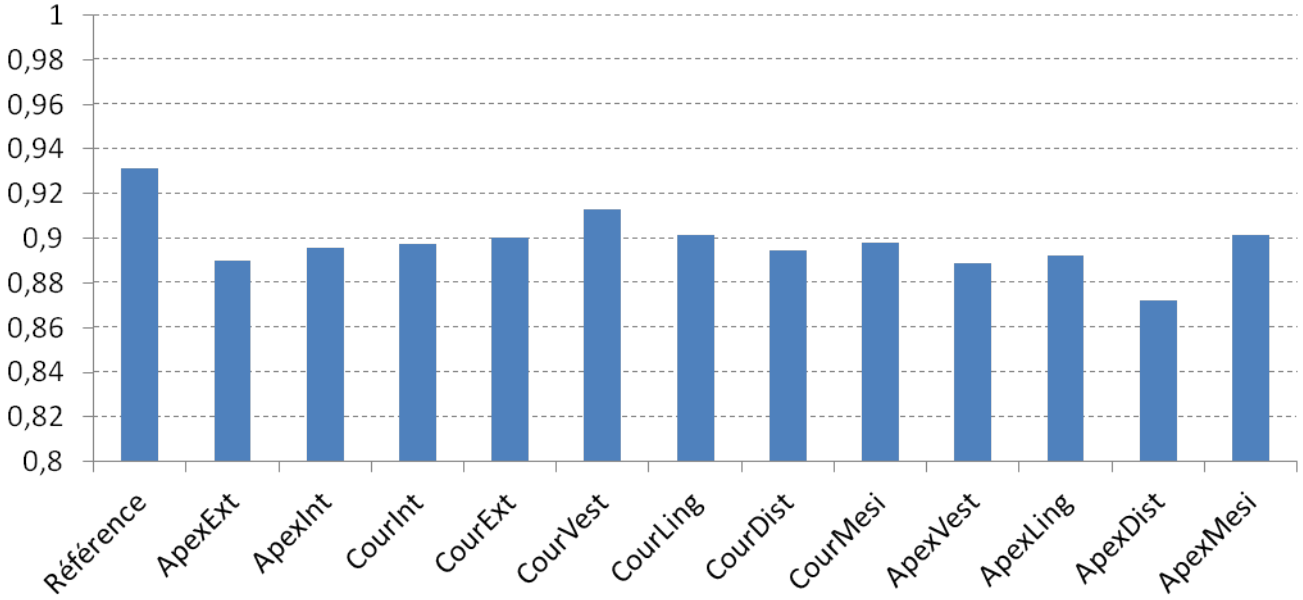


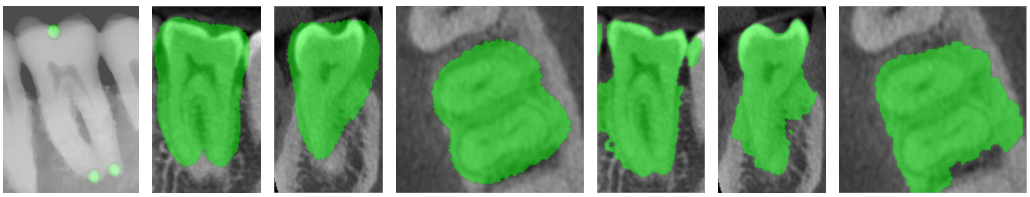
FIGURE 3.16 – Impact du placement des points donnés par l'utilisateur sur la segmentation finale (score de Dice). Le placement correct est noté "Référence". Les autres cas correspondent à une erreur de placement de 3 mm et sont : point coronal trop éloigné de la surface (CourExt) ou trop dans la couronne (CourInt), points d'apex trop éloigné (ApexExt) ou trop dans la racine (ApexInt), points coronal ou d'apex décalés dans le sens vestibulaire, lingual, distal et mésial respectivement (CourVest, CourLing, CourDist, CourMesi ApexVest, ApexLing, ApexDist et ApexMesi). Les paramètres du recalage sont : 100 itérations, $0,001$ de pas minimal, $0,001$ de tolérance au gradient. Les paramètres de la coupure de graphe sont : $\alpha = 1$, $\beta = 5$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$.

Recalage

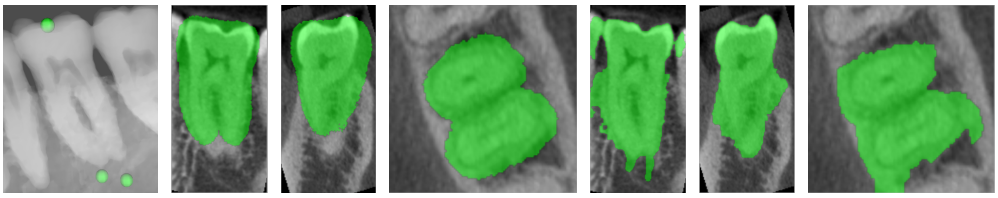
La seconde étape, le recalage, ne possède que peu de paramètres : le nombre d'itérations, la variation minimale du gradient lors de l'optimisation et le pas minimal de déplacement dans l'espace des formes. Grâce à l'initialisation donnée par les points de l'utilisateur, le nombre d'itérations nécessaire est systématiquement inférieur à une centaine, et la sensibilité au gradient et le pas minimal ne s'avèrent pas des facteurs critiques dans le positionnement du modèle d'apparence (voir figure 3.18). Il est aussi possible de procéder à un recalage multi-échelles, usuellement pour accélérer le processus dans ses premières itérations. Grâce à l'initialisation, cette approche n'est ici pas nécessaire. Cette étape est donc robuste, ce qui apparaît clairement dans la figure 3.19, où le coefficient de Dice final varie peu quelles que soient les valeurs prises par les paramètres.

Segmentation

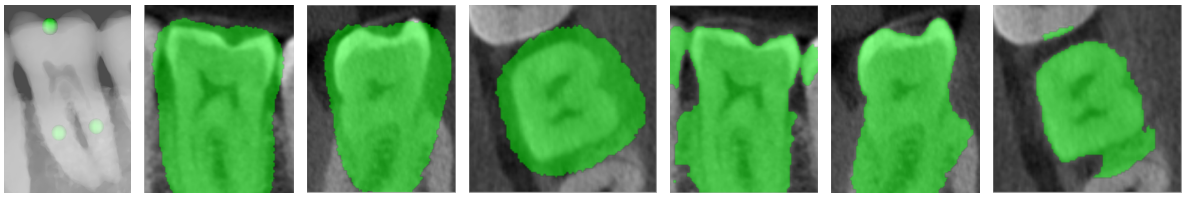
La partie de la méthode qui possède le plus de paramètres est la segmentation par coupure de graphe. Dans les figures 3.20 à 3.23 est illustré le comportement de chaque terme en fonction de son poids lorsque les autres poids ont une valeur fixe, trouvée expérimentalement comme donnant de bons résultats, c'est-à-dire $\alpha = 1$, $\beta = 5$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$. L'étude de la valeur de σ (figure 3.20) montre que plus sa valeur est importante, plus la coupe s'effectue sur les premiers contours autour des marqueurs. Il est important de noter que σ contrôle la tolérance accordée à l'écart des gradients (voir figure 3.11), donc ce comportement est attendu : si σ est faible, la gaussienne fonction de l'écart d'intensité est très étroite et donnera un terme qui ne pénalisera aucun écart de gradient et permettra une coupure aux gradients induits par le bruit et donc une segmentation bruitée. À l'opposé, si la gaussienne est très large, tous les gradients seront pénalisés identiquement, et la segmentation s'arrêtera aux premiers gradients autour des marqueurs. Il doit être adapté selon



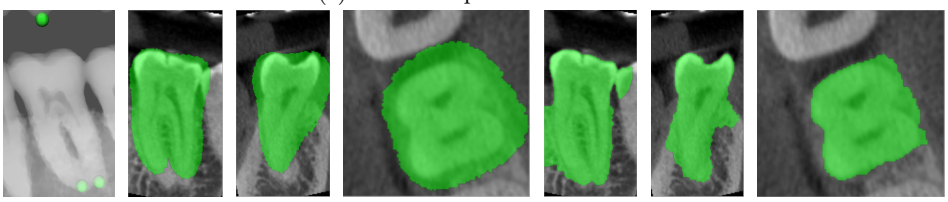
(a) Sélection correcte de points



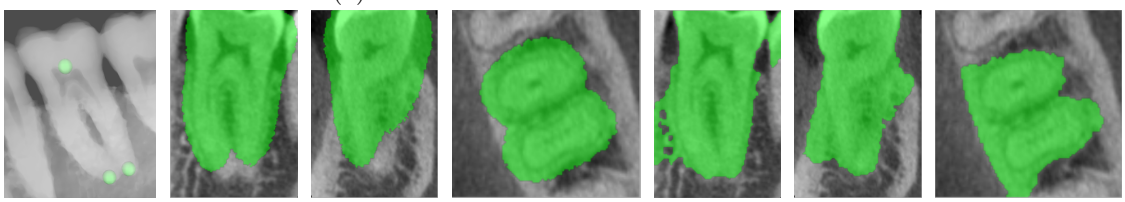
(b) Points d'apex en dehors des racines



(c) Points d'apex à mi-racine

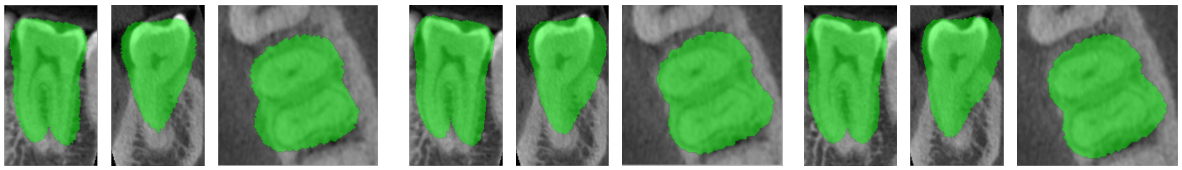


(d) Point de la couronne hors de la surface



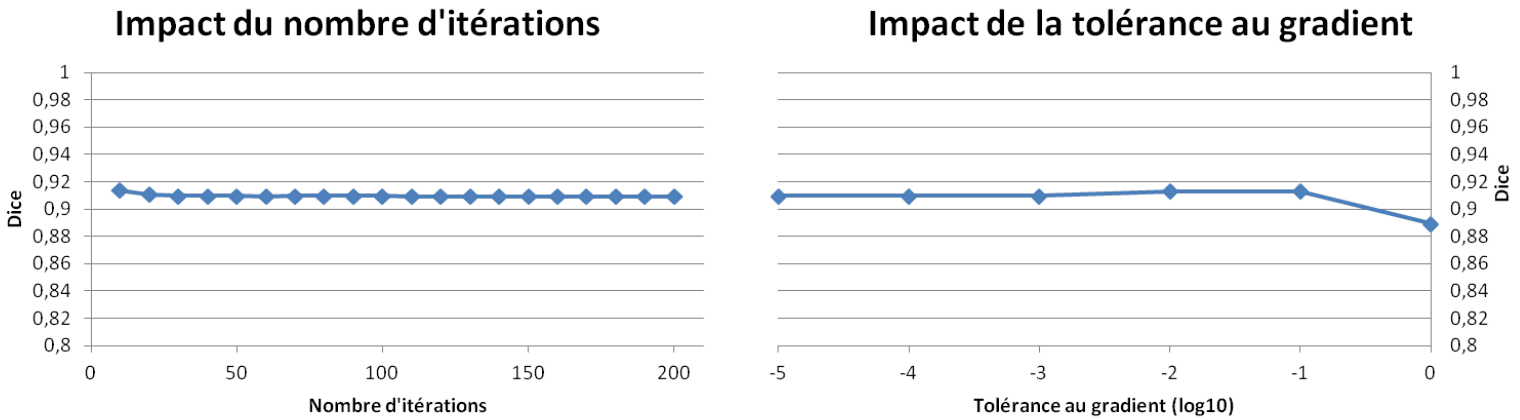
(e) Point de la couronne mal positionné au niveau de la chambre pulpaire

FIGURE 3.17 – Illustration de l'influence de la position des points donnés par l'utilisateur. Dans l'ordre des lignes : sélection correcte, points d'apex en dehors des racines, points d'apex trop haut dans les racines, point de la couronne trop éloigné de la surface, point de la couronne au niveau de la chambre pulpaire. Chaque ligne présente successivement les résultats du recalage (vignettes 2 à 4) et de la coupure de graphe (vignettes 5 à 7).



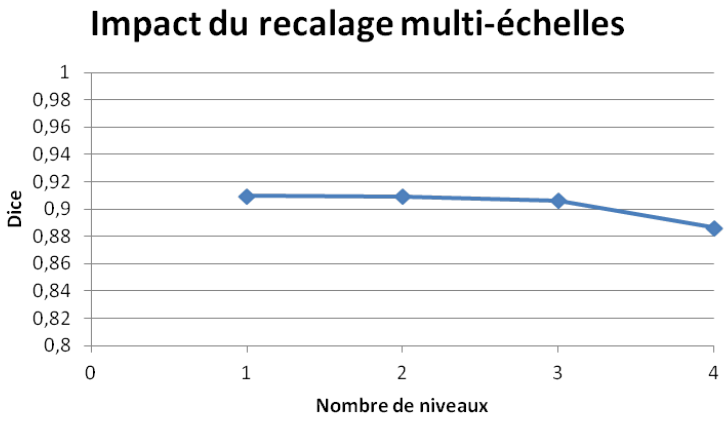
(a) Gradient minimal $= 0,1$ (b) Gradient minimal $= 0,01$ (c) Gradient minimal $= 0,001$

FIGURE 3.18 – Illustration de l'influence de la sensibilité au gradient dans la précision du recalage. Le nombre d'itérations requis pour obtenir ces résultats est respectivement 0 (l'initialisation est suffisamment correcte), 4 et 13. L'amélioration peut se noter au niveau de l'alignement de la forme dans la vue axiale.

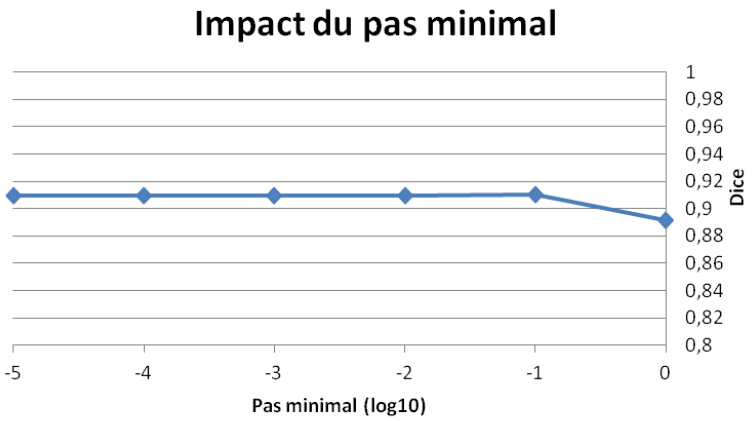


(a)

(b)



(c)



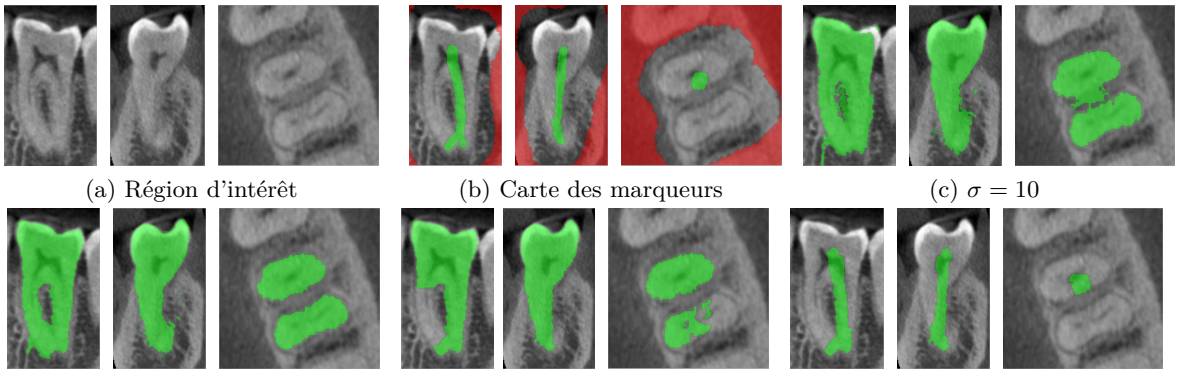
(d)

FIGURE 3.19 – Impact des paramètres de recalage sur la segmentation finale. Les valeurs fixes par défaut sont 100 itérations, tolérance au gradient et pas minimal de 0,001, un seul niveau de recalage, $\alpha = 1$, $\beta = 5$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$. (a) Score de Dice moyen en fonction du nombre d'itérations. (b) Score de Dice moyen en fonction de la tolérance au gradient. (c) Score de Dice moyen en fonction du nombre de niveaux d'échelle. (d) Score de Dice moyen en fonction du pas minimal accepté.

l'échelle des intensités, car une différence d'intensité de 50 sur une gamme de [0,255] n'a pas la même signification que sur une gamme de [0,65535] par exemple. Pour obtenir un comportement similaire en présence de données hétérogènes en intensité, ce qui n'est pas notre cas, il peut être nécessaire de procéder à une normalisation en intensité des images.

Si nous fixons $\sigma = 100$ et que nous faisons varier γ nous obtenons le comportement illustré en figure 3.21. Plus la valeur augmente, plus la régularisation imposée à la segmentation au niveau des gradients est forte. Là encore, le comportement correspond à ce qui peut être déduit de la formule du terme associé B : à pénalité de gradient constante (c'est-à-dire à σ constant), plus la surface de la segmentation est faible, moins le nombre d'arêtes du graphe coupées par l'optimisation est important, réduisant l'énergie. Le poids augmente donc la qualité de segmentation (voir 3.25 (b)) jusqu'à ce qu'il restreigne la segmentation à l'intérieur de la dent, faisant chuter les valeurs du coefficient de Dice. Notons que les fuites au niveau de la couronne et au niveau de l'os cortical sont progressivement supprimées, mais que dans le cas présenté, les racines étant très proches, nous englobons l'espace inter-radiculaire. Le poids γ est important pour obtenir une segmentation contenue à la dent ciblée au niveau des intensités similaires.

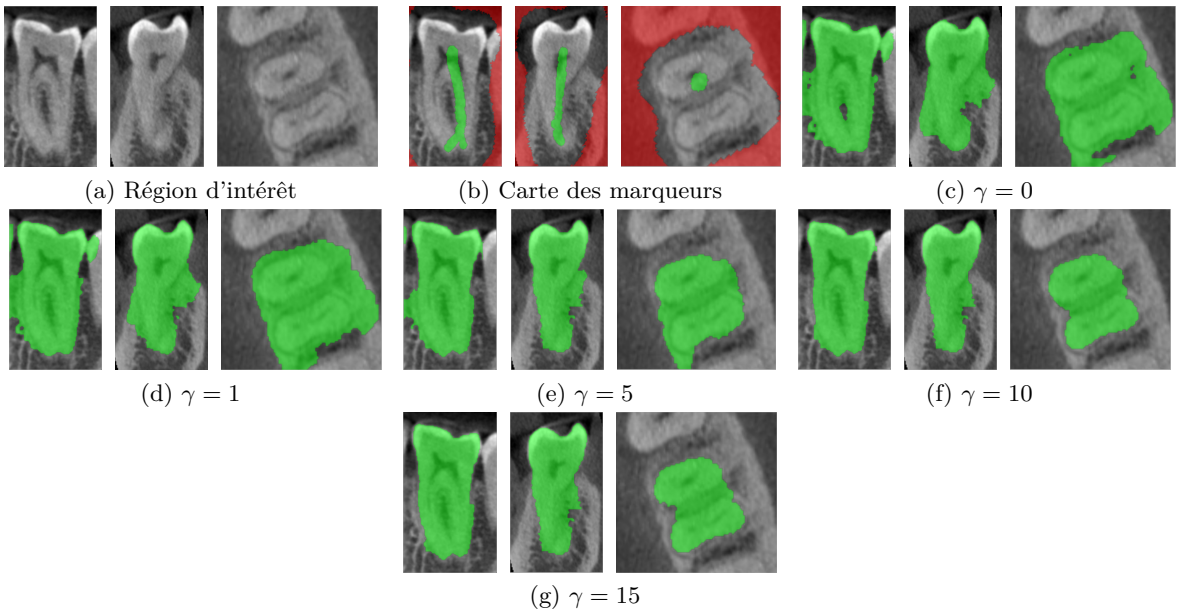
Dans le cas de α (associé au terme R) (figure 3.22), nous voyons qu'en son absence, une partie de la



(a) Région d'intérêt (b) Carte des marqueurs (c) $\sigma = 10$

(d) $\sigma = 25$ (e) $\sigma = 50$ (f) $\sigma = 100$

FIGURE 3.20 – Illustration de l'influence du paramètre σ dans le terme de régularisation reposant sur le gradient. Les autres paramètres sont $\alpha = 1$, $\beta = 5$, $\gamma = 15$, et $\xi = 1$.



(a) Région d'intérêt (b) Carte des marqueurs (c) $\gamma = 0$

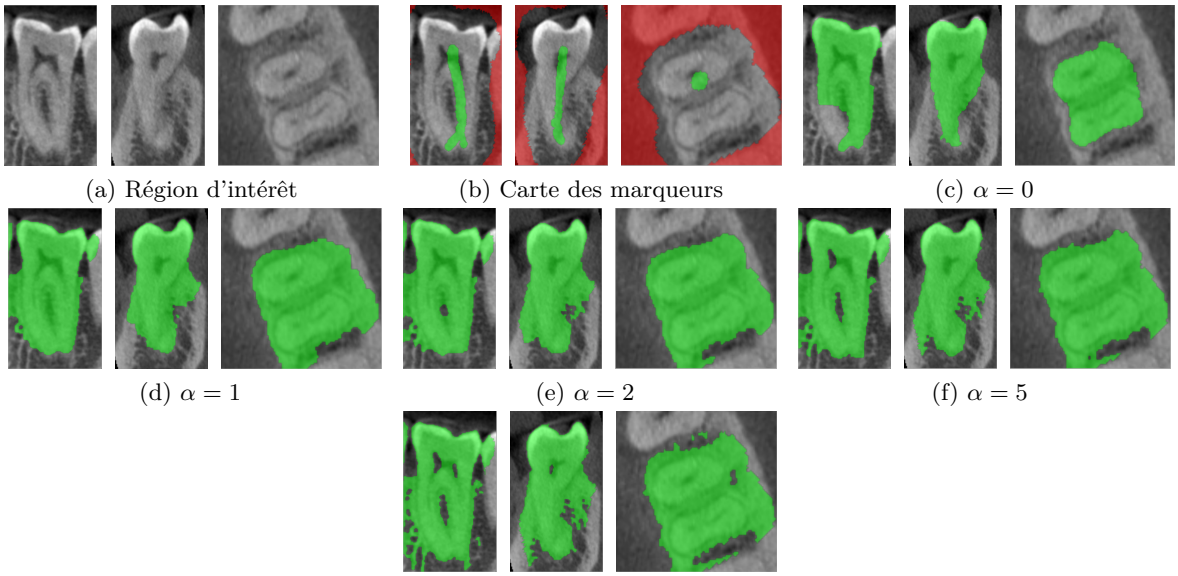
(d) $\gamma = 1$ (e) $\gamma = 5$ (f) $\gamma = 10$

(g) $\gamma = 15$

FIGURE 3.21 – Illustration de l'influence du terme B de coupes aux gradients lorsque son poids γ varie. Les autres paramètres sont $\alpha = 1$, $\beta = 5$, $\xi = 1$ et $\sigma = 100$.

racine est perdue puisque le marqueur du plus court chemin a fusionné avec celui de la racine adjacente : sans vraisemblance, seuls les gradients et la forme comptent. Son utilisation permet d'inclure dans la segmentation cette racine, mais nous voyons aussi apparaître des fuites au niveau des couronnes et de l'os cortical. Lorsque le terme est fortement pondéré, la segmentation prend alors les trabéculations internes de l'os et tend à exclure les canaux radiculaires, comportement similaire à un seuillage en intensité entre structures dures et tissus mous. Ce défaut était attendu, puisque la vraisemblance s'appuie sur des densités de probabilités qui favorisent l'os (puisque d'intensité similaire à la dentine, voir partie 3.1). Ce poids permet de s'assurer que toute la dent sera segmentée lorsque les marqueurs ne sont pas suffisants pour couvrir l'étendue de l'objet, au risque d'une segmentation des structures d'intensité proche si trop important (voir 3.25 (c)). Il sera donc utilisé dans notre cas avec une faible valeur.

Afin de pénaliser les fuites, nous avons introduit le terme D. Lorsque son poids β augmente (figures 3.23 et 3.25 (d)), nous pouvons voir que les fuites régressent, jusqu'à ce que la segmentation tende vers la forme moyenne : reposant sur la carte de distance de cette dernière, il force la segmentation à lui ressembler en faisant correspondre les étiquettes du fond et de l'objet leurs espaces respectifs dans



(a) Région d'intérêt (b) Carte des marqueurs (c) $\alpha = 0$

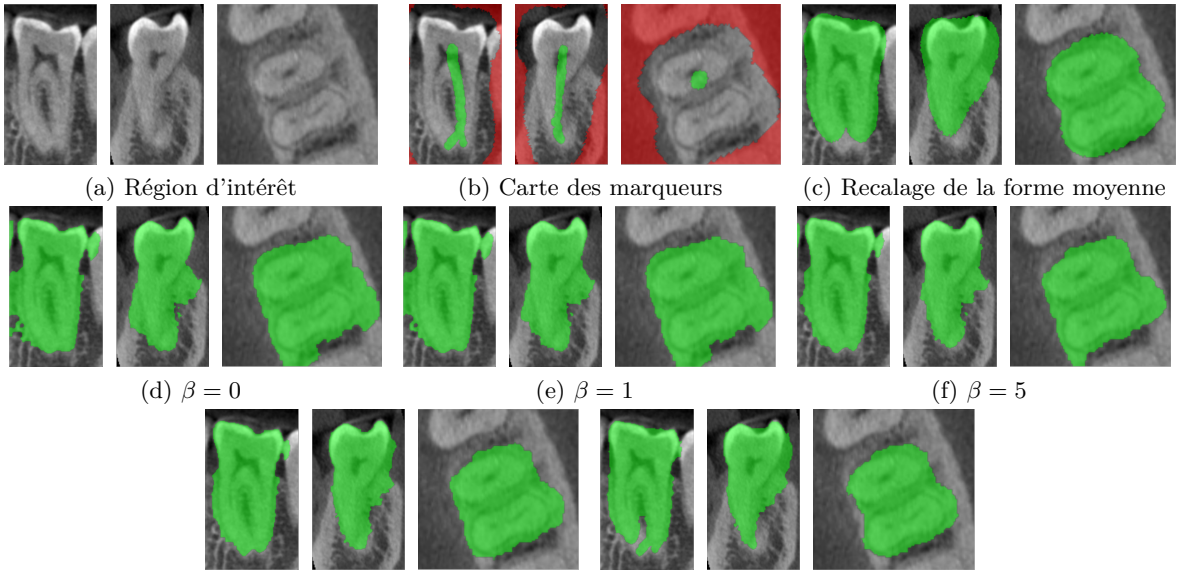
(d) $\alpha = 1$ (e) $\alpha = 2$ (f) $\alpha = 5$

(g) $\alpha = 10$

FIGURE 3.22 – Illustration de l'influence du terme R d'a priori sur l'intensité lorsque son poids α varie. Les autres paramètres sont $\beta = 5$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$.

la carte de distance. Il favorise donc l'apparition d'une mauvaise segmentation à l'intérieur de cette forme en prenant le pas sur le terme a priori d'intensité. Par exemple, nous voyons sur la vue coronale (deuxième vignette) une fuite au niveau de l'air entre le cas (d) et (h). Ce terme est intéressant lorsque l'information d'intensité est peu pertinente, notamment en présence de métal au niveau de la couronne. De même, le terme T agit de manière similaire (figures 3.24 et 3.25 (e)) à la différence qu'il est mis en balance avec le terme du gradient, ce qui permet de renforcer la séparation lorsque la délimitation est ténue.
La pertinence de ces termes est conditionnée par la forme sur laquelle ils reposent et donc par la bonne correspondance entre cette dernière et l'image. Dans notre cas, la forme moyenne est une bonne estimation pour les dents antérieures, et approximative dans le cas des molaires, une forme adaptée à l'image procurerait donc probablement des résultats encore meilleurs.

L'impact des différents termes sur la segmentation finale en termes de valeur du coefficient de Dice est présenté dans la figure 3.25. L'impact de σ (figure 3.25 (a)) est conforme à ce que nous avons observé visuellement, s'il est trop faible ou trop fort il ne joue pas son rôle de sélection entre les gradient. La valeur optimale ici est de 100. Nous pouvons voir que le poids γ (figure 3.25 (b)) permet d'améliorer sensiblement les résultats et qu'une valeur de 15 est optimale. La vraisemblance d'intensité (figure 3.25 (c)) présente une courbe très particulière en adéquation avec nos observations visuelles : sa présence permet d'augmenter significativement les résultats, mais plus elle pèse, plus les valeurs se dégradent. La valeur optimale est donc de 1. L'apport des contraintes de forme (figures 3.25 (d,e)) est plus ténu, mais des valeurs de $\beta = 5$ et $\xi = 1$ permettent d'augmenter le Dice de l'ordre de 0,03 et 0,01 respectivement. Les valeurs optimales correspondent donc à celles trouvées empiriquement, et donnent un coefficient de Dice moyen de 0,91.

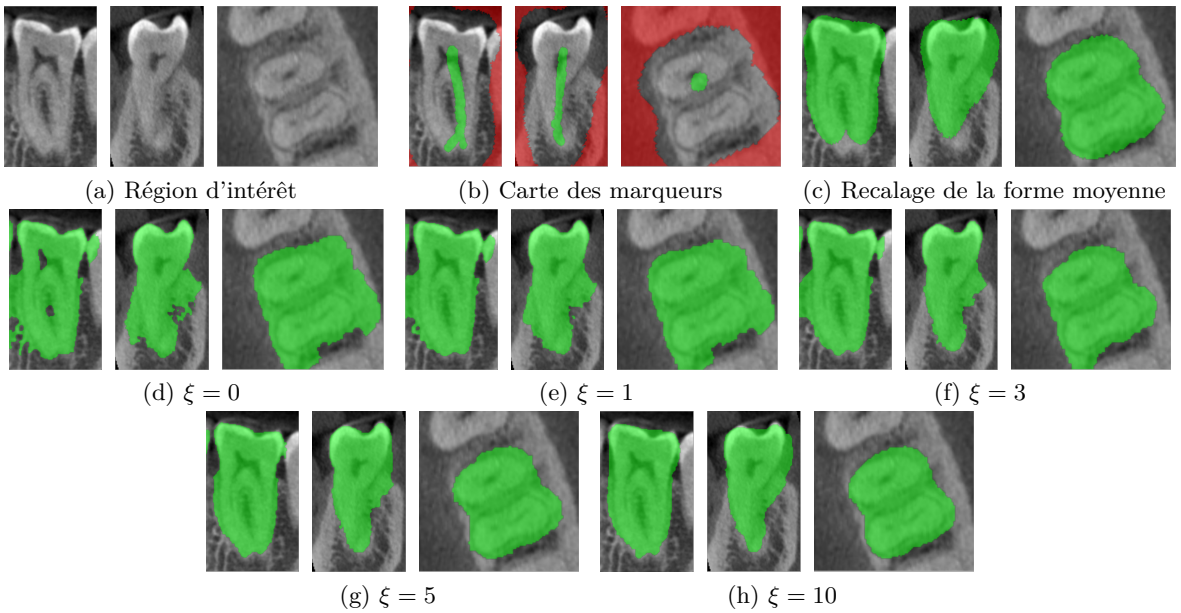


(a) Région d'intérêt (b) Carte des marqueurs (c) Recalage de la forme moyenne

(d) $\beta = 0$ (e) $\beta = 1$ (f) $\beta = 5$

(g) $\beta = 10$ (h) $\beta = 25$

FIGURE 3.23 – Illustration de l'influence du terme D d'a priori sur la distance lorsque son poids β varie. Les autres paramètres sont $\alpha = 1$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$.

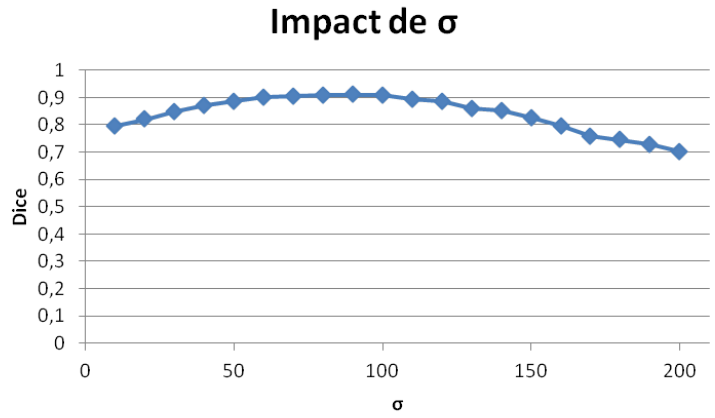


(a) Région d'intérêt (b) Carte des marqueurs (c) Recalage de la forme moyenne

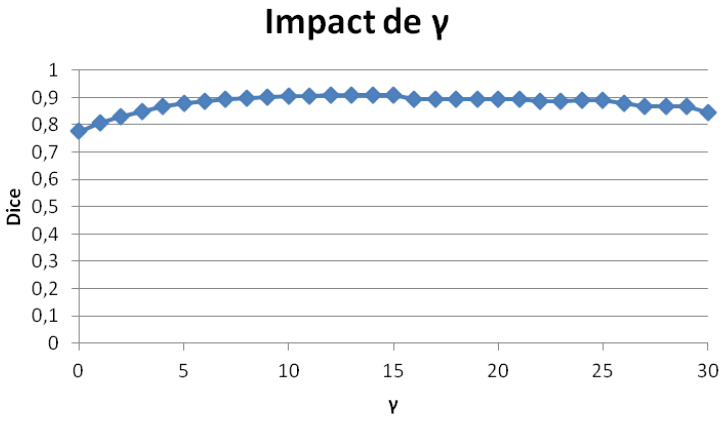
(d) $\xi = 0$ (e) $\xi = 1$ (f) $\xi = 3$

(g) $\xi = 5$ (h) $\xi = 10$

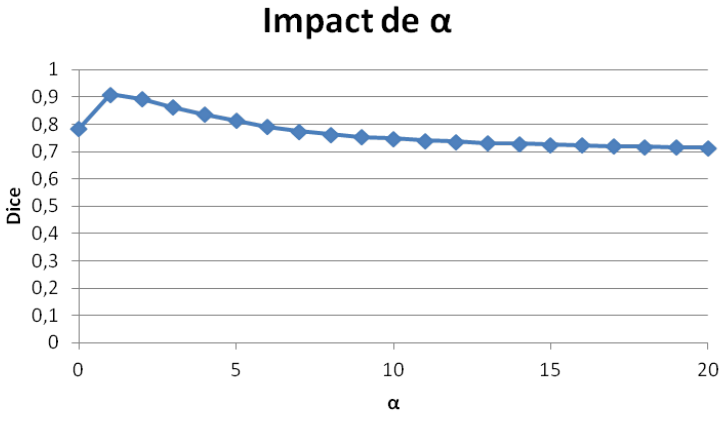
FIGURE 3.24 – Illustration de l'influence du terme T lorsque son poids ξ varie. Les autres paramètres sont $\alpha = 1$, $\beta = 5$, $\gamma = 15$ et $\sigma = 100$.



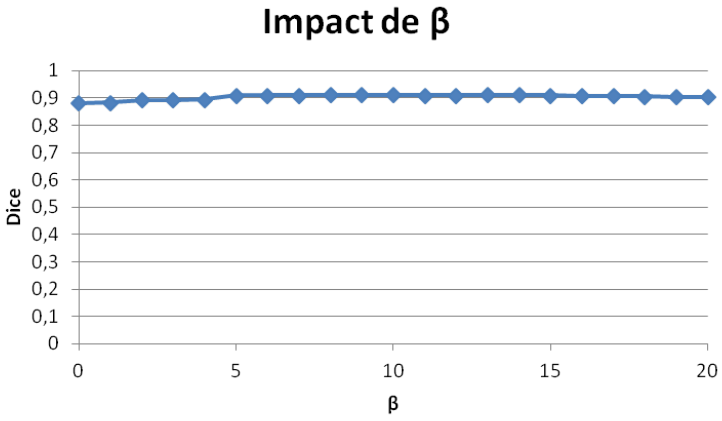
(a)



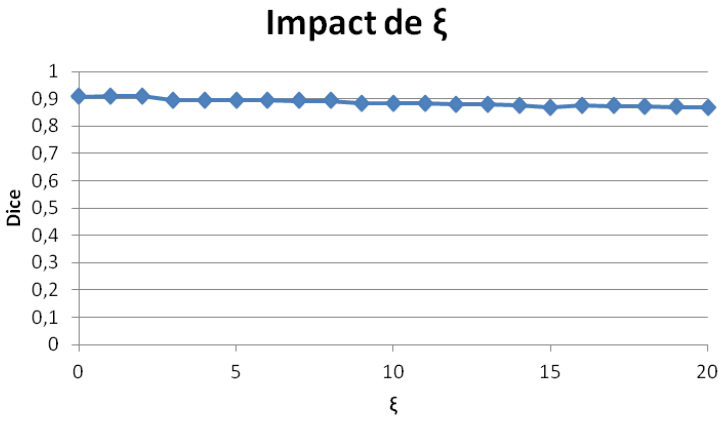
(b)



(c)



(d)



(e)

FIGURE 3.25 – Influence des paramètres de la segmentation par coupure de graphe sur la segmentation finale. Les valeurs fixes par défaut sont 100 itérations, tolérance au gradient et pas minimal de 0,001, un seul niveau de recalage, $\alpha = 1$, $\beta = 5$, $\gamma = 15$, $\sigma = 100$ et $\xi = 1$. (a) Score de Dice moyen en fonction de σ. (b) Score de Dice moyen en fonction de γ. (c) Score de Dice moyen en fonction de α. (d) Score de Dice moyen en fonction de β. (e) Score de Dice moyen en fonction de ξ.

3.4.3 Comparaison avec d'autres approches

Modèle déformable avec a priori de forme

La formulation utilisée ici est celle décrite dans [85] (voir section 2.1.2). Une contrainte de forme est ajoutée dans une formulation par ensemble de niveaux grâce à un modèle statistique pré-entraîné de manière similaire à celle que nous utilisons. À chaque itération du processus, la pose et la forme du modèle sont évaluées sur l'image des potentiels de gradient, puis le contour est déformé en fonction du gradient et de la position estimée du modèle. Cette méthode est donc très proche de la nôtre dans l'approche du problème en exprimant une contrainte de forme obtenue d'après un modèle de forme statistique comme un terme contraignant l'optimisation de l'énergie. Deux différences sont notables : d'une part, la contrainte est ici optimisée conjointement à la segmentation, alors qu'elle est fixée préalablement dans la méthode proposée, et d'autre part, les auteurs n'utilisent pas de contraintes sur l'intensité des régions, alors que nous employons une vraisemblance. L'implémentation utilisée est celle disponible dans la bibliothèque ITK.

Powerwatershed

Le *powerwatershed* [40] (expliqué dans la section 2.1.1) est une formulation énergétique du principe de la ligne de partage des eaux sur un graphe. Cet algorithme se trouve donc dans la même famille que la coupure de graphe ou encore que la marche aléatoire et présente en moyenne des performances robustes quelle que soit la distribution des marqueurs tout en ayant une complexité quasi-linéaire, ce qui le rend attractif pour le traitement de grands volumes, contrairement à la coupure de graphe qui se révèle très coûteuse lorsque la taille des volumes augmente. L'implémentation utilisée pour la comparaison est celle mise à disposition par les auteurs (http ://powerwatershed.sourceforge.net). Dans cette version, la formulation est restreinte à un terme de régularisation sur le gradient, de nature similaire à notre terme γ, à partir des marqueurs donnés par l'utilisateur.

Forêt aléatoire

Le principe général des forêts aléatoire a déjà été présenté dans le chapitre 2. Ici nous avons utilisé l'approche décrite dans [59] en utilisant comme caractéristiques d'apprentissage les descripteurs RISC (*Rotation invariant shape context*) ainsi que l'intensité du voxel. Les descripteurs RISC consistent dans notre cas en un ensemble de 20 bandes comprises entre des sphères concentriques centrées sur le voxel, de rayon aléatoire compris entre 0,3 et 5 mm. Dans chacune de ces bandes est calculé le pourcentage de voxels représentant une structure dure (définie empiriquement comme ayant une valeur supérieure ou égale à 2000). Le vecteur caractéristique de chaque voxel est donc constitué de 21 valeurs. L'apprentissage a été effectué sur 4 dentures complètes. L'implémentation utilisée est celle disponible dans la bibliothèque OpenCV.

Comparaison

Les résultats sont présentés dans la figure 3.26. Les résultats moyens du Dice sur l'ensemble de base sont les suivants :

— Forêt aléatoire avec descripteurs RISC : $0,5 \pm 0,173$

— Modèle déformable avec a priori de forme : $0,671 \pm 0,102$

— *Powerwatershed* : $0,855 \pm 0,131$

— Notre méthode : $0,91 \pm 0,04$

La forêt aléatoire produit les résultats les moins probants. Le classificateur tend à confondre l'os cortical avec la dent, tout en excluant les canaux radiculaires. Cette dernière méthode repose uniquement sur la présence de structures osseuses dans un voisinage concentrique, et nous pouvons voir que ce n'est pas suffisant dans notre contexte pour produire une segmentation de qualité. Notons aussi que le temps de calcul des descripteurs RISC est très élevé, environ 15 minutes par zone d'intérêt, dû au fait qu'elle contiennent quasiment exclusivement des structures dures. Il faudrait donc définir des

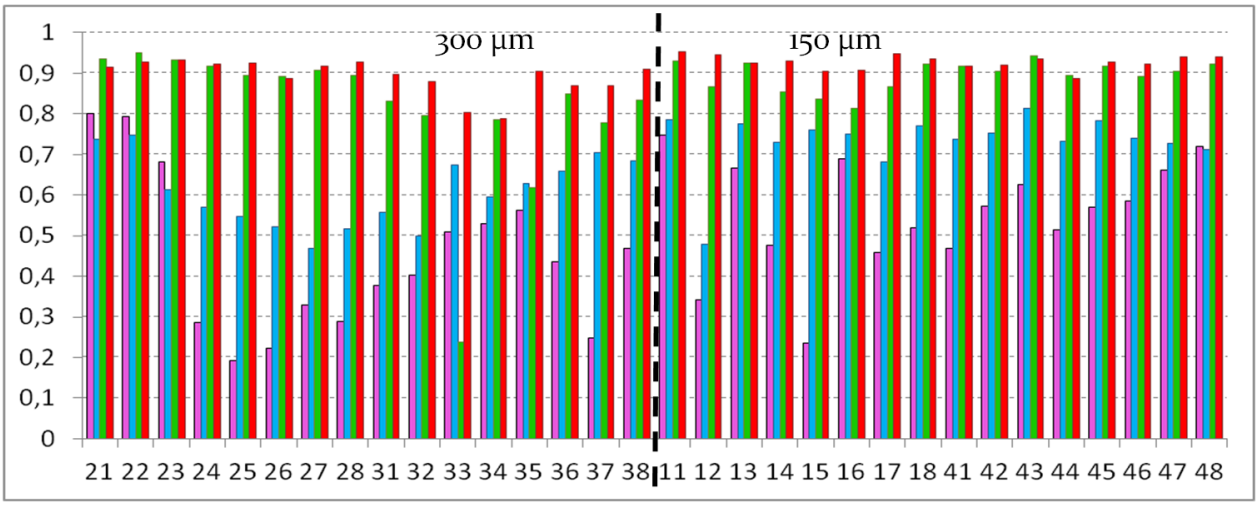


FIGURE 3.26 – Comparaison des résultats des approches sur notre base de test : forêt aléatoire avec descripteurs RISC [59] (violet), modèle déformable avec a priori de forme [85] (bleu), *powerwatershed* [40] (vert) et notre méthode [49] (rouge). Pour chaque numéro dentaire est indiqué le score de Dice associé. La première moitié des cas correspond à des échantillons issus de volumes avec une taille de voxels de 300 microns, contre 150 microns pour la seconde moitié. La perte de qualité des résultats à 300 microns est notable.

descripteurs plus aptes à capturer la spécificité des dents face aux os que ceux que nous avons choisi. La pertinence d'une approche par forêt aléatoire n'est pas remise en cause par cette comparaison.

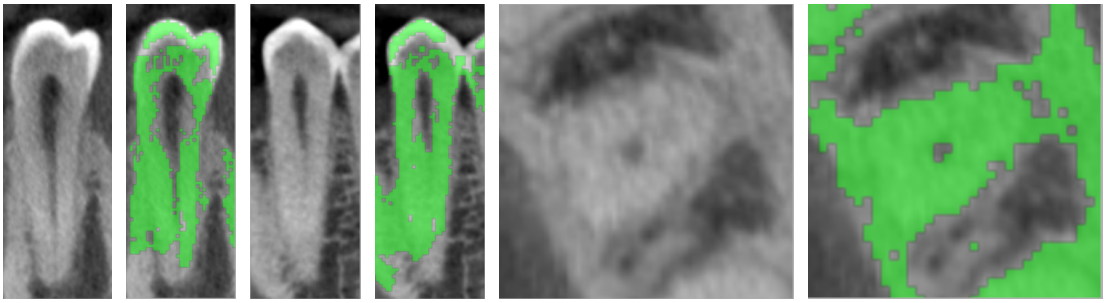
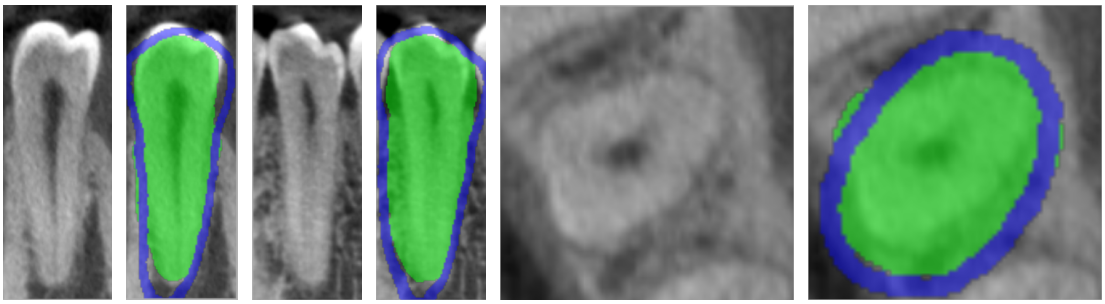


FIGURE 3.27 – Exemple de résultat produit par l'approche par forêt aléatoire avec descripteurs RISC. On note la prise en compte de l'os cortical dans la segmentation, et une absence du canal radiculaire.

L'approche par modèle déformable produit une segmentation qui peut être correcte (avec un score de Dice d'environ $0,8$), mais présente des problèmes de robustesse. Le principal problème provient de l'optimisation de la contrainte de forme. Celle-ci dépend de quatre termes : une vraisemblance de forme et une de pose, un terme de régularisation et enfin un terme d'attraction par le gradient. Ce dernier est le seul à lier la forme aux données de l'image, et s'avère souvent insuffisant pour fixer le modèle de forme aux contours réels de la dent. La présence de nombreux gradients plus importants tend à faire dériver la contrainte sur des frontières inexactes. Notons aussi que l'initialisation a un impact très important, non seulement sur la qualité de la segmentation mais aussi sur le temps de calcul, qui peut atteindre jusqu'à 3 minutes par dent.

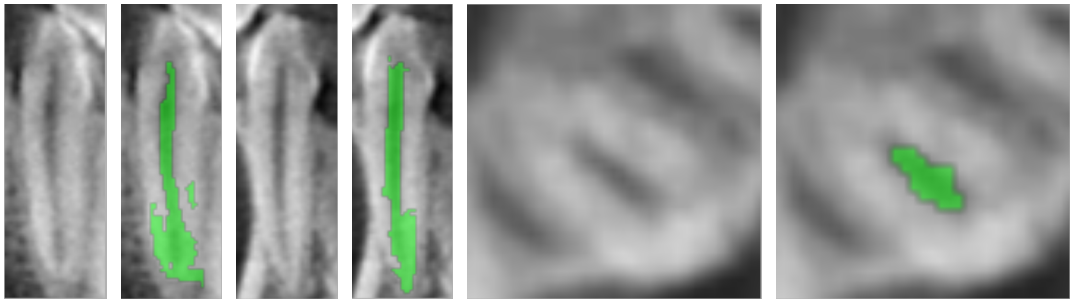
Enfin, le *powerwatershed* conduit dans la majorité des cas à des résultats de bonne qualité, très similaires à ceux obtenus par notre méthode, ce qui confirme l'importance des gradients dans la segmentation dentaire. Notons toutefois que parfois cette approche échoue à segmenter correctement, par exemple la dent numéro 33, présentée dans la figure 3.29. Dans ce cas, le marqueur initial de la dent étant contenu entièrement dans la lumière du canal, l'absence de contrainte sur la forme et



(a) (b) (c) (d) (e) (f)

FIGURE 3.28 – Exemple de résultat produit par l'approche par modèle déformable avec a priori de forme. La vue axiale permet de bien voir que la contrainte de forme (en bleu) se positionne sur des gradients importants qui ne correspondent pas aux contours dentaires.

la position de la dent conduit l'algorithme à ne segmenter quasiment que le canal, à cause du fort gradient entre ce dernier et la dentine. Nous souhaitons souligner cependant que le temps de calcul est très inférieur à celui de notre méthode, entre 2 et 6 secondes en moyenne selon la taille de la zone d'intérêt (c'est-à-dire selon le type dentaire) sur des volumes de 150 microns de résolution, ce qui représente un temps jusqu'à six fois inférieur à nos résultats. Au vu de la proximité méthodologique de ces deux approches, l'intégration de l'algorithme d'optimisation powerwatershed à notre formulation énergétique pourrait constituer une piste intéressante pour l'amélioration du service rendu dans le contexte clinique.



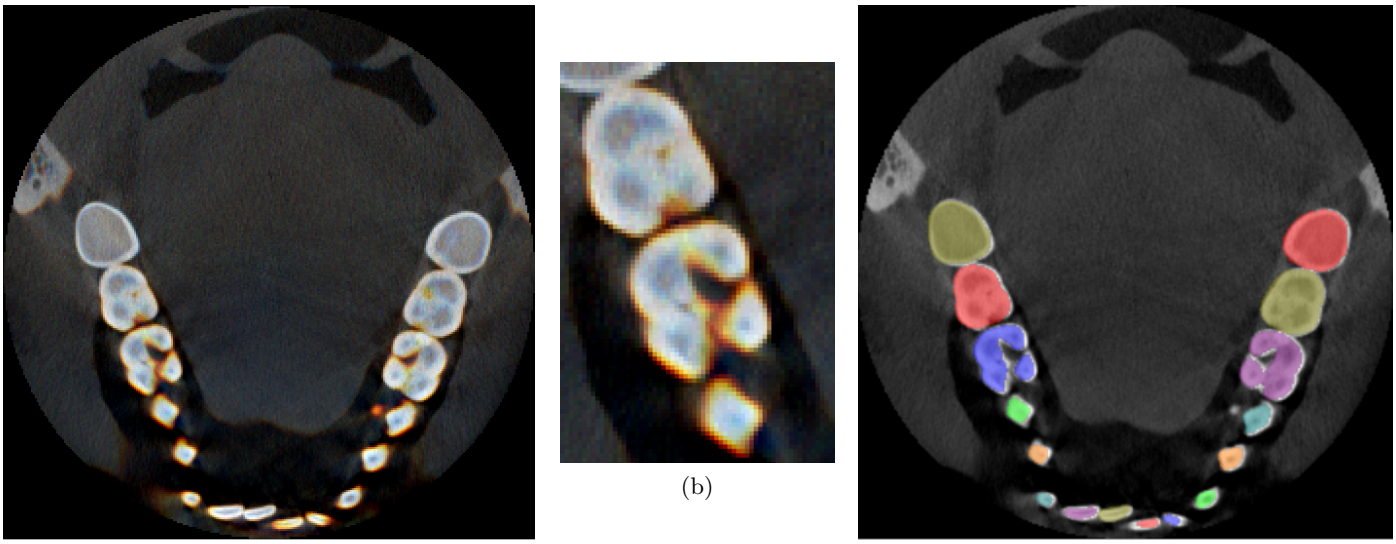
(a) (b) (c) (d) (e) (f)

FIGURE 3.29 – Cas particulier de la dent numéro 33 segmentée par *powerwatershed*.

Réseau de neurones convolutifs

L'approche employée ici est celle utilisée dans [127] comme décrite dans le chapitre 2 : l'architecture a été présentée en figure 2.6. Notre base d'images est constituée de 4 volumes (les patients 1 à 4, voir table 3.2), ce qui conduit à 1180 images 2D (fusionnant donc l'information de 3 coupes volumiques), séparées de manière aléatoire en 960 pour l'entraînement et 220 pour la validation. Chaque dent est annotée par son numéro dentaire, et nous entraînons le réseau à différencier chaque dent, contrairement aux expériences précédentes où nous cherchions à segmenter une dent face à son environnement.

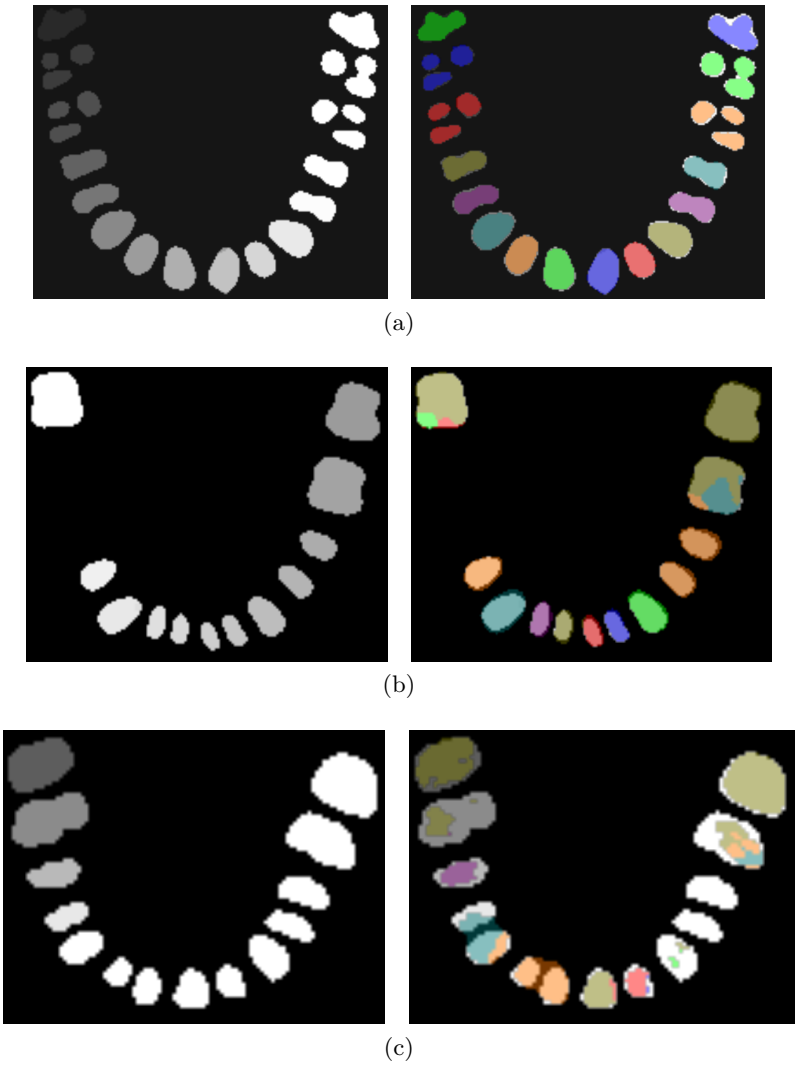
L'implémentation de cette approche est effectuée en Python à partir du code aimablement fourni par les auteurs de [127], reposant sur la bibliothèque Caffe. L'expérimentation a été réalisée sur une machine dotée d'une carte graphique Nvidia K80 (2×12 GB de mémoire) et 24 cœurs de processeurs (3,4GHz), du fait de la nécessité de faire appel à la programmation GPU pour obtenir des temps de calcul acceptables. Le temps d'apprentissage est d'environ deux jours, pour 150 000 itérations à un taux d'apprentissage de 10^{-9}, et cette étape consomme jusqu'à 11 GB de mémoire GPU. Le temps d'inférence est ensuite d'environ une seconde par coupe (soit environ entre cinq à dix minutes pour



(a) (b) (c)

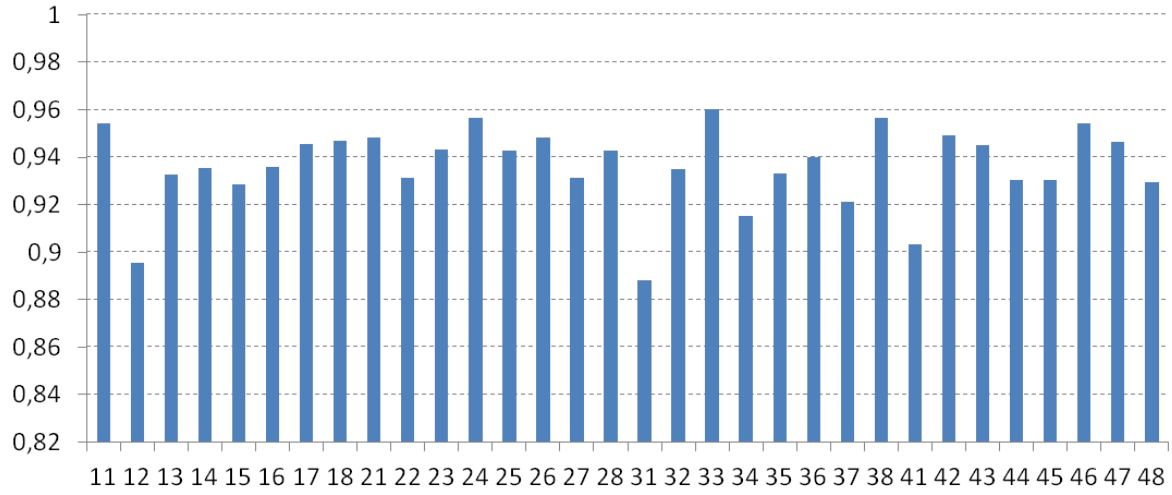
FIGURE 3.30 – Segmentation d'une image avec le réseau de neurones : (a) image d'entrée du réseau, (b) zoom sur une partie de l'image d'entrée, où l'on peut noter des zones de teinte bleue, qui correspondent à une information existante dans la coupe supérieure, et des zones de teinte rouge, correspondant à l'information provenant de la coupe inférieure, (c) image des annotations superposées à l'image d'entrée (les couleurs ne reflètent pas la numérotation dentaire).

un volume entier). Les résultats issus de la validation sont présentés en figure 3.32 (a) : le moins bon score de Dice obtenu est de 0,888, pour une moyenne de 0,936, ce qui témoigne d'une bonne qualité de segmentation (exemple en figure 3.31 (a)), du niveau de l'état de l'art. Cependant, lors du test sur une mâchoire partiellement édentée (absence des dents 38, 45, 46 et 48, voir figure 3.31 (b)) ne faisant pas partie des volumes initiaux d'apprentissage, les résultats sont beaucoup plus contrastés (voir figure 3.32 (b)). La seconde prémolaire gauche (35) est partiellement confondue avec la première prémolaire (34), la première molaire gauche (36) est elle aussi partiellement confondue avec la seconde molaire (37) et enfin la seconde molaire droite (47) est majoritairement étiquetée comme une première molaire (46), dent pourtant absente. Soulignons toutefois qu'abstraction faite de l'étiquette individuelle des dents, la délimitation entre les dents et leur environnement est très satisfaisante. Il serait donc intéressant d'expérimenter si nous pouvons obtenir des résultats plus précis en sacrifiant la discrimination du numéro dentaire (cette expérience n'a pas été menée par manque de temps). La dernière expérience (figure 3.32 (c)) a été réalisée avec des volumes acquis en grand champ. Ici les résultats montrent un réseau inadapté (figure 3.31 (c)), avec un coefficient de Dice moyen de $0,16$. Cette expérience montre que le réseau n'est pas capable de généraliser son analyse aux différentes résolutions de manière automatique, et souligne donc l'importance de la base d'apprentissage : son enrichissement avec des champs de vue et des résolutions différents ainsi que des cas pathologiques semble être nécessaire à la constitution d'un réseau plus performant, ce qui soulève le problème de la difficulté d'obtenir des données annotées.

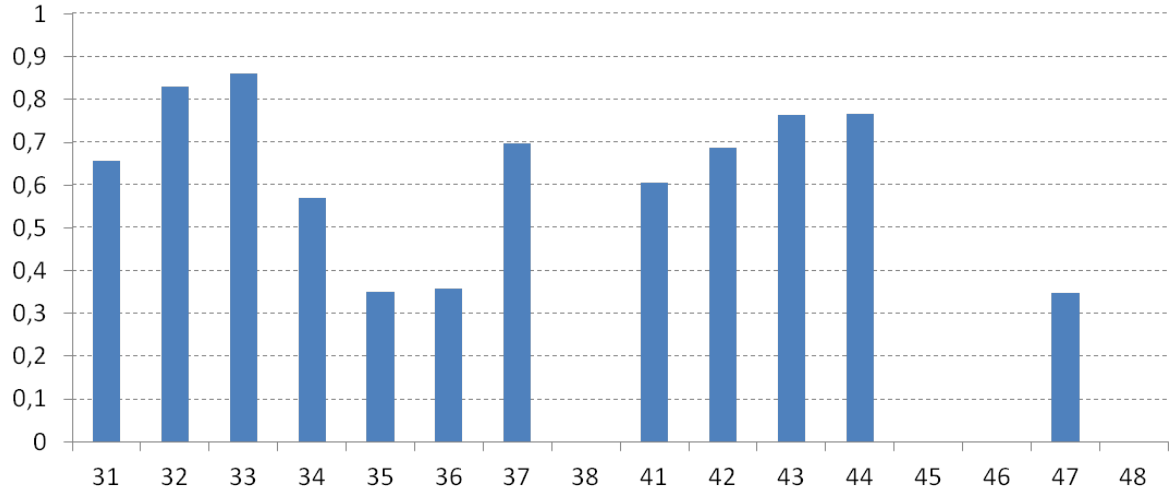


(a) (b) (c)

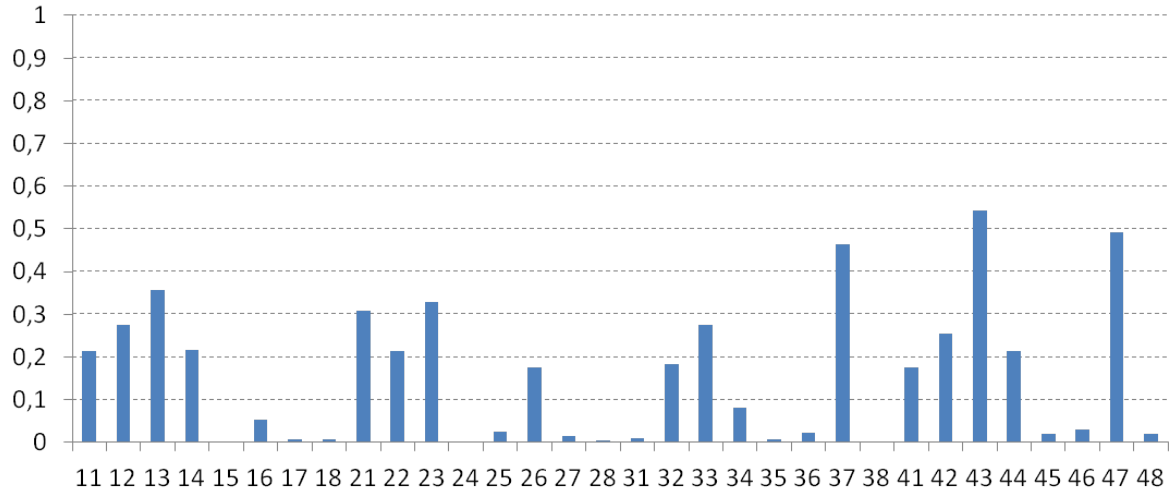
FIGURE 3.31 – Exemples de comparaison de la segmentation produite par le réseau superposée à la segmentation de référence. (a) Cas issu de la base de validation, (b) cas de la mâchoire édentée, (c) cas d'un volume en grand champ.



(a)



(b)



(c)

FIGURE 3.32 – Résultats du score de Dice obtenus par le réseau de neurones : (a) pour chaque numéro dentaire sur l'ensemble de validation, (b) sur une arche mandibulaire partiellement édentée, (c) sur un ensemble de volume en grand champ.

Comparaison avec une méthode interne de l'équipe

Nous comparons ici notre méthode avec une approche similaire dans son fonctionnement appelée "algorithme CV" dans la suite, déjà utilisée dans l'entreprise. Cette méthode est conçue pour la spécialité d'endodontie, et segmente les dents en faisant la distinction entre ses sous-parties (émail, dentine et canaux nommément). À partir d'un ensemble de points qui correspondent à la chambre pulpaire et à l'extrémité des canaux radiculaires, les canaux sont segmentés par une propagation de front [44] qui a pour critère l'intensité. Une fois les canaux obtenus, la dentine est segmentée en formulant le problème comme la minimisation d'une énergie et optimisée par coupure de graphe. Les termes de cette énergie sont équivalents à nos termes R (vraisemblance) et B (régularisation) à la différence notable que la vraisemblance repose sur une image du domaine spatial attendu pour la dentine, qui est obtenue par :

1. calcul de la matrice hessienne et création d'une image des éléments de forme planaire plus intense que leur environnement (méthode de [53]),
2. à partir des canaux précédemment segmentés, sélection uniquement de la dentine en éliminant toute partie non visible au sens morphologique depuis les points du canal, c'est-à-dire tous les points du domaine dont la géodésique aux points du canal intersecte le fond de l'image,

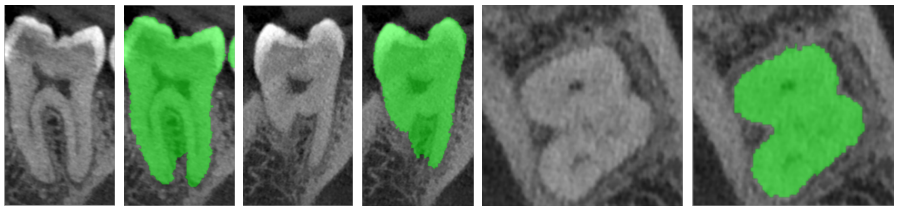
Cette méthode a démontré des résultats identiques aux nôtres (score de Dice moyen d'environ $0,95$) sur les volumes d'endodontie (de résolution isotrope de 75 microns), mais des réserves avaient été émises quand à son utilisation sur des résolutions moins fines dues à la perte d'information. La figure 3.33 illustre qu'effectivement, bien que nos approches soient très proches, la qualité d'image pour des volumes en grand champ n'est pas suffisante pour utiliser cette approche. Les deux raisons principales sont la perte de délimitation entre la dentine et l'os environnant, comme évoqué précédemment, et d'autre part la disparition des canaux dans l'image due à la combinaison de deux facteurs : un signal plus diffus du capteur et une taille anatomique tendant à devenir sub-voxellique [67], rendant la méthodologie inadaptée. Notez cependant que sur les volumes bien résolus, l'algorithme CV différencie mieux la zone inter-racines que la formulation proposée.

En résumé

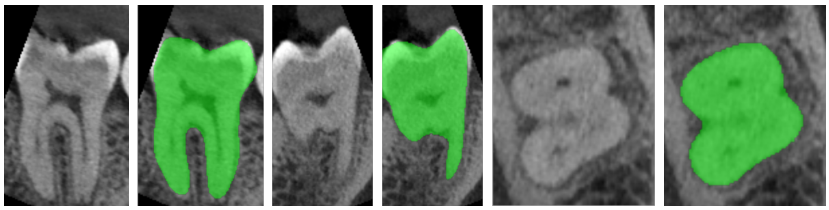
Notre méthode produit des résultats du niveau de l'état de l'art, et est facilement interprétable par l'utilisateur, mais tend à devenir coûteuse lors du traitement de grands volumes. Les performances démontrées par l'algorithme d'optimisation *powerwatershed* fournissent une piste intéressante d'amélioration de notre méthode. Nous avons aussi vu qu'une approche par apprentissage profond peut proposer de très bons résultats, mais souffre de deux problèmes majeurs à nos yeux. Premièrement, même avec un processus de transfert d'apprentissage, la diversité des images impose de constituer une base d'entraînement bien fournie, ce qui est problématique par le coût humain et temporel demandé par l'annotation des images. Deuxièmement, il est aujourd'hui difficile d'interpréter le fonctionnement d'un réseau, notamment pour un utilisateur dentiste, ce qui est problématique dès lors que ce dernier souhaite corriger les résultats. Une méthode pour prendre en compte le retour du praticien serait que le réseau réapprenne en fonction des corrections effectuées. Néanmoins, les ressources informatiques nécessaires à l'entraînement des réseaux (carte graphique de haut de gamme) ne sont pas compatibles avec les équipements habituellement présents dans les cabinets. C'est pourquoi nous pensons qu'à l'heure actuelle ce type d'approche, bien que prometteuse, n'est pas adaptée au contexte clinique dentaire.

3.4.4 Analyse de la performance sur des volumes de résolution 0,15 mm

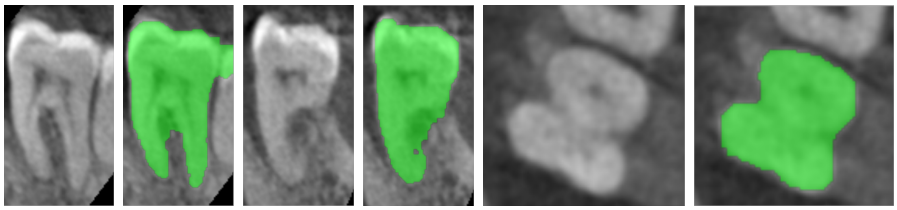
Le temps moyen de calcul partant de l'initialisation jusqu'à la segmentation finale, en excluant le temps pris par l'utilisateur pour indiquer les points et le numéro dentaire, est de 12 secondes pour des dents antérieures et jusqu'à 30 secondes pour des dents postérieures (la différence s'expliquant principalement par la taille des zones, qui impacte la partie de coupure de graphe), sur un ordinateur pourvu d'un processeur bi-cœur de fréquence 2,9 GHz. La consommation de mémoire du processus atteint 2 GB au maximum. Le nombre maximal d'itérations pour le recalage est fixé à 30 avec une



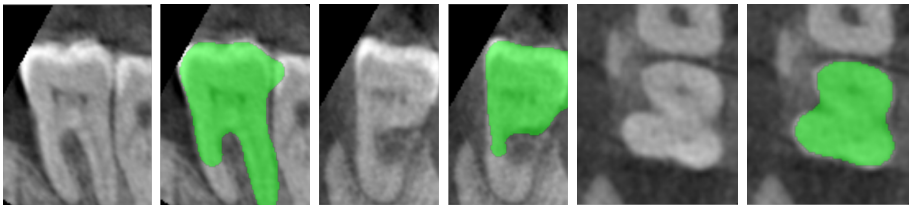
(a) Méthode proposée



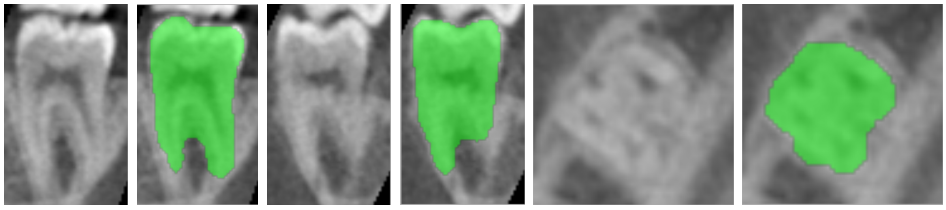
(b) Algorithme CV



(c) Méthode proposée



(d) Algorithme CV



(e) Méthode proposée

FIGURE 3.33 – Comparaison de la formulation proposée avec l'algorithme CV utilisé en interne. (a-b) Résultats sur un volume d'endodontie (taille de voxel isotrope de 75 microns) pour notre formulation puis pour le CV. (c-d) Résultats sur un volume de champ moyen (taille de voxel isotrope de 150 microns) pour notre formulation puis pour le CV. (e) Résultats sur un volume de grand champ (taille de voxel isotrope de 300 microns) : ce dernier cas met en échec la recherche des canaux dans l'algorithme CV, d'où l'absence de résultats. Chaque série de vignette présente successivement : la vue sagittale de l'image et de la segmentation, puis la vue coronale et enfin la vue axiale.

variation minimale du gradient de 0,001 lors de l'optimisation par descente de gradient. La constante de déformation C (équation 3.2) pour la génération de la contrainte spatiale est fixée à 3. À noter qu'ici nous n'avons pas retenu la zone résiduelle interne (élément bleu dans la figure 3.10) comme marqueur de la dent, cette dernière s'avérant une source d'erreur très sensible à la qualité du recalage. Le marqueur de la dent est donc uniquement constitué du chemin évalué entre les points donnés par utilisateur. Les paramètres de l'énergie sont quant à eux égaux à : $\sigma = 100$, $\alpha = 1$, $\beta = 5$, $\gamma = 15$ et

$\xi = 1$ (paramètres “optimaux”, issus de la section 3.4.2).

Les résultats moyens sur l'ensemble de la base sont présentés dans la table 3.4. Nous notons que les résultats apparaissent comme étant au niveau de l'état de l'art si nous nous référons à ce qui existe dans le domaine dentaire [58], avec un coefficient de Dice supérieur à 0,95 (voir figure 3.34). Le type de dent ne semble pas avoir d'incidence forte sur la capacité de segmentation de l'algorithme (voir figure 3.35) : si cela semble contre-intuitif car les modèles moyens des dents antérieures sont plus adaptés que celles des dents postérieures, donc avec une contrainte de forme plus pertinente, ces dernières tendent à être moins jointes entre elles, donc plus facilement isolables. Le post-traitement est bénéfique pour la majorité des cas (déplacement du nuage de points bleus vers de meilleurs résultats en comparaison du nuage rouge, figure 3.36). Cependant, soulignons une sensibilité moindre que la spécificité, ce qui indique que le processus tend à sous-estimer la segmentation réelle, mais en étant sûr que les éléments identifiés appartiennent effectivement à la dent.

	Coupure de graphe	Post-traitement
Sensibilité	0,969 ± 0,023	0,972 ± 0,023
Spécificité	0,998 ± 0,001	0,999 ± 0,001
Précision	0,998 ± 0,001	0,998 ± 0,001
Dice	0,953 ± 0,024	0,958 ± 0,023

TABLE 3.4 – Résultats moyens sur l'ensemble de la base pour la sensibilité, la spécificité, la précision et le coefficient de Dice .

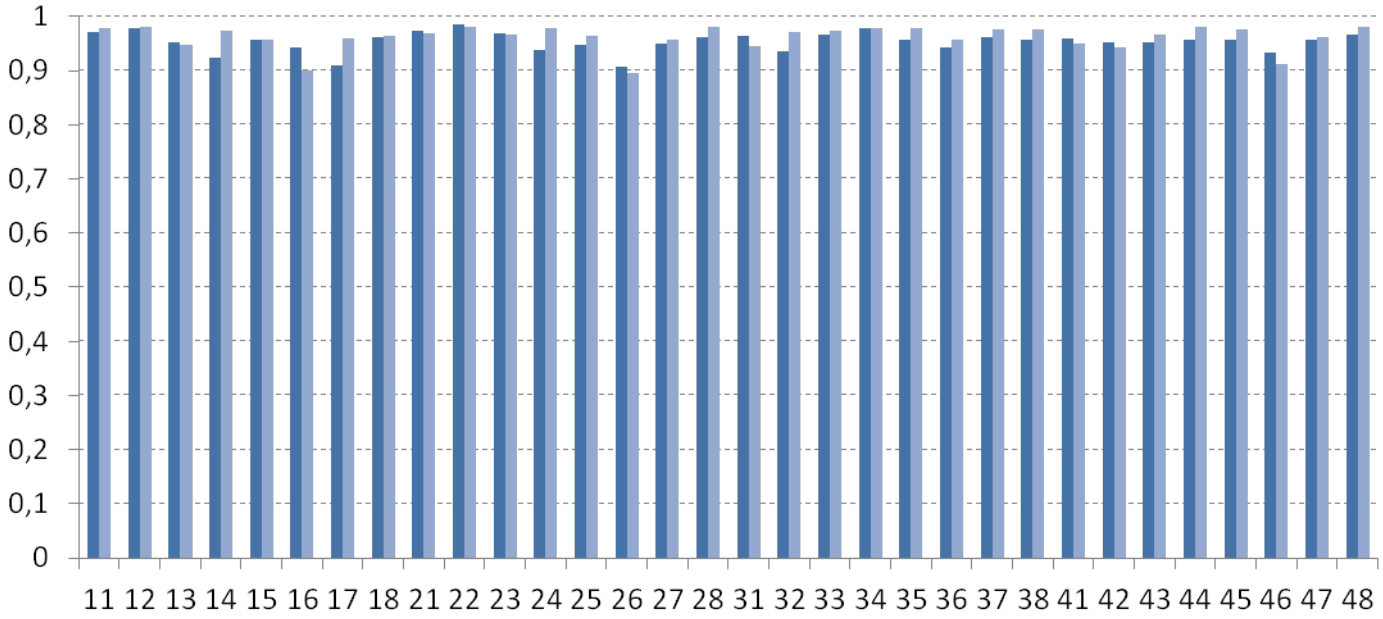


FIGURE 3.34 – Résultats du score de Dice pour chacun des 64 échantillons. La différence du ton de couleur indique un patient différent.

Comment expliquer ce comportement ? Si nous nous référons au graphe présenté en figure 3.36, nous nous apercevons que certains échantillons ont une sensibilité très inférieure à la moyenne et que le nuage principal a une variabilité assez importante sur cette mesure. Intéressons nous au couple de points repérés par le numéro 1 dans le graphe. Cela correspond à une image de molaire maxillaire illustrée dans la figure 3.37. Cette dent présente ce qui s'appelle un faux moignon (*inlay-core*) surmonté d'une couronne, le tout métallique, c'est-à-dire que la dent a été dévitalisée et comblée entièrement par une pièce en alliage servant d'ancrage à la couronne, ce qui en fait l'un des cas les plus extrêmes en termes de ratio métal/dent. La segmentation est mise en échec par les multiples faux contours

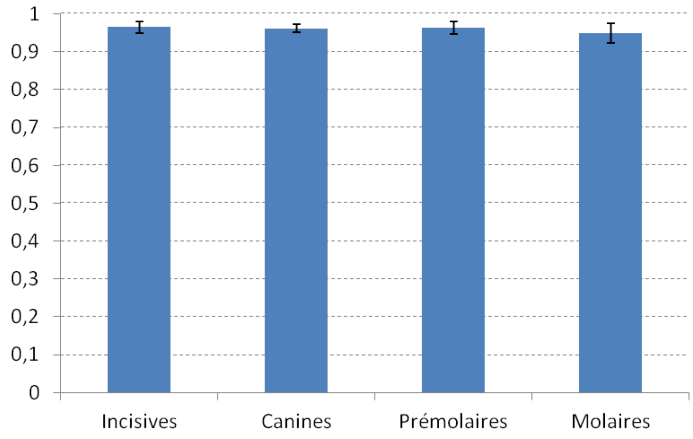


FIGURE 3.35 – Résultats du score de Dice moyen selon le type dentaire.

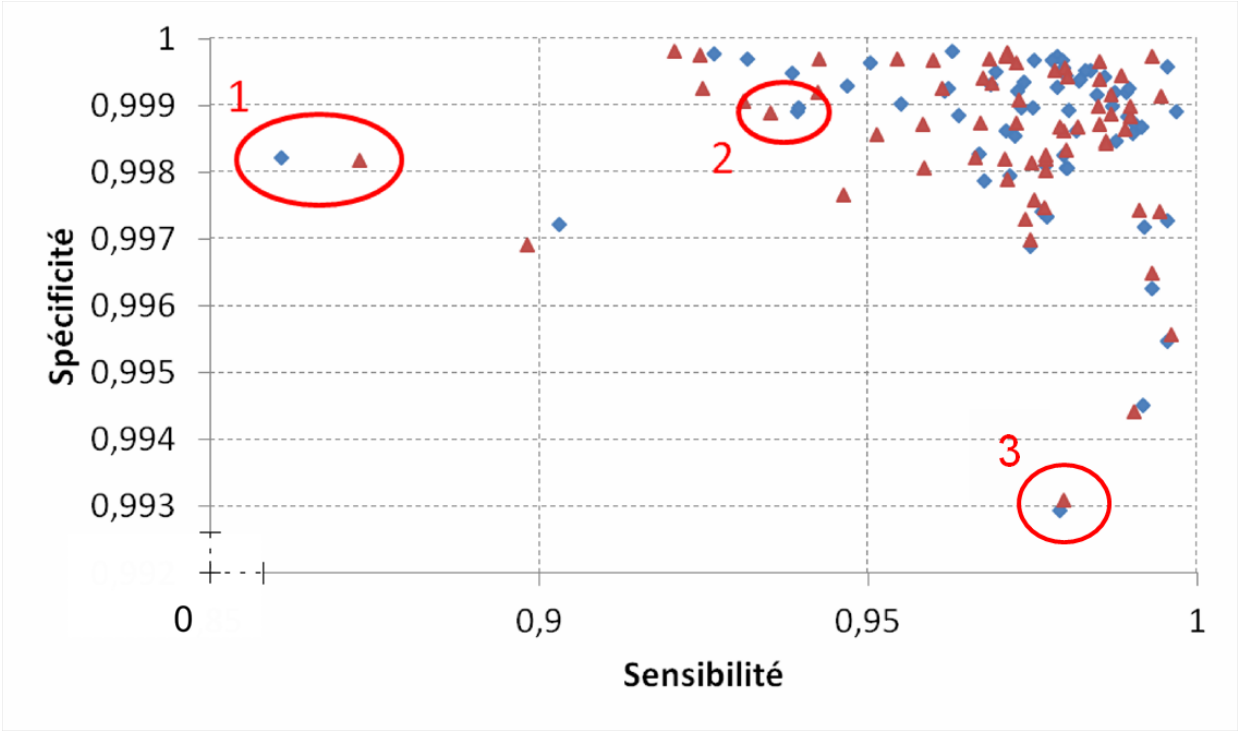


FIGURE 3.36 – Résultats de spécificité et de sensibilité pour chaque échantillon. Les couleurs différencient l'étape de segmentation : rouge pour la coupure de graphe et bleue pour le post-traitement.

induits dans l'image. Il manque de larges parties, principalement au niveau radiculaire, conduisant à une sensibilité de 0,873 pour ce cas et même 0,86 après post-traitement, due à la perte de bassins appartenant à la dent qui ne sont pas fusionnés avec le bassin central (lors de l'élimination des bassins de profondeur inférieure au seuil empirique de 0,15) à cause du fort relief induit par les faux contours. Dans ce cadre, il est important que l'utilisateur puisse avoir un retour et procéder à une correction manuelle de la segmentation, car s'il est possible de jouer sur les poids de l'équation, par exemple en favorisant la vraisemblance d'intensité, la récupération des parties manquantes se fera au détriment d'une mauvaise délimitation de la dent au niveau de l'os cortical. Une autre source de discordance potentielle est la qualité de la segmentation de référence, comme illustré par le cas numéro 2 du graphe, exposé en figure 3.38. Ici, la segmentation manuelle englobe toute une partie périphérique alors qu'il semble y avoir une délimitation avec le centre de la dent, là ou s'arrêtent nos résultats. La limite étant ténue et la qualité d'image limitée, il est difficile de savoir si cette partie externe fait réellement partie

de la dent ou bien s'il s'agit d'une gangue de calcification autour de la racine. Là encore, cela met en exergue le fait qu'un contrôle expert est nécessaire pour valider cliniquement tout résultat. Enfin, un dernier cas intéressant à étudier correspond au cas numéro 3, illustré dans la figure 3.39. Ici la proximité des sinus (espace aérien de très faible intensité) avec les racines de la molaire maxillaire conduit à une séparation entre la région du fond et de la dent au niveau de l'interface avec les sinus, ce qui intègre alors tout l'espace inter-radiculaire, expliquant la perte de spécificité observée.

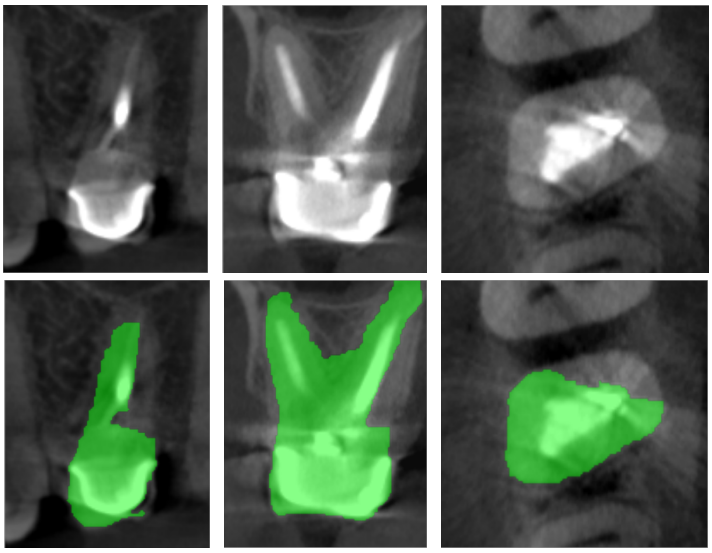


FIGURE 3.37 – Image correspondant à l'échantillon numéro 1 de la figure 3.36 et sa segmentation (sortie de la coupure de graphe). Remarquons la très forte dégradation induite par la présence de métal, notamment la présence de faux contours le long des raies hyper-intenses, qui met en échec la segmentation.

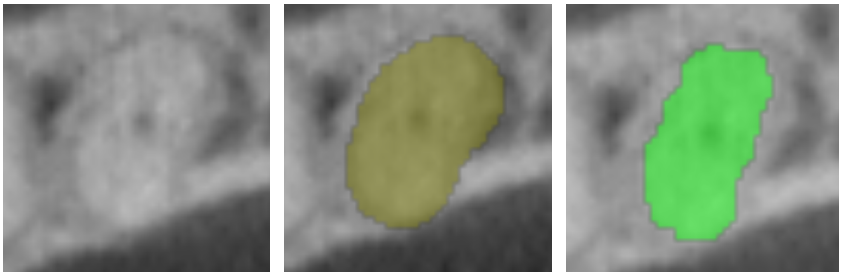


FIGURE 3.38 – Cas correspondant à l'échantillon numéro 2 de la figure 3.36 et sa segmentation (sortie de la coupure de graphe). Il est possible de voir que la segmentation référence (b) prend en compte toute une partie périphérique non retenue par la segmentation automatique (c).

3.5 Conclusion

Nous avons présenté dans ce chapitre une nouvelle méthode adaptée à la segmentation de la dentition dans les volumes patients acquis par CBCT, ce qui constitue une des trois contributions de cette thèse [49]. Cette dernière présente des résultats au niveau de l'état de l'art avec un score de Dice de $0,958$ pour des volumes de résolution isotropique de 150 microns. Elle repose sur l'utilisation d'un modèle statistique qui peut être obtenu même avec un faible nombre d'images annotées. La formulation permet de prendre en compte les modifications apportées par un utilisateur de manière aisée, et reste ouverte à une extension par des contraintes supplémentaires. La précision des résultats est jugée satisfaisante pour toutes les spécialités dentaires (hormis l'orthodontie, voir ci-après) et une compilation de résultats est disponible dans l'annexe B. Le code fonctionne sur des configurations

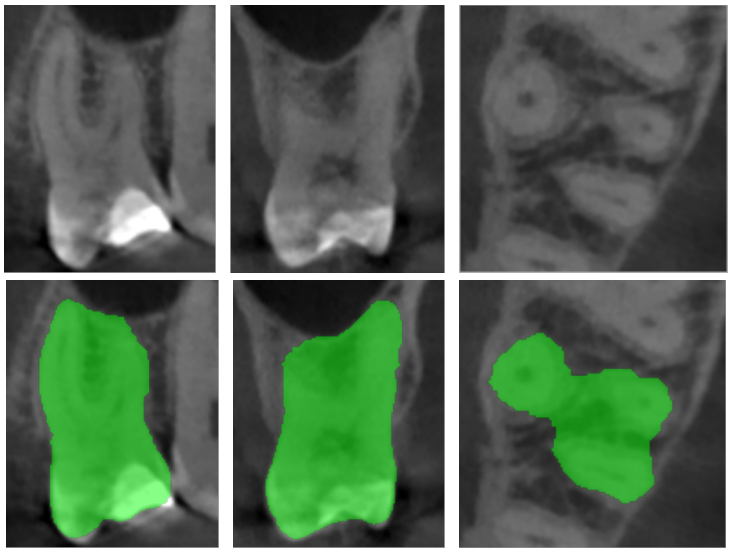


FIGURE 3.39 – Image correspondant à l'échantillon numéro 3 de la figure 3.36 et sa segmentation (sortie de la coupure de graphe). Remarquons le gradient très important induit par la proximité des sinus avec les racines de la molaire maxillaire qui conduit la segmentation à intégrer tout l'espace inter-radiculaire.

matérielles standard et est donc utilisable en cabinet. Nous avons cependant identifié des limitations importantes :

- Une sensibilité importante à la présence d'éléments métalliques, notamment lorsqu'ils induisent de faux gradients.
- Une baisse de la qualité de segmentation (coefficient de Dice de $0,91$, voir partie 3.4.3) lorsque la résolution des volumes est trop faible (supérieure à 200 microns, notamment utilisée en orthodontie), due à la disparition de la distinction entre os et dents dans les images.
- Une difficulté à traiter les cas de patients en occlusion due à l'invisibilité des interfaces entre couronnes opposées dans l'image.
- Une utilisation du modèle statistique qui entraîne une contrainte de forme reposant sur la forme moyenne, parfois inadaptée car trop éloignée des contours réels (principalement pour les molaires), ce qui réduit son apport (voir les bons résultats de l'approche powerwatershed sans contraintes de forme [40] dans la figure 3.26)
- Un temps de calcul cumulé pour la dentition complète encore important, plus de 10 minutes (en excluant le temps d'interaction de l'utilisateur).

Les pistes d'améliorations envisagées sont les suivantes :

- Rendre la contrainte de forme plus pertinente par l'addition de variables anatomo-morphologiques dans le modèle de forme, c'est-à-dire un modèle conditionnel [12].
- Réduire le temps de calcul en optimisant notre énergie avec l'algorithme powerwatershed.
- Réduire l'impact du métal par une utilisation d'algorithmes de réduction d'artéfacts métalliques (MAR).

Concernant le cas de l'occlusion et des volumes en grand champ, l'augmentation de la qualité des images est un facteur essentiel à l'amélioration des résultats. Il est aussi possible d'utiliser des informations provenant de modalités différentes (par exemple des scans de modèles en plâtre) pour contraindre la segmentation, comme présenté dans le chapitre 5. En parallèle, il peut être intéressant de réduire la quantité d'interactions demandée à l'utilisateur. Pour un cas sain, si l'utilisateur nous indique une dent et son numéro, nous devrions pouvoir déduire les positions des autres, et ainsi réduire le temps de sélection. Plus globalement, une information initiale sur un organe peut être pertinente pour les autres structures de la zone, ouvrant des possibilités de segmentations itératives ou multi-objets. Ces aspects seront discutés dans le chapitre suivant.

Chapitre 4

Vers une approche globale par un modèle structurel

Nous étudions maintenant la possibilité de procéder à une segmentation plus performante en tenant compte de plusieurs objets. Plus précisément nous étendons notre méthode de segmentation dentaire afin d'identifier toutes les dents présentes dans un champ de vue donné, pour tenter de résoudre les problèmes de fuites de segmentation par compétition. Puis, dans l'objectif de parvenir à une segmentation plus complète des volumes et des structures de la zone maxillo-faciale, nous proposons une approche globale de la segmentation. Afin d'appuyer le raisonnement, nous créons un modèle structurel des organes, qui regroupe les connaissances a priori que nous possédons sur les caractéristiques des éléments anatomiques. Ces informations sont de diverses natures : forme, apparence, ou encore organisation spatiale. Nous avons notamment défini les relations spatiales qui existent entre les organes, en nous appuyant sur la théorie des ensembles flous afin de prendre en compte l'aspect vague inhérent à la description de ces relations.

4.1 Extension multi-objets de la segmentation

4.1.1 Problématique

Une des faiblesses de la méthode présentée dans le chapitre 3 réside dans le fait que la segmentation d'un objet ne tient pas compte des résultats obtenus pour les autres objets. Or cette connaissance peut être pertinente pour améliorer la qualité de segmentation. Citons par exemple les erreurs de fuite de segmentation entre les couronnes dentaires qui pourraient être réduites si les dents adjacentes étaient connues. Deux approches complémentaires peuvent nous servir à réaliser cela. Dans un premier temps, lorsqu'un résultat est obtenu et validé (idéalement de manière interactive par le praticien), nous pouvons nous en servir comme d'une contrainte supplémentaire pour la segmentation suivante. Dans l'exemple des couronnes dentaires, nous pouvons forcer le résultat à ne pas se superposer au précédent (un exemple d'approche semblable est présenté dans le chapitre 5, section 5.3, à l'aide d'informations provenant d'un maillage de la surface dentaire). Dans un second temps, il est possible d'adapter l'algorithme de segmentation de manière à rechercher plusieurs objets, ce que nous présentons ci-après.

4.1.2 Nouvelle formulation de la segmentation

La nouvelle formulation de la segmentation s'appuie sur l'algorithme de coupure de graphe multi-étiquettes décrit dans [24, 81], nommé "alpha-expansion". Le choix s'est porté sur cet algorithme pour deux raisons : il permet d'obtenir une solution quasi-optimale en termes d'énergie (à un facteur connu près) au problème de minimisation à plusieurs étiquettes, contrairement à l'algorithme proche "alpha-beta swap" d'une part (tout en étant plus rapide car effectuant moins de calculs), et il est l'extension directe de l'algorithme utilisé pour notre segmentation mono-objet d'autre part. Nous espérons ainsi retrouver un comportement similaire entre les algorithmes mono-objet et multi-objets. Le principe général est de considérer chaque étiquette successivement face à l'ensemble restant et de résoudre une

coupure de graphe usuelle (avec deux nœuds terminaux, comme décrit dans le chapitre précédent). Il impose cependant des contraintes sur les termes utilisables. Plus précisément, pour toute énergie de type :

$$E(\mathcal{L}) = \sum_{v_i \in \mathcal{V}} E_{v_i}(\lambda_{v_i}) + \sum_{(v_i,v_j)\in\mathcal{E}} E_{v_i,v_j}(\lambda_{v_i}, \lambda_{v_j}),$$

où $\mathcal{L}$ représente une solution (partition), v_i, v_j sont des nœuds du graphe ($\mathcal{V}$ et $\mathcal{E}$ les ensembles des nœuds et des arêtes respectivement), et $\lambda_{v_i}, \lambda_{v_j}$ sont des étiquettes, à valeurs dans $\Lambda = \{0, ..., \lambda_A, ..., |\Lambda|\}$, il est nécessaire que les termes d'interaction binaire soient des métriques sur l'espace des étiquettes, c'est-à-dire satisfassent, pour toutes étiquettes $\lambda_A, \lambda_B, \lambda_C$ [24] :

$$E_{v_i,v_j}(\lambda_A, \lambda_B) = 0 \Leftrightarrow \lambda_A = \lambda_B, \tag{4.1}$$

$$E_{v_i,v_j}(\lambda_A, \lambda_B) = E_{v_i,v_j}(\lambda_B, \lambda_A) \geqslant 0, \tag{4.2}$$

$$E_{v_i,v_j}(\lambda_A, \lambda_B) \leqslant E_{v_i,v_j}(\lambda_A, \lambda_C) + E_{v_i,v_j}(\lambda_C, \lambda_B). \tag{4.3}$$

La fonction est alors dite “régulière”, condition nécessaire pour pouvoir être optimisée par coupure de graphe [81]. Si nous reprenons la formulation introduite dans le chapitre 3 :

$$E(\mathcal{L}) = \sum_{v_i \in \mathcal{V}} (\alpha R(\lambda_{v_i}) + \beta D(\lambda_{v_i})) + \sum_{(v_i,v_j)\in\mathcal{E}} (\gamma B(v_i, v_j) + \xi T(v_i, v_j)) \cdot \delta_{\lambda_{v_i} \neq \lambda_{v_j}},$$

$$\text{soit } E_{v_i}(\lambda_{v_i}) = \alpha R(\lambda_{v_i}) + \beta D(\lambda_{v_i}) \quad \text{et} \quad E_{v_i,v_j}(\lambda_{v_i}, \lambda_{v_j}) = (\gamma B(v_i, v_j) + \xi T(v_i, v_j)) \cdot \delta_{\lambda_{v_i} \neq \lambda_{v_j}},$$

nous pouvons voir que cette contrainte doit être respectée par le second terme. Pour rappel :

$$B(v_i, v_j) = \exp\left(-\frac{(\mathrm{Im}(v_i) - \mathrm{Im}(v_j))^2}{2\sigma^2}\right)\left(\frac{1}{\mathrm{dist}(v_i, v_j)}\right),$$

$$T(v_i, v_j) = |\phi\left(\frac{v_i + v_j}{2}\right)|.$$

Pour conserver T indépendant des étiquettes comme dans le cas mono-objet, nous avons modifié le terme en :

$$T(v_i, v_j) = \sum_{\lambda=0}^{|\Lambda|} |\phi_\lambda\left(\frac{v_i + v_j}{2}\right)|,$$

ce qui conserve la propriété de positivité (voir partie 3.3.1). Grâce à $\delta_{\lambda_{v_i} \neq \lambda_{v_j}}$, $E(\lambda_{v_i}, \lambda_{v_j})$ est égal à 0 si et seulement si $\lambda_{v_i} = \lambda_{v_j}$, et l'équation 4.1 est respectée. Le terme $E_{v_i,v_j}(\lambda_{v_i}, \lambda_{v_j})$ est positif mais aussi symétrique, de manière immédiate. L'équation 4.2 est donc respectée.
La dernière condition, l'inégalité triangulaire, impose que dans l'espace des étiquettes, il soit toujours moins coûteux (en termes d'énergie) de passer d'une étiquette λ_A à une étiquette λ_B directement, plutôt que de passer par une étiquette tierce, pour un couple de sommet donné [24]. En ré-écrivant le terme avec des étiquettes arbitraires pour un couple fixe de sommets tel que :

$$E_{v_i,v_j}(\lambda_A, \lambda_B) = \mathcal{K}_{v_i,v_j} \cdot \delta_{\lambda_A \neq \lambda_B},$$

où $\mathcal{K}_{v_i,v_j}$ est une constante ne dépendant pas des étiquettes pour un couple (v_i, v_j) donné tel que

$$\mathcal{K}_{v_i,v_j} = (\gamma B(v_i, v_j) + \xi T(v_i, v_j)),$$

nous devons alors vérifier que

$$E_{v_i,v_j}(\lambda_A, \lambda_B) \leqslant E_{v_i,v_j}(\lambda_A, \lambda_C) + E_{v_i,v_j}(\lambda_C, \lambda_B)$$

soit

$$\mathcal{K}_{v_i,v_j} \cdot \delta_{\lambda_A \neq \lambda_B} \leqslant \mathcal{K}_{v_i,v_j} \cdot (\delta_{\lambda_A \neq \lambda_C} + \delta_{\lambda_C \neq \lambda_B}),$$

Vérifions maintenant tous les cas :

— si $\lambda_A = \lambda_B$, alors le membre gauche est égal à 0,
— si $\lambda_A \neq \lambda_B$, alors le membre gauche est égal à $\mathcal{K}_{v_i,v_j}$,
— si $\lambda_A = \lambda_C$ ou $\lambda_B = \lambda_C$, le membre de droite est égal à $\mathcal{K}_{v_i,v_j}$,
— si les trois étiquettes sont différentes, alors $\mathcal{K}_{v_i,v_j} \leqslant 2\mathcal{K}_{v_i,v_j}$.

L'inégalité triangulaire est donc satisfaite, et $E(\lambda_{v_i}, \lambda_{v_j})$ est donc une métrique sur l'espace des étiquettes, ce qui implique finalement que notre formulation peut être optimisée par alpha-expansion.

D'un point de vue applicatif, la signification est légèrement modifiée par rapport au cas d'un objet unique, puisque la modification de T revient à faciliter la coupure du graphe si le point à mi-chemin entre v_i et v_j est un point de coïncidence des contours de l'ensemble des formes moyennes, et non plus seulement de la forme moyenne de l'objet considéré. L'idée sous-jacente est de profiter de l'accumulation des cartes pour repérer les frontières entre objets en corrigeant un éventuel mauvais recalage d'une partie des formes. Notons que cela est pertinent car les dents sont proches voir adjacentes, donc les contours peuvent se superposer. Dans le cas d'une recherche d'objets distants les uns des autres, la pertinence du terme reste à démontrer. Du reste, les dents sont pré-traitées de manière identique (extraction de zones d'intérêt puis recalage), seuls les éléments de la segmentation diffèrent. À partir des modèles moyens de forme recalés pour chaque dent, la carte des marqueurs est obtenue en utilisant la réunion des masques des objets.
L'équation 3.3 devient :

$$\text{Masque}(x) = \begin{cases} 1 & \text{si} \quad \forall l_\lambda > 0 \ \exists\, i : S^{\pm}_{l_\lambda,i}(x) \neq \mathcal{H}(-\bar{\mathbf{M}}_{\mathbf{l}_\lambda}(x)), \\ 0 & \text{sinon,} \end{cases}$$

et l'équation 3.4 devient :

$$\text{Marqueurs}(x) = \begin{cases} -1 & \text{si} \quad \text{Masque}(x) = 1, \\ 0 & \text{si} \quad \text{Masque}(x) = 0 \text{ et } \nexists l_\lambda > 0 : \mathcal{H}(-\bar{\mathbf{M}}_{\mathbf{l}_\lambda}(x)) = 1, \\ \lambda & \text{si} \quad \text{Masque}(x) = 0 \text{ et } \exists l_\lambda > 0 : \mathcal{H}(-\bar{\mathbf{M}}_{\mathbf{l}_\lambda}(x)) = 1. \end{cases}$$

Cela revient à dire que nous considérons comme zone inconnue (-1) pour la segmentation toute partie de l'espace qui peut être atteinte par la déformation d'une des dents impliquées, quelle qu'elle soit (représenté par le masque). Tout l'espace autour (c'est-à-dire à l'extérieur) est considéré comme appartenant au fond (0 dans les marqueurs) et toute partie de valeur nulle incluse dans ce masque est attribué à l'objet correspondant (λ). Sont ensuite ajoutés les marqueurs déduits de la recherche du plus court chemin pour chaque dent avec la valeur λ correspondante. L'information a priori de distance repose maintenant sur une collection de cartes ϕ_λ, celle du fond (ϕ_0) étant obtenue comme la transformée en distance du complémentaire de la réunion des formes moyennes recalées et binarisées des objets, soit :

$$\phi_0 = T_{dist}\left((\mathcal{H}(-\bar{\mathbf{M}}_{\mathbf{1}}(x)) \cup ... \cup \mathcal{H}(-\bar{\mathbf{M}}_{\mathbf{l}_\lambda}(x)))^C \right)$$

Cette approche nous assure que la solution trouvée, $\mathcal{L}^*$, a une énergie bornée par deux fois celle de la solution optimale $\widehat{\mathcal{L}}$ (preuve dans [24]), c'est-à-dire :

$$E(\mathcal{L}^*) \leqslant 2E(\widehat{\mathcal{L}}),$$

nous permettant d'obtenir une solution proche de la solution optimale en termes d'énergie, donc attendue comme satisfaisante en termes de domaine spatial. L'algorithme de l'alpha-expansion [24] est le suivant :

```
1. Initialisation avec une partition arbitraire  L
2. Réduction = 0
3. Pour chaque étiquette  λ_A ∈ Λ
   3.1. Trouver la solution  L̂  qui minimise  E(L̂)  par  λ_A-expansion
   3.2. Si  E(L̂) < E(L)  :
        Réduction = 1
        L = L̂
4. Si Réduction = 1 alors reprendre à 2.
5. Fin
```

4.1.3 Expérimentation

Afin d'étudier la performance de cette méthode, nous utilisons la base d'images de la partie 3.4.4, soit deux dentures adultes complètes (64 dents). Au lieu de restreindre la zone d'intérêt à une seule dent lors de la segmentation, nous étendons le champ de vue pour contenir les deux dents adjacentes, soit des zones contenant jusqu'à trois dents. L'analyse complète de la denture en une seule fois n'a pas été réalisée car les résultats présentés par l'analyse sur les zones d'intérêt étendues montrent déjà des limitations importantes en termes de qualité de segmentation et de ressources de calcul. Les paramètres sont identiques à ceux de la version mono-objet, soit :

— un nombre maximal d'itérations du recalage égal à 30,

— un pas minimal et une tolérance au gradient de 0,001 durant l'optimisation par descente de gradient,

— une constante de déformation C égale à 3,

— une pondération des termes de l'énergie par $\sigma = 100$, $\alpha = 1$, $\beta = 5$, $\gamma = 15$ et $\xi = 1$.

Le recalage n'étant pas modifié, les paramètres ont la même influence que dans le cas de l'objet unique. Quant aux paramètres de la segmentation, ils sont identiques à ceux de la section 3.4.4 pour la comparaison. Dans cette configuration, le calcul de la segmentation atteint une consommation de 6 GB, pour des temps compris entre 150 et 380 secondes (moyenne de 267 ± 92), le temps augmentant proportionnellement à la taille de la zone d'intérêt, donc au type de dent considéré dans le sens antéro-postérieur. Cela représente jusqu'à 8 minutes pour traiter 3 molaires, ce qui est très nettement supérieur au temps cumulé de segmentation individuelle (environ 1,5 minutes). Cela peut s'expliquer par l'algorithme de l'alpha-expansion (voir section 4.1.2), car la moindre réduction d'énergie entraînée par l'expansion d'une étiquette conduit à calculer une nouvelle passe (c'est-à-dire à ré-étudier toutes les étiquettes), ce qui augmente le temps de calcul notamment lorsque la solution tend à se stabiliser ou lorsque seule l'expansion d'une étiquette produit des résultats.

Les résultats sont présentés dans la table 4.1 et les figures 4.1 et 4.2. Les scores indiqués pour chaque échantillon sont calculés à partir de la matrice de confusion où sont cumulées les valeurs pour chaque objet présent dans la zone d'intérêt. Les scores moyens obtenus sont moins bons que lors de l'analyse individuelle, avec un Dice moyen de $0,901$ (table 4.1), soit un recul de $0,052$ (différence statistiquement significative avec un intervalle de confiance de 0,1% selon un test de Student). En regardant les valeurs pour chaque échantillon, nous constatons que les scores sont en recul et ce de manière homogène (nous n'observons pas de cas extrêmes). L'analyse des cas les moins performants (voir figure 4.3) indique que malgré une formulation similaire à celle du cas individuel, la méthode peut conduire à des erreurs de segmentation : la solution optimisée correspond à un minimum local et non pas global pour la fonction d'énergie, et cet écart est significatif dans l'image. Toutefois cette méthode est efficace pour individualiser les dents (c'est-à-dire pour marquer leurs points de contact), et les erreurs de segmentation telles que les fuites sur les couronnes adjacentes sont corrigées. Des exemples de résultats sont présentés dans la section annexe B.2.

	Coupure de graphe multi-objets	Coupure de graphe mono-objet
Sensibilité	0,904 ± 0,046	0,969 ± 0,023
Spécificité	0,995 ± 0,004	0,998 ± 0,001
Précision	0,990 ± 0,004	0,998 ± 0,001
Dice	0,901 ± 0,025	0,953 ± 0,024

TABLE 4.1 – Résultats moyens sur l'ensemble de la base pour la sensibilité, la spécificité, la précision et le coefficient de Dice (voir la section 3.4.1 pour la discussion sur la pertinence de la spécificité et de la précision).

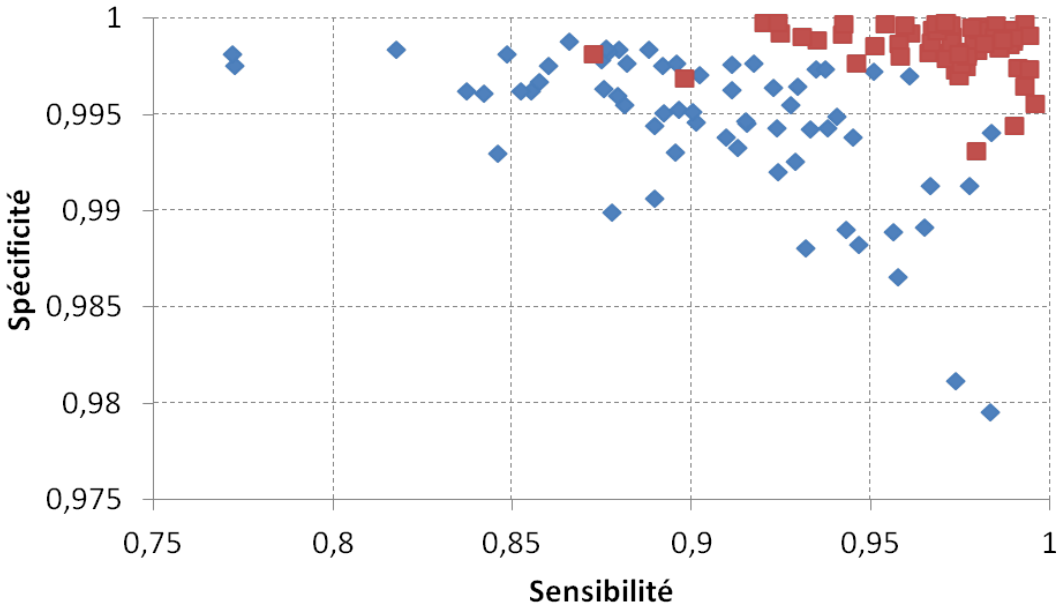


FIGURE 4.1 – Résultats de la segmentations multi-objets en termes de sensibilité et de spécificité. Le nuage des résultats multi-objets (bleu) est nettement en retrait comparativement au nuage des résultats mono-objet (rouge).

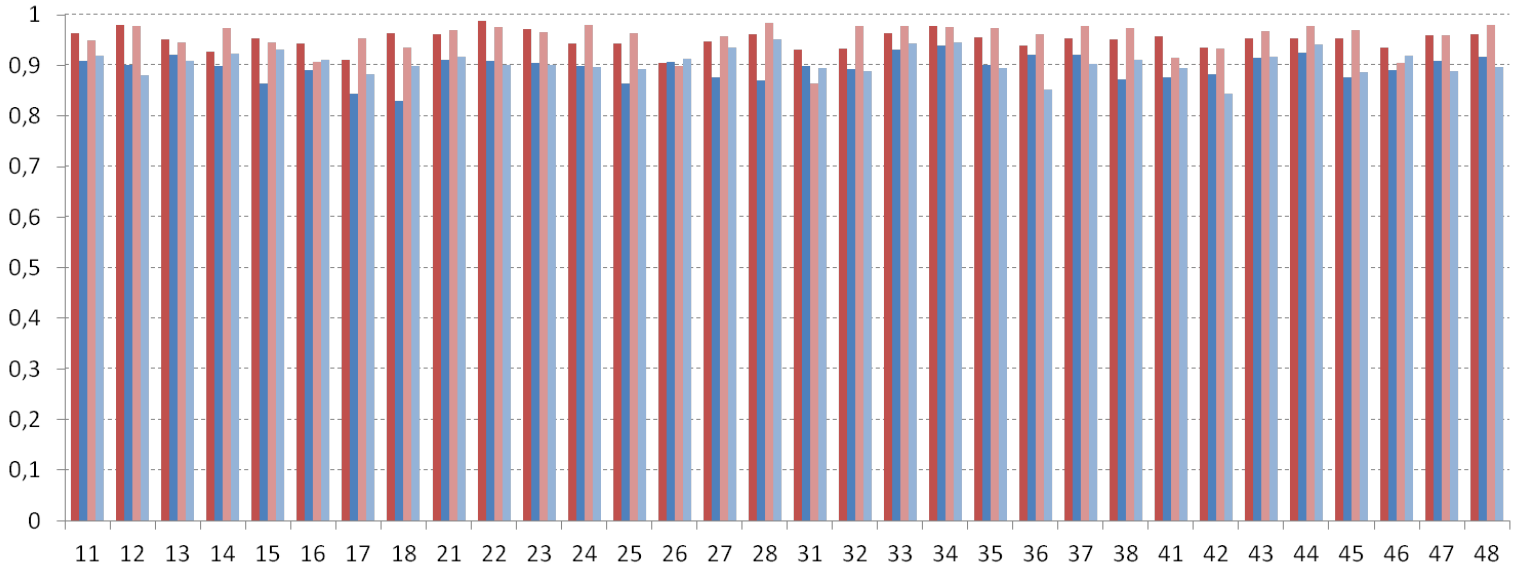
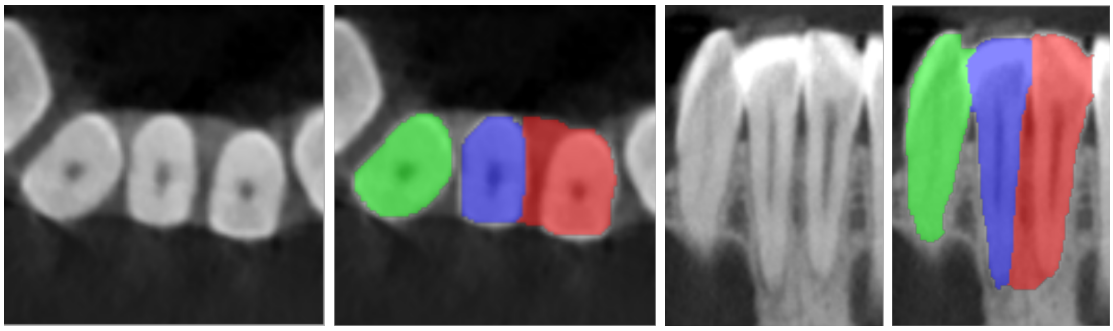


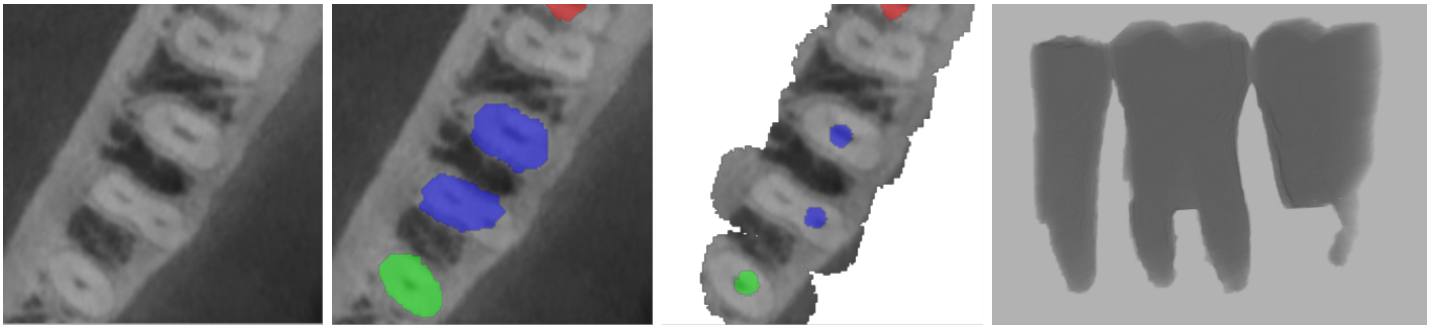
FIGURE 4.2 – Résultats du score de Dice pour chaque échantillon, en rouge pour le cas mono-objet et en bleu pour le cas multi-objets. La différence de ton d'une même couleur indique un patient différent. Il est possible de constater une régression de la qualité de segmentation pour la quasi-totalité des échantillons lors du passage à l'algorithme multi-objets.

4.1.4 Discussion

Les expériences font apparaître qu'en l'état l'extension multi-objets ne correspond pas à nos attentes d'une part parce que la précision de l'identification est en régression par rapport au cas individuel mais surtout cette modification s'accompagne d'une augmentation substantielle des ressources de calcul, notamment du temps. Nous estimons que si l'individualisation des dents est améliorée, du fait de la concurrence des segmentations, la plus-value n'est pas suffisante pour compenser la perte de pré-



(a)



(b)

FIGURE 4.3 – Illustration des cas erronés produits par la segmentation multi-objets. (a) Lorsque les racines des dents sont peu distantes, l'espace interdentaire est attribué à une des dents adjacentes par l'optimisation (cas de faible spécificité). (b) La solution sous-optimale trouvée conduit à ne pas segmenter toute la partie dentaire de la dent centrale malgré les marqueurs (3^{e} vignette). Le marqueur de la molaire en rouge est erroné : le plus court chemin du point de la racine mésiale est passé par la racine distale. Ce cas conduit à une chute de la sensibilité.

cision au niveau radiculaire et le surcoût temporel : une correction manuelle et interactive satisferait d'avantage le praticien qui aurait la certitude du résultat, à faire dent par dent, pour un coût temporel faible. L'intérêt augmente dans le cadre d'un processus automatique pour identifier l'ensemble des organes présents dans l'image, mais les travaux pour définir la méthodologie d'exploration (voir partie 4.4.1) ne sont pas suffisamment avancés à l'heure actuelle.

4.2 Modèle structurel

4.2.1 Motivation

La motivation principale pour la création d'un modèle structurel est de rassembler et d'exploiter toutes les informations que nous possédons à propos d'un contexte donné. Cet ensemble permet de lever certains obstacles à l'analyse des images. Une structure difficilement accessible à partir de l'information d'intensité peut être identifiée par la prise en compte de son environnement, de sa position ou de sa morphologie. Le modèle facilite la reconnaissance des structures dans les images en contraignant le problème.

4.2.2 Formalisation graphique et attributs

Structure générale et concepts

Nous avons choisi de représenter notre modèle sous la forme d'un graphe conceptuel, car ce formalisme permet de modéliser facilement les caractéristiques des structures, tout en présentant un aspect graphique qui rend sa compréhension aisée. Un extrait de notre modèle est présenté en figure 4.4. Cette

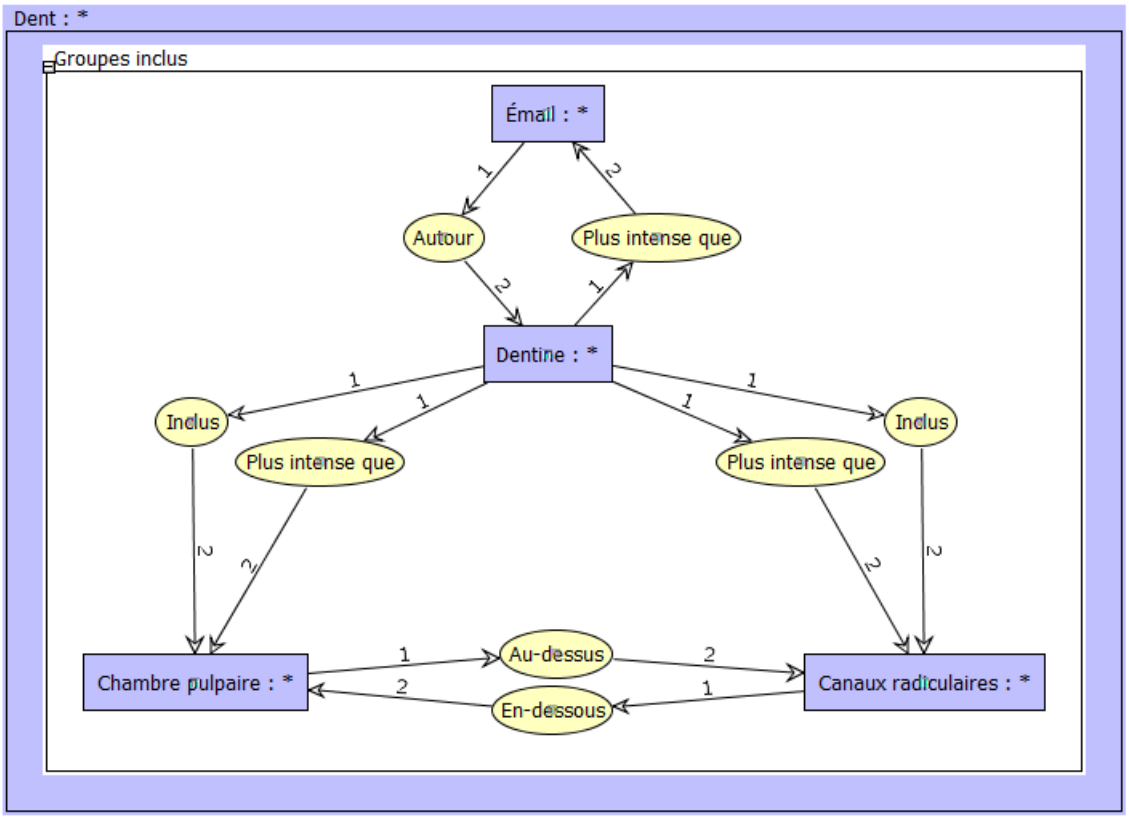


FIGURE 4.4 – Représentation graphique du modèle d'une dent. Les structures sont représentées par les formes rectangulaires tandis que les relations sont de forme ovoïde. Capture d'écran depuis l'application CoGui.

partie représente le concept d'une dent. Celle-ci est composée de deux types de sous-concepts : ceux, de forme rectangulaire, qui représentent les structures qui la composent (au nombre de quatre, l'émail, la dentine, la chambre pulpaire et les canaux radiculaires) ; et ceux, de forme ovale, qui traduisent une interaction entre les organes, et qui représentent une information a priori (leur modélisation est abordée à partir de la partie suivante). La relation peut être symétrique (par exemple "proche"), ce qui est représenté par des flèches à double sens, ou bien asymétrique ("plus intense que") et le sens de la flèche indique l'ordre d'interprétation. Par exemple, l'émail entoure la dentine et est plus intense. Les attributs propres à un organe particulier ne sont pas représentés graphiquement (texture, homogénéité, intensité moyenne, etc). Les concepts peuvent être emboîtés (ou inclus). Ainsi le concept d'arche dentaire comprend plusieurs concepts de dent (chaque numéro dentaire fait l'objet d'un concept spécifique). Le modèle complet est joint dans les annexes D (description des structures) et E (forme graphique) et est disponible sous la forme d'un fichier XML compatible avec la bibliothèque Cogitant (http ://cogitant.sourceforge.net), qui permet la manipulation et le raisonnement avec des graphes conceptuels.

Apparence

L'apparence d'un objet est caractérisée principalement par son intensité et sa texture (ou son absence). Concernant l'intensité, la valeur la plus triviale est la moyenne. Idéalement cet attribut est donc une valeur moyenne connue à l'avance. Comme expliqué lors de la description du modèle, le contexte ne nous permet pas d'indiquer une valeur d'intensité selon une échelle de référence. Nous avons donc créé une échelle qualitative divisée en quatre classes (faible, moyenne, forte et hyper-intense), en fonction des grandes catégories retrouvées dans les histogrammes (l'air, les tissus mous, les tissus durs et l'émail/le métal respectivement). Pour obtenir les valeurs dans un cas similaire au nôtre, le procédé le plus direct est d'utiliser des données annotées ou seuillées selon ces catégories, comme nous l'avons

fait pour les dents, obtenant ainsi des fonctions de vraisemblance. Cela conduit donc à rendre le modèle dépendant des données utilisées. Si le modèle est destiné à être utilisé sur des jeux de données avec des échelles variables (images provenant de différents constructeurs par exemple), il est important de normaliser les images, ou bien d'exprimer la valeur de l'intensité comme un pourcentage de l'échelle. Il est possible d'utiliser d'autres valeurs caractéristiques selon leur pertinence et selon l'importance de l'intensité dans le modèle : citons par exemple, l'écart-type (ou la variance), la médiane, ou encore le centre d'inertie (qui donne une information sommaire sur la répartition spatiale de l'intensité). La correspondance de nos classes sur nos données (voir partie 3.1) est indiquée dans la table 4.2. Notons que nous disposons, pour les dents uniquement, des distributions d'intensités qui sont plus précises que ces valeurs générales.

Classe	Intensité moyenne sur nos données
Faible	36 ± 94
Moyenne	964 ± 159
Forte	1575 ± 236
Hyper-intense	3483 ± 1002

TABLE 4.2 – Table de correspondance entre les classes qualitatives d'intensité et leur valeur dans nos données. Pour rappel, les données issues des scanners Carestream sont encodées sur 16 bits, soit des valeurs d'intensités comprises entre 0 et 65535. L'impact de la présence de métal dans la dernière classe entraîne une variabilité très importante.

La formalisation de la texture est moins immédiate. Plusieurs méthodes sont possibles pour caractériser la variation spatiale d'intensité et chacune fournit un ensemble différent de valeurs caractéristiques. Par exemple, pour étudier la présence d'une texture dans le canal mandibulaire, nous avons choisi l'approche par matrice de co-occurrence et l'analyse en ondelettes (voir annexe C).

Forme

Les attributs de forme les plus simples peuvent être par exemple :

— la taille de la boîte englobante,

— l'aire de la surface,

— le centre de gravité,

— les caractéristiques de l'ellipsoïde englobant, ou les axes principaux,

— la topologie,

— l'élongation et/ou la compacité.

Enfin, la forme peut être modélisée implicitement, comme nous l'avons fait pour les dents (voir chapitre 3), par la construction d'un modèle statistique associé. L'utilisation d'un tel attribut repose toutefois sur l'hypothèse que la forme de l'élément suit une loi normale autour d'une forme moyenne, ce qui est généralement le cas pour les organes, mais elle peut être incorrecte : par exemple, il n'est pas certain que les formes prises par les arches dentaires suivent ce schéma. Dans le cadre de l'imagerie médicale, il est possible d'utiliser en sus les informations fournies par l'anatomie pour inclure des critères anatomo-morphologiques moyens liés à l'organe concerné, par exemple

— le nombre de racines, de cuspides, de canaux dentaires,

— l'angulation de la mandibule, l'espacement entre les branches montantes,

— le volume et le diamètre des voies aériennes,

— etc.

Il est possible par ailleurs d'intégrer ces critères pour conditionner le modèle statistique, afin d'augmenter la pertinence des formes obtenues [12].

Apport dans la méthode de segmentation proposée

Les informations issues de ce modèle sont intégrées à plusieurs niveaux dans la méthode de segmentation proposée :

— les densités de probabilités de l'intensité de chaque classe de dent permettent de définir le terme de vraisemblance,

— les marqueurs internes initiaux sont obtenus en utilisant la faible valeur d'intensité des canaux radiculaires et leur position centrale dans chaque racine (lorsqu'ils sont visibles),

— la forme particulière des dents traduite dans le modèle statistique permet la délimitation d'un espace contraint pour la segmentation ainsi qu'une contrainte lors de l'optimisation.

Perspectives

Les autres attributs n'ont pas été utilisés pour le moment, mais voici ceux qui ont été définis comme pertinents lors de la construction du modèle :

— la position : absolue ou relative, elle peut se référer à un espace physique fixé par rapport au patient, ou à une partition de l'image,

— relations d'intensité : il est possible de définir une contrainte à partir d'un contraste entre les structures,

— relations spatiales directionnelles : une définition morphologique intéressante en termes d'équivalence avec une image est proposée dans [19],

— relations spatiales de distances [15],

— inclusion,

— adjacence (voir l'approche morphologique dans [15]),

— encerclement ("autour de", en utilisant la définition dans [121]),

— ancrage (que nous conceptualisons comme le fait d'être en partie proche et en partie en recouvrement).

4.3 Extension de deux relations spatiales en 3D

L'étude de l'anatomie de la région maxillo-faciale a montré que la relation "le long de" et la relation d'alignement sont pertinentes pour décrire et reconnaître le canal mandibulaire et l'arche dentaire. Une partie de la thèse a donc consisté à étendre en 3D des définitions existantes. Nous présentons en premier lieu nos travaux sur la relation "le long de", qui ont fait l'objet d'une publication [48], puis nous décrirons l'extension de la relation d'alignement.

4.3.1 Relation "le long de"

La modélisation de cette relation en deux dimensions a été réalisée dans [118], et nous proposons ici une version étendue au domaine tri-dimensionnel.

Nous considérons que deux objets A et B sont le long l'un de l'autre dans le cas où au moins l'un des deux est allongé et si A et B sont côte à côte par rapport à la direction d'allongement. "Côte à côte" implique que les deux objets sont proches l'un de l'autre par rapport à leur environnement. Une telle définition conduit à une forme particulière pour l'espace se trouvant entre A et B, qui est de fait allongé aussi. Mesurer l'élongation de cette région permet d'évaluer le lien spatial, comme indiqué dans [118].

Cette approche nécessite d'avoir défini au préalable la notion "entre" pour modéliser l'espace entre deux objets. Un avantage de cette définition est que la relation est symétrique, en considérant les deux objets de manière équivalente. De plus, elle donne la possibilité d'avoir un degré de satisfaction, plutôt qu'une réponse binaire à la question : sont-ils le long l'un de l'autre ? Nous pouvons de plus choisir par cette approche de n'étudier la relation qu'entre certaines parties des objets, selon la distance entre elles. La méthode globale pour définir le degré de satisfaction comporte ainsi trois étapes [118] :

1. Calculer la région inter-objets β.
2. Définir l'élongation de β.
3. Calculer le degré de satisfaction de la relation.

Région inter-objets

Il existe plusieurs méthodes dans la littérature permettant de calculer la région β entre deux objets compacts A et B [18]. Nous avons choisi de reprendre la méthode dite de la "visibilité" [18] qui permet de gérer les cas difficiles comme les concavités. Cette méthode utilise les segments admissibles. Un segment $]x, y[$ est dit admissible si x et y appartiennent à A et B respectivement, et si le segment est inclus dans $A^C \cap B^C$ où ^{C} désigne le complémentaire de l'objet (remarquons que cela nécessite que x et y soient sur le contour des objets). La région β entre A et B est définie par :

$$\beta(A, B) = \cup\{\ ]x, y[\ |\ x \in A, y \in B,\ ]x, y[\ \subseteq\ A^C \cap B^C\}$$

Cette définition ne fait aucune hypothèse sur la dimension de l'espace, et peut être utilisée directement en 3D. Les deux limites principales de cette approche concernent les cas où un objet est beaucoup plus étendu que le second (se référer à [18] pour une définition adaptée), et les cas où la relation est étudiée entre des ensembles d'objets non connexes (par exemple entre deux rangées de dents).

Mesure de l'élongation

Une définition commune de l'élongation est de considérer l'inverse de la compacité [93] comme :

$$e_{3D}(\beta) = \frac{S^3(\beta)}{V^2(\beta)}$$

avec S et V désignant l'aire et le volume de β respectivement. Cette quantité est sans dimension, admet un minimum égal à 36π pour les formes sphériques et augmente pour les formes plus allongées. Cette définition est une bonne base car elle donne une mesure invariante par translation, par rotation et par homothétie. Nous définissons à partir de cette dernière un degré d'élongation $\alpha_1 = f_a\left(e_{3D}(\beta)\right)$ où f_a est une fonction croissante, par exemple $\frac{1-e^{-ax}}{1+e^{-ax}}$, avec $\alpha \in [0, 1]$. Normaliser le degré est important pour pouvoir comparer les différentes mesures d'élongation proposées dans la suite du document. Une autre manière de définir le degré est de normaliser e_{3D} par la valeur pour une sphère. Le degré est alors dans $[1, +\infty[$.

Cependant, la région β peut être allongée sans que A et B ne soient le long l'un de l'autre comme représenté dans la figure 4.5. Une contrainte supplémentaire est que β doit être allongé dans la même direction que les objets, ce qui peut être caractérisé par la surface de contact entre $A \cup B$ et β [118]. Toutefois, cela n'est pas suffisant en 3D, car si nous ne considérons que l'aire totale de β en contact avec les objets, il est impossible de distinguer entre un cas où deux pavés dont les côtés les plus longs se font face (cas (a) de la figure 4.6) et un cas où deux cylindres dont les bases se font face (cas (b) de la figure 4.6) si l'aire totale des surfaces en vis-à-vis est la même dans les deux cas.

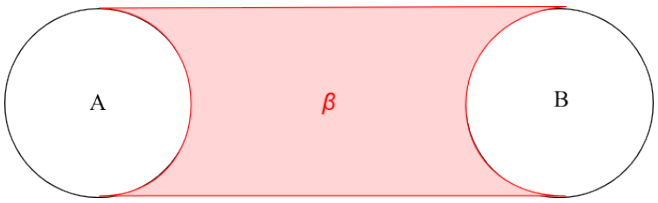


FIGURE 4.5 – Cas où la définition e_{3D} ne convient pas. La région β est allongée alors que ni A ni B ne le sont : ils ne sont pas le long l'un de l'autre.

Ce qui apparaît ici est le fait que même en 3D, être "le long de" implique que la région inter-objets doit être allongée dans une seule direction. Nous proposons donc une nouvelle définition afin d'inclure

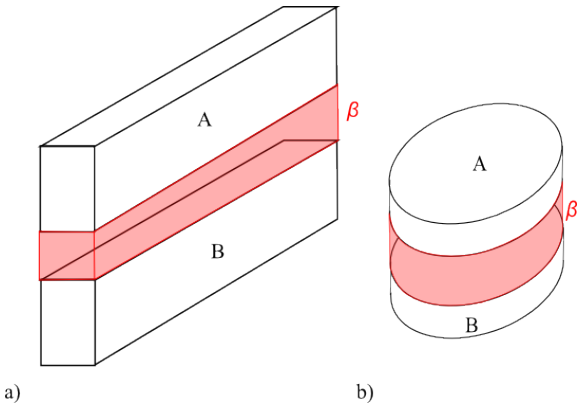


FIGURE 4.6 – Cas où l'aire des surfaces de contact n'est pas discriminante. La forme de cette interface doit être prise en compte.

cette contrainte :

$$\alpha_2 = f_a\left(\frac{(\gamma_A S_A^c(\beta) + \gamma_B S_B^c(\beta))^3}{V(\beta)^2}\right) \qquad \text{avec} \qquad \gamma_X = 1 - \frac{2\sqrt{S_X^c(\beta)\pi}}{P_X^c(\beta)} \tag{4.4}$$

où $S_X^c(\beta)$ est l'aire de la frontière entre β et l'objet X, et $P_X^c(\beta)$ le périmètre de celle-ci. Le terme γ_X est une mesure de l'élongation de la surface de contact. Lorsque celle-ci tend vers une forme circulaire, γ_X tend vers zéro car le ratio tend vers 1 (égal à 1 dans le cas d'un cercle plat parfait, ce qui est la forme la moins allongée possible), et augmente vers 1 si la forme est plus allongée, le ratio tendant vers 0. Cette définition permet de prendre en compte la forme de chaque objet séparément.

Dans certains cas plus spécifiques, tels que ceux exposés dans la figure 4.7, la distance entre les deux objets peut avoir une importance : la première configuration est plus proche de notre représentation intuitive de la relation que la seconde, mais le degré α_2 ne pourra pas les différencier si l'aire et le volume de β sont identiques. Nous incluons alors la distance avec la fonction $D_{AB}(z)$, ce qui est proposé dans [118] :

$$\nu(\beta) = \int_\beta D_{AB}(z)dz \quad \text{avec} \quad D_{AB}(z) = \text{dist}(z, A) + \text{dist}(z, B) \tag{4.5}$$

où $\text{dist}(z, A)$ et $\text{dist}(z, B)$ sont les distances d'un point de l'espace z à A et B respectivement, et ν est l'intégrale de la distance sur β. Nous définissons un nouveau degré qui inclut la distance comme :

$$\alpha_3 = f_a\left(\frac{(\gamma_A S_A^c(\beta) + \gamma_B S_B^c(\beta))^2}{\nu(\beta)}\right) \tag{4.6}$$

Notons que les exposants ont été modifiés afin de conserver l'adimensionnalité de la mesure. Un inconvénient de cette formulation est la perte de l'invariance au changement d'échelle provenant du fait que la mesure de distance dans la définition n'est pas normalisée. Cette invariance peut toutefois être retrouvée en normalisant la distance par sa valeur maximale.

Il est possible d'aller plus loin avec cette distance pour n'examiner la relation qu'entre les parties d'objets qui sont proches l'une de l'autre, ce qui peut être utile selon la forme des objets. Nous restreignons la région β en la seuillant à une distance d afin de ne garder que les points qui ont une valeur inférieure à d, ce qui définit β_d. Ce seuillage peut créer des trous, la cohérence doit alors être restaurée soit en les remplissant pour chaque composante connexe restante de β_d, soit en recherchant un seuil optimal autour du seuil initial qui ne produit que des composantes connexes pleines (lorsque ce seuil existe). Les degrés α_4 et α_5 sont alors définis en remplaçant β par β_d dans les équations (4.4) et (4.6), soit :

$$\alpha_4 = f_a\left(\frac{(\gamma_A S_A^c(\beta_d) + \gamma_B S_B^c(\beta_d))^3}{V(\beta_d)^2}\right)$$

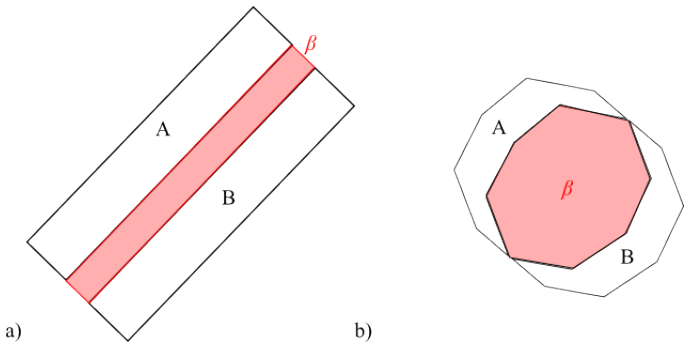


FIGURE 4.7 – Cas où la distance apporte une information supplémentaire sur la configuration des objets.

$$\alpha_5 = f_a\left(\frac{(\gamma_A S_A^c(\beta_d) + \gamma_B S_B^c(\beta_d))^2}{\nu(\beta_d)}\right)$$

Extension aux objets flous

L'utilisation du flou est intéressante car elle permet d'une part de rendre la mesure plus robuste, notamment lors de l'étude entre des objets définis imprécisément comme cela est fréquemment le cas dans des processus de segmentation automatiques en imagerie médicale, mais aussi parce qu'elle permet de prendre en compte la nature intrinsèquement vague de la relation.

L'inclusion du segment dans le complémentaire des objets flous est maintenant modélisée comme un degré d'inclusion [15], qui peut être exprimé par :

$$\mu_{\text{inc}}(]x,y[) = \inf_{z\in]x,y[} \min[1-\mu_A(z), 1-\mu_B(z)]$$

où μ_A et μ_B sont les fonctions d'appartenance des deux ensembles flous (par la suite nous assimilons un ensemble flou à sa fonction d'appartenance pour plus de simplicité). La région β est alors l'ensemble flou défini par [118] :

$$\mu_\beta(z) = \sup\{\mu_{inc}(]x,y[) \mid x \in \text{Supp}(\mu_A), y \in \text{Supp}(\mu_B), z \in]x,y[\}.$$

Cette définition a pour conséquence que le degré d'inclusion est égal à 1 quand le segment est entièrement visible, et égal à 0 si au moins un des points est caché. Pour un ensemble flou 3D μ, l'aire et le volume sont définis comme les intégrales de la norme du gradient et de la fonction d'appartenance :

$$S(\mu) = \int_{\text{Supp}(\mu)} |\nabla\mu(z)|dz \qquad V(\mu) = \int_{\text{Supp}(\mu)} \mu(z)dz$$

qui sont les extensions directes des définitions dans [105]. Pour définir les surfaces de contact entre la région β et un objet μ, nous utilisons l'intersection du support de β et de l'objet, soit :

$$F_\mu = (\ \text{Supp}(\mu) \cap \text{Supp}(\mu_\beta)\ ) \quad \text{ce qui donne} \quad S_\mu^c = \int_{F_\mu} |\nabla\mu(z)|dz. \tag{4.7}$$

L'équation (4.4) devient :

$$F\alpha_2 = f_a\left(\frac{(S_{\mu_A}^c(\mu_\beta) + S_{\mu_B}^c(\mu_\beta))^3}{V(\mu_\beta)^2}\right). \tag{4.8}$$

Les termes γ ont été supprimés puisqu'il n'est plus possible de définir le périmètre du fait que la frontière ne se réduit plus à une surface. Pour ajouter l'information de distance, nous considérons la distance minimale entre les supports de chaque objet plutôt que la longueur des segments admissibles comme proposé dans [118] :

$$D_{\mu_A\mu_B}(z) = \inf\{||\overrightarrow{xy}|| \mid x \in \text{Supp}(\mu_A), y \in \text{Supp}(\mu_B), z \in]x,y[\}.$$

Degré	Formule	Description
α_1	$f_a\left(\frac{S^3(\beta)}{V^2(\beta)}\right)$	Inverse de la compacité. Varie de 0 pour les formes sphériques à 1 pour les formes tubulaires.
α_2	$f_a\left(\frac{(\gamma_A S_A^c(\beta)+\gamma_B S_B^c(\beta))^3}{V(\beta)^2}\right)$	Ratio prenant en compte uniquement l'interface entre β et les objets. Les aires sont pondérées par un score d'élongation favorisant les interfaces allongées.
α_3	$f_a\left(\frac{(\gamma_A S_A^c(\beta)+\gamma_B S_B^c(\beta))^2}{\nu(\beta)}\right)$	Similaire à α_2 mais le volume est pondéré par la distance absolue minimale aux objets.
α_4	$f_a\left(\frac{(\gamma_A S_A^c(\beta_d)+\gamma_B S_B^c(\beta_d))^3}{V(\beta_d)^2}\right)$	Similaire à α_2 avec β seuillée sur la distance pour ne prendre en compte que les parties proches des objets.
α_5	$f_a\left(\frac{(\gamma_A S_A^c(\beta_d)+\gamma_B S_B^c(\beta_d))^2}{\nu(\beta_d)}\right)$	Similaire à α_3 avec β seuillée sur la distance pour ne prendre en compte que les parties proches des objets.
$F\alpha_2$ - $F\alpha_5$		Versions floues de α_2 - α_5. Les objets et β sont flous.

TABLE 4.3 – Résumé des degrés d'élongation proposés.

L'équation (4.5) est modifiée en pondérant $D_{\mu_A\mu_B}(z)$ par $\mu_\beta(z)$ [118] :

$$\nu(\mu_\beta) = \int_{\text{Supp}(\mu_\beta)} \mu_\beta(z) D_{\mu_A\mu_B}(z)\ dz. \tag{4.9}$$

Nous définissons ensuite $F\alpha_3$ en remplaçant le volume de l'équation (4.8) par celui donné dans l'équation (4.9). Utiliser un seuillage sur la distance comme suggéré pour α_4 et α_5 résulterait en une perte de la nature floue de la région. Pour l'éviter, nous utilisons une fonction décroissante $g : \mathbb{R} \longrightarrow [0,1]$ de $D_{\mu_A\mu_B}$, et β_d devient un ensemble flou :

$$\mu_{\beta_d}(z) = \mu_\beta(z) g(D_{\mu_A\mu_B}(z)) = \mu_\beta(z)(1 - f_{a_1}(D_{\mu_A\mu_B}(z)))$$

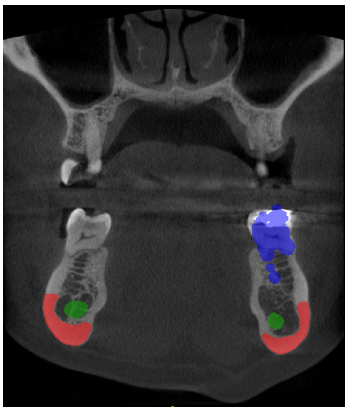
où f_{a_1} est une fonction du même type que f_a, c'est-à-dire une fonction de normalisation. Pour obtenir la frontière de β_d, nous remplaçons β par β_d dans l'équation (4.7). Pour définir $F\alpha_4$ et $F\alpha_5$, β est remplacé par β_d dans $F\alpha_2$ et $F\alpha_3$ respectivement. La table 4.3 résume les définitions proposées.

Validation

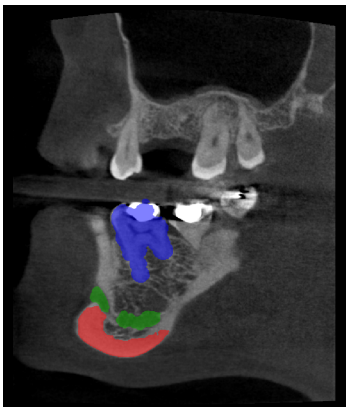
L'étude de la performance des degrés proposés est réalisé en deux phases sur des structures au voisinage du canal mandibulaire, car nous conceptualisons cet organe comme répondant aux critères de la relation "le long de". sur un ensemble de cinq objets que nous avons segmentés manuellement dans les volumes de trois patients. Pour chacun d'entre eux sont identifiés : les canaux mandibulaires droits (abrégé RMC) et gauches (LMC), la première molaire mandibulaire gauche ("36" dans la classification dentaire, notée simplement 36 ci-après), et le "plancher" mandibulaire qui correspond à la partie inférieure de l'os cortical de la mandibule, noté LMF côté gauche, et RMF côté droit. Une illustration de ces éléments est présentée dans la figure 4.8. Notons que ces segmentations sont grossières, la forme de l'espace inter-objet étant peu modifiée tant que la morphologie générale des objets est respectée.

Le degré de satisfaction de la relation "le long de" est calculé en utilisant $\alpha_i, i = 1, ..., 5$ pour chaque paire de structures. Nous attendons les meilleurs résultats entre les canaux et les planchers mandibulaires, pour chaque côté, puisque ce sont deux structures allongées dans la même direction. Comparer la dent 36 au plancher et au canal gauche devrait aussi donner un degré significatif puisque la dent est à côté d'une structure allongée. Tous les autres cas devraient avoir un score bas.

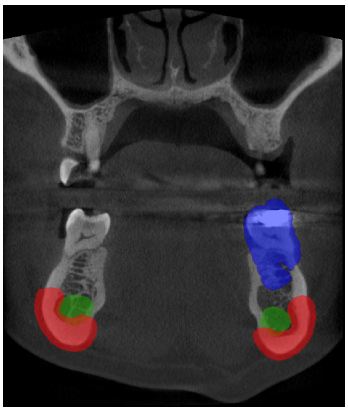
Les résultats sont visibles en figure 4.9. Le coefficient α_1 montre un haut degré pour quasiment toutes les paires de structures excepté pour le canal mandibulaire et le plancher mandibulaire gauche, ce qui peut apparaître comme contre-intuitif. Cela peut être expliqué par la distance entre les structures qui donne une région β allongée comme discuté en début de section 4.3.1. Comme attendu, cette



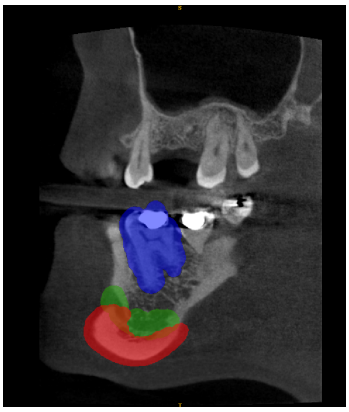
(a) Objets segmentés superposés sur une coupe coronale.



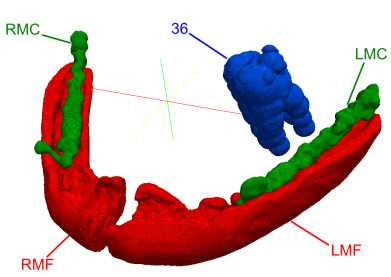
(b) Objets segmentés superposés sur une coupe sagittale.



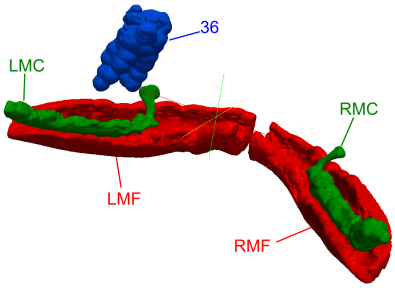
(c) Objets flous sur la coupe coronale.



(d) Objets flous sur la coupe sagittale.



(e) Vue 3D de face.

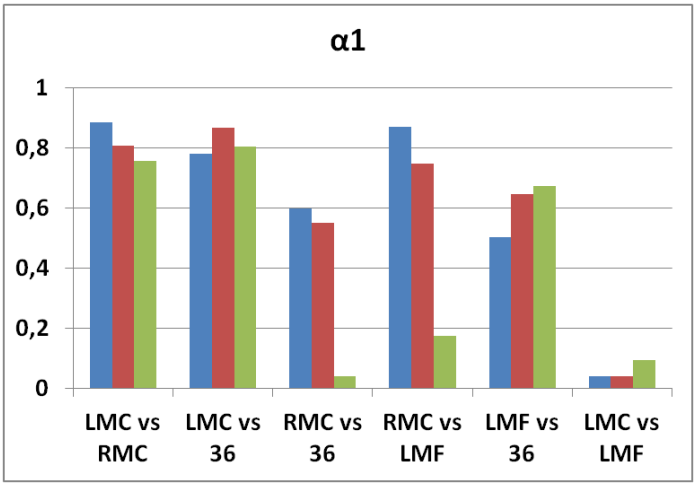


(f) Vue 3D de dos.

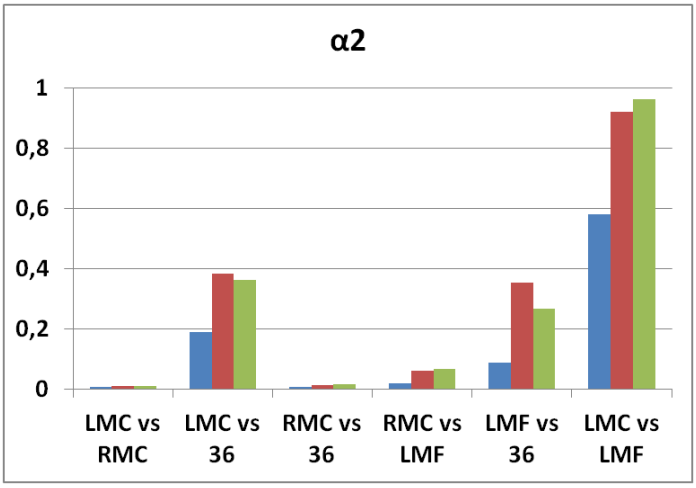
FIGURE 4.8 – Objets manuellement segmentés de la région maxillo-faciale. Les canaux mandibulaire sont en vert, les planchers mandibulaires en rouge et la dent 36 est en bleu.

version ne convient pas à notre objectif. Concernant les résultats de α_2, nous observons que le degré est fort lors de la comparaison des canaux et planchers du même côté, ce qui est intuitif, mais aussi dans une moindre mesure lors de la comparaison de la dent avec le canal ou le plancher, ce qui est sensé au regard de notre hypothèse même si la dent n'est pas une structure allongée. Pour toutes les autres comparaisons, le degré est très faible ($< 0,08$). Cette version distingue donc bien les diverses situations. L'ajout de l'information de distance ne modifie pas beaucoup les résultats, comme le montrent les profils similaires de α_3 et α_2. Toutefois, l'utilisation d'un seuil sur la distance semble être utile. Nous voyons avec α_4 que prendre en compte uniquement les parties proches des objets peut révéler des profils différents (par exemple LMF vs 36 pour les patients 1 et 3). Cela peut être utile pour les objets de forme complexe ou pour étudier seulement de petites parties des objets. Cette version garde des degrés significatifs, et nous pouvons voir que le plus fort score est toujours obtenu pour le canal mandibulaire face au plancher. Nous observons le même comportement avec α_5. Les seuils pour ces tests ont été choisis empiriquement comme la médiane de l'étendue des valeurs de distance.

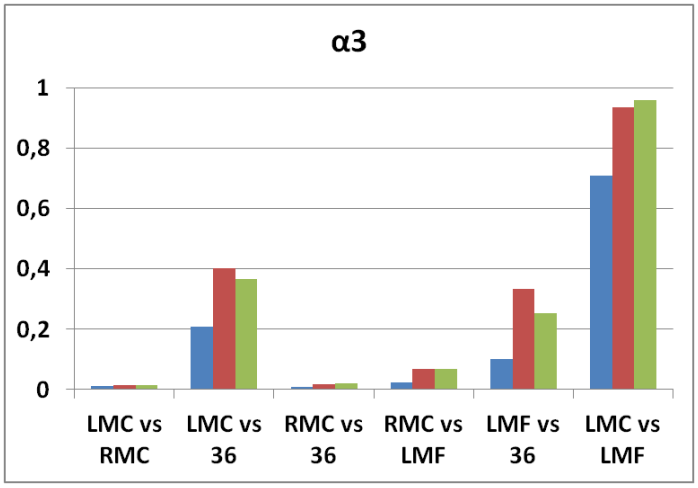
Pour simuler une segmentation floue des objets, les segmentations manuelles sont considérées comme le noyau de la fonction d'appartenance ($\mu(x) = 1$), et des points ont été ajoutés avec une valeur d'appartenance décroissante de 1 à 0 en fonction de la distance au noyau sur une marge d'1,5 mm (0 à cette distance). Comme indiqué précédemment, nous sommes intéressés par le formalisme des ensembles flous pour sa capacité à prendre en compte la variabilité morphologique inter-individus et les imprécisions de segmentation, notamment car cela peut conduire à des structures chevauchantes (ce qui est parfois le cas entre le canal et le plancher). Les résultats utilisant $F\alpha_2$ jusqu'à $F\alpha_5$ sont montrés en figure 4.10. La mesure $F\alpha_2$ présente le même profil que α_2, avec le plus fort degré pour les cas MC/MF, et des degrés significatifs pour les cas LMF/36 et LMC/36. L'ajout de l'information de distance ne donne pas de résultats différents comme le démontre $F\alpha_3$. Pour $F\alpha_4$ et $F\alpha_5$, l'utilisation de la distance permet là aussi de ne choisir que les structures proches pour faire une analyse locale. Pour ces tests, la fonction de normalisation pour β_t est la même que celle pour la normalisation des



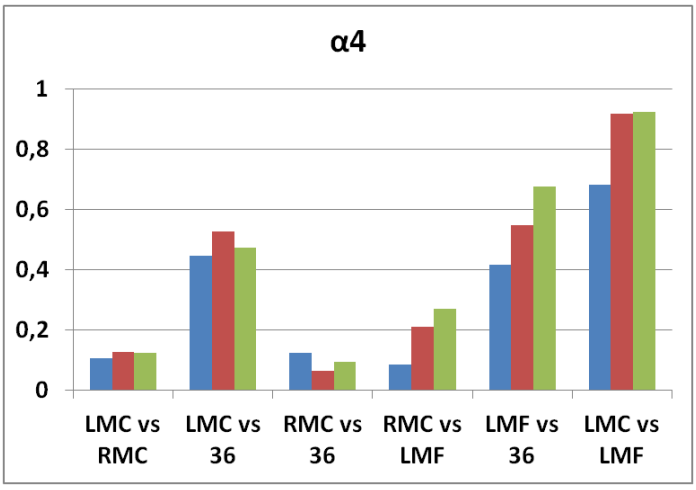
(a) Résultats pour α_1



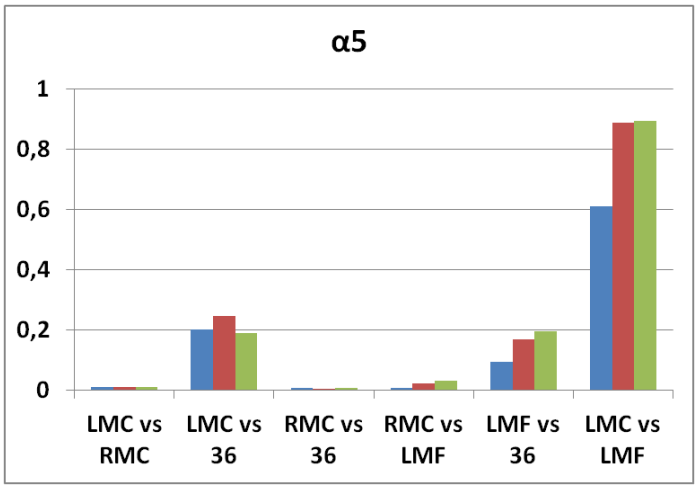
(b) Résultats pour α_2



(c) Résultats pour α_3



(d) Résultats pour α_4

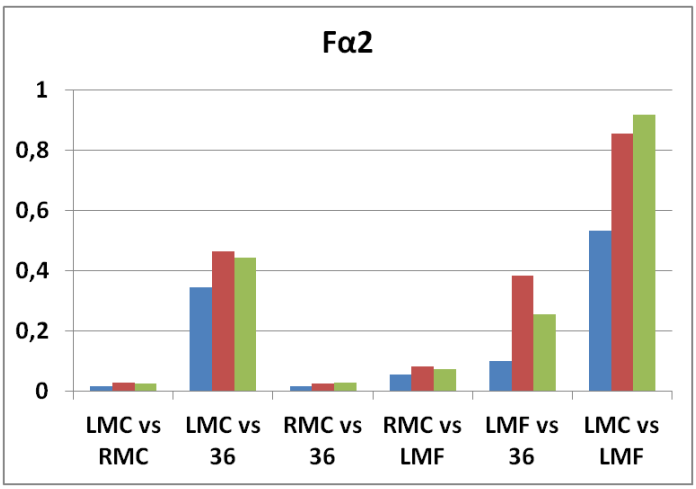


(e) Résultats pour α_5

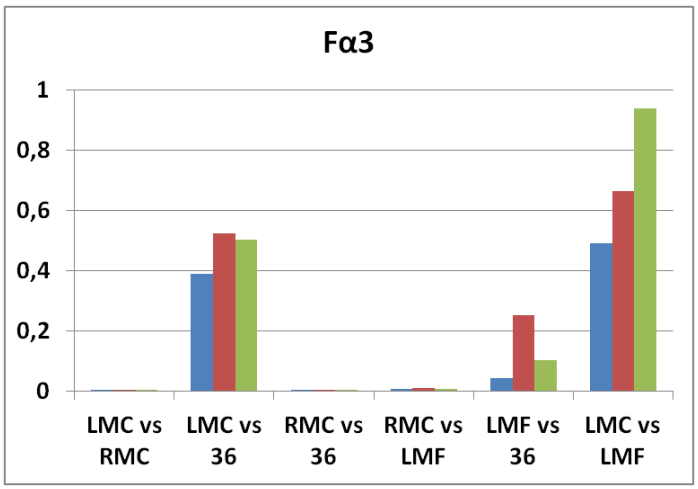
FIGURE 4.9 – Résultats obtenus sur les objets nets. La couleur différencie les patients.

degrés (avec le paramètre a égal à $0,05$). Le formalisme des ensembles flou permet d'avoir des degrés similaires lors de la comparaison des dents à des structures allongées, alors que la version nette est bien plus sensible à la position et à l'orientation de la dent.

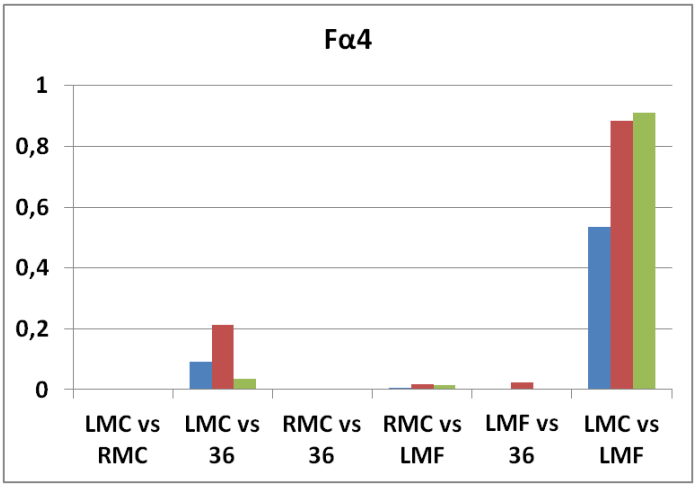
En fonction de ces résultats, nous avons sélectionné α_2 et $F\alpha_2$ comme les meilleures versions puisqu'elles sont adaptées à l'ensemble d'objets considéré. Un autre avantage est leur moindre coût



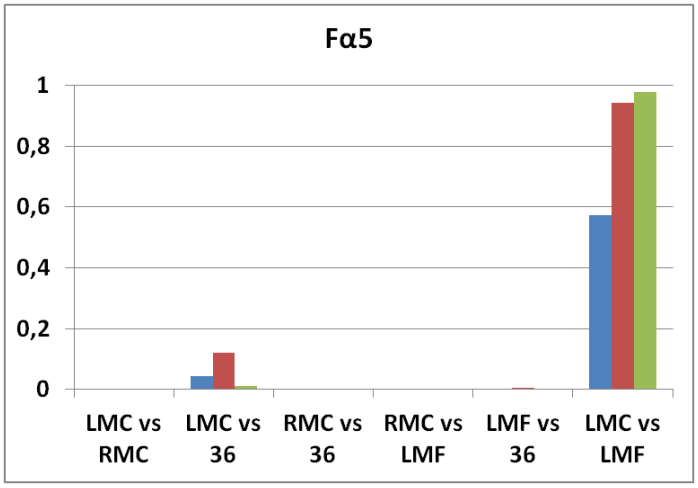
(a) Résultats pour $F\alpha_2$



(b) Résultats pour $F\alpha_3$



(c) Résultats pour $F\alpha_4$



(d) Résultats pour $F\alpha_5$

FIGURE 4.10 – Résultats obtenus pour les objets flous.

de calcul, car elles ne calculent pas de carte de distance, qui n'ajoute pas d'information pertinente dans notre cas. Toutefois, pour d'autres applications, l'utilisation des versions seuillées de β (telles que α_4) devrait être considérée. Pour nous assurer de la robustesse des versions sélectionnées, nous les testons sur trois autres patients supplémentaires, dont les planchers mandibulaires droits ont aussi été segmentés. Les résultats sont résumés dans la figure 4.11. Ces résultats confirment les premiers tests, distinguant correctement les cas en fonction de leur arrangement spatial, et indiquant une bonne robustesse au regard de la variabilité anatomique. De ce fait, nous proposons une classification pour le degré de satisfaction comme suit : une valeur entre 0 et 0,1 indique que les deux structures ne sont pas liées par une relation "le long de", entre 0,1 et 0,6 les deux structures partagent un lien spatial proche, au moins l'une des deux étant allongée, et de 0,6 à 1 les structures peuvent être dites le long l'une de l'autre. Il est évident que comparer deux degrés à l'intérieur d'une même catégorie reste sensé, puisqu'un degré de 0,5 représente une relation plus forte qu'un degré de 0,3. Dans le même esprit, un degré égal à 0,58 ne devrait pas être éliminé lorsqu'est recherché un fort lien "le long de" juste parce qu'il n'appartient pas à la "meilleure" catégorie.

Aspects pratiques

Il est difficile d'utiliser cette relation dans une mécanique exploratoire, puisque le calcul demande de connaître l'espace entre les deux objets. L'utilisation d'une telle relation spatiale est donc plutôt destinée à servir de contrôle de la qualité de segmentation, et de validation du processus. Une fois un

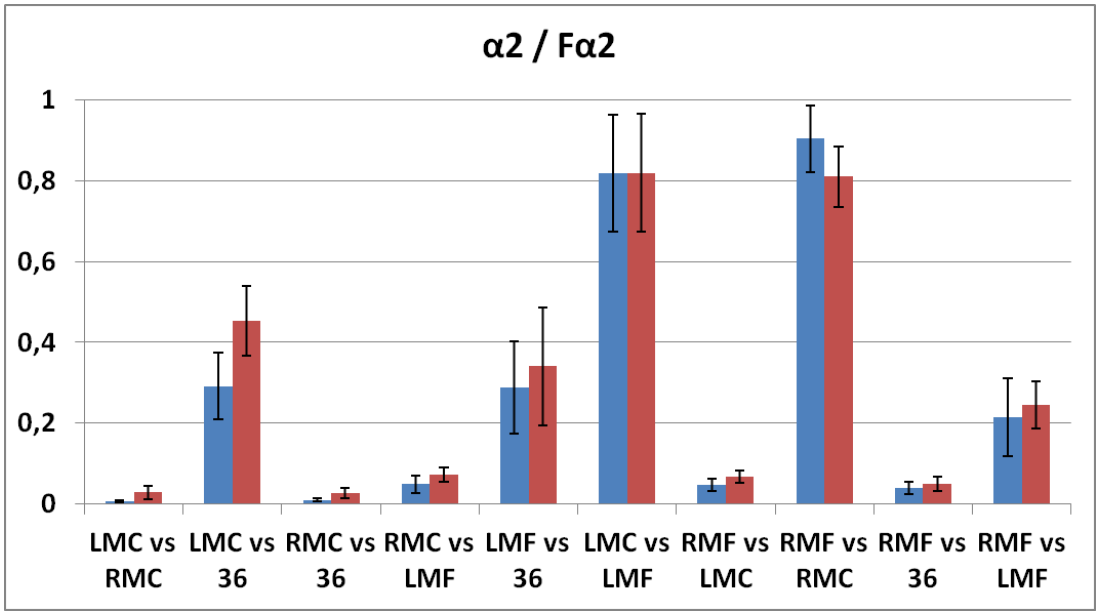


FIGURE 4.11 – Moyenne des scores obtenus pour les degrés sélectionnés (α_2 et $F\alpha_2$) sur 6 patients. Le premier ensemble de résultats correspond à α_2 (bleu) et le second à $F\alpha_2$ (rouge). L'écart-type est montré sous la forme de barres d'erreur.

objet délimité, s'il est lié par une relation "le long de" avec une structure voisine également identifiée, le calcul du score permet d'obtenir une information sur la pertinence des résultats : si le score est faible, les segmentations impliquées devront être contrôlées. Notez cependant qu'un score élevé n'est pas suffisant seul pour valider les résultats. Deux objets peuvent être erronés avec des fuites de segmentation dans la même direction spatiale en induisant des fortes valeurs de réponse au score. Lorsque l'identification des structures est effectuée en même temps que la délimitation, la relation peut aussi servir pour faciliter l'étiquetage des segmentations.

4.3.2 Relation d'alignement

L'alignement est une relation particulière au sens où elle implique un ensemble d'objets. Deux grandes approches existent pour étudier l'alignement. Premièrement nous pouvons rechercher une droite intersectant tous les objets ou bien un angle θ tel que chaque objet est situé dans la direction θ ou $(\theta + \pi)$ de son voisin. Nous nous intéressons à la seconde approche, qui utilise les histogrammes d'angles. Plus précisément, nous étendons la notion d'histogrammes d'orientation [122] à l'espace tri-dimensionnel. La stratégie globale est de trouver si l'histogramme présente le même profil entre tous les objets d'un même groupe. Pour ce faire, les histogrammes sont calculés sur l'ensemble des points d'un objet face à l'ensemble des points de son vis-à-vis.

Histogrammes d'angles et d'orientation

En 3D ; une direction est définie par deux angles [19] : la longitude ($\theta \in [0, 2\pi[$) et la latitude ($\delta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$). Dans la suite, A et B sont deux objets, x et y deux points appartenant à A et B respectivement. Cela conduit à des histogrammes d'angles bi-dimensionnels définis par :

$$H^A(B)(\theta,\delta) = \frac{\sum_{x\in A, y\in B | \angle(\mathrm{proj}_{XY}(\overrightarrow{xy}),\overrightarrow{u_X})=\theta \text{ et } \angle(\overrightarrow{xy},\mathrm{proj}_{XY}(\overrightarrow{xy}))=\delta} 1}{\max_{\Theta\in[0,2\pi[,\Delta\in[-\frac{\pi}{2},\frac{\pi}{2}]} \sum_{x\in A, y\in B | \angle(\mathrm{proj}_{XY}(\overrightarrow{xy}),\overrightarrow{u_X})=\Theta \text{ et } \angle(\overrightarrow{xy},\mathrm{proj}_{XY}(\overrightarrow{xy}))=\Delta} 1}$$

où A est l'objet de référence et B l'objet cible, $\mathrm{proj}_{XY}(\overrightarrow{xy})$ est la projection orthogonale de $\overrightarrow{xy}$ sur le plan XY et $\overrightarrow{u_X}$ le vecteur unitaire de l'axe X. La division par le maximum des fréquences permet l'obtention d'un histogramme normalisé. Comme évoqué dans la section 4.3.1, il est souvent utile de

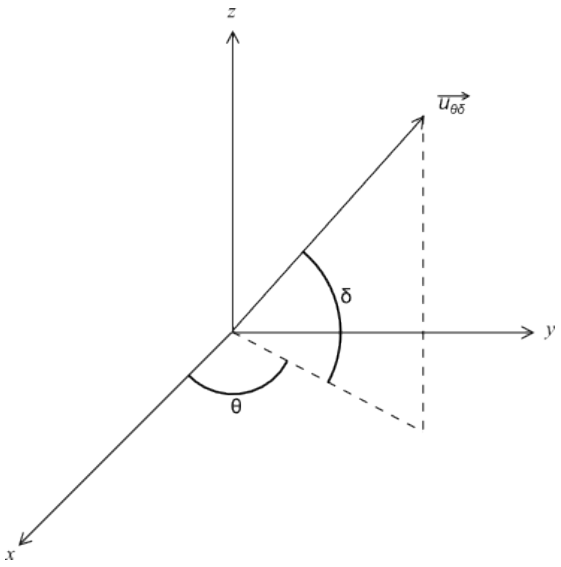


FIGURE 4.12 – Direction dans un espace 3D.

considérer des objets flous. Lorsque les objets sont flous, ils sont définis par leur fonction d'appartenance ($\mathbb{Z}^3 \rightarrow [0,1]$) ce qui conduit à :

$$H^A(B)(\theta,\delta) = \frac{\sum_{x,y\in\mathbb{Z}^3 | \angle(\text{proj}_{XY}(\overrightarrow{xy}),\overrightarrow{u_X})=\theta \text{ et } \angle(\overrightarrow{xy},\text{proj}_{XY}(\overrightarrow{xy}))=\delta} t[\mu_A(x),\mu_B(y)]}{\max_{\Theta\in[0,2\pi[,\Delta\in[-\frac{\pi}{2},\frac{\pi}{2}]} \sum_{x,y\in\mathbb{Z}^3 | \angle(\text{proj}_{XY}(\overrightarrow{xy}),\overrightarrow{u_X})=\Theta \text{ et } \angle(\overrightarrow{xy},\text{proj}_{XY}(\overrightarrow{xy}))=\Delta} t[\mu_A(x),\mu_B(y)]}$$

où t est une t-norme (le minimum par exemple). Le principal défaut de ces histogrammes est de considérer les objets de manière différente, avec un objet cible et un objet de référence, ce qui ne traduit pas la réciprocité de la relation.

Nous utilisons alors l'histogramme d'orientation [120], qui est un histogramme où les angles sont calculés de manière à avoir une symétrie entre les deux objets. Pour ce faire, θ est calculé modulo π sur un support de longueur π et δ est la valeur absolue de la latitude sur un support de longueur $\frac{\pi}{2}$, soit :

$$O(A,B)(\theta,\delta) = \frac{\sum_{x,y\in\mathbb{Z}^3 | \text{ mod } (\angle(\text{proj}_{XY}(\overrightarrow{xy}),\overrightarrow{u_X}),\pi)=\theta \text{ et abs}(\angle(\overrightarrow{xy},\text{proj}_{XY}(\overrightarrow{xy})))=\delta} t[\mu_A(x),\mu_B(y)]}{\max_{\Theta\in[0,\pi[,\Delta\in[0,\frac{\pi}{2}]} \sum_{x,y\in\mathbb{Z}^3 | \text{ mod } (\angle(\text{proj}_{XY}(\overrightarrow{xy}),\overrightarrow{u_X}),\pi)=\Theta \text{ et abs}(\angle(\overrightarrow{xy},\text{proj}_{XY}(\overrightarrow{xy})))=\Delta} t[\mu_A(x),\mu_B(y)]}.$$

Notez que la définition implique que $O(A,B) = O(B,A)$.

Comparaison d'histogrammes

La stratégie la plus simple pour rechercher un alignement à partir des histogrammes consiste à calculer l'ensemble de ces derniers pour chaque paire d'objets puis à les comparer. Parmi différentes mesures de similarité possibles [14, 131], nous utilisons ici la hauteur maximale de l'intersection entre les histogrammes :

$$\text{sim}(O(A,B),O(C,D)) = \max_{\theta\in[0,\pi[,\delta\in[0,\frac{\pi}{2}]} t[O(A,B)(\theta,\delta),O(C,D)(\theta,\delta)]$$

Généralisée à un groupe d'objets, cette similarité vaut :

$$\text{sim}(O(A_0,A_1),...,O(A_i,A_j),...,O(A_{N-1},A_N)) = \max_{\theta\in[0,\pi[,\delta\in[0,\frac{\pi}{2}]} t(O(A_i,A_j)(\theta,\delta),\ i,j=0...N, i\neq j).$$

Toutefois, comme développé dans [120], la comparaison tend à donner de faibles valeurs de similarité même si les histogrammes sont proches, cela étant dû à la comparaison angle par angle, alors que dans ce contexte, se référer à un angle (θ,δ) recouvre en réalité un angle approximatif dû à la discrétisation

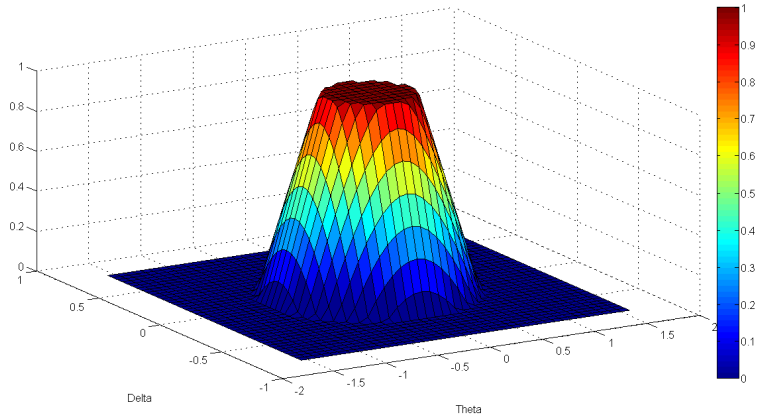


FIGURE 4.13 – Forme de l'élément structurant utilisé pour la dilatation des histogrammes d'orientation.

de l'espace. Pour pallier ce problème, une dilatation floue est réalisée sur les histogrammes afin de propager les hautes valeurs aux valeurs d'angles similaires. L'élément structurant pour réaliser la dilatation est défini comme suit :

$$v_0(\theta,\delta) = \begin{cases} 1 & \text{si} \quad \sqrt{\theta^2+(2\delta)^2} \leq t_1, \\ \frac{t_2-\sqrt{\theta^2+(2\delta)^2}}{t_2-t_1} & \text{si} \quad t_1 < \sqrt{\theta^2+(2\delta)^2} \leq t_2, \\ 0 & \text{si} \quad \sqrt{\theta^2+(2\delta)^2} > t_2. \end{cases}$$

où t_1 et t_2 sont les paramètres contrôlant la forme de l'élément structurant (celui-ci correspond à un cône plus ou moins évasé dont le sommet est tronqué, voir figure 4.13), ce qui permet de régler la propagation des valeurs d'angles. Ils sont directement liés à l'imprécision lors du calcul d'angles sur un espace discret. Plus les points considérés sont proches, plus l'approximation sera grande et inversement. Il est donc intéressant d'avoir une idée de la distance séparant les objets impliqués dans l'alignement supposé. Dans le cas où tous les objets se trouveraient à une distance similaire, il est proposé dans [120] d'utiliser ces paramètres :

$$t_1 = \arcsin\left(\frac{2}{d_{average}}\right),$$

$$t_2 = \arcsin\left(\frac{4}{d_{average}}\right),$$

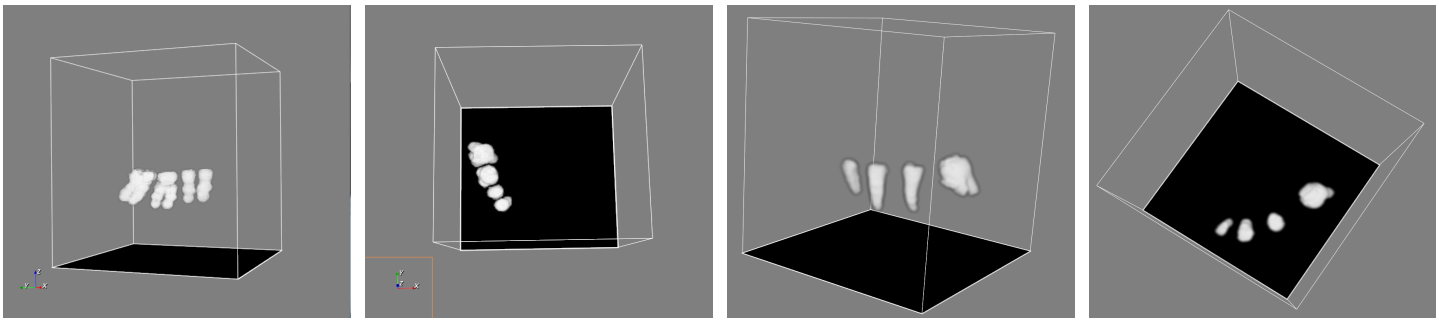
où $d_{average}$ est la distance moyenne en pixels entre les barycentres des objets voisins.

Recherche d'alignement

Il est possible de concevoir l'alignement comme des objets sur une ligne, potentiellement courbe (par exemple les dents sur l'arche dentaire), et c'est ici la notion d'alignement local qui est importante, où comme des objets se trouvant sur une droite de l'espace, ce qui est l'alignement strict (ou global). Dans les deux cas, il est nécessaire que les objets soient à une distance proche les uns des autres pour que la propriété apparaisse.

À partir de ces constatations, Vanegas [120] définit l'alignement local comme un ensemble d'objets connectés par une relation de voisinage $Neigh$ où chaque objet est aligné par rapport à ses voisins directs. La relation $Neigh$ est satisfaite entre deux objets A et B lorsque $A \cap B = \emptyset$ et $A \cup N(B) \neq \emptyset$ où $N(B)$ définit un voisinage de l'objet B, par exemple une distance euclidienne. Le degré d'alignement local peut alors être obtenu par similarité sur l'ensemble des groupes voisins d'un ensemble S d'objets ($|S| \geq 3$), soit :

$$\mu_{local}(S) = \min_{A,B,C:Neigh(A,B)\text{et}Neigh(B,C)} sim(O(A,B),O(B,C)).$$



(a) Vue sagittale-coronale des dents (objets nets).

(b) Vue axiale des dents (objets nets).

(c) Vue sagittale-coronale des dents (objets flous).

(d) Vue axiale des dents (objets flous).

FIGURE 4.14 – Agencement spatial des dents utilisées pour tester l'alignement.

À l'inverse, un alignement global possède une direction selon laquelle est disposé l'ensemble des objets le composant. Il serait donc possible de calculer l'ensemble des histogrammes d'orientation sur toutes les paires possibles d'objets et de les comparer, néanmoins cela souffre d'une grande sensibilité au moindre écart à la droite idéale, ce qui peut arriver fréquemment selon la forme des objets considérés. Pour assouplir la mesure, et puisque la direction est un élément partagé par l'ensemble du groupe, les histogrammes sont calculés entre chaque individu face au reste du groupe. Le degré d'alignement global se définit alors comme :

$$\mu_{global}(S) = sim(O(A_0, S \setminus \{A_0\}), ..., O(A_N, S \setminus \{A_N\})).$$

où $S = \{A_0, ..., A_N\}$. Un groupe S est dit localement/globalement aligné à un degré $\mathcal{D}$ lorsque $\mu_{local/global}(S) \geq \mathcal{D}$. Soulignons que tout groupe globalement aligné l'est localement aussi. Nous pouvons donc rechercher l'ensemble des groupes globalement alignés comme un sous-ensemble des groupes localement alignés.

Afin d'illustrer ces définitions, un test a été réalisé sur les dents avec un ensemble d'objets nets, et un ensemble d'objets flous. Ces ensembles peuvent être vus dans la figure 4.14. Pour l'ensemble net (figure 4.14 (a,b)), le degré d'alignement global est de $0,92$, et le calcul de l'alignement local détermine ici un seul groupe (pour un voisinage de 30 pixels) avec un degré de $0,81$ lorsque la dilatation d'histogramme n'est pas employée, contre un degré égal à 1 avec dilatation. Pour l'ensemble d'objets flous (figure 4.14 (c,d)), le degré d'alignement global est de $0,67$. Nous retrouvons localement un seul groupe, pour un degré de $0,52$ sans dilatation ($0,61$ avec). Nous n'avons pas employé la dilatation dans le cas des alignements globaux car la construction d'histogrammes face à l'ensemble du groupe produit déjà des histogrammes larges.

Aspect pratiques

Cette relation spatiale peut être utilisée à la fois comme validation a posteriori de la cohérence de la segmentation d'un ensemble d'objets, par exemple dans le cas d'une arche dentaire, mais peut aussi être utilisée de manière exploratoire. Si nous reprenons le cas dentaire, dès lors que trois dents sont identifiées, il est possible de créer un ensemble flou spatial qui identifie une zone de l'espace, et qui peut être utilisé pour définir un périmètre de recherche pour l'élément suivant, ou comme support pour un terme d'attache aux données dans une formulation énergétique, par exemple telle qu'employée dans [5]. Cet ensemble peut représenter un alignement global (ce sera alors l'équivalent d'une direction) ou un alignement local : pour une valeur donnée du degré de satisfaction de la relation et un voisinage, l'ensemble représentera les emplacements possibles pour l'extension du groupe. Nous estimons que cette définition permettrait de réduire le nombre d'interactions demandé à un utilisateur qui souhaite segmenter une denture, en déduisant l'emplacement dentaire à partir d'un nombre réduit de dents initiales.

4.4 Conclusion

4.4.1 Utilisation pratique du modèle structurel

Une fois le modèle établi, il peut être utilisé pour contraindre un processus de segmentation. Dans l'optique d'analyser un volume entier de la manière la plus autonome possible, nous avons principalement deux méthodes : soit procéder à l'identification et la segmentation de manière conjointe et globale, en formalisant les connaissances comme un réseau de contraintes à satisfaire [98, 120] ; soit partant d'un point d'entrée (une information de l'utilisateur, ou un élément facilement identifiable), procéder de manière itérative en parcourant le graphe structurel. Pour notre application, nous envisageons d'utiliser la seconde méthode [52]. D'une part, il est possible d'obtenir l'initialisation grâce à une interaction de l'utilisateur ou grâce à de l'information provenant des autres modalités (numérisations d'empreintes ou scans intra-oraux par exemple), et d'autre part, il est possible de n'explorer qu'une partie du modèle. Cela est intéressant du point de vue du coût de calcul, car seuls les éléments choisis seront traités, mais aussi du point de vue de la robustesse. En effet, considérant notre contexte, il est probable que le contenu de l'image soit modifié par rapport au cas "sain" par des structures surnuméraires ou manquantes, dégradées, ou par la présence de pathologies. Dans de tels cas, l'utilisation du modèle comme réseau de contraintes est fortement impactée par ces modifications, c'est-à-dire que le modèle ne va plus correspondre au contenu, donc l'information apportée est à exploiter avec précaution. L'approche itérative permet d'en limiter l'impact si la modification n'intervient pas dans la partie du modèle explorée. Le revers de ce choix est la nécessité d'établir une stratégie d'évaluation des résultats successifs et d'exploration, ce pour deux raisons. Premièrement, si le traitement d'une structure est de mauvaise qualité, la suite du processus est compromise puisque reposant sur l'étape précédente, nous propageons alors une erreur qui s'amplifie. Deuxièmement, si la segmentation échoue, il faut identifier un autre chemin dans le graphe pour atteindre notre objectif. Néanmoins, l'adaptation des méthodes de [52] à notre contexte reste à faire au moment de l'écriture. Il sera aussi nécessaire de préparer l'adaptation du modèle aux cas des pathologies et de l'application pédiatrique. Dans le premier cas, il faut ajouter le concept de la pathologie dans le modèle, et ce faisant réévaluer la validité des relations qui l'entourent. Les pathologies ayant généralement un impact différent selon chaque patient, ces travaux rejoignent la problématique de la spécialisation du modèle aux cas individuels, et des propositions existent dans la littérature [4]. L'adaptation pédiatrique est plus problématique : si la majeure partie des structures conserve des relations identiques au cours de la croissance (les fosses nasales seront toujours au-dessus des maxillaires), l'arcade dentaire présente une évolution très marquée, impactant notamment les relations spatiales. Une piste pourrait être de parvenir à constituer un ensemble de modèles conditionnés par la maturation de l'arcade, et d'utiliser celui le plus proche du stade du patient traité, mais il faut alors définir une méthode apte à quantifier l'évolution dentaire chez le patient.

4.4.2 Amélioration de l'efficacité de la segmentation multi-objets

Les défauts de notre approche actuelle sont directement liés à la méthode de résolution. Il est nécessaire de modifier cette dernière pour envisager de pouvoir utiliser une approche multi-objets au niveau dentaire. Pour cela, la piste qui nous semble la plus prometteuse et la plus aisée à mettre en place est d'adapter la méthode d'optimisation du power-watershed [40] à notre formulation. La minimisation est alors globale même dans le cas multi-objets, et les performances de calculs sont nettement supérieures à la coupure de graphe pour de grands volumes. Lorsque des performances satisfaisantes auront été obtenues, l'étape suivante sera de tenter de diminuer la quantité d'interactions demandée à l'utilisateur, jusqu'ici identique au cas mono-objet (par exemple en initialisant automatiquement les marqueurs des dents adjacentes). La deuxième évolution prévue est d'inclure les contraintes relationnelles (notamment) spatiales telles que précédemment définies (partie 4.2.2) pour continuer à renforcer la précision de la segmentation.

Chapitre 5

Application : Outil de segmentation de l'arche dentaire à destination des praticiens

Ce chapitre présente la mise en œuvre de la méthode de segmentation dentaire décrite aux chapitres 3 et 4 sous la forme d'un prototype logiciel fonctionnel et prêt à l'emploi, destiné à être utilisé directement par les praticiens, ce que nous estimons être la contribution finale de cette thèse.

5.1 Mise en œuvre

Nous avons choisi d'implémenter notre méthode de segmentation dans le langage de programmation C++, pour sa rapidité par rapport à des langages dits "interprétés" tels que MATLAB ou Python, ainsi que pour la grande base d'algorithmes existants mis à disposition soit dans les bibliothèques logicielles soit directement par les développeurs. Nous nous sommes notamment appuyés sur les bibliothèques ITK (www.itk.org) pour le traitement d'images, VTK (www.vtk.org) pour la visualisation et FLTK (www.fltk.org) pour l'interface graphique. L'algorithme de coupure de graphe utilisé est celui mis à disposition librement par les auteurs de [23] (pub.ist.ac.at/~vnk/software.html). Dans sa forme actuelle, le logiciel est sur Microsoft Windows, mais le code est indépendant et notre solution peut être portée sur tous les autres systèmes à architecture x86 existants (Unix, Linux, MacOSX). Pour aborder la construction du prototype, nous avons pour hypothèse qu'il peut être distribué à n'importe quel praticien dentaire, sans viser de spécialisation en particulier, et qu'il peut servir à analyser des dentures complètes tout autant que des arches partielles. Les spécifications visées sont donc les suivantes :

- Volumes d'entrée aux formats MetaImage (.mhd et .mha, utilisés par ITK et VTK) ainsi que DICOM (.dcm).
- Volumes avec une taille de voxel isotrope comprise entre 0,075 mm et 0,3 mm.
- Temps de segmentation pour une denture complète d'environ 5 minutes.
- Pas de paramètres à régler pour l'utilisateur.
- L'utilisateur doit pouvoir évaluer le résultat de la segmentation et la corriger le cas échéant.
- L'utilisateur doit pouvoir exporter ses travaux.
- Segmentation de chaque dent au cas par cas guidée par l'utilisateur (une dent à la fois).
- Traitement des dentures adultes et adaptabilité aux cas d'enfants.

Le flux de travail est présenté dans la figure 5.1. Le processus comprend quatre étapes principales. La première consiste à charger en mémoire les données sur lesquelles l'utilisateur veut travailler. Pour le moment, le logiciel ne permet que l'utilisation de données locales, c'est-à-dire stockées sur la machine de l'utilisateur ou dans le réseau local (interne). L'adaptation aux images distantes disponibles dans des systèmes PACS (systèmes d'archivage et de transmission d'images, permettant la gestion et la distribution des images) est une évolution souhaitable. Cette mise en place s'effectue au travers d'une

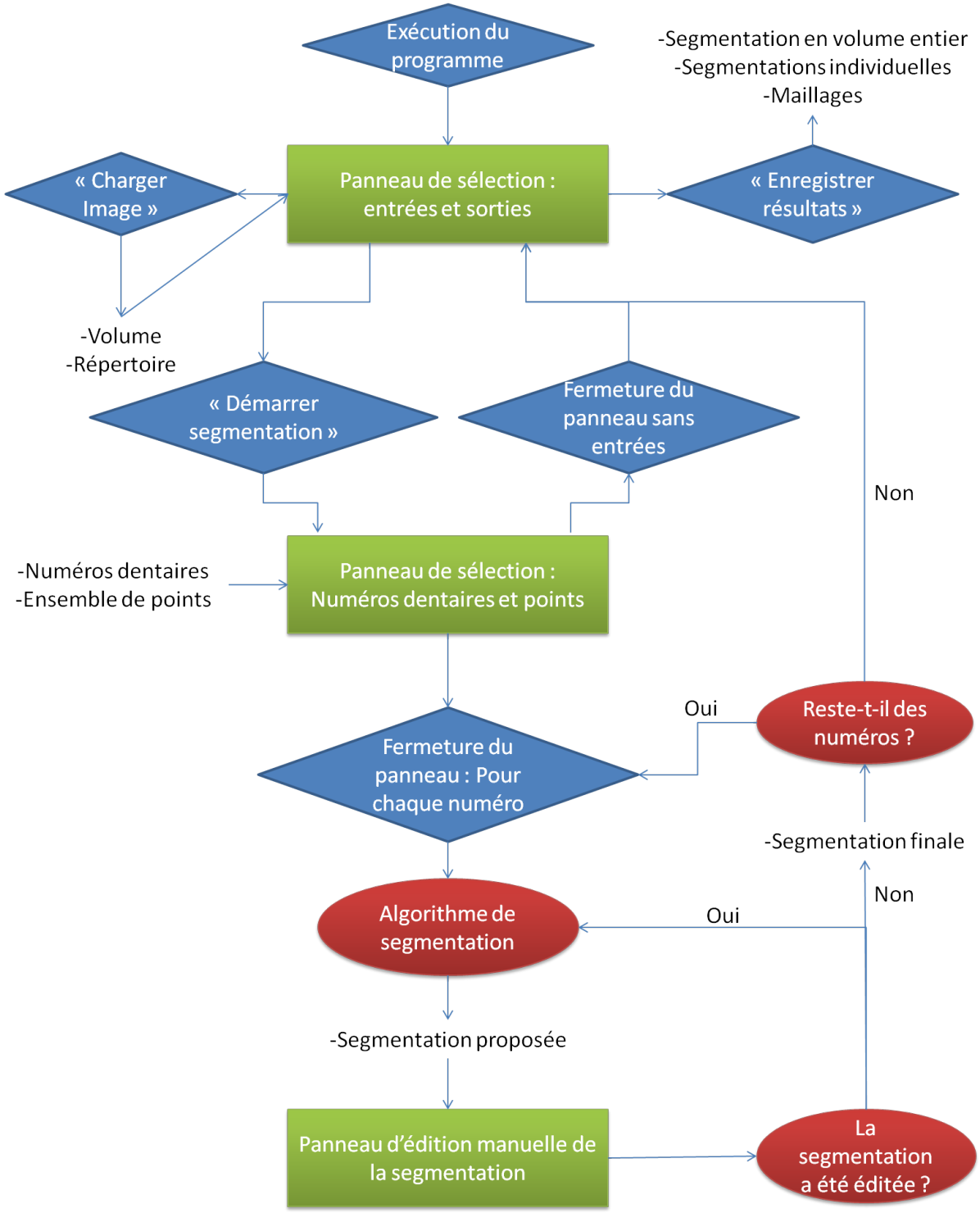


FIGURE 5.1 – Flux de travail utilisé dans le prototype logiciel. Les numéros dentaires et les ensembles de points sont les éléments nécessaires à la segmentation, donnés par l'utilisateur.

première fenêtre qui permet la gestion des entrées et des sorties du programme. L'utilisateur navigue à travers son système de fichiers pour trouver le volume souhaité. Ce dernier doit se présenter sous la forme d'un fichier DICOM (extension .dcm, format standardisé d'imagerie médicale de référence, dicom.nema.org) ou sous la forme d'un fichier MetaImage (extension .mhd ou .mha, format de base des bibliothèques ITK et VTK, représentant le volume en données binaires adjoint d'une description d'en-tête). L'avantage du fichier DICOM tient au fait qu'en plus de l'image, il contient un ensemble d'informations supplémentaires liées au patient (comme le nom, l'âge, le sexe), aux conditions d'ac-

quisitions (comme la date d'acquisition, les paramètres d'acquisition, le capteur et sa résolution, ou le champ de vue) à la provenance de l'image (comme le type de scanner, le nom de la clinique ou encore le nom du praticien), etc. Bien que ces méta-données puissent apporter des informations pertinentes pour l'analyse de l'image, notre logiciel n'en fait pas usage pour le moment. Une fois le volume chargé, la seconde étape est de sélectionner les dents que l'utilisateur souhaite segmenter. Pour ce faire, nous avons besoin qu'il indique deux informations pour chaque dent. Premièrement, son numéro dentaire, ce qui nous permet d'utiliser un modèle de forme approprié (voir chapitre 3), ainsi qu'un ensemble de points correspondant à la couronne et aux apex racinaires, grâce à une fenêtre dédiée où est visualisée l'image. Vient ensuite la troisième partie, qui correspond à la segmentation à proprement parler des dents sélectionnées. Pour chaque numéro dentaire indiqué, nous utilisons notre algorithme semi-automatique pour proposer à l'utilisateur une segmentation de la dent. Il est important que le praticien puisse disposer d'un retour pour juger de la qualité du résultat, et le corriger si besoin. Cela est réalisé au travers d'une dernière fenêtre qui représente la segmentation. Si l'utilisateur juge nécessaire de retravailler le résultat semi-automatique, il pose des marqueurs qui correspondent soit à l'extérieur soit à l'intérieur de l'objet, et cette information est utilisée par l'algorithme pour recalculer la segmentation. Ce cycle est répété jusqu'à satisfaction de l'utilisateur. Une fois la segmentation de cette dent achevée, la suivante est traitée de la même manière. Une fois que tous les numéros dentaires indiqués par l'utilisateur sont traités, le programme revient sur la fenêtre des entrées et sorties. En dernière étape, l'utilisateur configure alors le lieu où il souhaite enregistrer ses fichiers et sauvegarde les résultats à l'aide d'un bouton. Le logiciel écrit alors plusieurs types de fichiers :

— la segmentation de toutes les dents dans l'espace du volume (fichier MHD),

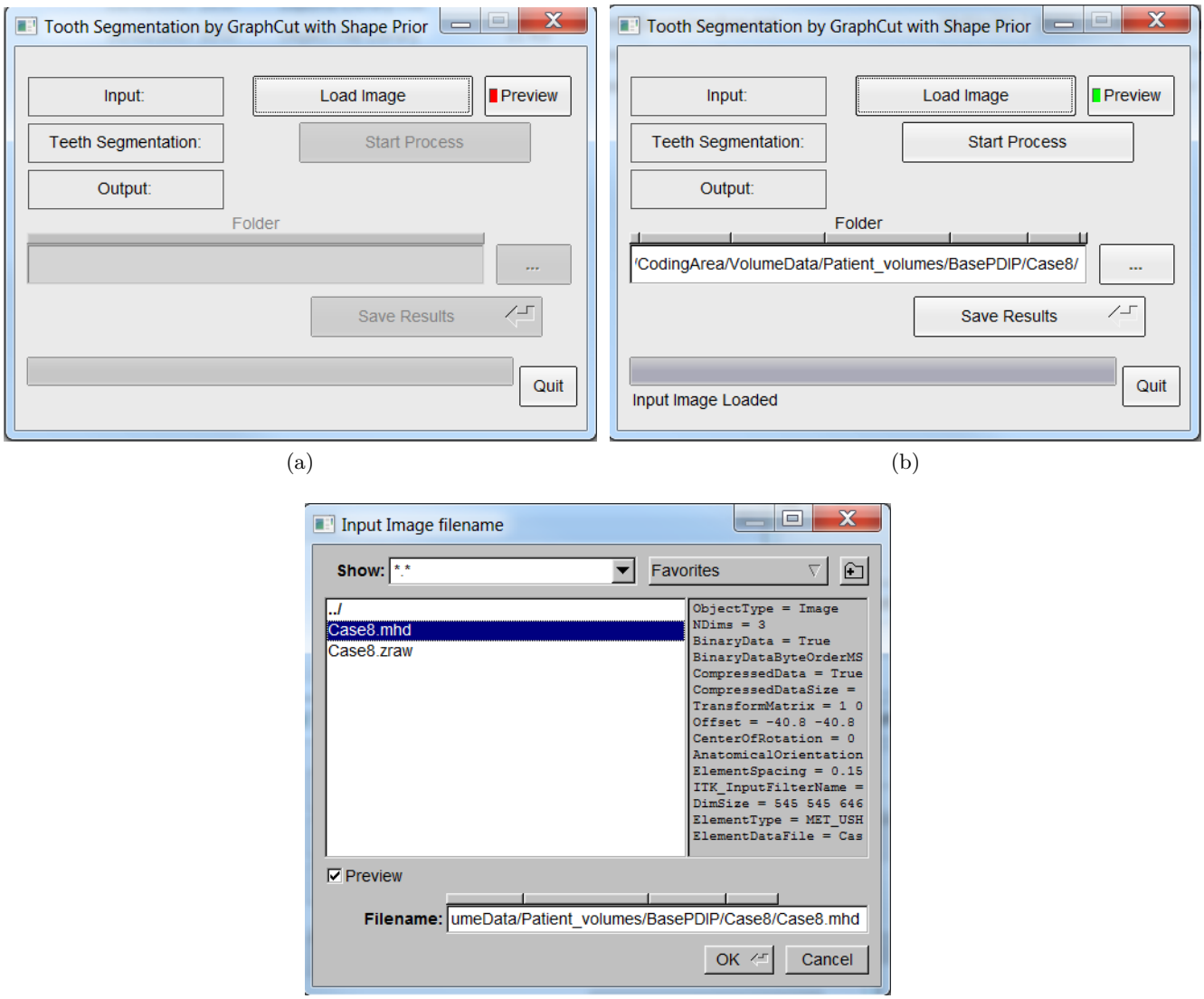
— la segmentation de chaque dent séparément dans son volume d'intérêt (fichiers MHD),

— le maillage individuel de chaque dent dans son volume d'intérêt (fichiers STL).

Notons que si l'utilisateur souhaite recommencer des segmentations, il peut réouvrir la fenêtre de sélection et réaliser de nouvelles dents, ou même celles qu'il a déjà effectuées (ce qui efface la segmentation précédente pour repartir à neuf). La segmentation d'une denture complète s'effectue en environ 20 à 30 minutes selon la difficulté du cas pour un utilisateur expérimenté, ce qui est loin de l'objectif des 5 minutes, mais présente une amélioration notable face aux outils actuellement usités chez Carestream, qui demandent entre 45 minutes à 1 heure pour la même tâche. Soulignons qu'en l'état, le programme ne permet pas de commencer un processus de segmentation, d'enregistrer les résultats puis de corriger a posteriori dans un second temps. Cette contrainte est néanmoins à relativiser, puisqu'il est possible de segmenter un petit ensemble de structures à chaque séance, puis de fusionner les résultats de manière aisée : tous les résultats issus d'un même volume se trouvent dans le même repère physique, ce qui rend la fusion triviale.

5.2 Interface graphique et outils de correction manuelle

Lors de la réalisation de l'interface graphique nous avons cherché à la rendre concise, claire, pratique et minimisant les risques d'erreurs de l'utilisateur. Dans la figure 5.2 (a) est présentée la fenêtre initiale du programme lors de son exécution. Elle présente les grandes étapes du processus (entrées, segmentation, sorties) de manière successive pour inciter l'utilisateur à suivre l'ordre attendu des actions. Notons que seuls les éléments correspondant à l'étape des entrées sont activés pour insister sur l'action à effectuer, ainsi que pour parer à toute manipulation inattendue. Lorsque l'utilisateur clique sur le bouton de chargement de l'image, la fenêtre (c) s'ouvre. Il s'agit d'un explorateur de fichier conventionnel, qui permet de naviguer dans le système pour sélectionner le volume voulu. Lorsque ce dernier est choisi, l'explorateur se ferme, le programme charge l'image et la fenêtre initiale prend alors l'aspect présenté dans la figure (b) (la barre au pied de la fenêtre est une barre de progression montrant l'avancement des processus). Le bouton d'accès au processus de segmentation est maintenant activé, ainsi que la partie des sorties. Par défaut, l'emplacement d'écriture des sauvegardes est celui du fichier d'entrée, mais l'utilisateur peut le modifier, soit manuellement, soit par l'intermédiaire d'un explorateur de fichiers. Le bouton de sauvegarde est activé aussi, bien qu'il ne produise aucun résultat si rien n'est segmenté. Lorsque l'utilisateur commence la segmentation, la fenêtre de sélection s'ouvre (figure 5.3).



(a) (b)

(c)

FIGURE 5.2 – Interface graphique correspondant à la première étape de traitement.

Au centre se retrouve un ensemble de cases à cocher, qui représentent les numéros dentaires (selon la norme ISO). Ils sont disposés comme si un dentiste regardait son patient, selon la représentation habituelle utilisée pour la nomenclature. Sont ensuite affichées les trois vues correspondant aux plans orthogonaux habituels du volume 3D, soit le plan coronal (en haut à gauche), axial (en bas à gauche) et sagittal (en bas à droite). Ces plans permettent la sélection 3D des points par un double clic de l'utilisateur. Soulignons que l'ordre est important : il est nécessaire d'indiquer en premier point celui de la couronne (pour chaque dent), puis ceux des racines (dans un ordre quelconque). Le dernier quart est occupé par un rendu volumique des structures dures (les os et les dents) qui sert à visualiser le placement des points. L'évolution de la sélection est visible dans les figures 5.4 et 5.5, où deux dents (les 41 et 42, soit les deux incisives mandibulaires droites) ont été sélectionnées. L'utilisateur, lorsqu'il a fini sa sélection, doit ensuite fermer cette fenêtre pour déclencher le processus de segmentation. La première dent est alors segmentée par notre algorithme. La barre de progression indique l'avancée des traitements. Une fois ceux-ci terminés, la fenêtre de contrôle et d'édition manuelle s'ouvre (figure 5.6). De la même manière que pour la sélection, on retrouve la disposition des plans et du rendu 3D, au niveau de la zone d'intérêt de la dent, avec en superposition la segmentation proposée par notre algorithme. Dans le coin inférieur gauche se trouvent les réglages disponibles pour l'édition, à savoir le rayon du pinceau d'édition en coordonnées physiques et si l'édition s'effectue en trois dimensions ou

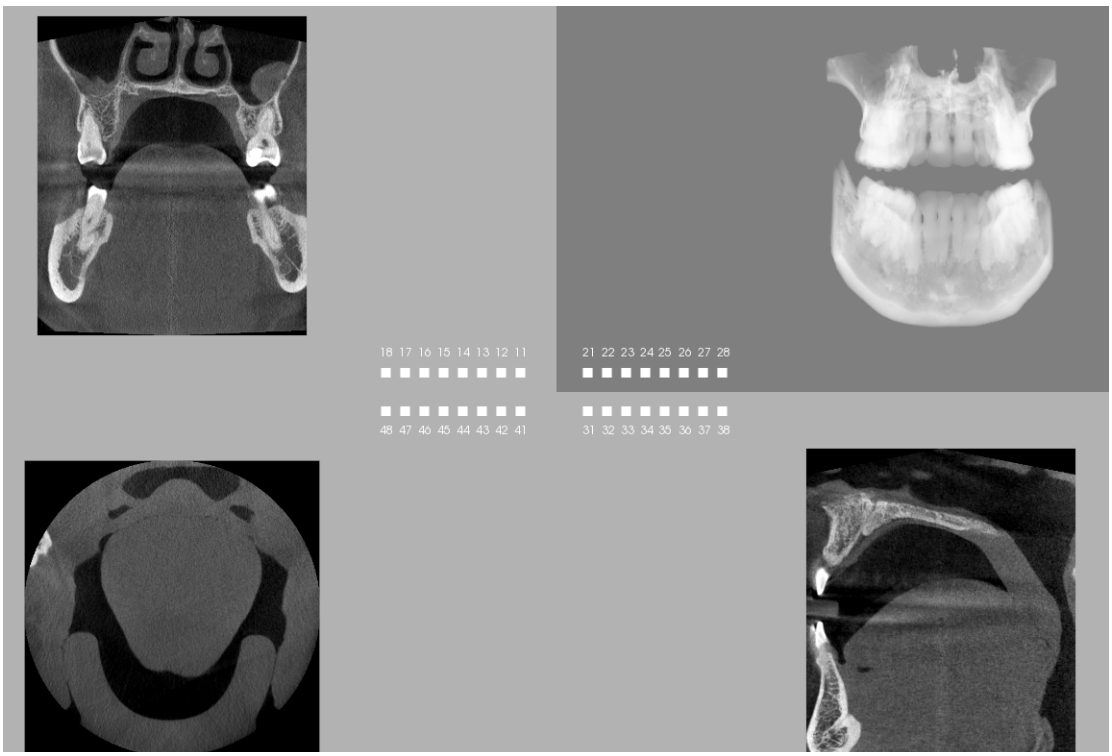


FIGURE 5.3 – Interface graphique correspondant à la seconde étape de traitement : état initial.

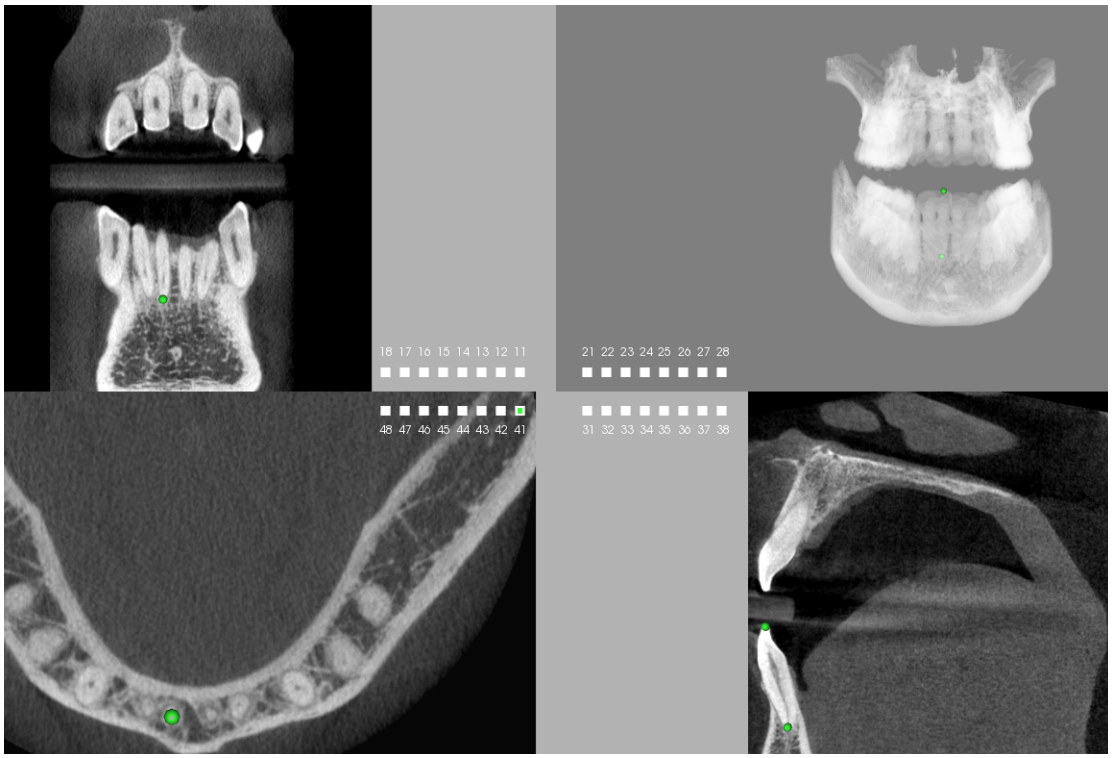


FIGURE 5.4 – Interface graphique correspondant à la seconde étape de traitement : sélection d'une première dent, l'utilisateur coche le numéro dentaire, puis indique les points.

en deux dimensions selon le plan de la coupe utilisée. Dans le cas présent, la segmentation est correcte. L'utilisateur peut donc simplement fermer la fenêtre pour indiquer de passer à la dent suivante. Une fois qu'une dent est terminée, l'algorithme segmente la suivante dans l'ordre de sélection, ici la dent 42 (figure 5.7). Dans ce cas, la segmentation proposée est partiellement erronée. Au niveau de la

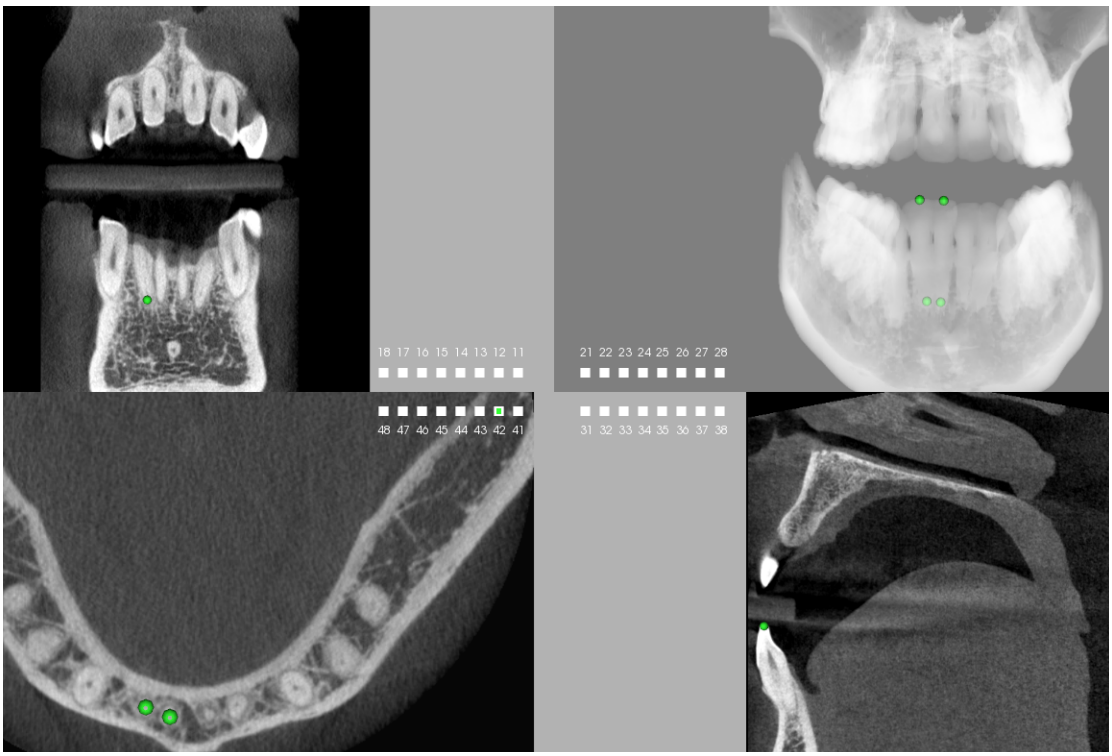


FIGURE 5.5 – Interface graphique correspondant à la seconde étape de traitement : sélection d'une autre dent, de la même manière.

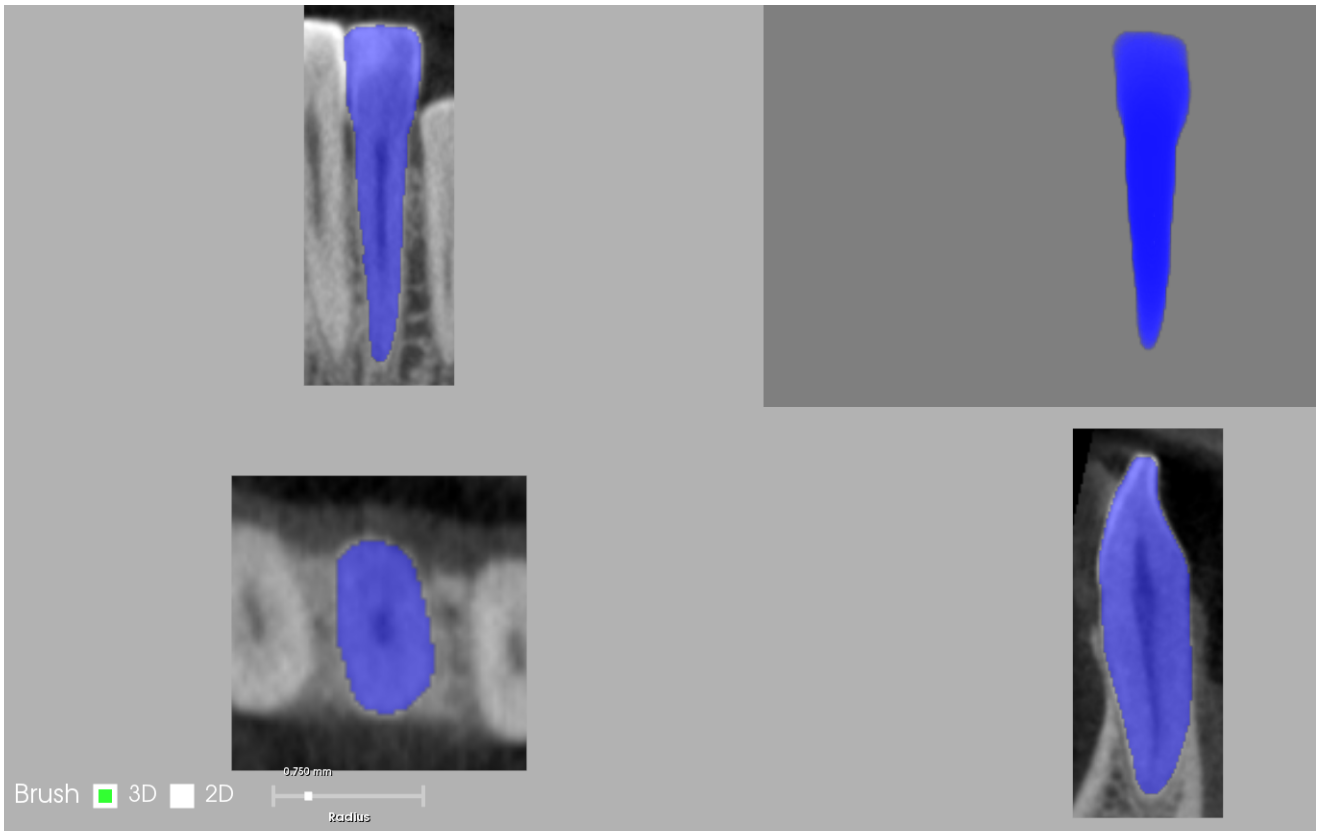


FIGURE 5.6 – Interface graphique correspondant à la troisième étape : segmentation initiale de la dent 41.

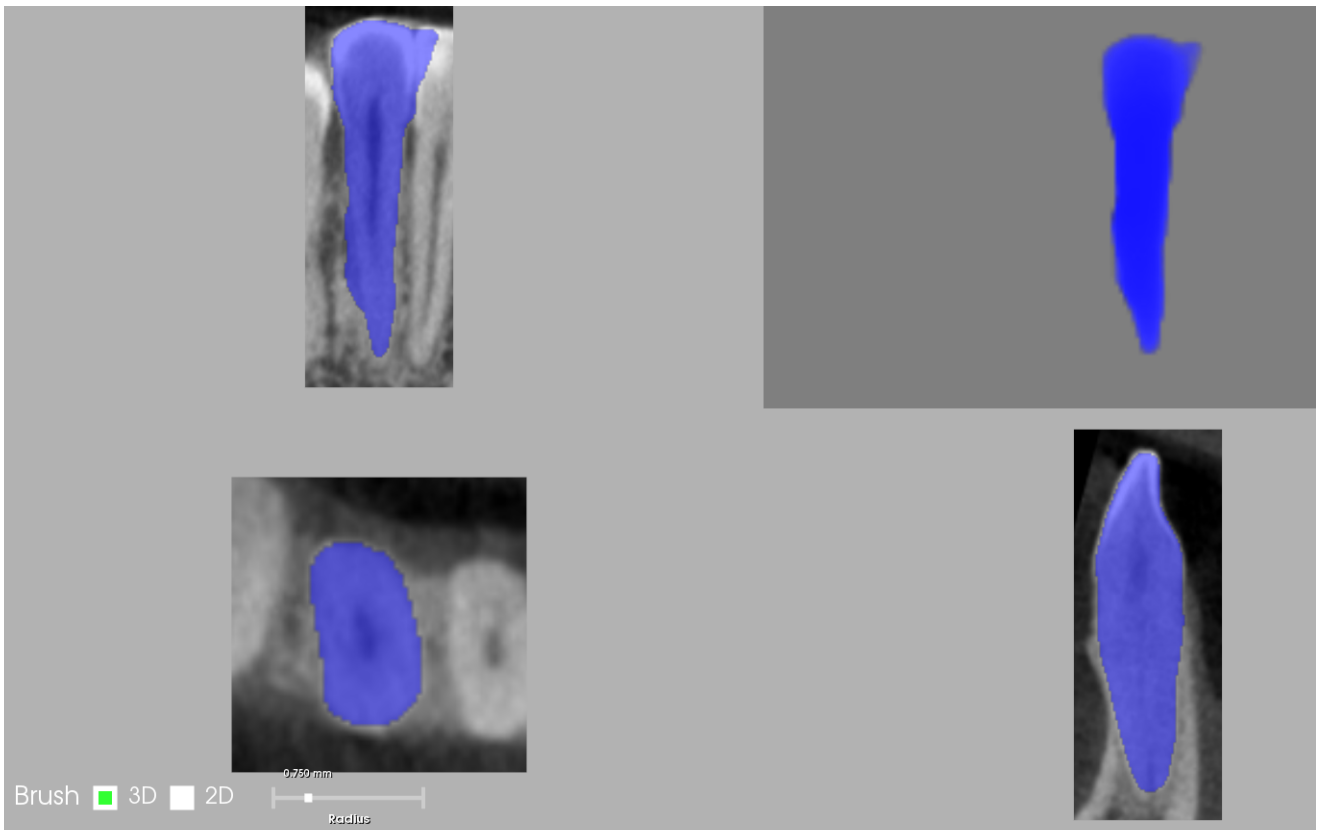
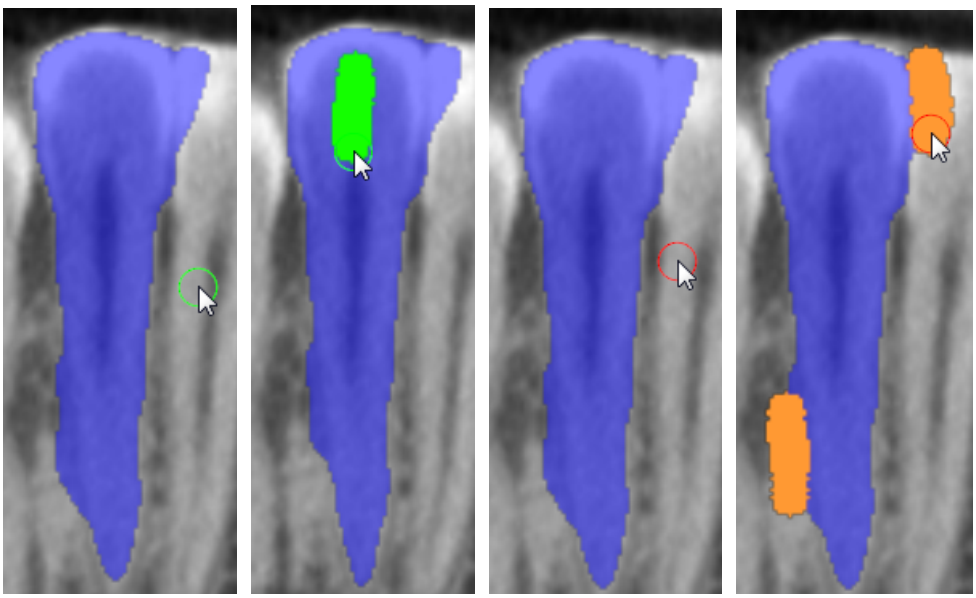


FIGURE 5.7 – Interface graphique correspondant à la troisième étape : segmentation initiale de la dent 42.

couronne, il y a une fuite sur la dent adjacente, ainsi qu'un léger débordement dans la partie inférieure de la racine. L'utilisateur procède alors à une correction. Lorsque ce dernier utilise les touches "Ctrl" et "Shift", apparaissent les outils d'ajout de marqueurs internes et externes respectivement (voir figure 5.8). Par un clic maintenu et déplacé s'ajoutent alors les marqueurs en question. L'édition est répercutée sur l'ensemble des vues. Lorsque la correction est terminée (figure 5.9), il suffit alors de fermer la fenêtre pour relancer l'algorithme avec ces nouvelles informations. Une nouvelle itération de segmentation est faite afin de proposer un résultat amélioré, comme présenté en figure 5.10. Ici, le résultat est satisfaisant. Si tel n'est pas le cas, la correction peut être affinée de manière similaire itérativement jusqu'à satisfaction. Une fois toutes les dents sélectionnées correctement segmentées, l'utilisateur se retrouve sur la fenêtre originelle, lui indiquant que son analyse est terminée et qu'il peut soit recommencer à traiter des dents, soit enregistrer les segmentations (voir figure 5.11). Le dossier de résultats comporte l'image de segmentation globale, ainsi qu'un dossier pour chaque dent individuelle contenant l'image de segmentation, ainsi que le maillage qui en est issu (figure 5.12).

Outil pour développeur

Une version plus modulable du programme a été réalisée : elle permet d'accéder au paramétrage et de le modifier à la volée, pour permettre de l'adapter au contexte d'utilisation, dans un souci de facilité d'utilisation, pour permettre à des experts du traitement d'images de tester l'algorithme sur de nouveaux types de données. Le flux de travail reste le même, mais la fenêtre principale est plus complète, telle qu'elle est montrée dans la figure 5.13. Cette application a pour avantage de permettre un test rapide de nouveaux jeux de paramètres, de visualiser et d'enregistrer les résultats intermédiaires ainsi que de réaliser des comparaisons à une vérité terrain, sans nécessiter de modifier le code et de devoir recompiler le programme. Chaque encadré correspond à une étape de l'algorithme : après avoir sélectionné les dents et leurs numéros (grâce au bouton "Select Teeth" qui ouvre un panneau similaire à celui de la figure 5.3), la zone d'intérêt est extraite, ce qui donne accès au recalage. Une fois le



(a) (b) (c) (d)

FIGURE 5.8 – Aspect des outils de correction manuelle : l'ajout de marqueurs internes est signalé par une couleur verte, tandis que ceux externes sont oranges (rouge pour l'outil).

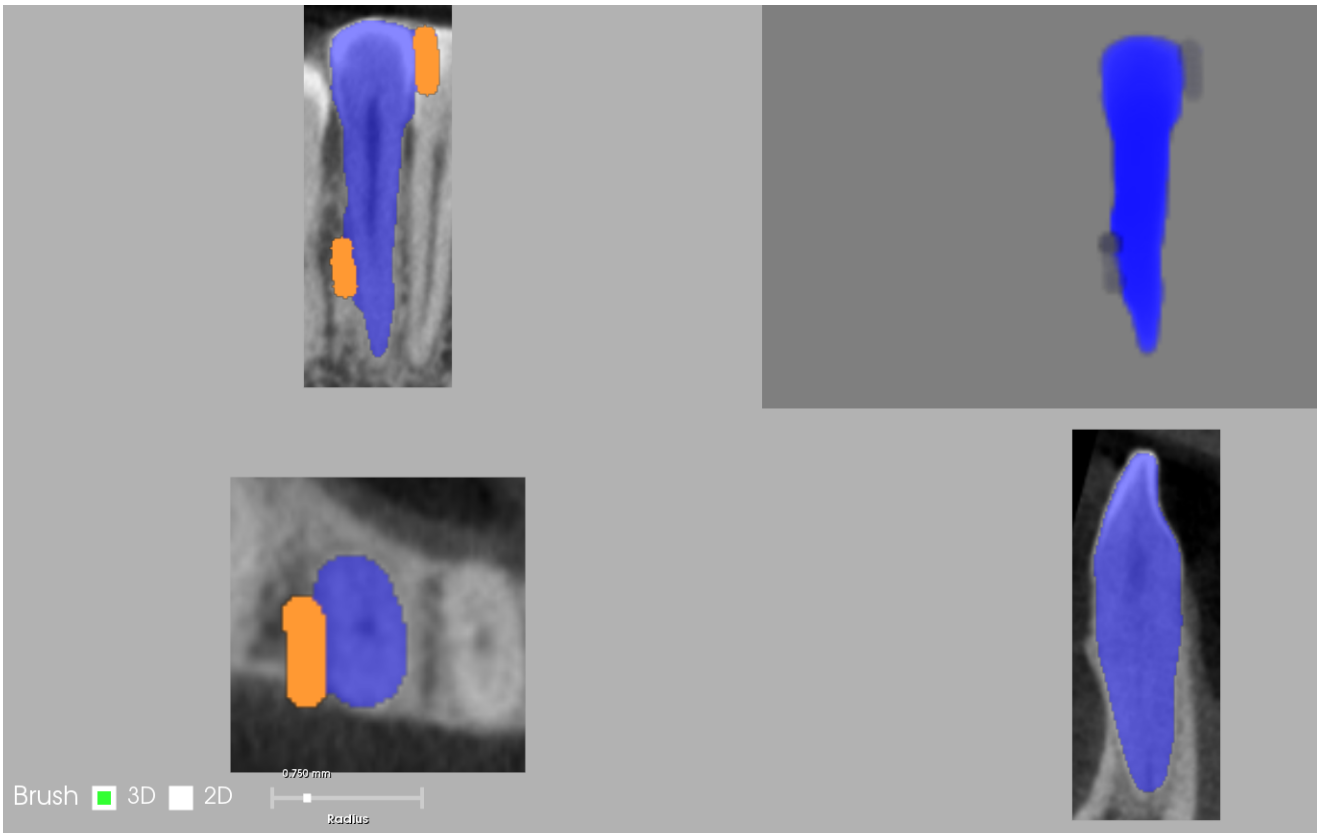


FIGURE 5.9 – Interface graphique correspondant à la troisième étape : correction du résultat de la dent 42.

recalage achevé, la coupure de graphe est effectuée, avant de passer au post-traitement (voir partie 3.3.3). Le dernier encadré correspond au choix des éléments à enregistrer.

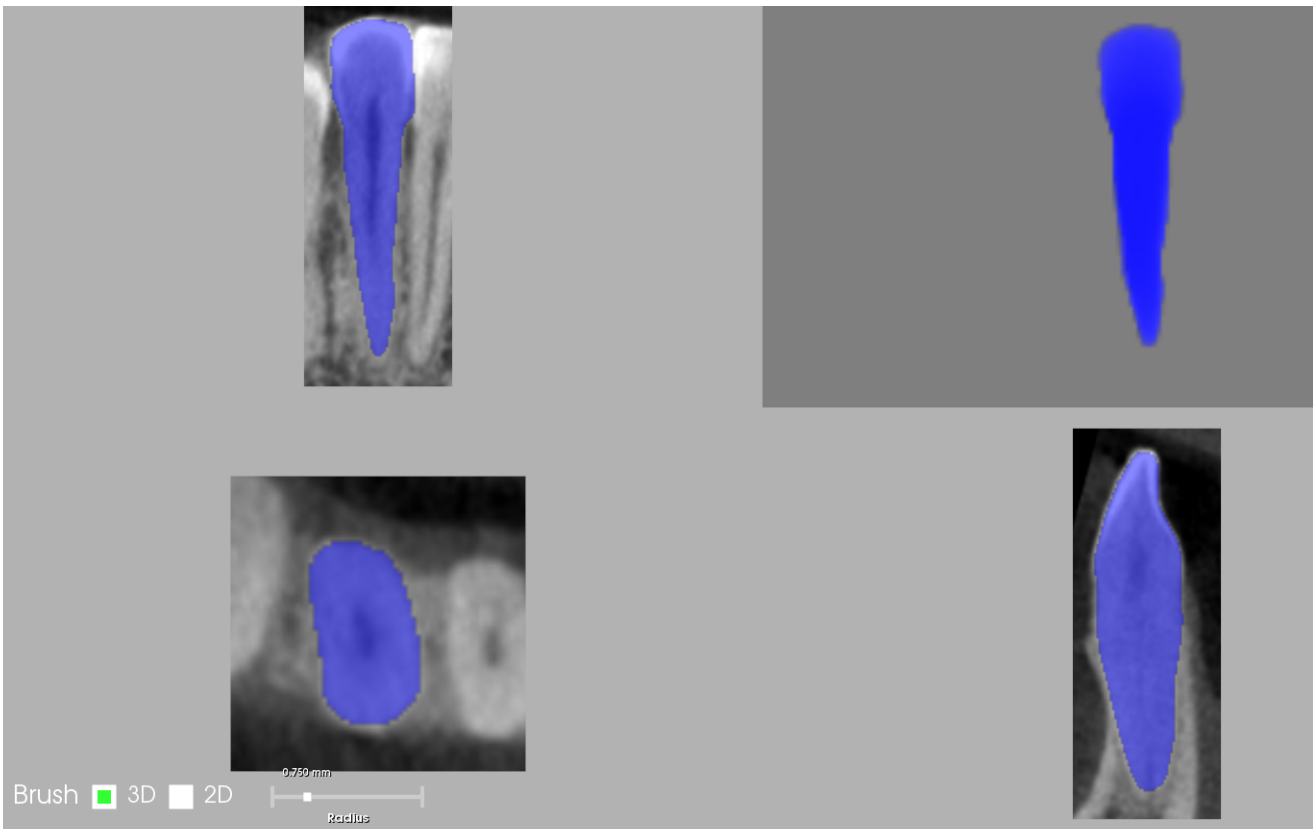


FIGURE 5.10 – Interface graphique correspondant à la troisième étape : segmentation finale de la dent 42.

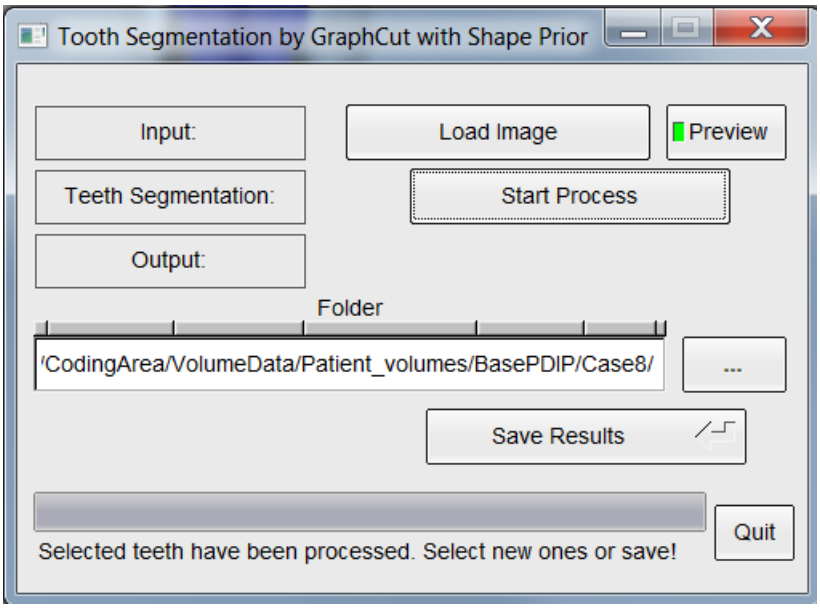


FIGURE 5.11 – Interface graphique correspondant à la quatrième étape : nouvelle analyse ou sauvegarde des données.

5.3 Applications cliniques

Au moment de l'écriture, la validation clinique auprès des experts référents n'a pas encore été réalisée. Néanmoins, les applications possibles pour notre outil sont variées, car la segmentation dentaire intervient dans l'ensemble des spécialités dentaires. Carestream est en relation avec le radiologue

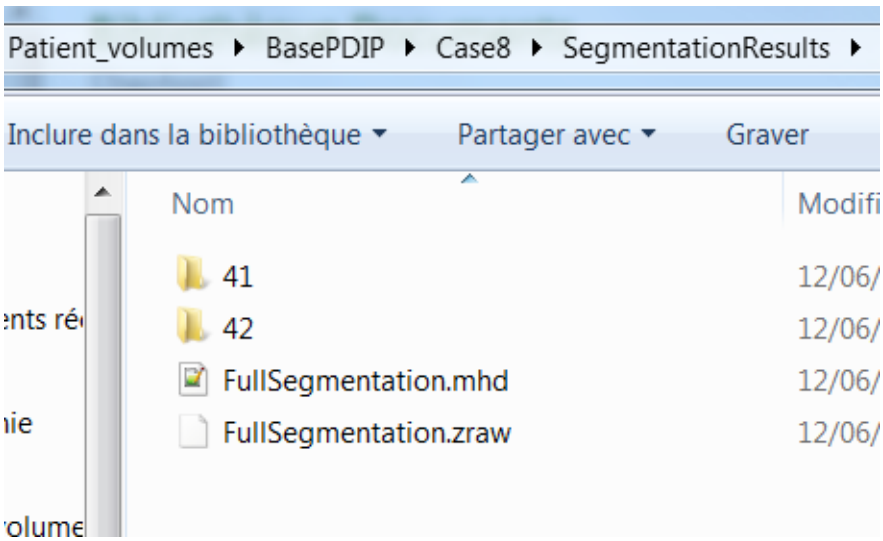


FIGURE 5.12 – Organisation des résultats enregistrés.

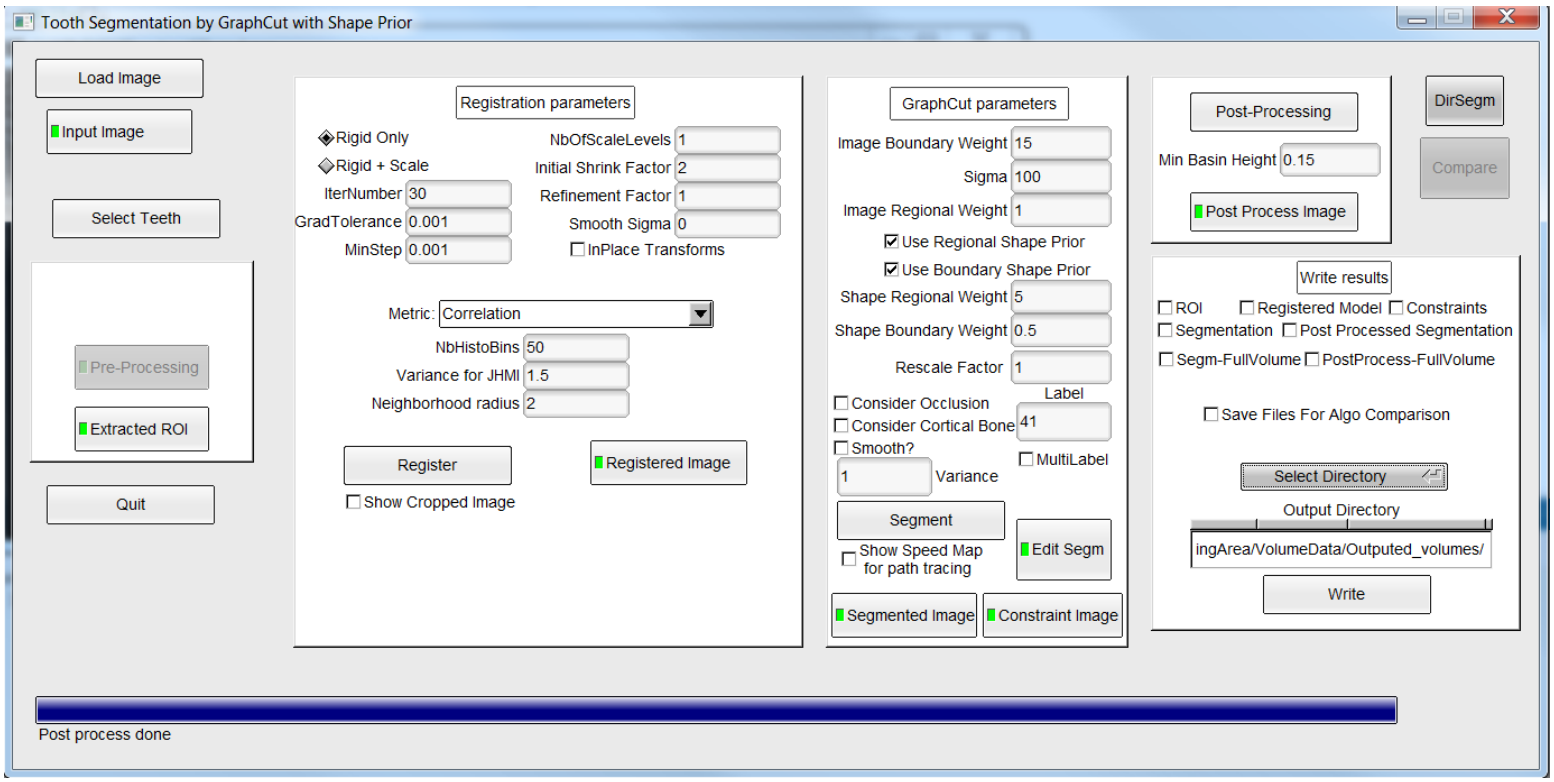


FIGURE 5.13 – Fenêtre de contrôle de l'outil pour développeur.

Jacques TREIL, qui travaille à développer et à démocratiser l'utilisation de la céphalométrie 3D [50]. La céphalométrie s'attache au calcul des mesures anatomiques de la tête, qui permettent de réaliser des diagnostics et des planifications de soins. Actuellement, la majorité des mesures céphalométriques (environ une vingtaine) sont toujours réalisées sur un cliché 2D latéral, ce qui complique l'interprétation à cause de la superposition des structures. L'ajout de profondeur permet d'envisager l'identification de nouvelles structures anatomiques, pour des mesures plus pertinentes et plus informatives, conduisant à un diagnostic plus complet. Afin de faciliter cette tâche, le Dr. TREIL à réalisé un modèle qui intègre ces nouvelles informations sous une forme synthétique afin de pouvoir réaliser des comparaisons entre patients, ou de suivre l'évolution d'un même patient. En particulier, les dents sont représentées dans le modèle par leur centre d'inertie et leurs directions principales. L'acquisition de ces informations passe

par leur segmentation, qui est actuellement faite de manière semi-automatique par un outil chronophage (entre 45 minutes et 1 heure pour une arcade complète). Nous projetons donc de proposer notre outil, intégrant les calculs automatiques d'axes d'inerties et de barycentre, afin d'optimiser la création du modèle. L'objectif est de 5 minutes pour une arcade complète.

Un autre débouché en développement est celui de la segmentation pour l'endodontie. Pour rappel, cette spécialité concerne la santé de la pulpe dentaire et de l'espace péri-apical pour le traitement des abcès et la dévitalisation des dents. Une fois la dent délimitée, il devient plus aisé d'identifier et de segmenter les structures qui la composent, telles que l'émail ou les canaux radiculaires. Le but premier est de pouvoir proposer une procédure de soin pour la dévitalisation :

— évaluation de l'atteinte des tissus,

— proposition des limes optimales, avec un diamètre et une longueur adaptées, ainsi que la quantité nécessaire de consommables (comme la gomme de comblement),

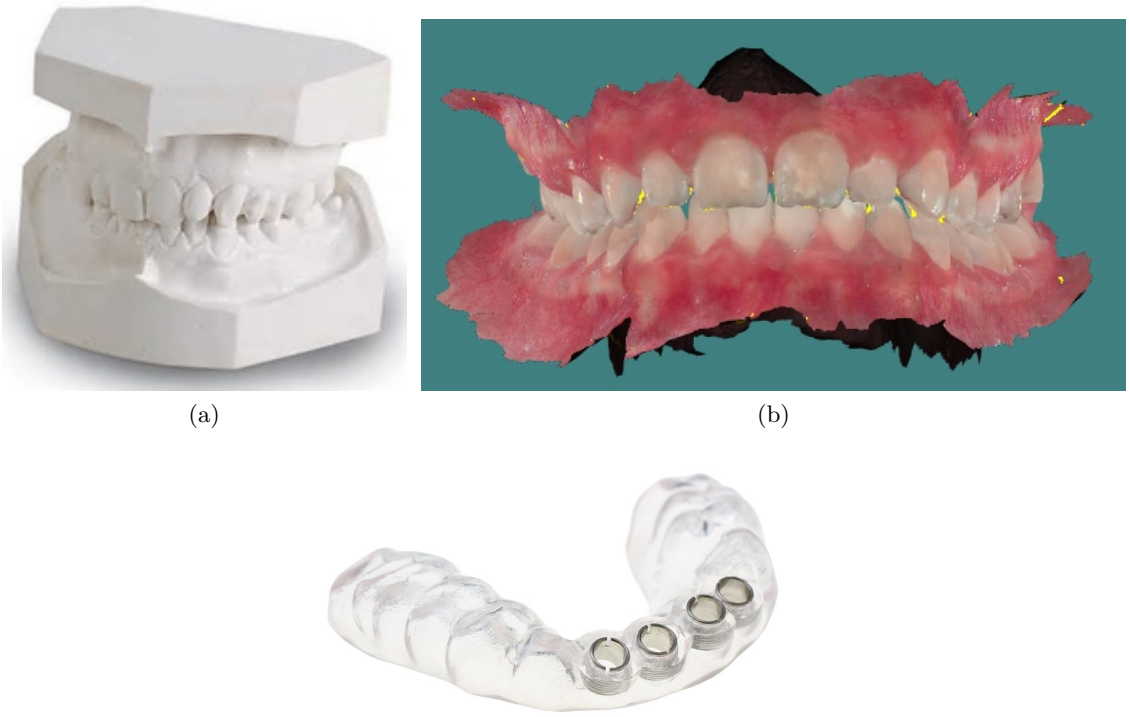
— proposition du planning opératoire (point d'entrée, axe de forage ou encore épaisseur de tissu à éliminer),

— simulation de la procédure et rendu final.

Enfin, le dernier grand projet est celui du "PDIP" (*Patient Driven Implant Planning*), qui vise le domaine de l'implantologie. Dans l'optique de faciliter la pose d'implants, ce projet vise à fusionner les informations issues des volumes de patients et celles provenant de la numérisation des modèles d'étude (appelés couramment "plâtres", voir figure 5.14 (a)) ou du scan intra-oral (reconstruction surfacique des arches acquise à partir d'une caméra numérique intra-orale, (b)). Le maillage est recalé automatiquement dans l'image CBCT puis les couronnes et la gencive qui le composent sont segmentés. Du côté de l'image, les dents sont aussi segmentées, en tenant compte des délimitations introduites par le maillage. De cette manière est obtenue une segmentation plus précise des dents grâce à une bonne délimitation pour les structures externes grâce au maillage, notamment pour les volumes en occlusion (figure 5.15), où l'image tend à avoir une interface entre les couronnes difficilement identifiable. Le praticien peut alors créer une couronne virtuelle en fonction de l'environnement dentaire, puis prévoir un implant adapté (voir figure 5.16). Une planification optimale pour sa pose peut alors être définie en fonction des segmentations obtenues et il est possible de réaliser automatiquement un guide chirurgical (voir figure 5.14 (c)).

5.4 Perspectives

Si globalement le prototype remplit les critères que nous nous étions fixés, deux étapes importantes sont encore à franchir. Premièrement, la validation clinique est nécessaire pour s'assurer que la forme graphique utilisée est suffisamment souple et pratique pour répondre aux problématiques rencontrées par les dentistes. Il nous faut donc déployer ce programme auprès d'experts qui pourront tester et nous donner un retour à la fois sur la partie prise en main et sur la performance de l'algorithme. Deuxièmement, le temps nécessaire à la réalisation d'une denture complète est encore trop loin de l'objectif des cinq minutes visées. Deux pistes peuvent permettre de progresser sur cet aspect : d'un côté, notre procédure de sélection des dents par l'utilisateur est certes adaptée pour permettre une utilisation dans toutes les spécialités dentaires, mais elle tend à être fastidieuse pour une denture complète, il serait donc pertinent de prévoir une modification de l'interface et/ou de réaliser une extraction automatique des points actuellement donnés par les utilisateurs pour alléger la charge de travail. La seconde piste est d'améliorer le temps de calcul de la segmentation de chaque dent : cela peut passer soit par une optimisation de la structure du code (ce qui est appelé "remaniement") soit par l'application de méthodes de parallélisation, qui permettent de tirer profit de plusieurs cœurs de processeurs. Notons qu'il nous paraît plus pertinent de nous occuper en priorité de l'aspect utilisation et interface, en réduisant le temps nécessaire à l'utilisateur, car le service rendu est meilleur si une interaction est peu requise, quitte à laisser les calculs s'effectuer, que si le praticien doit rester disponible durant tout le processus. Enfin, il est intéressant de permettre l'accès et l'utilisation de données distantes, notamment celles des systèmes d'archivage et de transmission d'images (*PACS*), ou encore celles des bases publiques en ligne. À terme, ce prototype pourra être fusionné à la suite logicielle

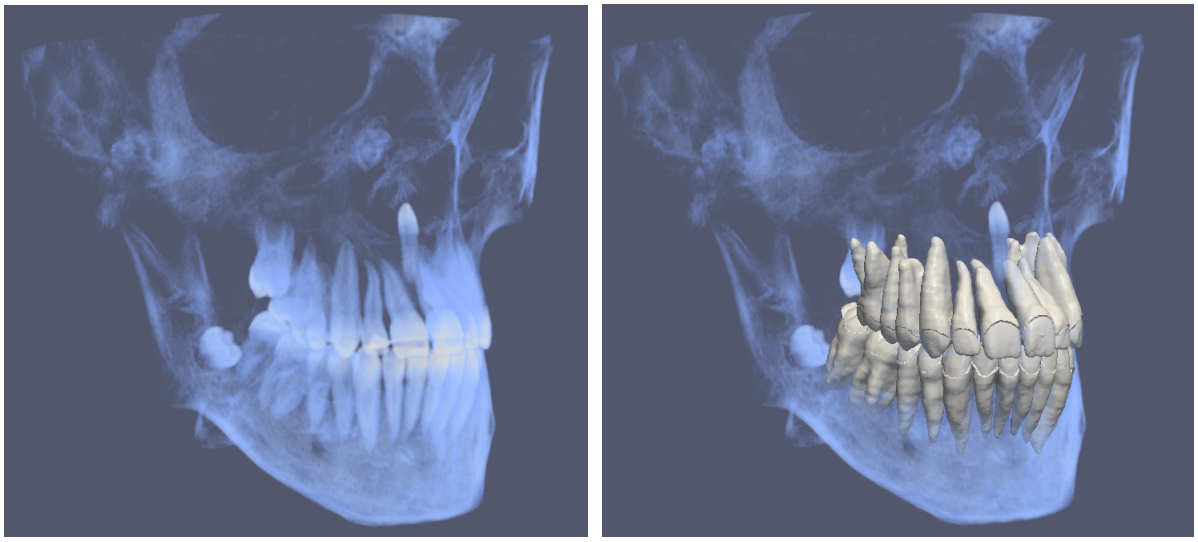


(a) (b)

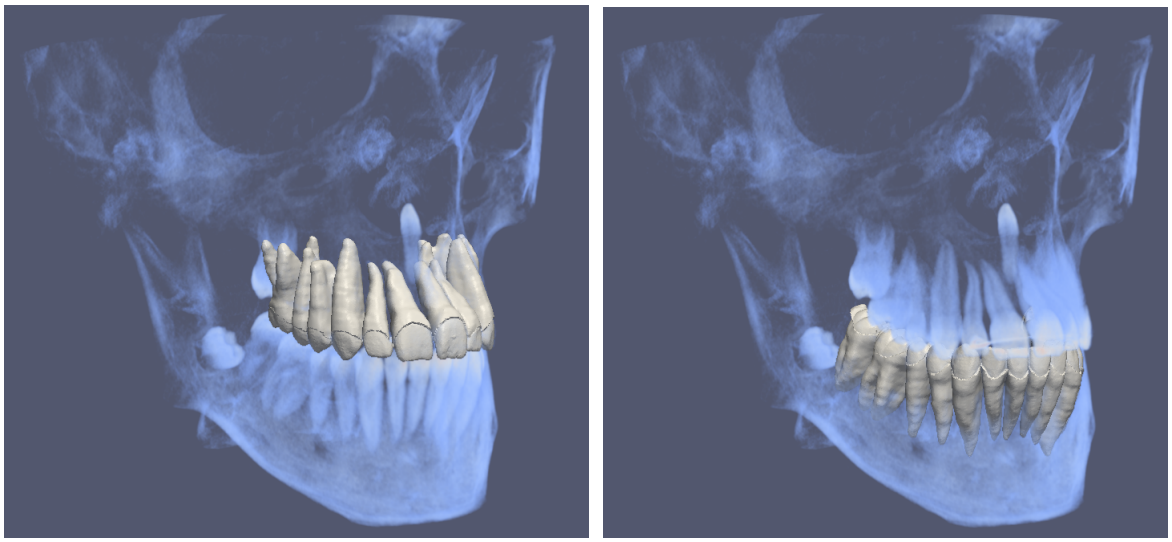
(c)

FIGURE 5.14 – Exemple de : (a) modèle d'étude dentaire, ou "plâtre", (b) scan intra-oral, (c) guide chirurgical, qui s'adapte à la denture afin de faciliter la pose d'implant.

existante de Carestream, afin de proposer cette méthode à l'ensemble de la base d'utilisateurs, que ce soit comme étape d'une fonctionnalité plus évoluée (calcul automatique d'indicateurs cliniques par exemple) ou comme outil directement au service du praticien.

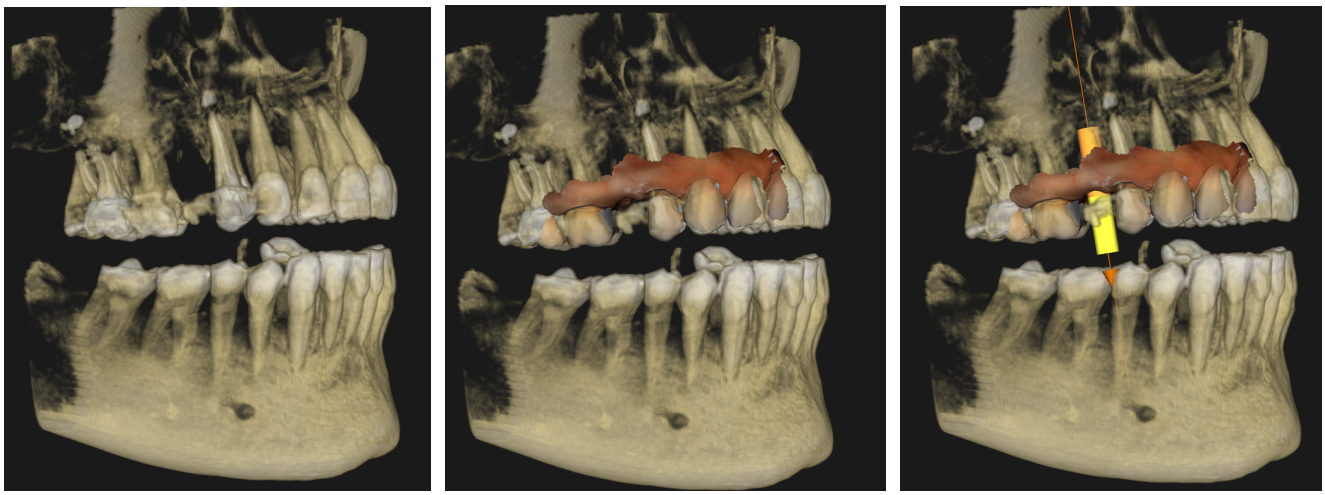


(a) (b)

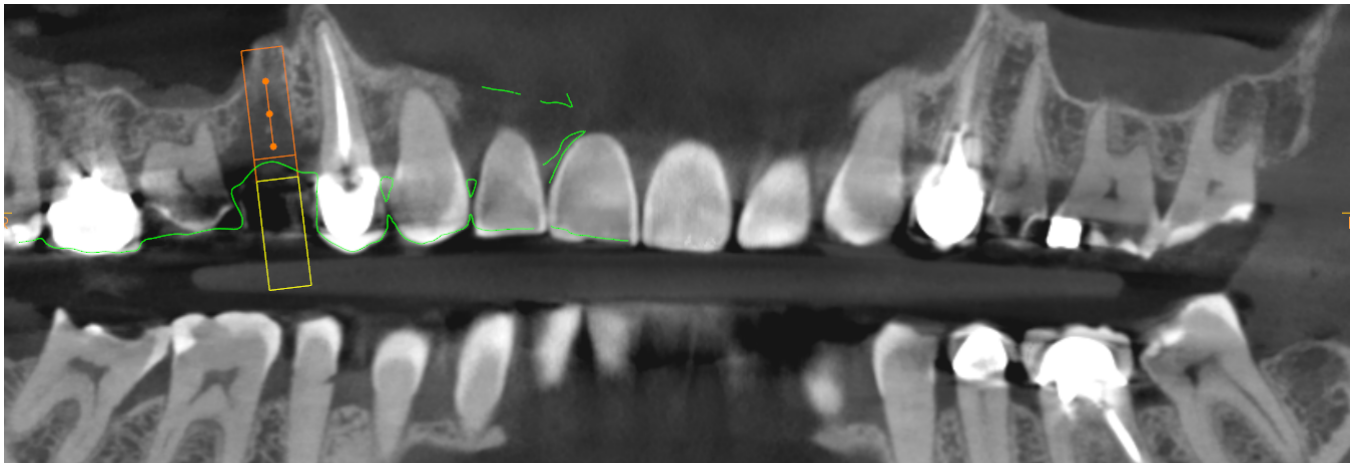


(c) (d)

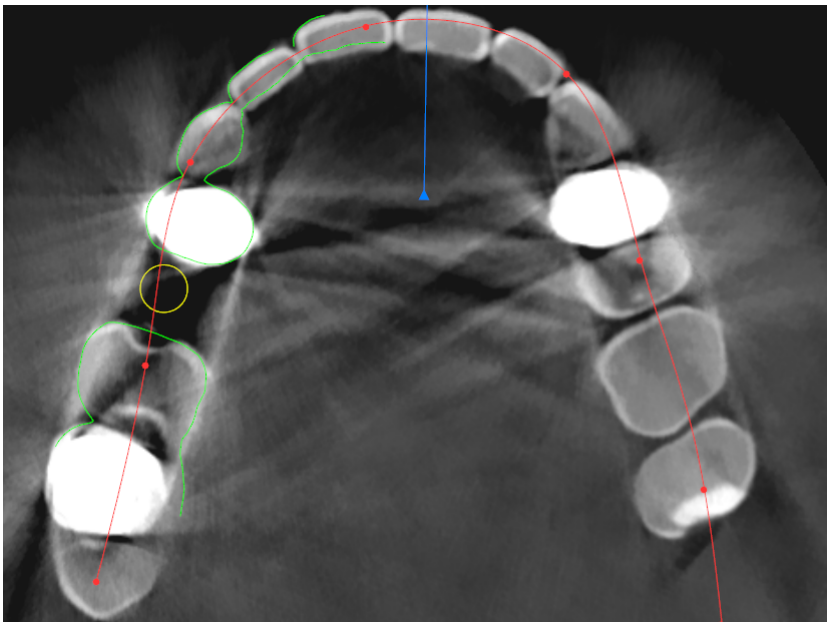
FIGURE 5.15 – Exemple de segmentation dentaire contrainte par le maillage recalé dans le cadre du projet PDIP (*Patient Driven Implant Planning*). Toutes les dents sont correctement segmentées grâce à la connaissance de l'interface entre les couronnes provenant du maillage.



(a) (b) (c)



(d)



(e)

FIGURE 5.16 – Illustration du projet PDIP. (a-c) Vue 3D des structures dures (os et dents), avec recalage du maillage du scan intra-oral puis ajout d'un implant. (d) Vue pseudo-panoramique, où est représentée en vert la limite du maillage qui épouse correctement la forme des dents, ainsi que l'implant en jaune-orangé. (e) Vue axiale, reprenant les codes de couleurs de la vue pseudo-panoramique.

Chapitre 6

Conclusion et perspectives

6.1 Apport de la thèse

6.1.1 1er axe : une méthode de segmentation dentaire

Nous avons présenté dans le chapitre 3 une nouvelle méthode de segmentation semi-automatique d'une dent dans des images volumiques de patients acquis en CBCT, reposant sur l'utilisation d'un modèle statistique de la forme des différentes catégories dentaires pour la création de contraintes a priori, utilisées dans un schéma d'énergie formulée sur un graphe, et optimisée par des coupures de graphe. Les imperfections de délimitation sont corrigées par un traitement final utilisant de la morphologie mathématique pour aboutir à la segmentation finale. Il est demandé à l'utilisateur d'indiquer, pour chaque dent, deux éléments : un ensemble de points, contenant un point sur la couronne, ainsi qu'un point à l'extrémité de chaque racine, et le numéro dentaire correspondant. Cette méthode est jugée compatible avec la pratique clinique (par les experts cliniques Carestream) car nécessitant moins de temps que les méthodes pré-existantes, et utilisable sur les configurations courantes d'ordinateur dans les cabinets (c'est-à-dire un processeur bi-cœurs, 4 GB de RAM et sans GPU dédié). Cette contribution a été publiée [49] et présentée oralement lors de la conférence ISBI2017.

Limites de la segmentation

Si, pour la plupart des cas, notre méthode permet d'obtenir des résultats du niveau de l'état de l'art avec un indice de Dice moyen de $0,91$, la segmentation tend à devenir erronée dès lors que la résolution et la qualité d'images sont en baisse (particulièrement pour des tailles de voxels de 200 microns ou plus), car les structures permettant une délimitation efficace des dents (notamment le ligament parodontal) ne sont plus visibles : le score moyen de Dice est par exemple d'environ 0,95 pour des volumes de taille de voxels isotrope de 150 microns. Le deuxième élément ayant un fort impact sur la qualité des résultats est la présence de métal dans les images. Selon la reconstruction volumique (création de l'image à partir des projections angulaires) effectuée, le métal induira des faux contours de manière plus ou moins importante. La segmentation s'appuyant en partie sur les gradients, ces faux contours conduiront à perturber la délimitation à proportion de leur importance. Si nous avons montré qu'il était possible de segmenter une denture, la problématique du métal n'a donc pas été résolue et reste entière. D'un point de vue méthodologique, nous avons aussi souligné que si l'emploi de contraintes a priori de forme permettait d'obtenir une bonne segmentation, notre utilisation du modèle statistique pour générer les contraintes de forme n'était pas optimale : en effet ces dernières reposent sur la forme moyenne uniquement, alors que les modes de déformations servent à générer un espace de segmentation possible. Plus la différence entre l'objet et la forme moyenne sera importante, plus la pertinence de ces contraintes sera faible. Pour améliorer cela, nous avons évoqué la possibilité de recourir à des modèles statistiques conditionnels, nous permettant de générer une forme adaptée à l'image grâce à des mesures anatomo-morphologiques, pour assurer la pertinence de la contrainte, mais ces travaux n'ont pas été commencés au moment de l'écriture. Pour finir, si les dentistes apprécient d'avoir un contrôle sur le processus de segmentation et que le choix d'une méthode interactive n'est pas remis en cause dû à la difficulté de repérer automatiquement les points dentaires (notamment

radiculaires), la quantité d'interaction demandée peut devenir un frein, surtout lors de l'analyse d'une denture complète (ce qui conduit à positionner environ une cinquantaine de points dans l'espace), et nous estimons que la question de l'interaction n'est pas close.

6.1.2 2e axe : un modèle structurel maxillo-facial

Le chapitre 4 contient la description d'un modèle structurel contenant les organes identifiés comme étant d'intérêt clinique et apparaissant dans les images, représenté sous la forme de graphe conceptuel pour faciliter la compréhension des relations inter-organes. Ce modèle a pour vocation de constituer le socle d'un système à base de connaissances, en intégrant toutes les informations dont nous disposons sur les structures et sur les relations qui les lient, de manière à disposer d'éléments permettant de circonscrire chaque problème et de créer de nouvelles méthodes pour analyser le contenu de l'image. Chaque élément (nœud) du modèle renvoie à un ensemble de données sous la forme de relations spatiales, de modèles statistiques, de forme, d'intensité ou encore d'apparence. Nous avons à cette occasion étendu au domaine tridimensionnel deux définitions de relations spatiales, que sont l'alignement et "le long de", cette dernière ayant fait l'objet d'une publication [48] ainsi qu'une présentation orale lors de la conférence ICIAP2015. L'alignement a été identifié comme une caractéristique des arcades dentaires, tandis que le canal est considéré comme longeant l'os mandibulaire. Ces relations spatiales permettent notamment la représentation des liens de position sous la forme de cartes spatiales utilisables pour conditionner les méthodes d'identification/segmentation.

Usage manuel du modèle

Dans l'état actuel, le modèle réalisé est comparable à une encyclopédie qui permet de caractériser d'un point de vue structurel ce qui se trouve dans les images des patients. L'apport de cette connaissance est donc tributaire de son inclusion manuelle dans des approches qui restent à définir. Ce modèle nécessite donc d'être entretenu et enrichi pour apporter une réelle plus-value. De plus, il n'est défini que pour les cas sains, or nous savons que les pathologies peuvent conduire à une modification substantielle des caractéristiques des organes et de leur positionnement.

Validation de l'apport

Nous n'avons pas expérimenté suffisamment pour prouver l'apport concret de l'utilisation d'un tel modèle, en particulier concernant les relations spatiales. La question de l'extraction des structures difficilement identifiables de manière directe, telles que le canal mandibulaire, n'a donc pas été résolue et reste ouverte. De même, l'apport du formalisme des ensembles flous, utilisé lors de la modélisation des relations spatiales pour intégrer leur caractère vague, n'est pas démontré ici, mais il l'a été dans d'autres applications (voir [3, 17, 52, 75, 98]).

6.1.3 3e axe : un logiciel pour praticiens

La dernière contribution, objet du chapitre 5, est la création d'un logiciel intégrant notre méthode de segmentation dentaire. Ce logiciel est pourvu d'une interface graphique qui permet la segmentation des dents dans les volumes de patients de manière intuitive et pratique pour tout utilisateur issu du monde médical. Un outil de correction interactive de la segmentation a été adjoint à notre algorithme pour permettre la validation manuelle des résultats. Un soin particulier a été apporté à la conception du logiciel de manière à faciliter son utilisation : la méthode est paramétrée pour correspondre au plus grand nombre d'images possible, le processus est pensé pour répondre aux besoins des différentes spécialités dentaires, tandis que l'interface et ses contrôles imitent ceux des autres logiciels existants afin de ne pas déstabiliser l'utilisateur. Les résultats obtenus sont facilement accessibles, permettant l'usage du prototype dans des procédures d'analyses plus complexes.

Absence de validation clinique

Au moment de l'écriture de ce mémoire, nous n'avons pas eu le temps de distribuer le prototype logiciel aux dentistes partenaires. Notre méthode de segmentation ainsi que l'ergonomie d'utilisation

n'ont donc pas été testées en clinique. Ce manque de retour de la part des praticiens signifie que malgré le fait que nous ayons veillé à nous mettre à la place de l'utilisateur, des défauts sont à envisager. Les points essentiels qui doivent être contrôlés sont de deux types. Premièrement, d'un point de vue méthodologique, est-ce que la qualité de segmentation est suffisante ? À quel degré de précision est-il nécessaire d'aller pour l'usage qu'en ont les différents dentistes ? Le temps de calcul est-il raisonnable ? Deuxièmement, d'un point de vue pratique, est-ce que la forme logicielle répond aux attentes et aux besoins ? Combien de temps l'utilisateur doit-il réellement prendre pour réussir son analyse, et est-ce compatible avec la pratique clinique quotidienne ? Pour chacune de ces questions, le critère important sera de savoir si cela entrave l'adoption de la méthode, définissant les pistes d'améliorations prioritaires.

6.2 Perspectives

6.2.1 La segmentation en CBCT maxillo-facial

Les pistes qui restent à explorer pour la segmentation sont nombreuses. Au niveau de l'identification des dents, les prochaines étapes sont d'étendre l'utilisation et la pertinence des a priori de forme par une meilleure adaptation à l'image, en créant un nouveau modèle statistique intégrant des mesures anatomo-morphologiques (le volume de la chambre pulpaire, le diamètre au collet ou le nombre de cuspides par exemple) en employant une approche similaire à celles décrites dans [12, 13] (où ces caractéristiques sont calculées sur l'image étudiée afin de conditionner la forme issue du modèle), et d'améliorer la segmentation multi-objets de manière à la rendre utile (moins d'interaction, en temps similaire). Pour ce faire, nous estimons prometteur l'adaptation de notre formulation à une optimisation par l'algorithme du *powerwatershed*, qui présente de bien meilleures performances d'un point de vue calculatoire que la coupure de graphe, et ce proportionnellement à la taille de l'image. Une autre de piste serait d'utiliser la méthode de segmentation mono-objet de manière séquentielle en incluant à chaque itération des contraintes données par les segmentations déjà effectuées (par exemple de non recouvrement). Notons qu'en parallèle, la délimitation de chaque dent nous permet d'envisager l'analyse de ses parties constitutives, à savoir l'émail, la dentine et ses canaux, ce qui intéresse la spécialité d'endodontie. Une amélioration de l'apport de la méthode serait de réfléchir à la chaîne de traitement plus globalement, par exemple en incluant une détection préalable des cas difficiles (notamment en présence de métal) pour alerter l'utilisateur, qui pourrait alors décider de directement segmenter manuellement pour gagner du temps ou ajouter des marqueurs supplémentaires dans la segmentation.

Au niveau du canal mandibulaire, nos espoirs résident dans une approche par contraintes. Si nous parvenons à segmenter les structures environnantes (denture, os mandibulaire) il est possible de réduire la zone de recherche ainsi que de définir ses extrémités pour parvenir, si ce n'est à sa segmentation, au moins à son estimation et à la création d'une zone de danger si les racines dentaires sont trop proches, ce qui est une information essentielle pour toute intervention sur la partie inférieure de la denture. Une approche par modélisation similaire à celle de [77] (modèle statistique puis recherche de chemin) enrichie avec la denture, et définissant de meilleurs critères de suivi de chemin (notamment reposant sur les relations spatiales, empêchant l'intersection avec l'os cortical ou les racines dentaires) est aussi une piste intéressante, mais nous permet de soulever un autre point critique pour l'analyse des images : les données annotées.

En effet, à ce jour, il n'existe aucune base d'images publique concernant notre application, et les images segmentées manuellement par des experts sont très rares, et souvent incomplètes (c'est-à-dire que les annotations ne concernent qu'une fraction des structures d'intérêt). Dans ce travail, nous n'avions qu'une vingtaine de volumes avec annotations d'experts (15 avec dents annotées, 6 pour le canal mandibulaire), ce qui est très nettement insuffisant pour représenter efficacement la diversité existante des morphologies, sans même évoquer les pathologies. L'obtention de ces ressources est, à nos yeux, une condition *sine qua non* pour faire progresser la recherche dans le domaine, notamment avec les récents progrès fait par les méthodes d'apprentissage profond pour la segmentation. Il reste donc un travail important à effectuer, à la fois dans le recrutement d'experts, dans l'utilisation des méthodes d'augmentation de données, et surtout dans la communication publique pour souligner l'intérêt de santé publique à la constitution de bases d'images et de challenges dans le domaine (similaire à ce qui

existe pour les pathologies cérébrales).

Notamment, la disponibilité de segmentations manuelles par plusieurs experts des mêmes données permettrait d'effectuer une analyse approfondie de la variabilité inter-utilisateur et de la précision attendue des résultats, ce qui soulignerait mieux l'apport et le positionnement des méthodes proposées, et donnerait éventuellement des pistes d'améliorations.

Finalement, nous avons observé que la qualité d'image est un élément critique quant aux possibilités de reconnaissance et d'extraction des structures. Une évolution du procédé d'acquisition est actuellement en cours, avec de nouveaux capteurs qui réduisent le flou et le bruit obtenu par une meilleure collimation (ce qui empêche la dispersion des photons, réduisant le bruit) et de meilleurs scintillateurs, c'est-à-dire les matériaux utilisés pour la capture et la traduction du signal photonique en signal électrique, ce qui réduit le flou. Cette amélioration nous semble être un élément supplémentaire qui conduira à de meilleurs résultats. Des algorithmes de reconstruction volumique intégrant la réduction des artéfacts métalliques sont aussi une voie importante pour l'obtention d'images de qualité.

6.2.2 Vers une conception globale du problème d'analyse d'images CBCT dentaires

Avec la constitution de notre modèle structurel nous avons ébauché les premières étapes d'une approche plus générale de l'analyse des images de patients : utiliser les connaissances acquises dans d'autres domaines sur le contenu pour enrichir les méthodes d'extraction. En formalisant toutes les connaissances a priori que nous possédons de notre problème, il est alors possible de contraindre efficacement le problème de reconnaissance et/ou segmentation pour aboutir à de meilleurs résultats. Cela revient à expliciter les relations qui existent implicitement dans les approches par atlas, tout en améliorant la souplesse du modèle, puisqu'il est alors possible de relâcher les contraintes (de forme, d'aspect, de position, etc.) selon l'application.

D'un point de vue méthodologique, les deux voies que nous avons envisagées pour aller plus loin sont directement inspirées par la littérature : dans [98] le modèle est considéré comme un réseau de contraintes que l'image doit satisfaire, alors que dans [52] est décrite une approche itérative qui consiste à parcourir le modèle comme un graphe en résolvant une segmentation à chaque étape. La seconde semble dans notre cas plus adaptée pour plusieurs raisons. Dans un premier temps, notre modèle s'attache à décrire l'ensemble de l'espace maxillo-facial, alors que nous savons que le champ de vue réel des images est variable selon le type d'acquisition. Utiliser le modèle comme réseau de contraintes demanderait alors d'identifier précisément quelle sous-partie du modèle est contenue dans l'image, ce qui n'est pas trivial automatiquement, et fastidieux pour un utilisateur. À l'inverse, une approche itérative, sous condition de posséder une initialisation (donnée par un utilisateur ou obtenue automatiquement), permet un parcours qui s'arrête automatiquement en l'absence de structures à segmenter : cela passe par une méthode de choix de parcours et de contrôle des résultats, qui est de toute façon nécessaire dans cette approche, pour éviter la propagation d'erreur et permettre la résolution des cas d'impasse. De même, il est possible de n'explorer qu'une sous-partie du modèle si l'analyse ne porte que sur une fraction de l'image. Le point le plus critique de l'approche itérative est la définition de la séquence de parcours et des critères d'évaluation : dans [52] ces derniers reposent principalement sur des correspondances entre ensembles flous issus des relations spatiales ; dans notre cas, il nous semble utile d'essayer d'enrichir ces critères avec des informations de forme et d'apparence compte tenu des structures particulières du domaine dentaire.

En parallèle, il est nécessaire de réaliser l'enrichissement et l'adaptation des connaissances aux cas pathologiques, notamment pour les pathologies déformantes (modifications des relations spatiales et des formes), ou ayant un fort impact radiométrique. Dans le champ d'application dentaire, elles sont peu nombreuses et plutôt localisées, mais le cas de l'édentation (totale ou partielle) entraîne une perte d'éléments centraux du modèle. En revanche, deux spécificités induites par l'application dentaire nous semblent être de bons candidats pour l'extension du modèle : les cas pédiatriques et les dispositifs médicaux. Dans le premier cas, de nombreuses images sont acquises sur des patients préadultes, notamment en orthodontie, et cela entraîne nombre de différences notables dans l'anatomie (par exemple la denture temporaire). Dans le second cas, les procédures de soin entraînent la pose

de nombreux éléments exogènes, constitués en forte proportion de métal. Bien qu'il soit difficile de prévoir l'impact réel dans l'image de tels dispositifs, il n'en reste pas moins qu'il est possible de prévoir une dégradation de la pertinence du modèle selon l'emplacement de pose : une couronne métallique conduira invariablement à une dégradation de la forme coronale dans l'image ; un appareillage dentaire (ensemble de bagues et d'arcs orthodontiques) dégradera l'information au niveau de la face vestibulaire de la denture. La fréquence et la relative standardisation de ces dispositifs augmentent l'intérêt de leur intégration dans la réflexion sur le contenu des images.

La validation du modèle devra reposer sur deux axes principaux, à savoir d'une part une discussion avec les praticiens, qui seront aptes à indiquer quelles sont les structures et les relations qui leur semblent pertinentes pour l'identification des organes, et d'autre part une étude quantitative de l'apport des contraintes aux résultats de segmentation, pour identifier les plus importantes.

6.2.3 Développement des solutions applicatives

Le dernier aspect ouvert est celui du transfert aux cliniciens. Si nous avons pu réaliser un prototype logiciel, ce dernier n'a pas encore été confronté à l'usage clinique. Il est cependant essentiel de prendre en compte le retour des praticiens pour parvenir à améliorer le service rendu. Les deux axes d'améliorations sont d'une part ce qui relève de la méthode, de sa précision et de la qualité des résultats, et d'autre part ce qui relève de l'ergonomie et de son utilisabilité. Nous considérons notamment que si la méthode semble être d'une précision satisfaisante, l'interaction initiale de l'utilisateur doit être améliorée et allégée car trop fastidieuse dans sa forme actuelle pour une denture complète. Compte tenu de l'anatomie, il serait par exemple possible de positionner automatiquement la vue des images en fonction du numéro dentaire sélectionné, ou bien de proposer des points prédéfinis en fonction des points déjà positionnés, pour gagner du temps lors de l'étape initiale. Parallèlement à cette phase de test, l'utilisation d'outils issus du domaine des Interfaces Homme-Machine (IHM) [124], reposant sur l'analyse des réponses physiologiques et psychologiques humaines face à la machine, pourrait constituer une piste pour améliorer la forme prise par notre application. En dehors ce prototype, l'intégration de l'algorithme dans l'application de visualisation de Carestream permettra une plus large distribution, et ce dans un environnement pré-défini éprouvé et habituel pour les utilisateurs.

Annexe A

Approche morphologique de la segmentation dentaire

A.1 Arbres de composantes

Une des premières tentatives de segmentation de la denture fut réalisée en utilisant une approche morphologique par zones plates (zones de même intensité). Cette méthode repose sur la décomposition de l'image selon l'intensité, constituant des composantes connexes de taille décroissante, la racine étant le domaine de l'image. La représentation habituelle de cette décomposition est celle d'un arbre, d'où le nom d'arbres des composantes. Deux décompositions usuelles sont l'arbre maximum (max-tree) et minimum (min-tree) [7, 97, 107]. Le premier représente le domaine de l'image comme une composante de faible intensité, tandis que les feuilles de l'arbre seront les fortes intensités. L'arbre minimum est l'objet dual, considérant les composantes de faible intensité comme les feuilles. Ces constructions sont bien adaptées à la segmentation d'éléments qui sont très intenses (pour le max-tree) ou peu intenses (min-tree) mais qui ne présentent pas d'inclusion. Un exemple sur une image synthétique est présenté dans la figure A.1 (a-c) : avec le max-tree, il est aisé de segmenter le cadre blanc, puisqu'il constitue une feuille de l'arbre. De la même manière, le min-tree permet d'identifier le cadre noir. Considérons maintenant que l'élément blanc est un objet sur fond noir, qui contient des parties variables : les arbres précédents ne permettent pas de segmenter cet objet aisément, puisqu'il ne correspond à aucun nœud. Il existe pour cela une autre construction, reposant sur la relation d'inclusion, appelée l'arbre des formes [62], et illustrée en figure A.1 (d). Dans ce cas, le second nœud de l'arbre correspond au carré blanc plein.

A.2 Conception initiale du problème

Nous nous sommes intéressés à cette approche puisque l'apparence de la dent semble suivre un schéma adapté, à savoir une composante d'émail (hyper intense) qui contiendrait de la dentine (intensité forte), dans laquelle est incluse la chambre pulpaire et ses canaux (intensité moyenne). La dent étant séparée de son environnement par l'air au niveau des couronnes et par le ligament parodontal dans sa partie radiculaire (voir figure A.2), nous avons pour hypothèse qu'il est possible d'extraire un nœud dans l'arbre des formes (qui devrait correspondre à la zone plate de l'émail) correspondant à la segmentation recherchée. De plus les algorithmes existants de construction d'arbres permettent un traitement rapide, même lorsque les images possèdent un nombre important de voxels [28].

Afin de tester cette hypothèse nous procédons de la manière suivante :

1. extraction de zones d'intérêt de dimensions 2cm $\times$ 2cm $\times$ 3,5 cm,
2. modification de la dynamique de l'image en saturant les intensités inférieures à 2000 (tissus mous et air) et supérieure à 5000 (émail et métal), pour ne conserver que la gamme où les interfaces entre objets apparaissent,
3. réduction de la quantification sur une gamme $[0, 255]$,
4. calcul de l'arbre des formes,

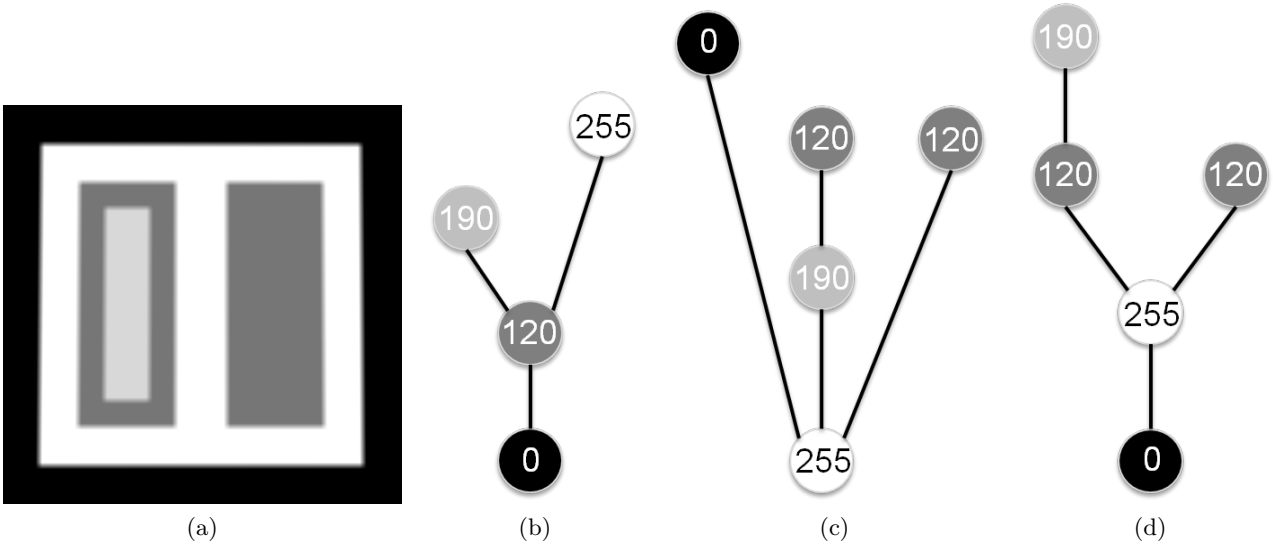


FIGURE A.1 – Exemple d'arbres de composantes : (a) image synthétique, (b) arbre maximum, (c) arbre minimum et (d) arbre des formes.

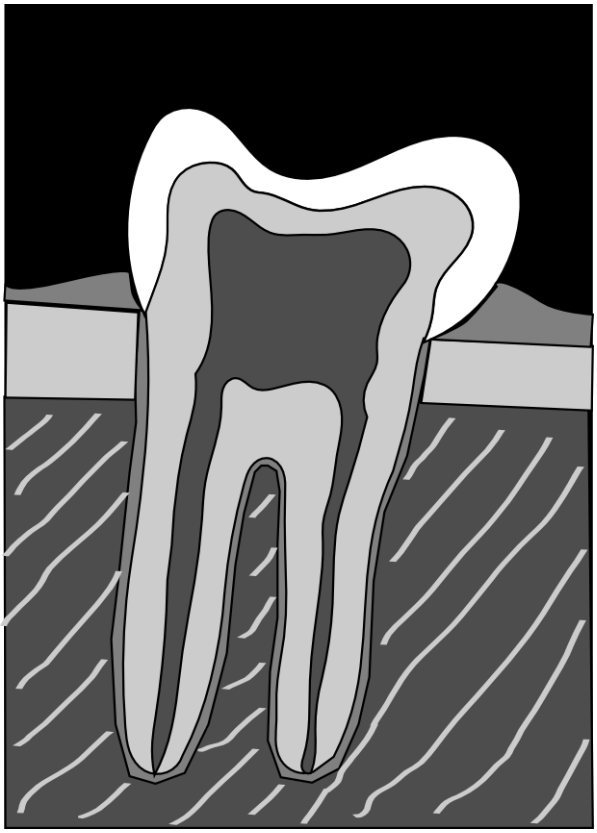


FIGURE A.2 – Représentation simplifiée de l'aspect d'une dent (molaire) dans nos images. Si la dent est correctement séparée de l'os cortical adjacent par le ligament parodontal, l'identification de la composante connexe correspondante à l'émail devrait correspondre à une segmentation de la dent.

5. filtrage de l'arbre par un critère de volume : élimination de tous les nœuds représentant moins de 10% du volume de l'image.

L'analyse visuelle des arbres (alors transformés en chaîne par le filtrage) a montré que pour la plupart des cas, il n'était pas possible d'identifier un nœud contenant la dent bien constituée sans récupérer

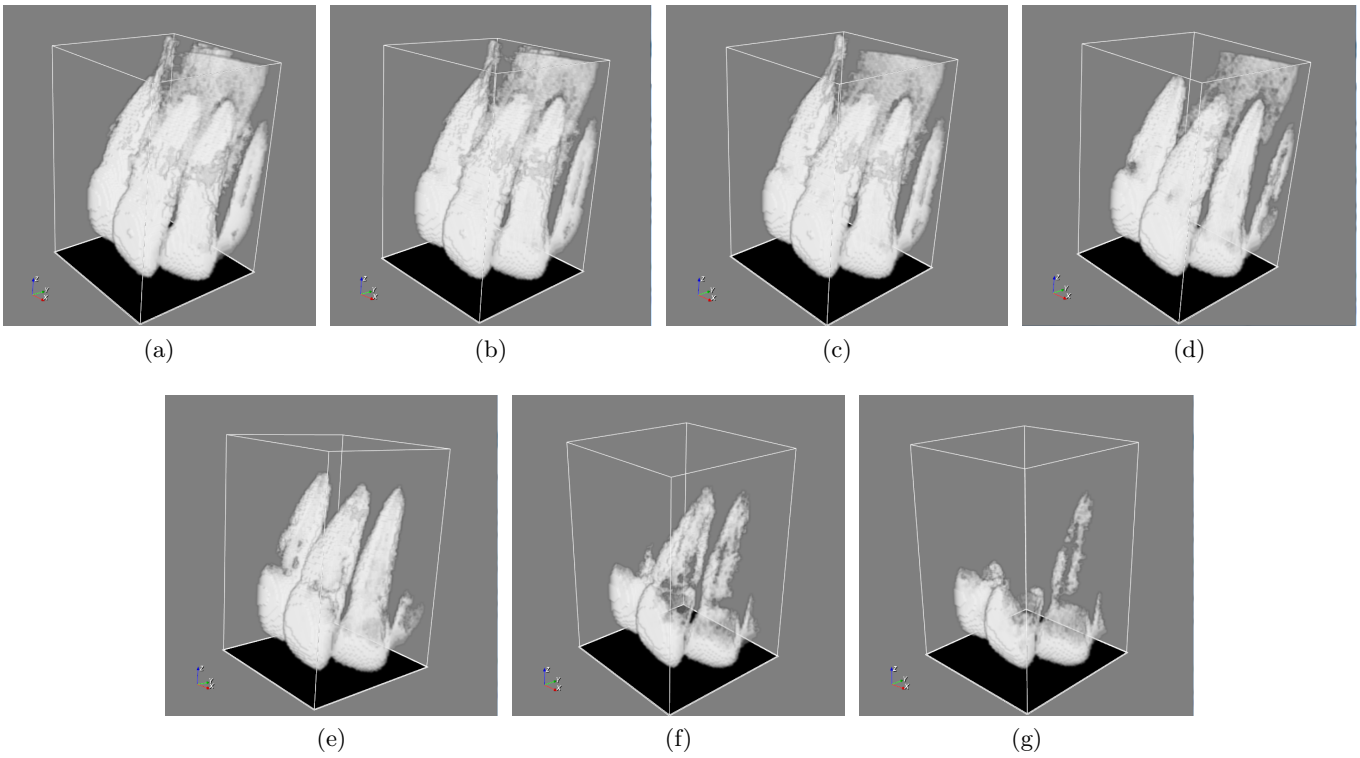


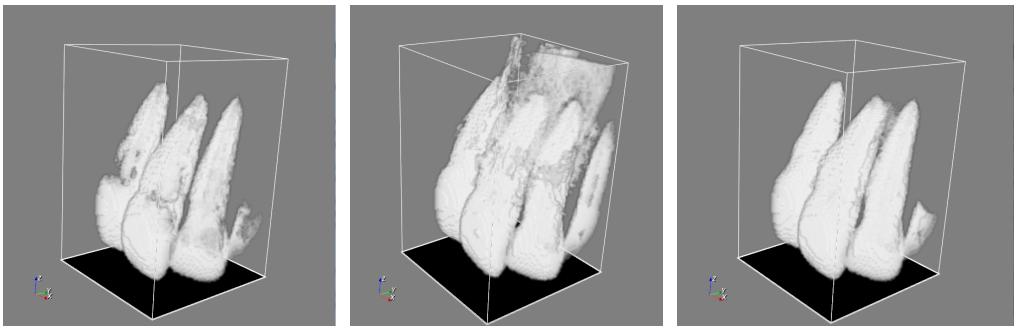
FIGURE A.3 – Représentation des composantes connexes de nœuds extraits d'un arbre de forme pour une incisive maxillaire, dans un parcours de la racine aux feuilles. Notez qu'au fur et à mesure, les parties appartenant aux structures environnantes disparaissent, mais la dent se dégrade.

des éléments indésirables, notamment des parois d'os cortical (voir un exemple en figure A.3). La segmentation directe à partir de l'arbre n'est donc pas possible car il n'existe pas de nœud adapté.

A.3 Révision de l'approche

La segmentation directe par sélection d'un nœud n'étant pas possible, nous avons essayé d'utiliser l'information contenue dans les nœuds les plus proches de la feuille pour filtrer les composantes connexes plus profondes. En demandant à l'utilisateur de choisir deux composantes, une qui représente au mieux la dent sans éléments indésirables (le "marqueur"), et une où la dent est entière et à nettoyer (le "masque"), nous procédons à une fermeture morphologique avec une sphère de 2 mm de rayon du marqueur, ce qui permet d'obtenir une composante connexe sans dégradation et qui représente approximativement la dent. L'intersection de cette fermeture et du masque donne une version nettoyée de la dent (voir figure A.4). Pour faciliter le choix des deux composantes, nous proposons une sélection restreinte à l'utilisateur. Pour ce faire, seules les composantes dont la variation de volume par rapport à son parent est significative sont présentées. Nous calculons le critère de perte de volume pour chaque nœud, et les candidats sont ceux dont cet attribut est supérieur à $1,7$ fois la moyenne. Le processus global est résumé dans la figure A.5.

Cette méthode possède deux inconvénients principaux. Le premier est qu'il n'y a pas de contrainte sur la forme du résultat, c'est-à-dire que si des éléments indésirables sont adjacents à la dent dans le masque, l'intersection avec le marqueur ne les éliminera pas. Deuxièmement, si le marqueur initial est très dégradé, notamment au niveau radiculaire, la fermeture ne pourra pas capturer toute l'étendue de l'organe, et de ce fait ne permettra pas de le récupérer. Cette approche, bien que rapide et peu coûteuse, met en avant le fait qu'il n'est pas possible de se reposer uniquement sur l'intensité pour obtenir une segmentation correcte des dents.



(a) (b) (c)

FIGURE A.4 – Illustration du filtrage des composantes de dents par une composante dégradée. La fermeture effectuée sur (a) donne un marqueur suffisamment précis pour pour récupérer les voxels de la dent de (b). (c) est le résultat de l'intersection.

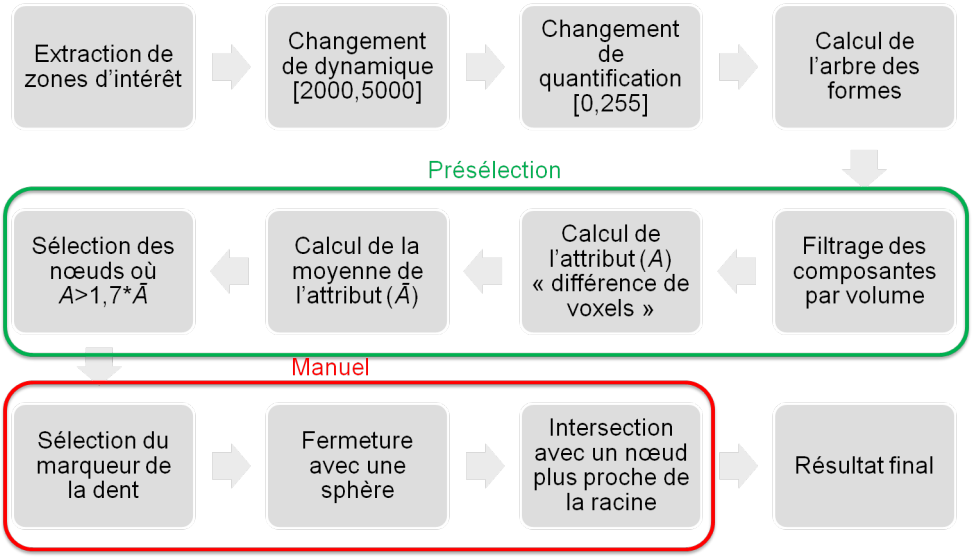


FIGURE A.5 – Processus de segmentation d'une dent.

A.4 Vers une approche par forme

Dans l'optique de se servir de la forme particulière des dents pour améliorer leur identification, nous avons établi des modèles statistiques, décrits dans le chapitre 3, constitués d'une forme moyenne et des principaux modes de déformation. Nous avons alors voulu nous servir de ce modèle pour rechercher et approcher la dent dans l'arbre des formes selon le processus suivant :

```
Zones_dIntérêt = Extraction(VolumeInitial, PointsUtilisateur)
pour chaque Image dans Zones_dIntérêt:
    Image = ChangementDynamique(Image)
    Image = ChangementQuantification(Image)
    Arbre = CalculArbreDesFormes(Image)
    Arbre_Filtré = FiltrageParVolume(Arbre, 0.1)

    pour chaque Nœud dans Arbre_Filtré:
        ComposanteBinaire = CréerImageDeComposante(Arbre_Filtré, Nœud)
        ComposanteFiltrée = ReconstruireDent(ComposanteBinaire)
        ModèleRecalé = RecalageRigide(ComposanteFiltrée, FormeMoyenne)
        Score = EstimationDeCorrespondance(ComposanteFiltrée, ModèleRecalé)
        ListeScores <—— Score
```

```
    MeilleurScore = 0

    pour chaque Score dans ListeScores:
        si Score > MeilleurScore:
            MeilleurScore = Score
            NœudIdéal = index_Score

    ComposanteIdéaleBinaire = CréerImageDeComposante(Arbre_Filtré,NœudIdéal)
    ComposanteIdéaleFiltrée = ReconstruireDent(ComposanteIdéaleBinaire)
    Approximation_Dent = OptimisationModèle(ComposanteIdéaleFiltrée,ModèleStatistique)
    Segmentation = Masquage(ComposanteIdéaleFiltrée, Approximation_Dent)
```

La première partie correspond au pré-traitement déjà décrit dans les parties précédentes. Le changement d'approche intervient dans la manière de rechercher un nœud idéal et d'en tirer une segmentation.

Sélection du nœud idéal

Une fois l'arbre construit et filtré, nous parcourons les nœuds de la racine aux feuilles pour identifier celui qui sera le plus représentatif de la dent. Une fois la composante connexe mise sous la forme d'une image binaire, nous tentons d'isoler la dent en la reconstruisant à partir d'une coupe sans éléments indésirables en procédant de la manière suivante :

1. extraction d'une coupe axiale aux deux tiers de la hauteur de la zone d'intérêt, ce qui correspond à la zone du collet dentaire lorsque le point utilisateur est correctement positionné, qui permet usuellement de bien isoler la dent de son environnement,
2. sélection de la composante connexe (2D) avec le barycentre le plus proche du centre de l'image, et ajout à la reconstruction finale,
3. comblement et dilatation morphologique de la coupe par un élément circulaire de rayon 1, donnant un masque,
4. sélection de l'intersection du masque et de la coupe de hauteur h-1 (et h+1), et ajout à la reconstruction finale,
5. répétition des étapes 3 et 4 en incrémentant la hauteur jusqu'à ce que toutes les coupes axiales aient été traitées.

Une telle reconstruction (voir figure A.6) permet d'éliminer la plupart des parties qui ne constituent pas du tissu dentaire, mais échoue pour les zones directement adjacentes dans les coupes 2D (voir par exemple la partie d'os cortical reconstruite). Une fois la composante ainsi filtrée, nous procédons à un recalage rigide de la forme moyenne binarisée du modèle correspondant à l'étiquette indiquée par l'utilisateur (notez qu'il s'agit là d'une information demandée en supplément des points par rapport aux approches précédentes). Ce recalage est optimisé par descente de gradient et la fonction de coût est l'erreur quadratique moyenne définie comme :

$$MSE(\mathrm{Im}, \mathcal{T}(\bar{\mathbf{M}})) = \frac{\sum_{x\in\Omega}(\mathrm{Im}(x) - \mathcal{T}(\bar{\mathbf{M}})(x))^2}{|\Omega|},$$

où $\mathcal{T}$ est la transformation de pose dont nous optimisons les paramètres, ici au nombre de six (trois pour la translation et trois pour la rotation), Im l'image de la composante binaire, $\bar{\mathbf{M}}$ l'image de la forme moyenne binaire et Ω le domaine de l'image. L'optimisation s'arrête à 100 itérations ou pour un changement de gradient inférieur à 10^{-4}. Sous l'hypothèse que notre recalage conduise à un positionnement approximativement correct du modèle (c'est-à-dire au niveau de la dent), nous évaluons l'adéquation entre la forme moyenne et la composante connexe en calculant le score suivant :

$$Score = \frac{V_I}{V_E} + \lambda\frac{V_I}{V_M},$$

avec

$$V_E = \max(\sum_{x\in\Omega}(\mathcal{T}(\bar{\mathbf{M}})^C \cap Im)(x)\ ,\ 1\ )$$

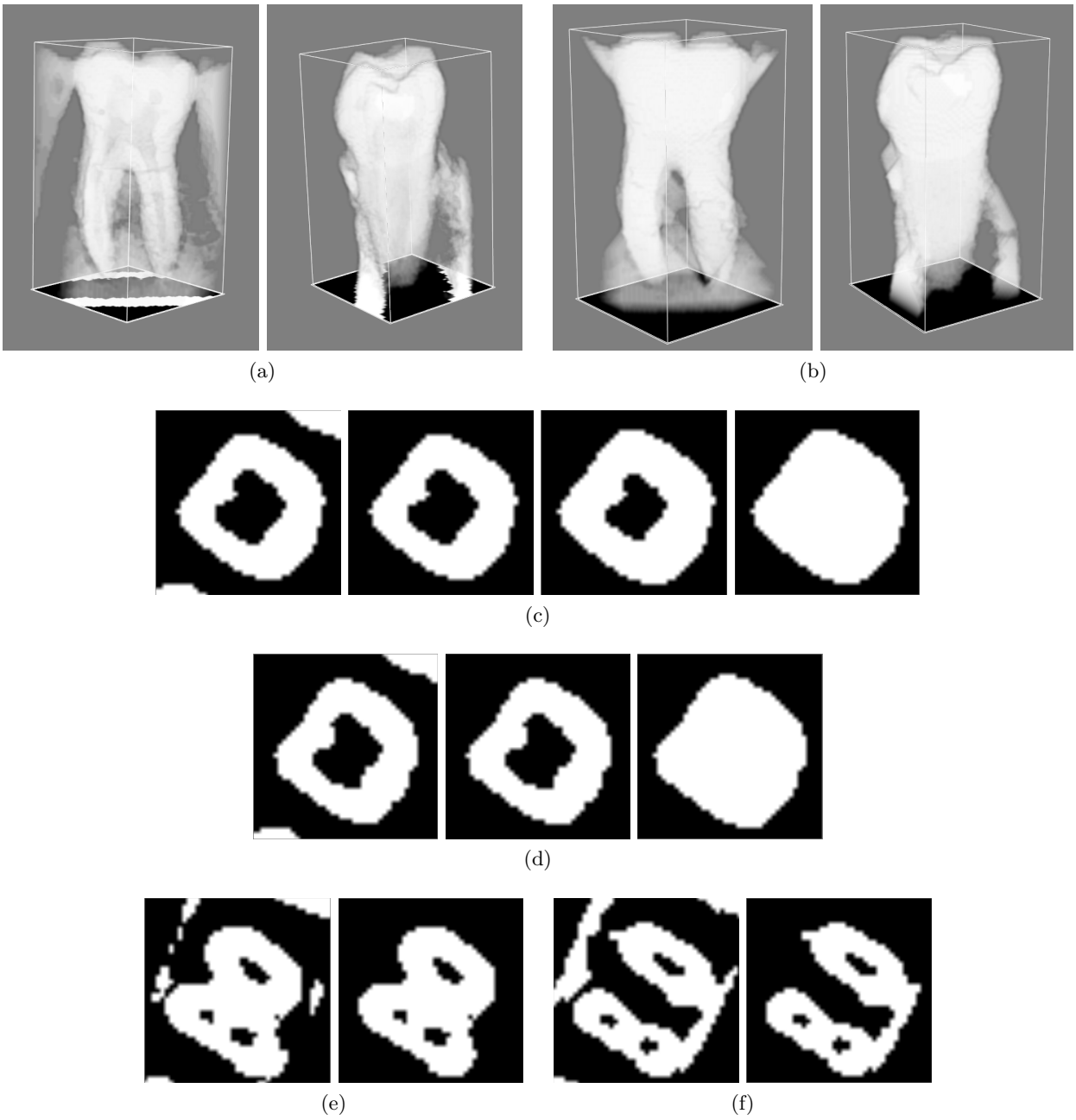
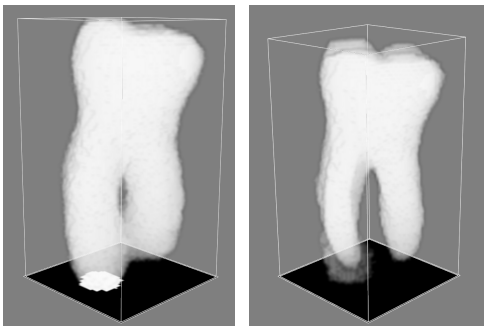


FIGURE A.6 – Illustration de la reconstruction effectuée sur les composantes connexes de l'arbre pour éliminer un maximum de structures indésirables. (a) Image binaire d'une composante connexe. (b) Résultat final de la reconstruction. (c) Coupe initiale et son traitement (sélection de l'élément central, dilatation et remplissage) donnant le masque de comparaison avec la coupe suivante. (d) Coupe adjacente, intersection avec le masque et masque pour itération suivante. (e) Exemple de résultat dans une coupe plus proche des racines. (f) Illustration de l'échec du filtrage lorsque l'élément indésirable est directement adjacent.

$$V_I = \sum_{x \in \Omega} (\mathcal{T}(\bar{\mathbf{M}}) \cap Im)(x)$$

$$V_M = \sum_{x \in \Omega} \mathcal{T}(\bar{\mathbf{M}})(x)$$

Le premier ratio compare le volume (en termes de nombre de voxels) de la composante qui se trouve à l'intérieur du modèle au volume qui en est exclus. Il est compris entre 0 et le volume du modèle, et plus sa valeur est importante plus la composante correspond au modèle. Le second terme est un terme



(a) (b)

FIGURE A.7 – Optimisation finale du modèle statistique (a) et approximation finale de la segmentation (b) par masquage de la composante.

pénalisant les composantes entièrement incluses dans le modèle mais de petite taille, il est compris entre 0 et 1, pondéré par λ. Plus le score est important, plus la composante ressemble au modèle, ce que nous considérons comme nœud idéal. Soulignons ici une approximation importante, il s'agit du modèle moyen, et qu'il est donc important que ce dernier soit bien représentatif de la dent, ce qui n'est pas trivial, notamment pour les dents postérieures (molaires) qui présentent une forte variabilité de forme en comparaison des dents antérieures. Une fois le score calculé pour tous les nœuds de l'arbre, le meilleur est sélectionné pour l'étape suivante.

Approximation de la segmentation

L'image de la composante du nœud idéal est préparée de la même manière que précédemment puis nous approchons la segmentation par une mise en correspondance du modèle statistique, en cherchant une réalisation qui représente au mieux le contenu de l'image. Pour cela nous procédons à une optimisation alternée entre les paramètres de pose (la transformation géométrique rigide) et ceux de forme (les coefficients des modes de déformation, voir équation 3.1). Le processus global est le suivant :

1. estimation de la pose de la forme moyenne (par exemple avec les résultats de l'étape précédente),
2. optimisation initiale des paramètres de forme,
3. optimisation de la pose,
4. optimisation de la forme,
5. répétition des étapes 3 et 4 jusqu'à convergence du critère de satisfaction.

L'optimisation de la pose est identique à celle décrite dans la partie précédente. Celle des paramètres de forme est effectuée de manière similaire (même fonction de coût et même algorithme de descente de gradient), avec les différences suivantes :

— nombre d'itérations maximal de 200,

— tolérance au gradient de 10^{-5}.

— pondération des paramètres en fonction du poids représenté par le mode de déformation.

Ce dernier point signifie que si le premier mode de déformation représente 70% des variations observées lors de la construction du modèle, le deuxième 20% et que le dernier ne représente que 10%, leurs échelles dans l'espace des paramètres seront pondérées par $\frac{1}{7}$,$\frac{1}{2}$ et $\frac{1}{1}$ respectivement. Ainsi, le paramètre associé au premier mode aura tendance à varier plus facilement que ses suivants, dans le but de produire des formes plus probables. Le processus est itéré tant qu'une passe produit un changement (une passe étant une itération de chaque optimisation) ou lorsque que le nombre maximal de passes est atteint (ici fixé à cinq). L'image ainsi obtenue est finalement utilisée pour masquer l'image de la composante, constituant la segmentation (voir figure A.7).

Expérimentation

Afin de tester la viabilité de notre méthode, nous avons utilisé une denture complète (32 dents) d'un volume en moyen champ de résolution isotrope de 150 microns. Les résultats sont présentés dans la figure A.8 : le score de Dice moyen pour l'ensemble de la denture est de $0,781 \pm 0,065$, pour des valeurs comprises entre $0,589$ et $0,876$. Ces résultats, bien que corrects, ne sont pas au niveau de l'état

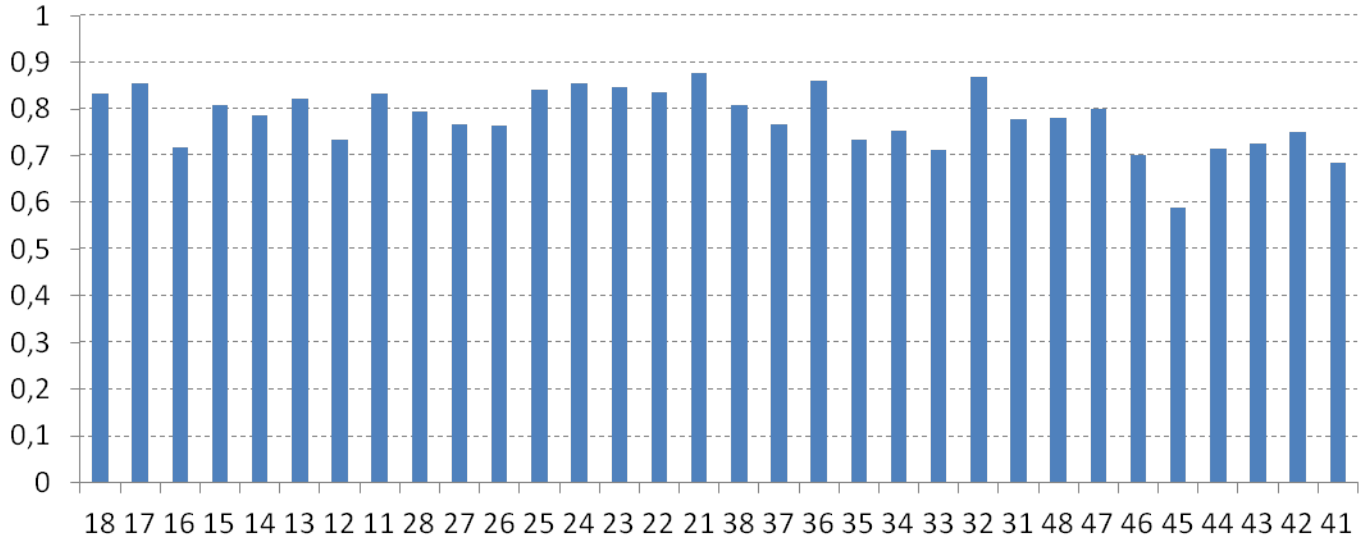


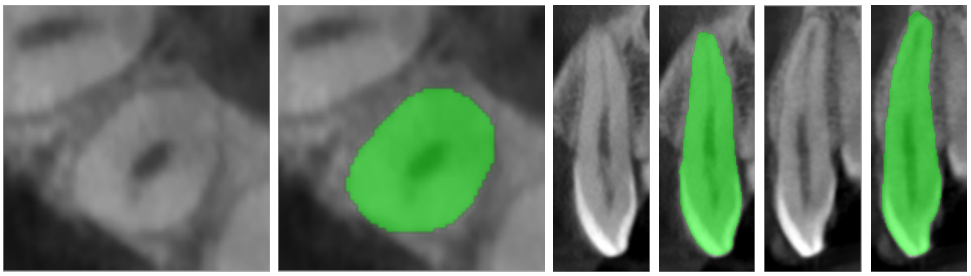
FIGURE A.8 – Scores de Dice obtenus pour chaque dent par notre approche utilisant sur le modèle statistique. Le score moyen est de $0,781 \pm 0,065$.

de l'art (scores de Dice $> 0,9$). Une des principales faiblesses provient de la méthode de sélection du nœud idéal. En effet, pour près de la moitié des dents ici traitées (14 sur 32), l'application de la méthode sur l'ensemble de l'arbre démontre qu'il existe un nœud donnant de meilleurs scores de Dice (écarts allant jusqu'à $0,1$). De plus, comme il est possible de le voir en figure A.7, l'optimisation du modèle de parvient pas à séparer correctement l'os cortical adjacent de la dent, et conduit souvent à la récupération de pans de parois osseuses. Cette approche nécessite donc des améliorations pour être exploitable, notamment au niveau de la méthode de sélection du nœud idéal. Une méthode de correction du résultat, automatique ou manuelle, est aussi une condition pour parvenir à de meilleures segmentations.

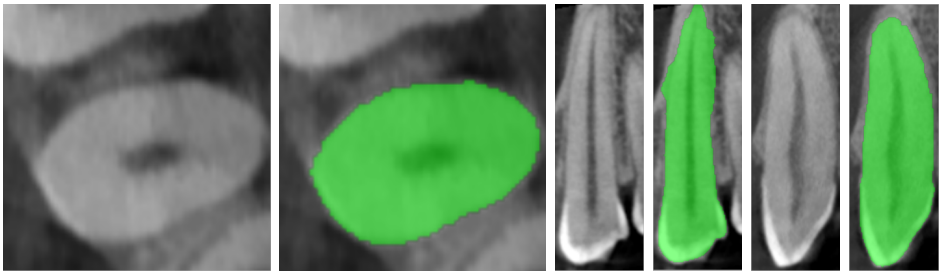
Annexe B

Exemples de résultats de segmentation

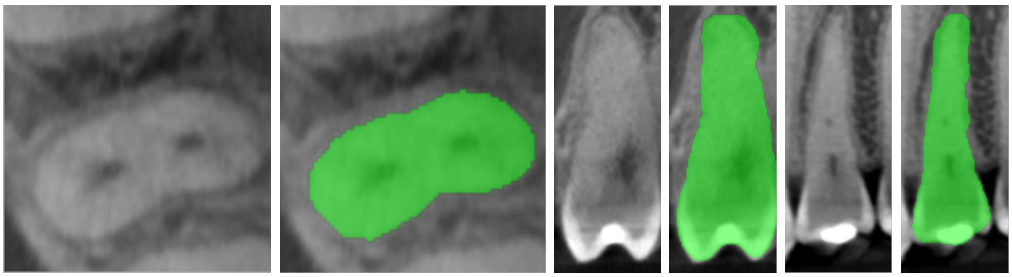
B.1 Segmentation mono-objet



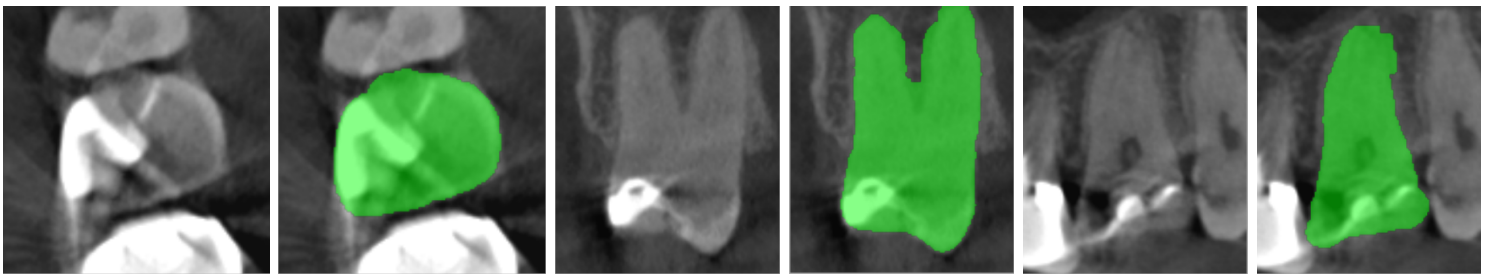
(a) Incisive 12. Taille de voxel de 150 microns.



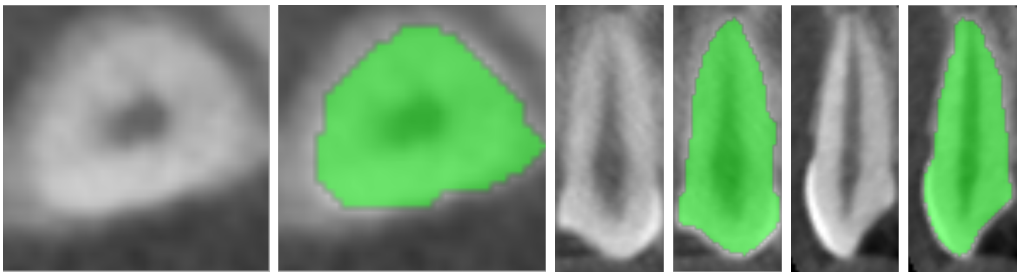
(b) Canine 13. Taille de voxel de 150 microns.



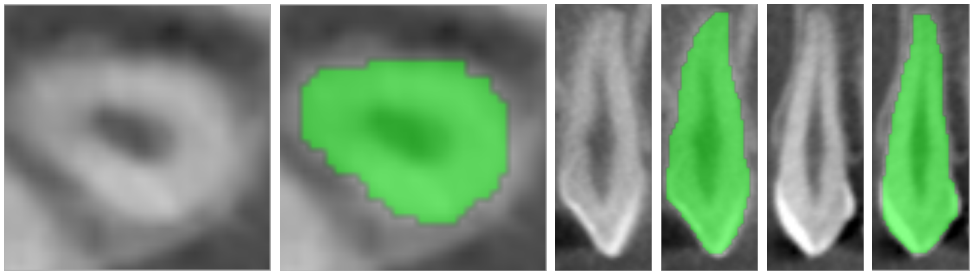
(c) Prémolaire 14. Taille de voxel de 150 microns.



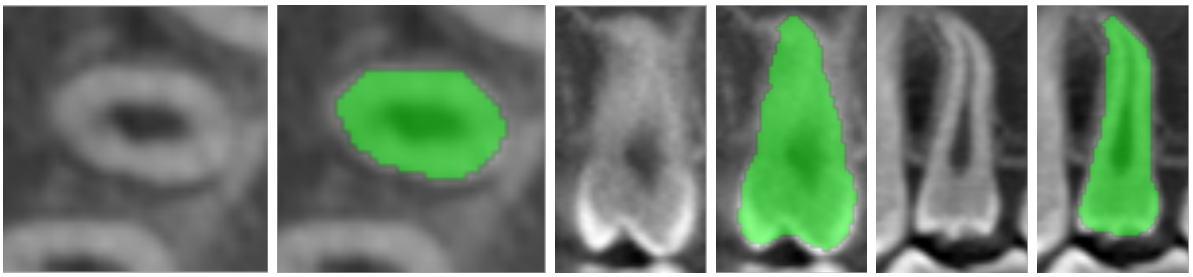
(d) Molaire 17. Taille de voxel de 150 microns.



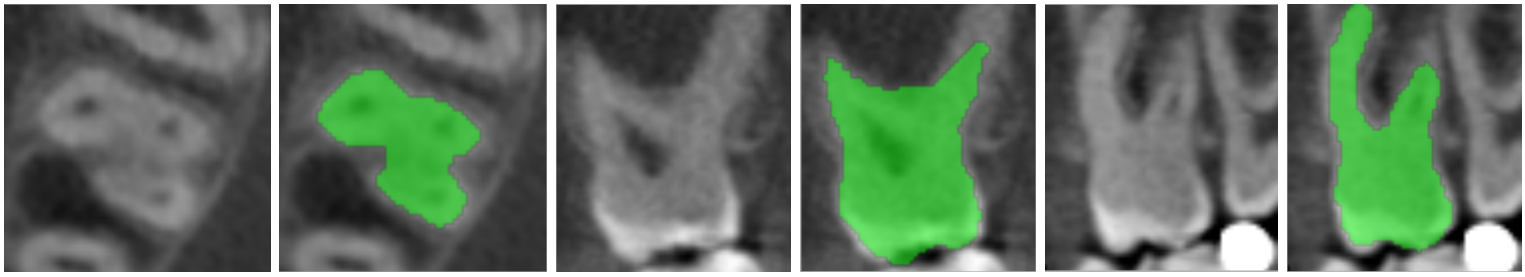
(e) Incisive 21. Taille de voxel de 300 microns.



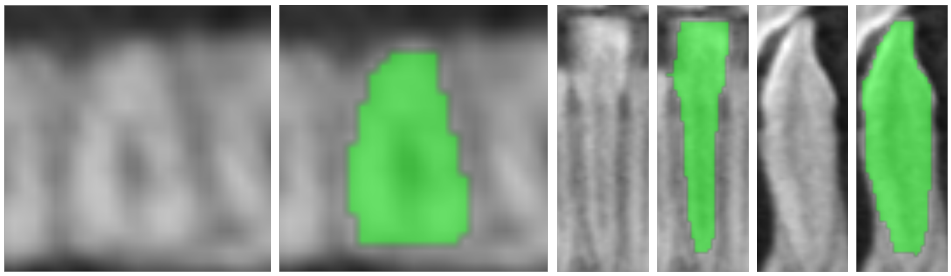
(f) Canine 23. Taille de voxel de 300 microns.



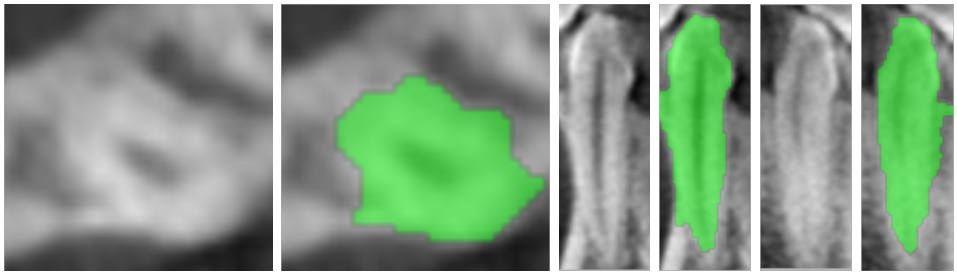
(g) Prémolaire 25. Taille de voxel de 300 microns.



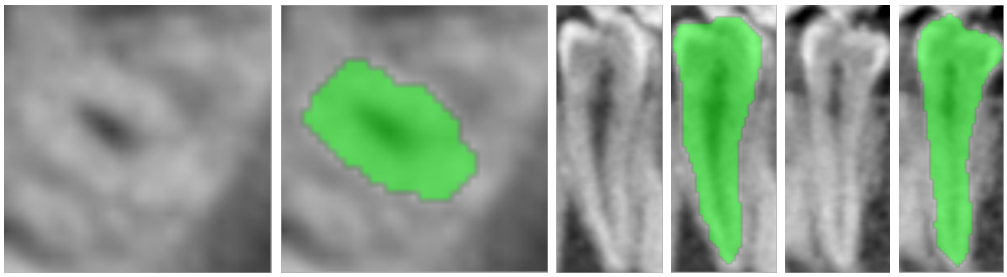
(h) Molaire 26. Taille de voxel de 300 microns.



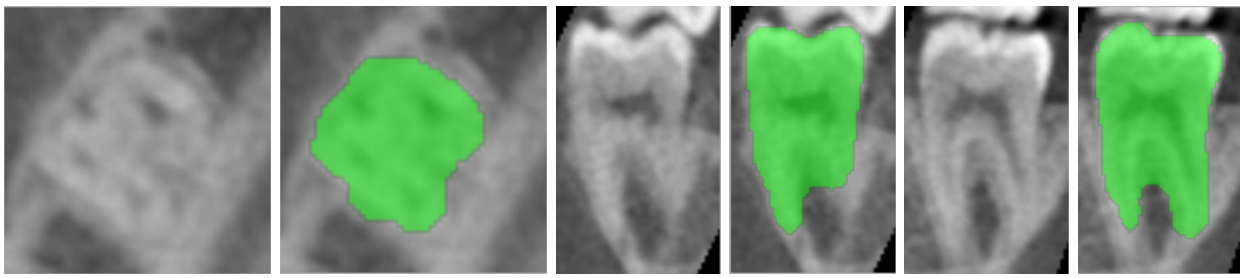
(i) Incisive 32. Taille de voxel de 300 microns.



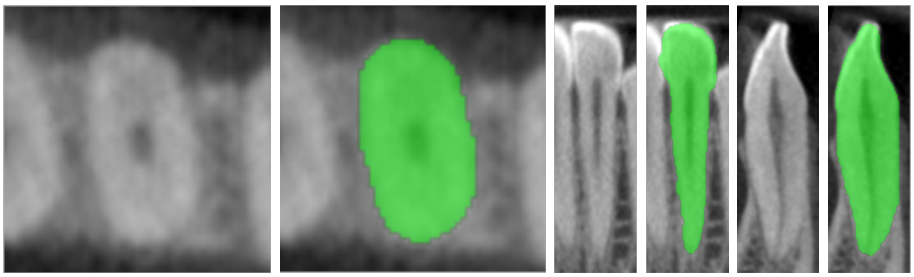
(j) Canine 33. Taille de voxel de 300 microns.



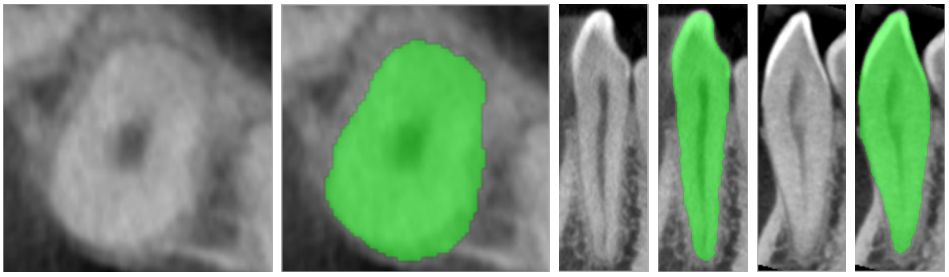
(k) Prémolaire 34. Taille de voxel de 300 microns.



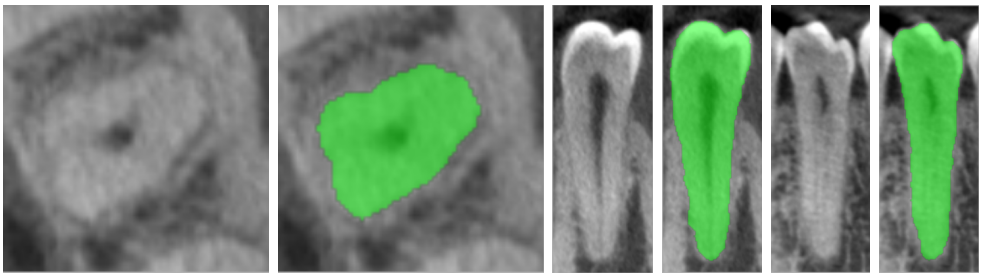
(l) Molaire 37. Taille de voxel de 300 microns.



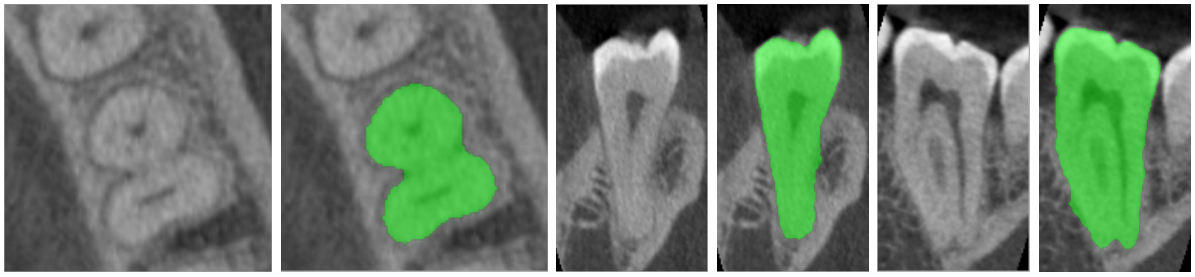
(m) Incisive 41. Taille de voxel de 150 microns.



(n) Canine 43. Taille de voxel de 150 microns.

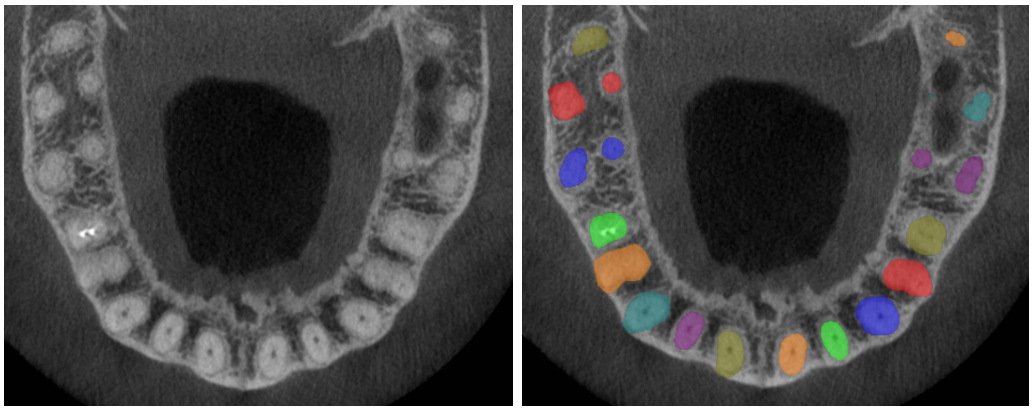


(o) Prémolaire 45. Taille de voxel de 150 microns.

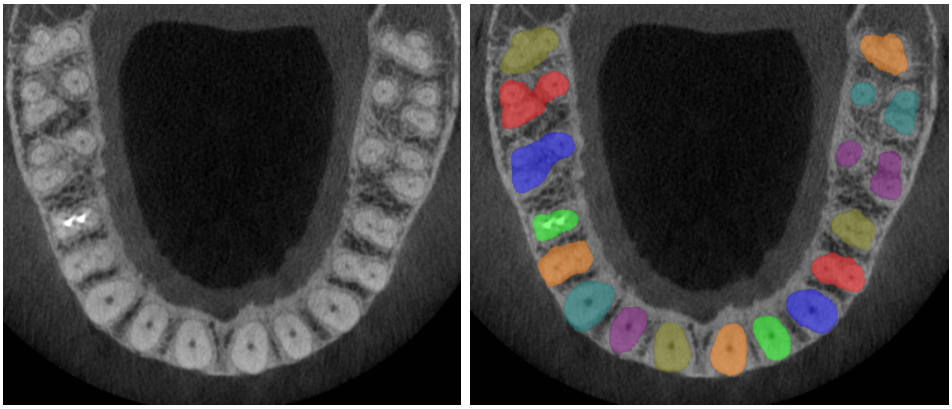


(p) Molaire 47. Taille de voxel de 150 microns.

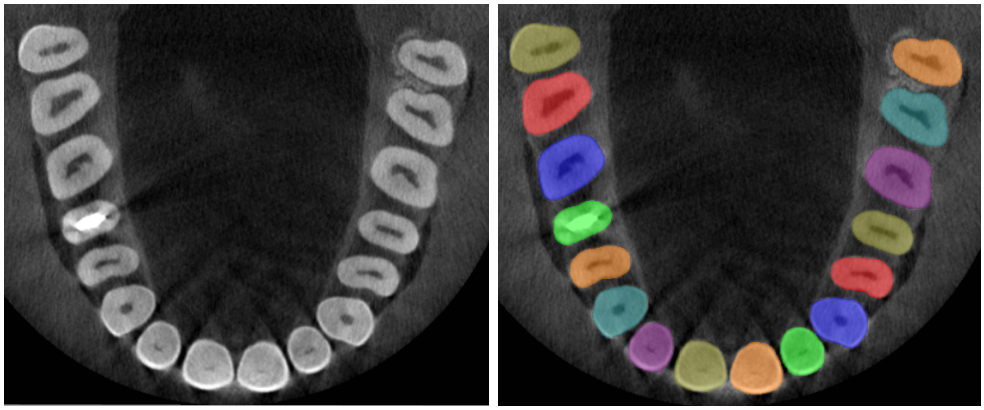
FIGURE B.1 – Résultats obtenus par dent individuelle avec la méthode présentée dans le chapitre 3. Pour chaque série est présentée la coupe axiale puis la coupe coronale et enfin la coupe sagittale, chacune avec leur résultat.



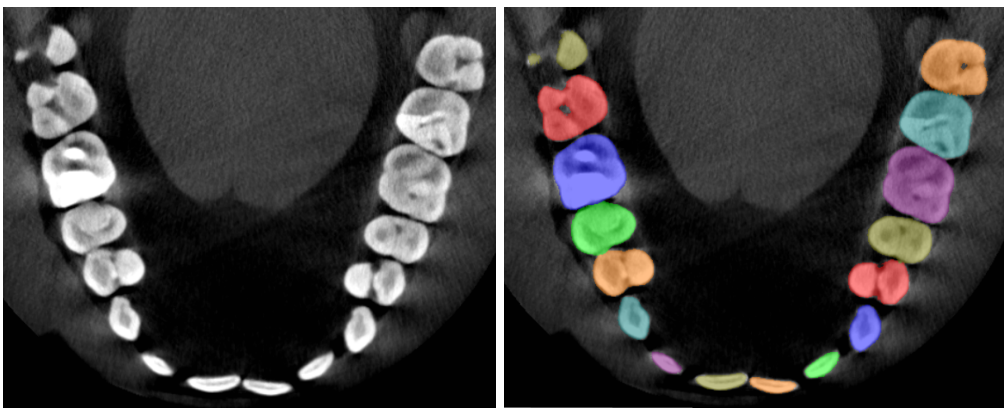
(a) Coupe au niveau des racines des dents maxillaires.



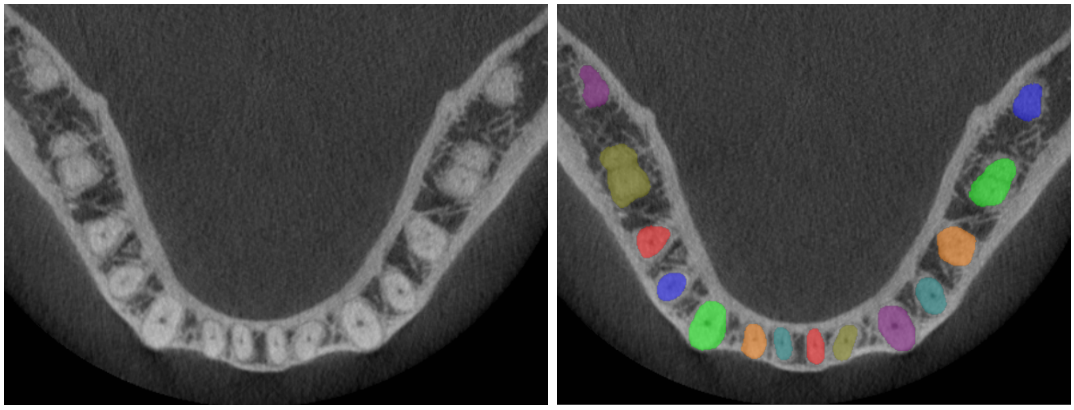
(b) Coupe intermédiaire des dents maxillaires.



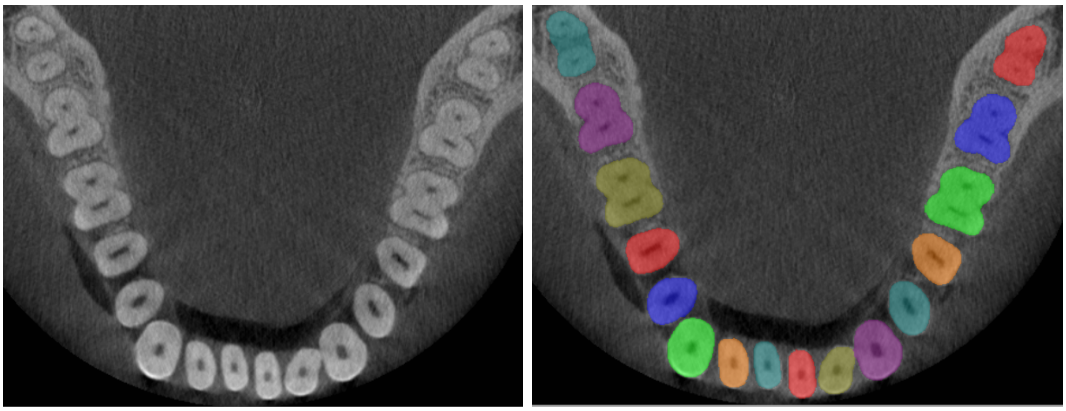
(c) Coupe intermédiaire des dents maxillaires.



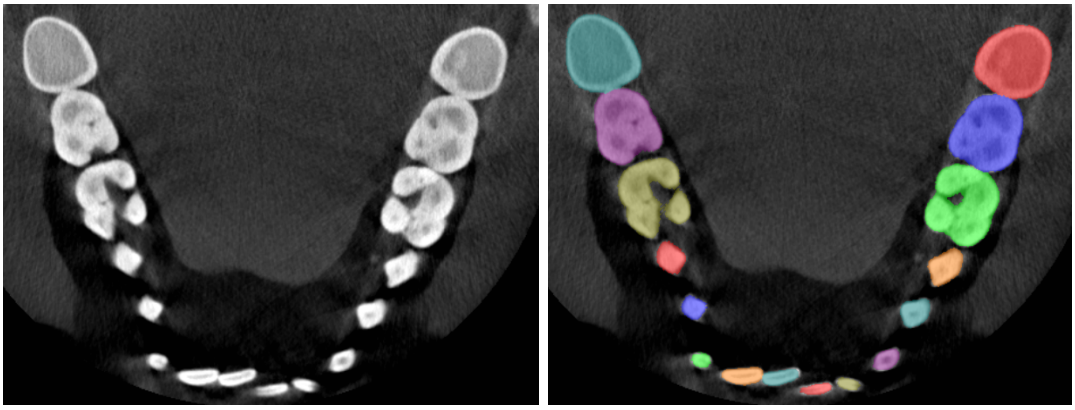
(d) Coupe au niveau des couronnes des dents maxillaires.



(e) Coupe au niveau des racines des dents mandibulaires.

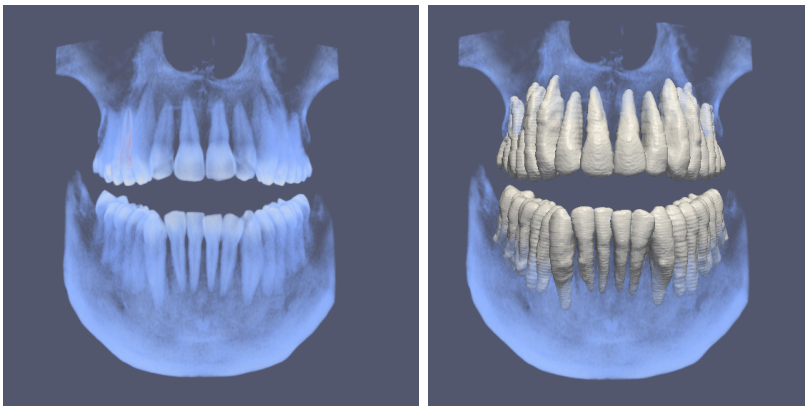


(f) Coupe intermédiaire des dents mandibulaires.

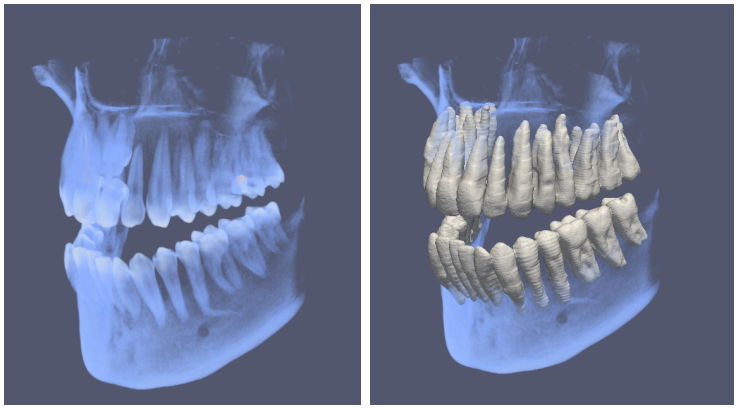


(g) Coupe au niveau des couronnes des dents mandibulaires.

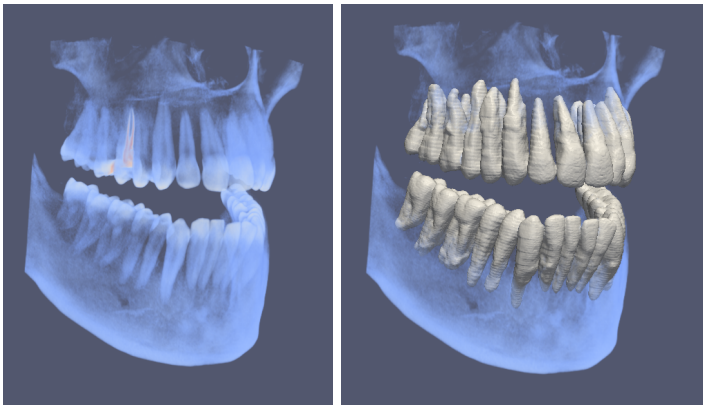
FIGURE B.2 – Superposition d'une segmentation complète d'un volume par la méthode présentée dans le chapitre 3.



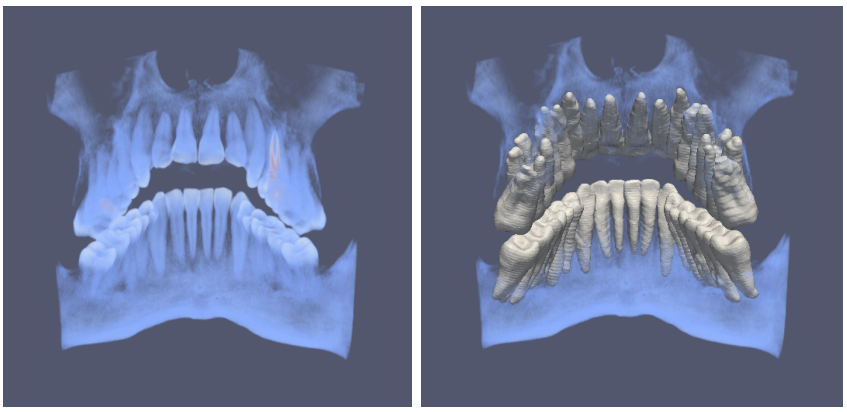
(a) Rendu de face.



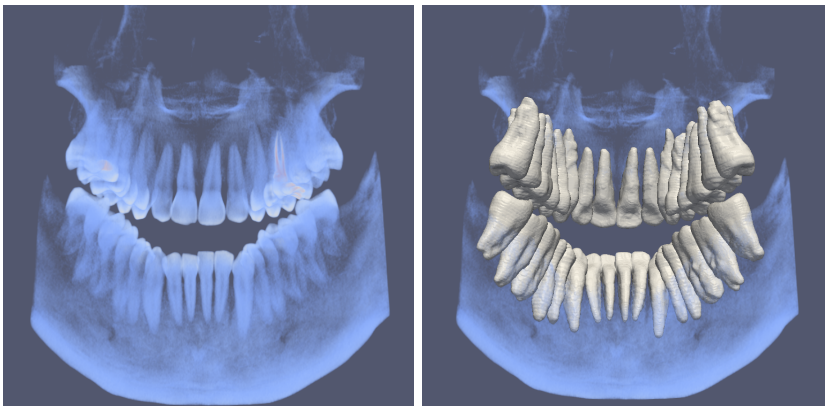
(b) Rendu de profil gauche.



(c) Rendu de profil droit.



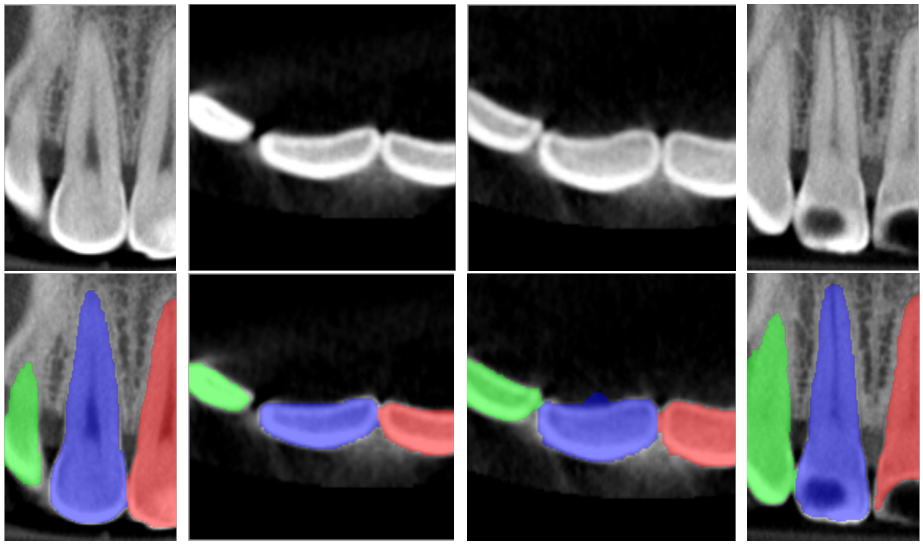
(d) Rendu de dos avec vue sur les interfaces coronales des dents mandibulaires.



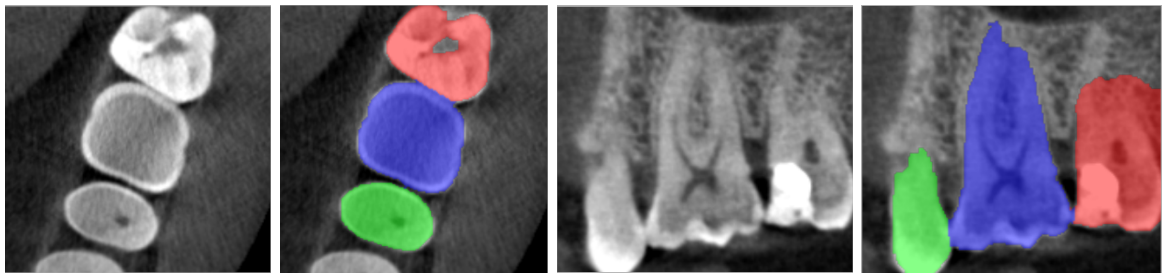
(e) Rendu de dos avec vue sur les interfaces coronales des dents maxillaires.

FIGURE B.3 – Rendu 3D des résultats de segmentation d'un volume complet avec la méthode présentée dans le chapitre 3.

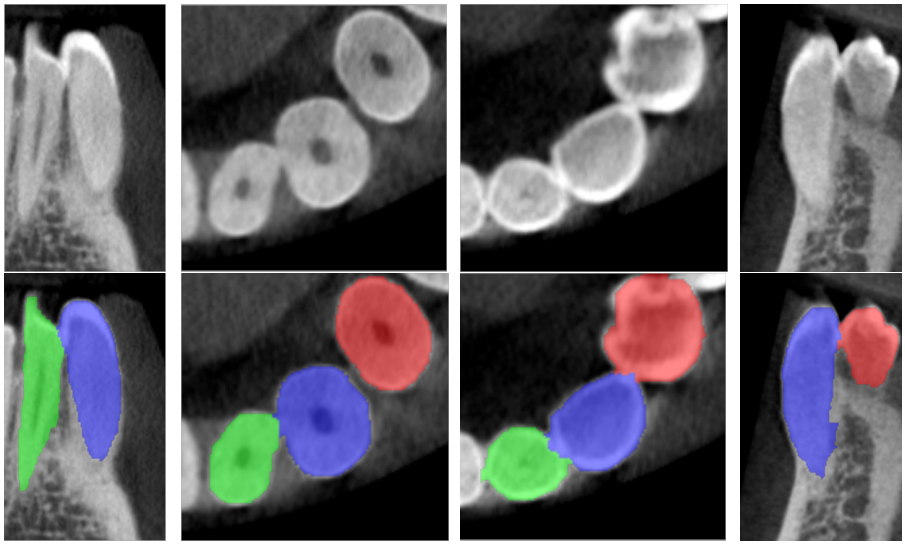
B.2 Segmentation multi-objets



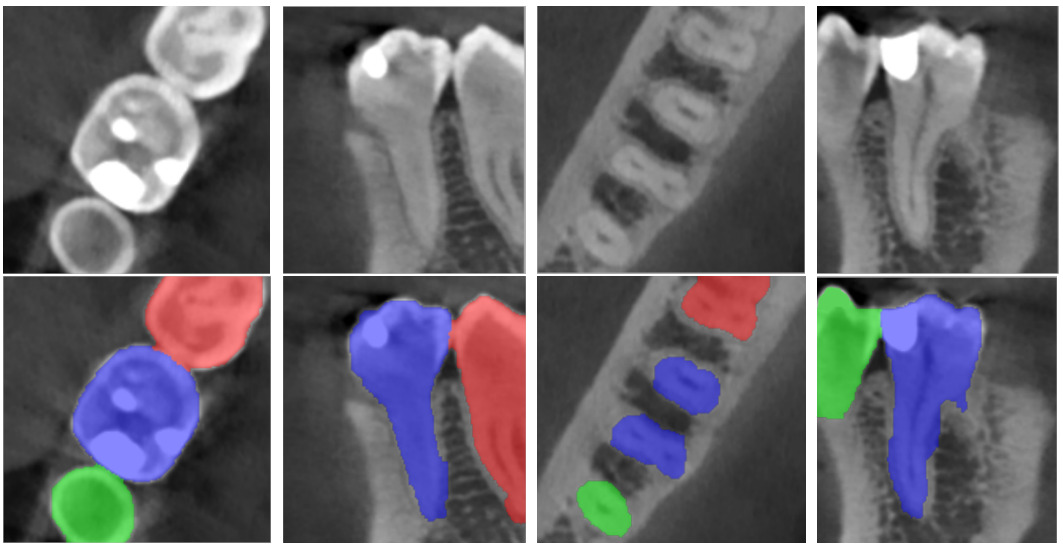
(a) Incisive 11. Taille de voxel de 150 microns.



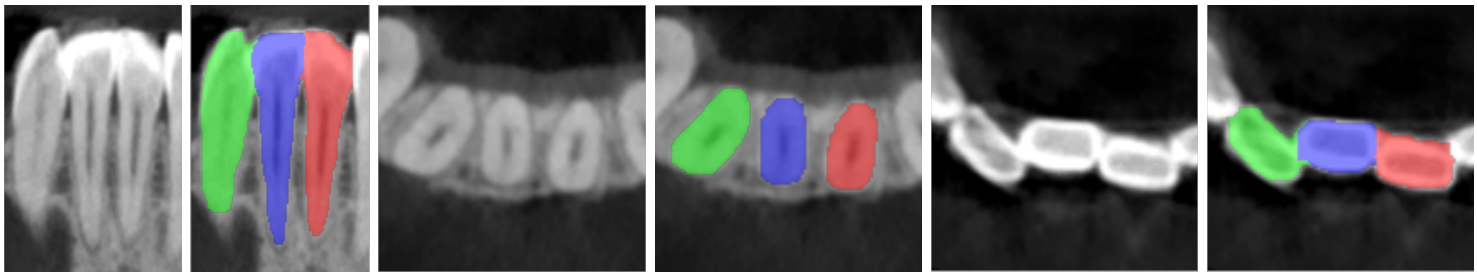
(b) Molaire 26. Taille de voxel de 150 microns.



(c) Canine 33. Taille de voxel de 150 microns.



(d) Molaire 36. Taille de voxel de 150 microns.



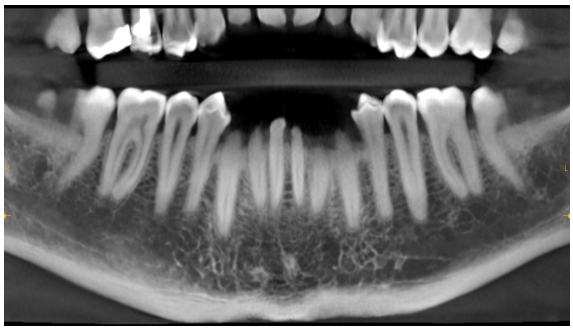
(e) Incisive 41. Taille de voxel de 150 microns.

FIGURE B.4 – Exemples de résultats avec la méthode proposée dans le chapitre 4. Notons que la prise en compte des dents adjacentes permet de mieux identifier les contours de chaque dent, notamment au niveau des couronnes, grâce à la compétition entre chaque étiquette. Nous présentons des coupes axiales ainsi que des coupes sagittales ou coronales selon l'axe permettant de mieux distinguer les interfaces entre les dents.

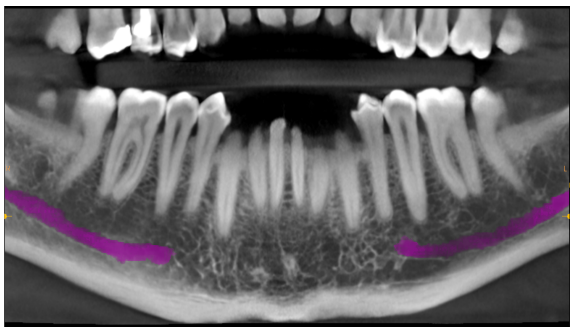
Annexe C

Caractérisation du canal mandibulaire

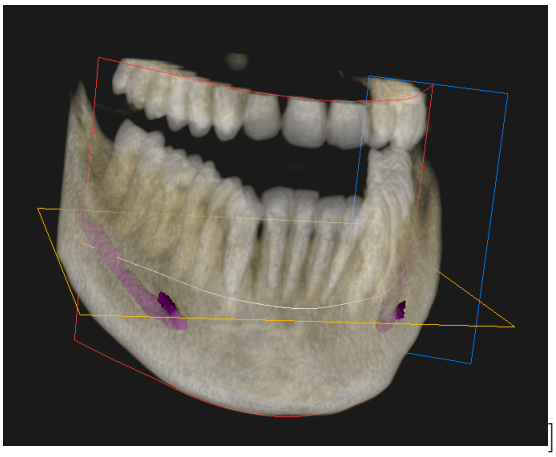
Nous avons procédé à une étude des caractéristiques du canal mandibulaire dans nos images CBCT dans le but d'évaluer la faisabilité d'une segmentation directe. Ce dernier est particulièrement intéressant pour plusieurs raisons. Tout d'abord, l'intérêt clinique de le localiser est bien connu. En raison de son rôle, qui est d'accueillir le nerf mentonnier, et de sa position, souvent juste à l'aplomb des dents, les dentistes ont besoin de le visualiser afin de ne pas l'endommager lors des procédures chirurgicales sous peine de dysthésie (diminution ou exagération de la sensibilité) ou de douleurs neuropathiques pour le patient. Deuxièmement, ce canal est une structure peu aisément identifiable sur des données CBCT. Son apparence tend à se confondre avec celle de son environnement, et il n'est parfois repérable que sur un cliché pseudo-panoramique. Par exemple sur la figure C.1, notez que les canaux ne sont visibles que grâce à la fine enveloppe d'os qui le recouvre. Cette dernière n'est toutefois pas toujours visible. Les canaux mandibulaires ont été segmentés manuellement sur 6 volumes de patients, provenant de deux types de scanners afin de servir de référence.



(a) Image pseudo-panoramique.



(c) Segmentation manuelle du canal superposée sur la vue pseudo-panoramique.



(b) Visualisation 3D de l'agencement.

FIGURE C.1 – Identification du canal sur les données CBCT.

C.1 Analyse de l'intensité

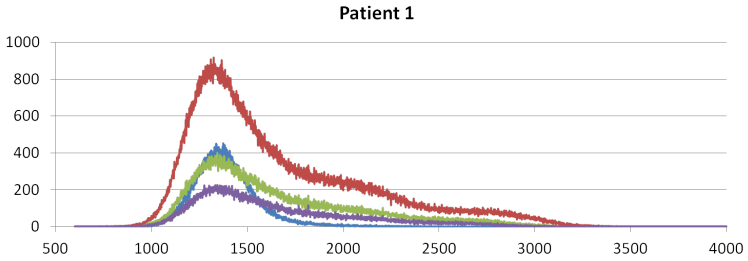
Les histogrammes de ces canaux ainsi que de leur environnement sont étudiés pour fournir une première estimation de l'impact de l'intensité. L'environnement est sélectionné sous la forme de marges de 0,5 mm, 1 mm et 1,5 mm autour de la segmentation manuelle. Les résultats sont présentés en figure C.2. Deux choses sont nettement visibles : premièrement les histogrammes du canal et de son environnement ont un recouvrement très important, ce qui indique bien que l'environnement est d'intensité globale semblable au canal. Deuxièmement, l'histogramme de l'environnement semble présenter une partie inexistante pour le canal, aux valeurs supérieures à 1700. Cette partie de l'histogramme correspond à l'os cortical, et les différences entre les patients correspondent au fait que le canal longe plus ou moins longtemps la paroi mandibulaire, ainsi qu'à son embouchure autour du foramen d'une part, et que les trabéculations sont plus ou moins denses d'autre part. Le patient 6 présente de très fortes valeurs car c'est un patient où la zone trabéculaire de la mandibule est fortement calcifiée, ressemblant à de l'os cortical. Si cette information est intéressante, elle ne saurait suffire à réaliser de bonnes segmentations, car le canal longe l'os principalement dans ses sections terminales, or nous savons que la partie difficile à segmenter est la partie centrale, au centre de la zone trabéculaire.

Afin d'étudier la lumière du canal, nous générons le squelette de la segmentation manuelle puis procédons à des dilatations successives avec un élément structurant sphérique de 2 voxels de rayon, puis nous analysons l'intensité des différentes strates. Les résultats sont résumés dans la figure C.3. Pour tous les patients, les histogrammes présentent la même forme. Nous en concluons que la lumière du canal est homogène en intensité, ne fournissant pas d'informations à même de permettre sa délimitation.

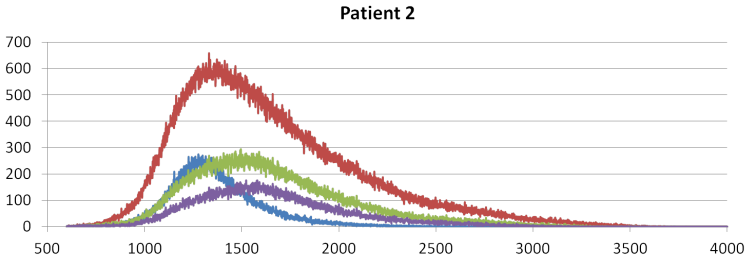
Pour résumer les deux expériences précédentes, la seule différence entre le canal et son environnement direct d'un point de vue des distributions des niveaux de gris provient de la présence variable d'os cortical et de trabéculations à proximité. Comme nous savons ces deux facteurs très variables selon les images et les protocoles d'acquisition, nous ne pouvons pas en faire un critère discriminant.

Cas	Moyenne du canal	Moyenne de l'environnement (1,5 mm)
Patient 1	1391 ± 178	1659 ± 461
Patient 2	1367 ± 244	1650 ± 467
Patient 3	1236 ± 244	1630 ± 595
Patient 4	1181 ± 182	1513 ± 560
Patient 5	1452 ± 148	1668 ± 476
Patient 6	1211 ± 180	1724 ± 453

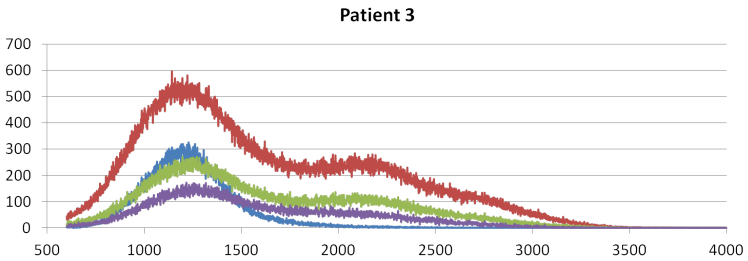
TABLE C.1 – Table récapitulative des moyennes d'intensités du canal et de son environnement à 1,5 mm dans nos cas d'étude.



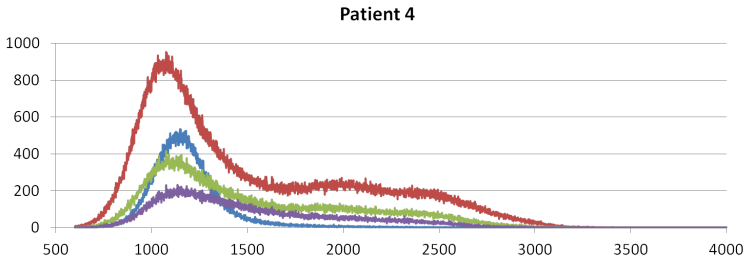
(a)



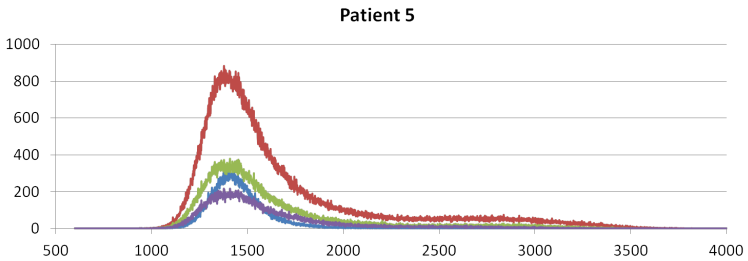
(b)



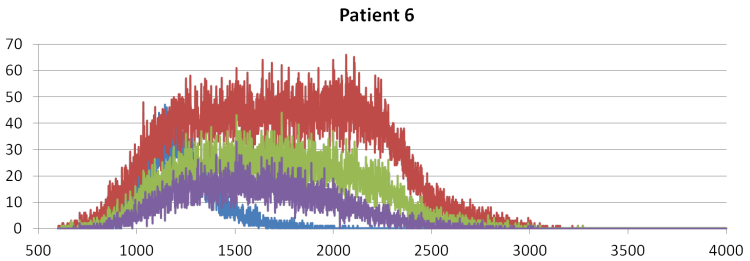
(c)



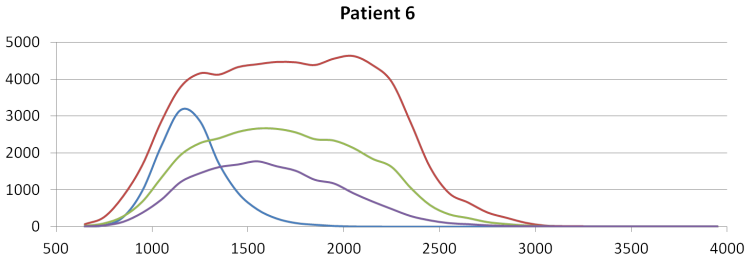
(d)



(e)

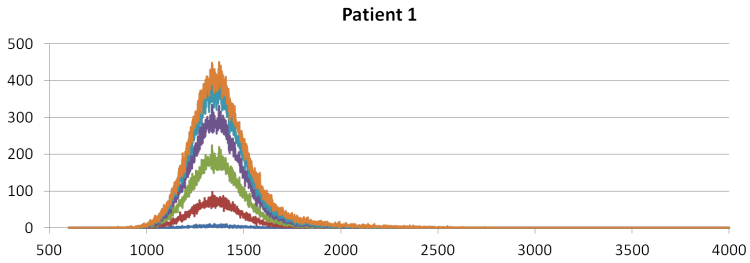


(f)

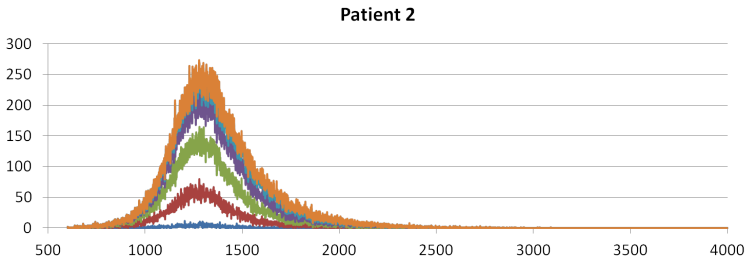


(g) Courbe lissée pour plus de clarté.

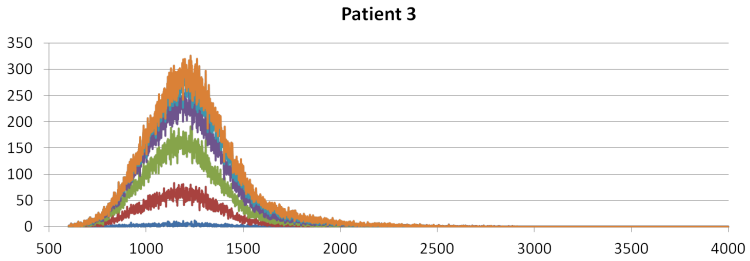
FIGURE C.2 – Étude des histogrammes du canal mandibulaire (courbe bleue) et de son environnement. Les trois épaisseurs considérées sont 0,5 mm (courbe violette), 1mm (courbe verte) et 1,5mm (courbe rouge). Les moyennes et écarts-types du canal et de l'épaisseur 1,5mm sont indiquées dans la table C.1.



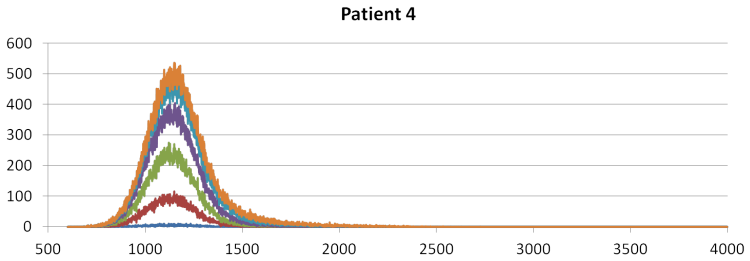
(a)



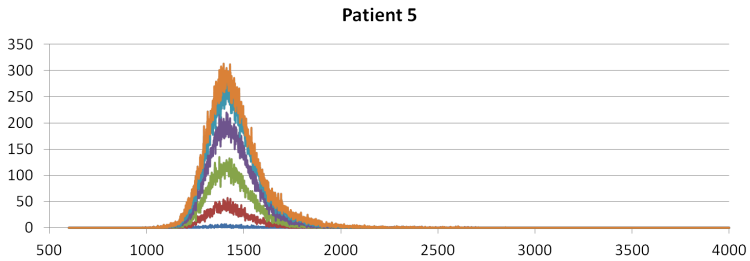
(b)



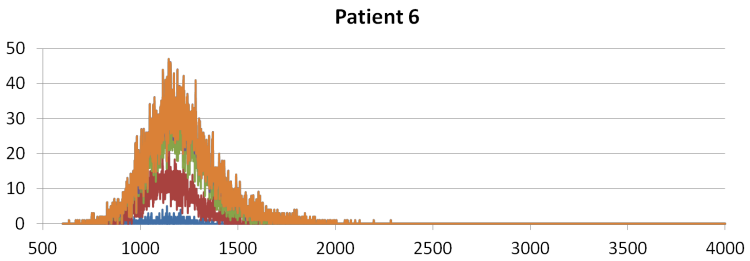
(c)



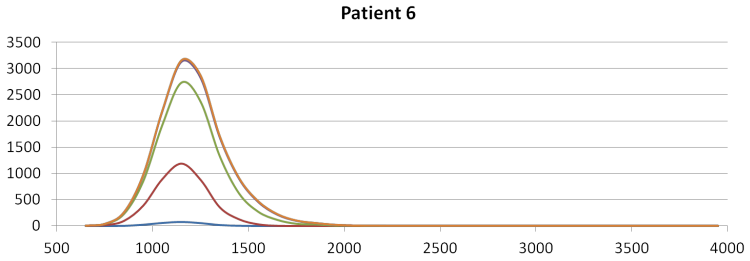
(d)



(e)



(f)

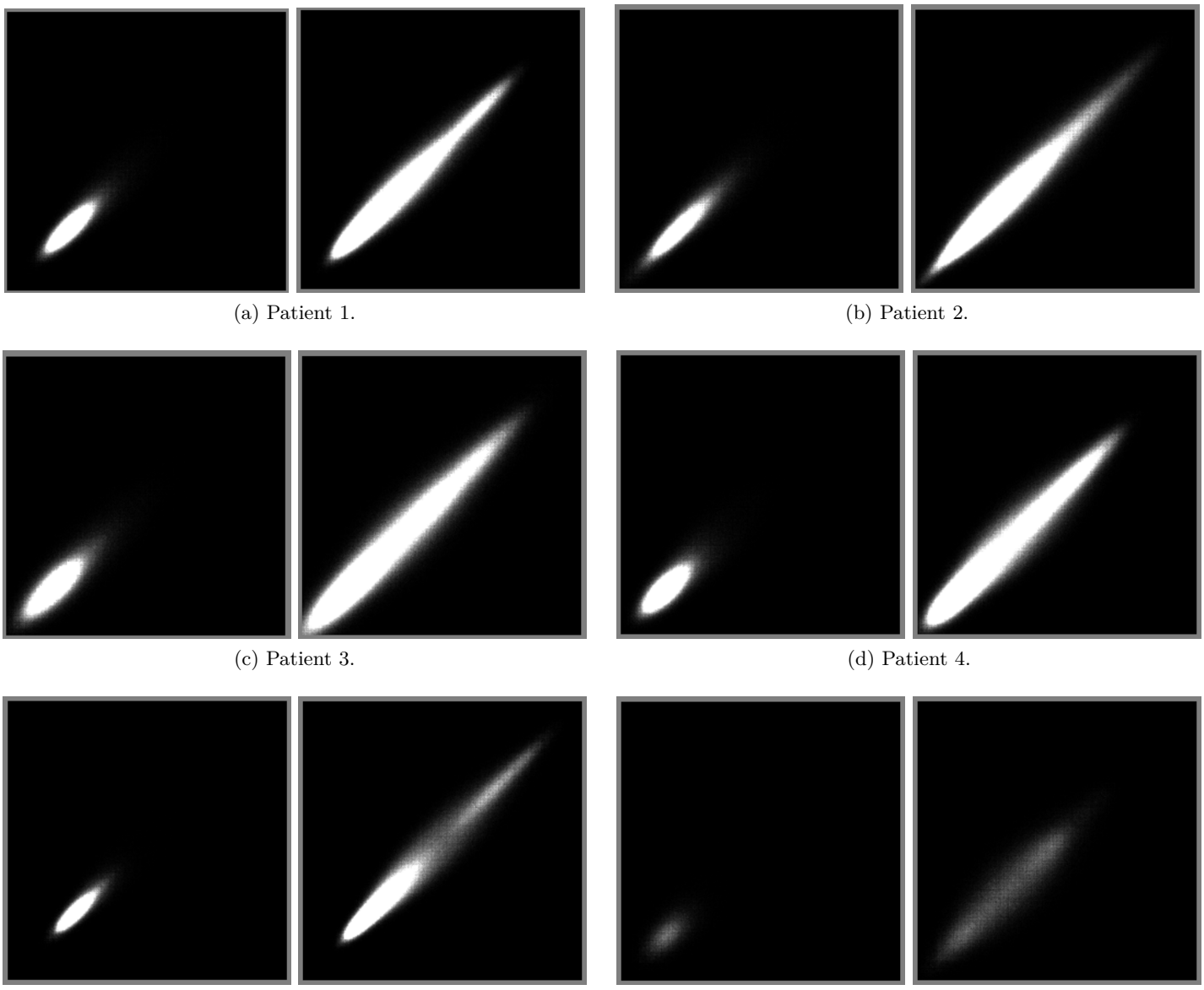


(g) Courbe lissée pour plus de clarté.

FIGURE C.3 – Étude des histogrammes des différentes strates du canal mandibulaire. Le squelette (courbe bleue) est généré à partir de la segmentation manuelle, puis dilaté successivement par un élément structurant sphérique de 2 voxels de rayon jusqu'à retrouver la segmentation initiale (courbe orange). L'ordre des couleurs est : bleue, rouge, verte, violette, turquoise puis orange.

C.2 Analyse de la texture par matrice de co-occurrences

Puisque l'intensité ne nous permet pas d'isoler les canaux, nous avons recherché la présence de textures discriminant les structures et leur environnement. Les matrices de co-occurrences [37, 70, 71] permettent de synthétiser la distribution locale des intensités en comptant la fréquence d'apparition d'un couple de valeurs de niveaux de gris selon un décalage spatial donné. Il est alors possible de calculer des caractéristiques (énergie, entropie, etc.) qui résument le contenu de la matrice. Nous avons comparé la segmentation manuelle du canal à son environnement pour une épaisseur de 1,5



(a) Patient 1. (b) Patient 2.

(c) Patient 3. (d) Patient 4.

(e) Patient 5. (f) Patient 6.

FIGURE C.4 – Aspect comparé des matrices de co-occurrences, avec celle du canal (gauche) et de l'environnement (droite). Pour chaque patient, le modèle de répartition est similaire. La différence de taille correspond à la plus grande diversité de niveaux de gris dans la partie environnementale que dans le canal.

mm. La matrice de co-occurence est calculée pour trois écarts spatiaux : un pas de 1 voxel dans chaque direction (soit $[1,0,0]$, $[0,1,0]$ et $[0,0,1]$). Dans la figure C.4 sont présentées les images des matrices ainsi obtenues. Notez que la matrice pour le canal présente le même modèle que celle pour l'environnement, à savoir une répartition des fortes valeurs autour de la diagonale, ce qui signifie que les intensités varient doucement et que deux voxels proches partagent une intensité similaire. Dans le cas contraire, l'image de la matrice présenterait des zones intenses éloignées de la diagonale. L'aspect semble donc indiquer qu'il n'existe pas de texture marquée que ce soit au niveau de la lumière du canal ou de son environnement. D'un point de vue quantitatif, nous avons calculé et utilisé quatre valeurs caractéristiques de ces matrices :

— l'énergie, $feat_1 = \sum_{i,j} M(i,j)^2$, entre $[0,1]$, est une mesure de l'uniformité de la texture, et vaut 1 pour une image constante,

— l'entropie, $feat_2 = -\sum_{i,j} M(i,j) \cdot \log_2 M(i,j)$, est une mesure du désordre indiquant si la texture

est proche de l'aléatoire statistique, et est donc corrélée négativement avec l'énergie (vaut 0 pour une image constante),

— l'homogénéité, $feat_3 = \sum_{i,j} \frac{M(i,j)}{1+(i-j)^2}$, entre $[0, 1]$, mesure la dispersion autour de la diagonale de la matrice, indiquant si la texture est douce (peu de variations) ou marquée, et tend vers 1 pour une image constante,

— l'inertie (ou contraste), $feat_4 = \sum_{i,j}(i-j)^2 \cdot M(i,j)$, corrélée négativement avec l'homogénéité, est la mesure locale de l'amplitude de la différence des intensités entre voxels, comprise entre 0 (pour une image constante) et le carré du nombre de niveaux de gris considérés,

où M est la matrice de co-occurrence indexée par i, j (et normalisée par la somme de ses valeurs). Les résultats pour les patients étudiés se trouvent dans la figure C.5. D'après un test de Student à 1% d'erreur, seule la différence entre les énergies est significative. Or cette dernière peut s'expliquer par l'aspect beaucoup plus concentré de la répartition des valeurs de la matrice du canal (ce qui conduit à une énergie plus importante), due à un nombre de niveaux de gris plus restreint.

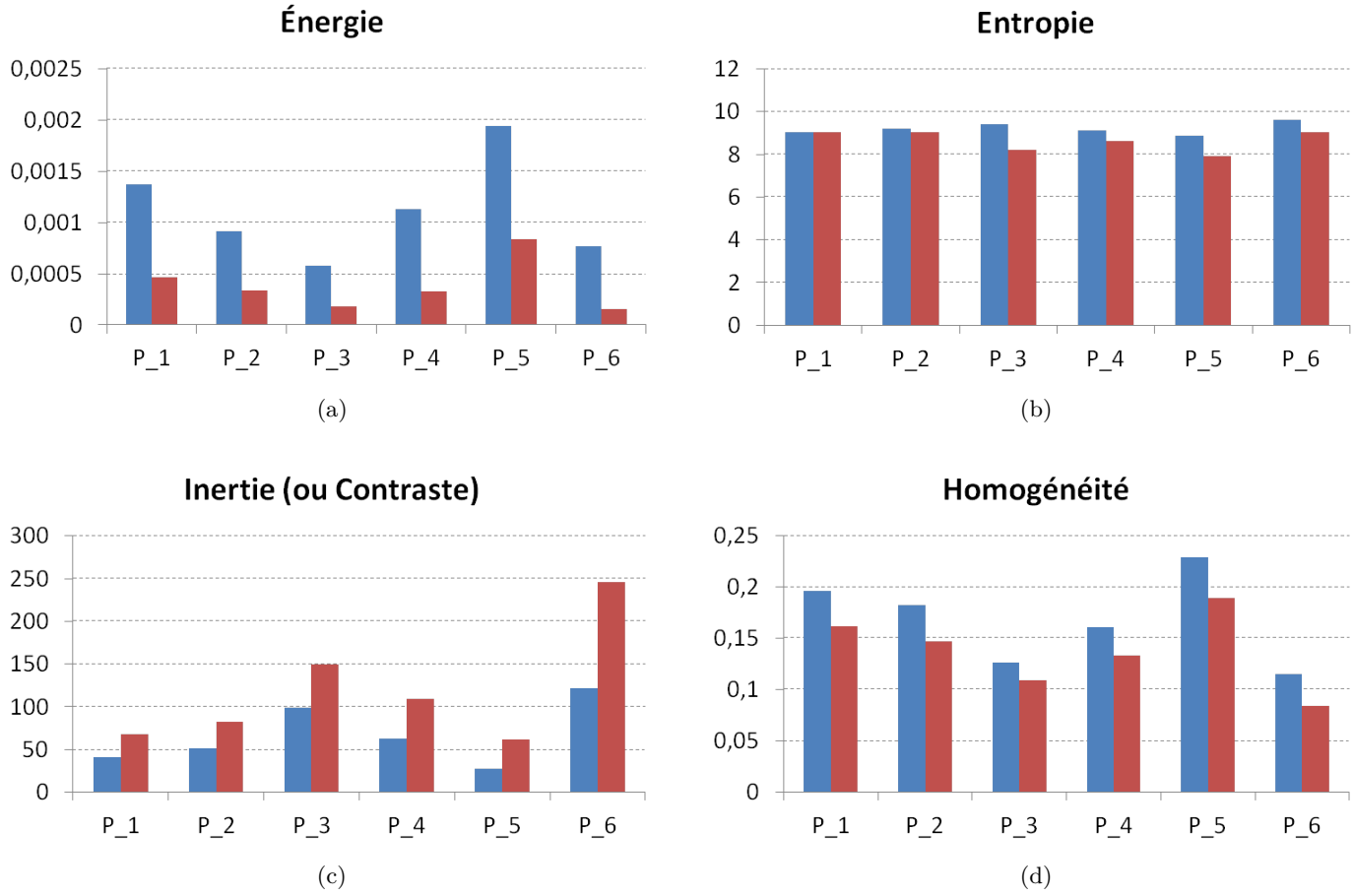


FIGURE C.5 – Valeurs caractéristiques calculées pour les matrices des zones d'intérêt (en bleu pour le canal, en rouge pour son environnement). Seule la différence d'énergie est statistiquement significative (test de Student à 1%). (a) Énergie, (b) entropie, (c) inertie, (d) homogénéité.

Il est possible que le fait de prendre un pas spatial si faible masque une potentielle texture variant plus lentement. Cependant, la forme de nos zones d'intérêt étant fortement anisotrope (tube et cylindre), il est difficile d'augmenter le pas sans manquer d'échantillons dans les directions transverses. Il nous apparaît donc qu'il n'est pas possible de différencier le canal par une analyse de sa texture par matrice de co-occurrences.

C.3 Analyse de la décomposition en ondelettes

Pour finir, nous avons voulu caractériser la canal à travers sa réponse aux ondelettes. L'analyse multi-échelles par ondelettes [88, 109] repose sur la décomposition de l'image en une série de coefficients qui constituent la réponse à une famille (la base) de fonction dérivée d'une fonction initiale, l'ondelette mère. Les différentes ondelettes correspondent à un décalage spatial et/ou une modification d'amplitude de l'ondelette mère. Les coefficients obtenus lors de la décomposition permettent alors de caractériser le contenu de l'image. Une différence de réponse entre le canal et son environnement permettrait alors d'envisager des méthodes pour l'isoler. L'algorithme ici utilisé est celui de [115] pour la décomposition discrète pyramidale avec élévation (*lifting scheme*) : un ensemble de filtres passe-bas et passe-haut directionnels sont utilisés en combinaison pour effectuer l'analyse. Pour chaque niveau de décomposition, nous obtenons huit images, selon le filtre utilisé pour chaque direction. Nous avons choisi de réaliser une décomposition en trois niveaux, avec pour ondelette mère la CDF-5/3 (voir figure C.6), ce qui nous donne 21 ensembles de coefficients. Pour synthétiser l'information nous en comparons l'énergie (de même formulation que dans la partie précédente, soit la somme quadratique des coefficients) pour chaque patient (voir figure C.5).

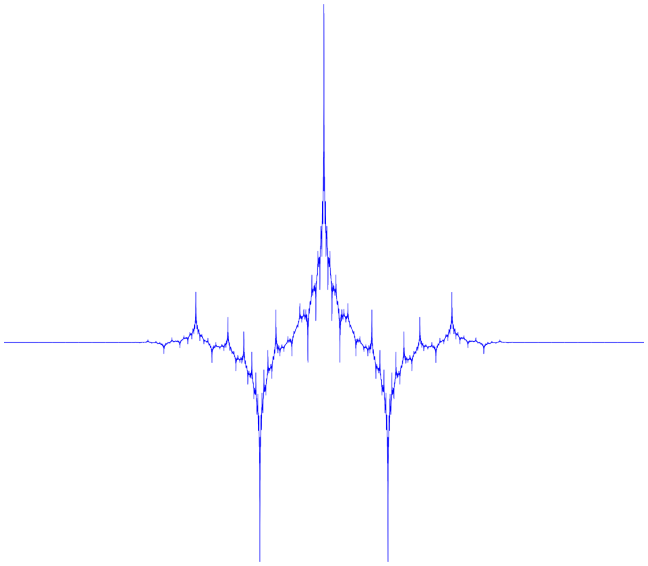
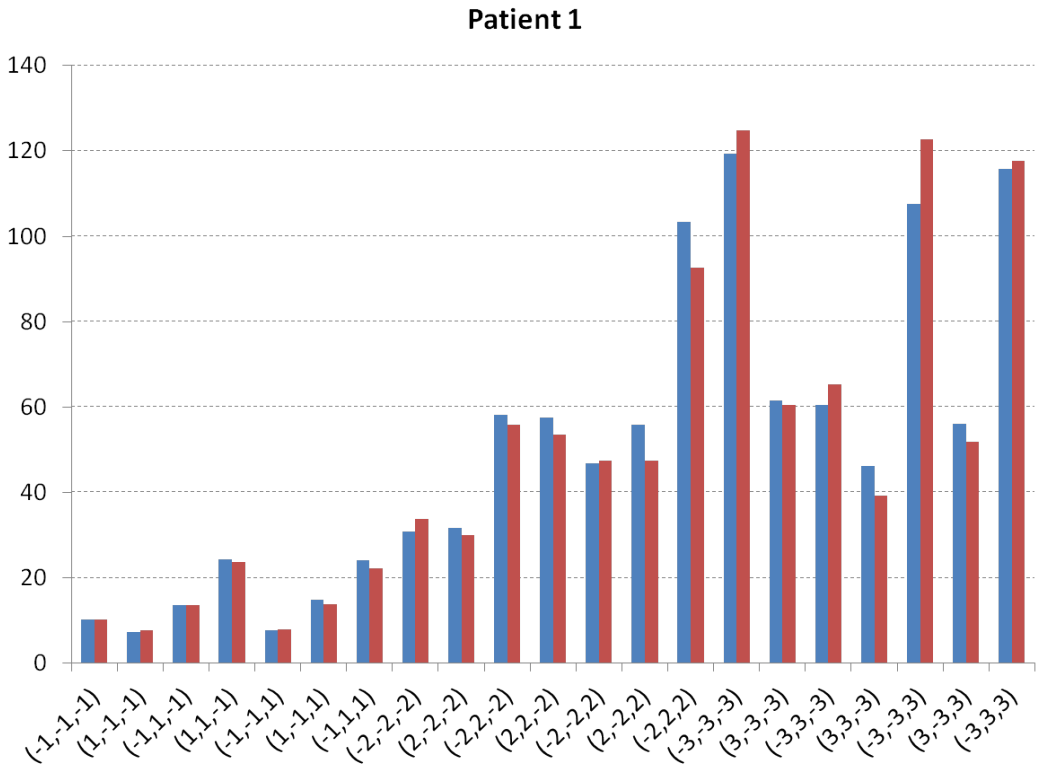
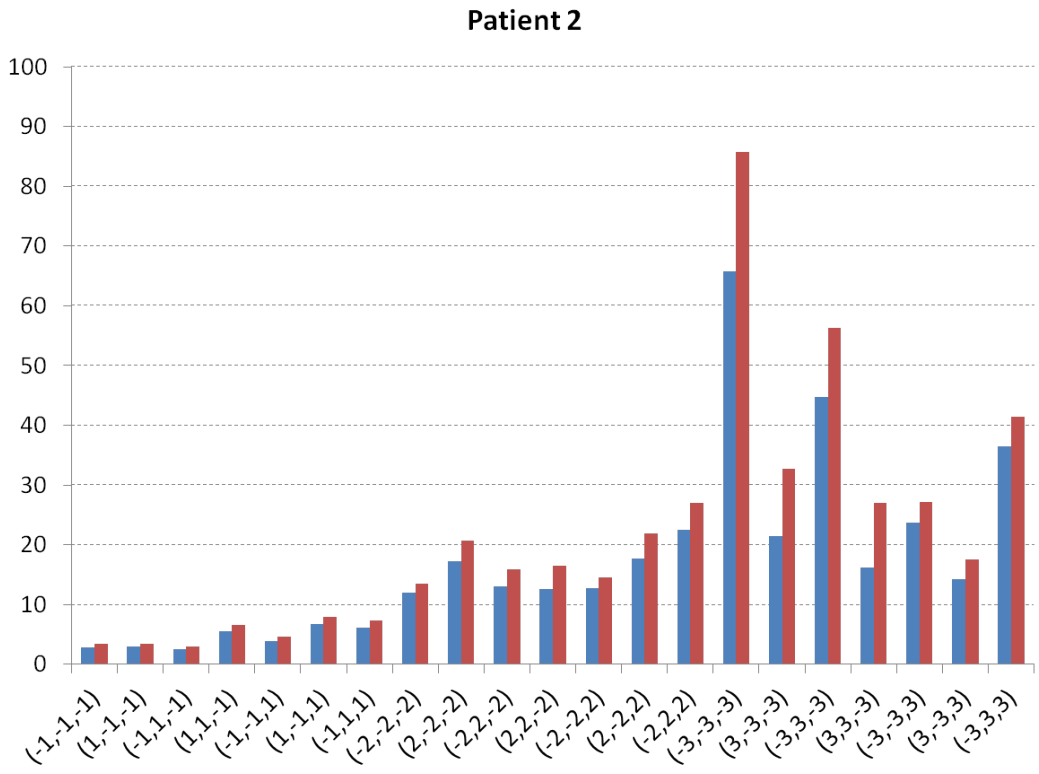


FIGURE C.6 – Ondelette mère utilisée (CDF-5/3).

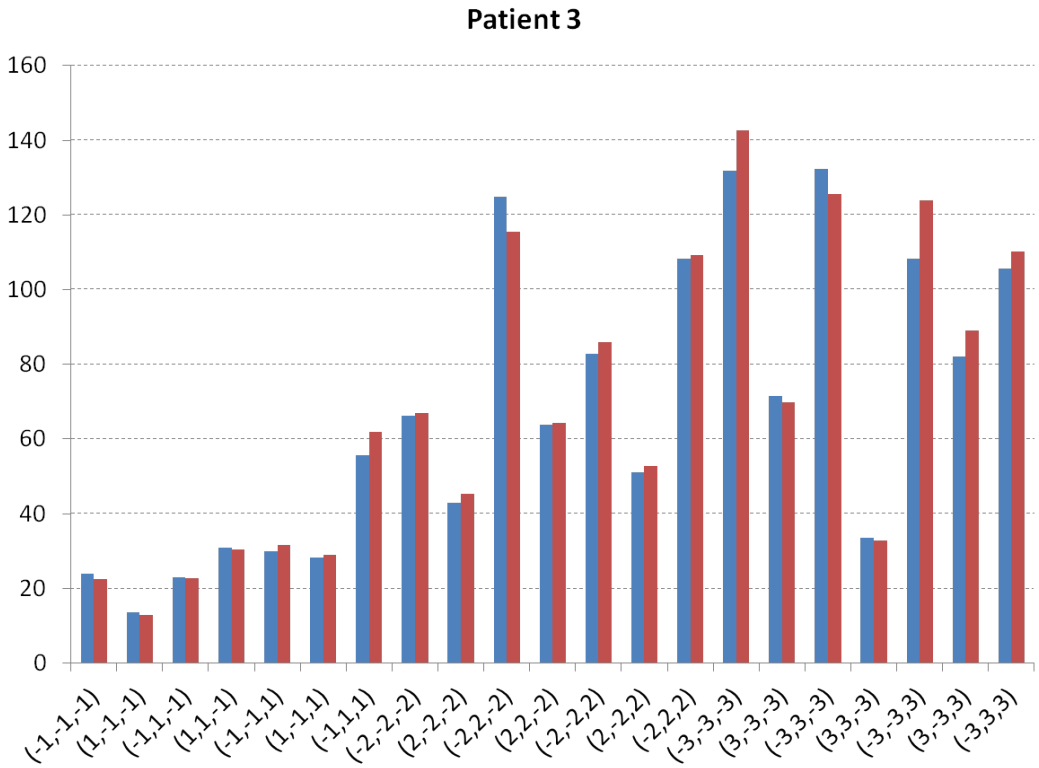
Pour chacun des cas, ce qui ressort en premier est la similitude des valeurs entre les deux zones, ce qui indique qu'il n'est pas possible de de différencier le canal par cette caractéristique. De plus, notez qu'il n'existe pas un modèle constant de réponse aux ondelettes : si les quatre premiers patients semblent montrer une augmentation de l'énergie proportionnelle au niveau de décomposition, ce n'est pas le cas des deux derniers, qui ont un profil différent. Nous ne sommes donc pas en mesure d'identifier le canal par cette analyse. Plusieurs éléments pourraient permettre d'améliorer cela. Premièrement, le calcul de valeurs caractéristiques autres que l'énergie pourrait affiner l'analyse des coefficients et mettre au jour des différences pertinentes. Deuxièmement, l'impact du choix de l'ondelette mère n'a pas été étudié, nous l'avons choisi principalement pour le fait qu'elle est symétrique, rendant l'analyse invariable par effet miroir, et faire varier celle-ci pourrait conduire à de meilleurs résultats.


Patient 1
140
120
100
80
60
40
20
0
(-1,-1,-1)
(1,-1,-1)
(-1,1,-1)
(1,1,-1)
(-1,-1,1)
(1,-1,1)
(-1,1,1)
(-2,-2,-2)
(2,-2,-2)
(-2,2,-2)
(2,2,-2)
(-2,-2,2)
(2,-2,2)
(-2,2,2)
(-3,-3,-3)
(3,-3,-3)
(-3,3,-3)
(3,3,-3)
(-3,-3,3)
(3,-3,3)
(-3,3,3)

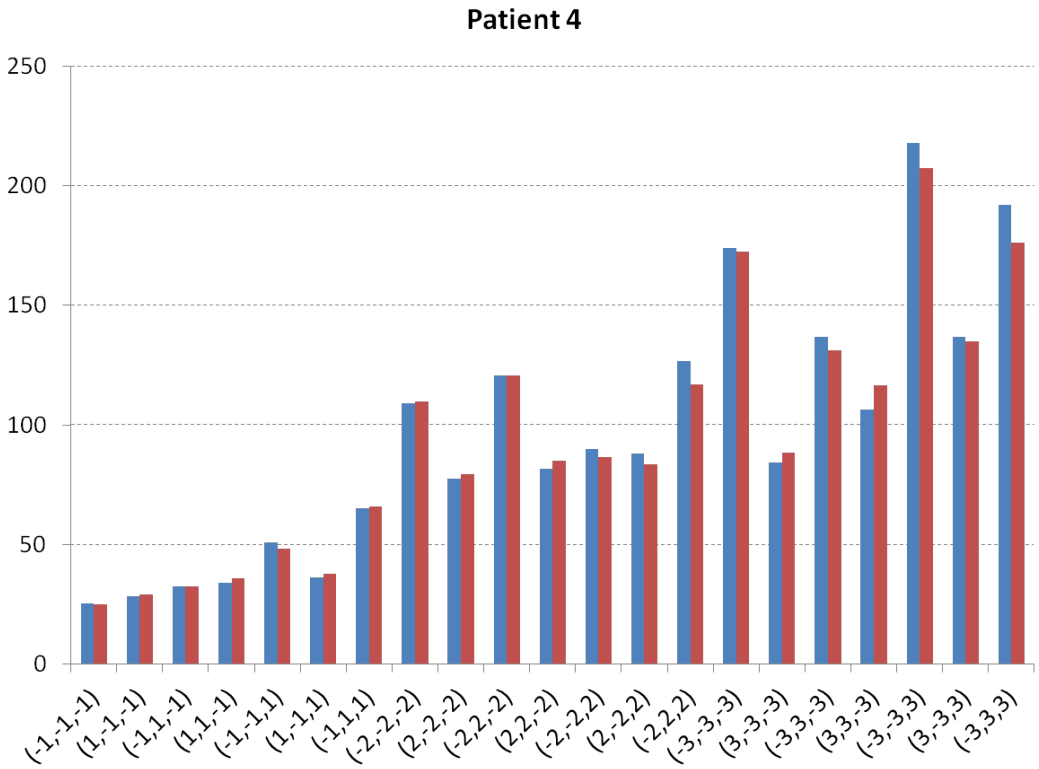
(a)


Patient 2
100
90
80
70
60
50
40
30
20
10
0
(-1,-1,-1)
(1,-1,-1)
(-1,1,-1)
(1,1,-1)
(-1,-1,1)
(1,-1,1)
(-1,1,1)
(-2,-2,-2)
(2,-2,-2)
(-2,2,-2)
(2,2,-2)
(-2,-2,2)
(2,-2,2)
(-2,2,2)
(-3,-3,-3)
(3,-3,-3)
(-3,3,-3)
(3,3,-3)
(-3,-3,3)
(3,-3,3)
(-3,3,3)

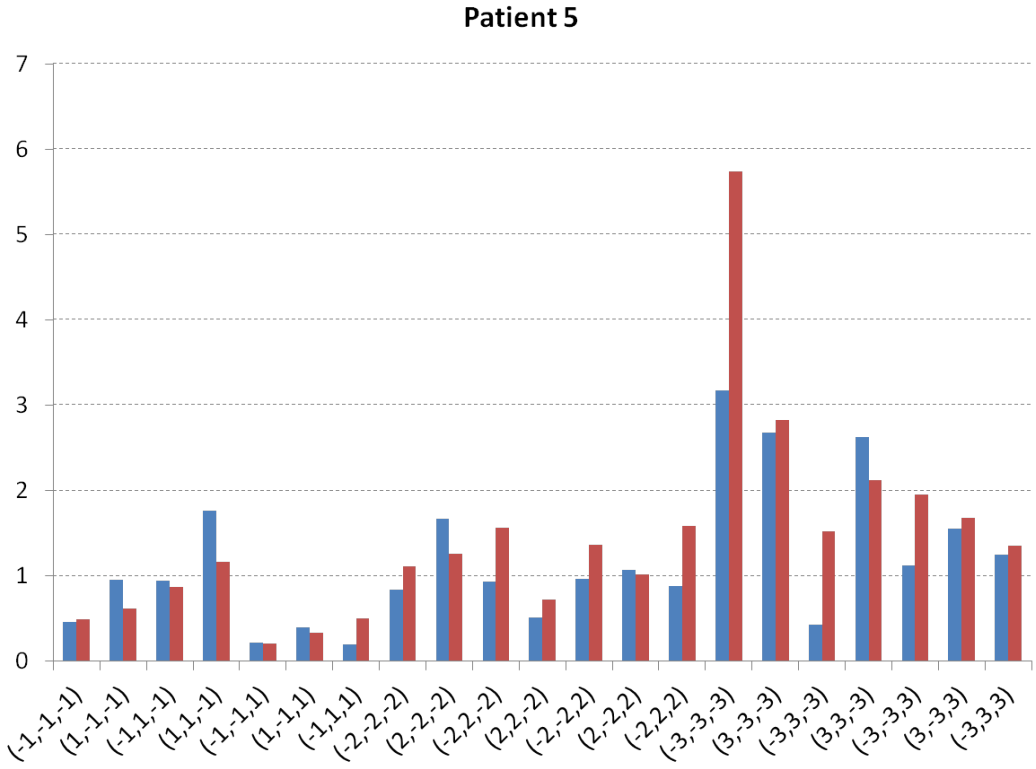
(b)


Patient 3
160
140
120
100
80
60
40
20
0
(-1,-1,-1)
(1,-1,-1)
(-1,1,-1)
(1,1,-1)
(-1,-1,1)
(1,-1,1)
(-1,1,1)
(-2,-2,-2)
(2,-2,-2)
(-2,2,-2)
(2,2,-2)
(-2,-2,2)
(2,-2,2)
(-2,2,2)
(-3,-3,-3)
(3,-3,-3)
(-3,3,-3)
(3,3,-3)
(-3,-3,3)
(3,-3,3)
(-3,3,3)

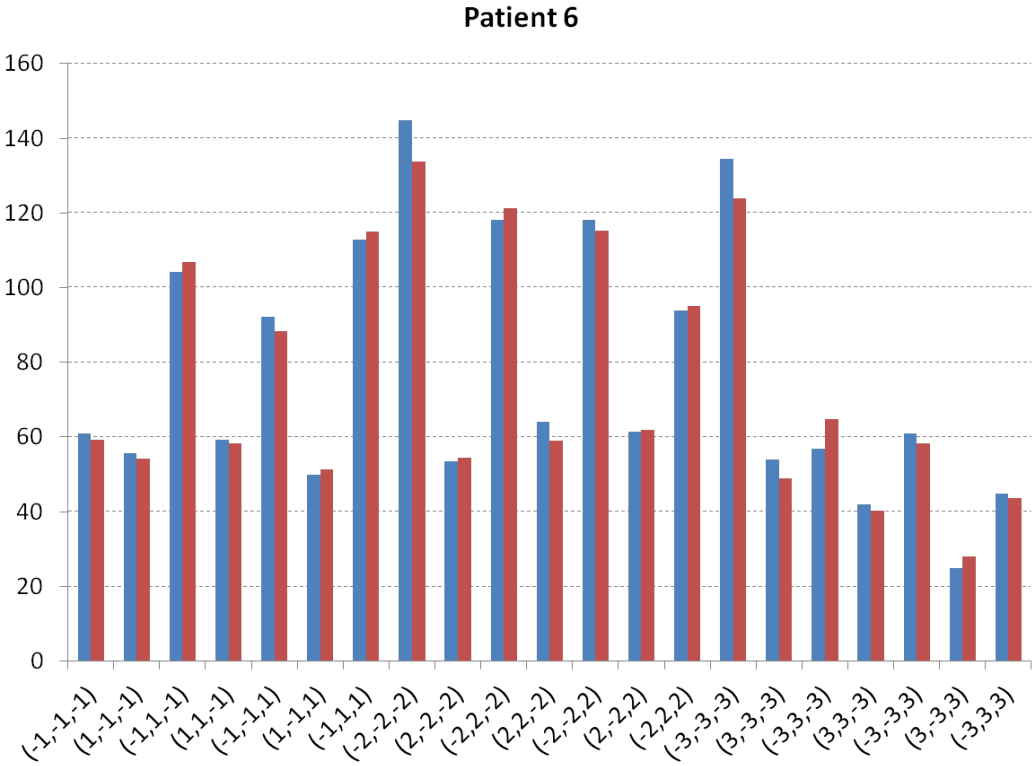
(c)


Patient 4
250
200
150
100
50
0
(-1,-1,-1)
(1,-1,-1)
(-1,1,-1)
(1,1,-1)
(-1,-1,1)
(1,-1,1)
(-1,1,1)
(-2,-2,-2)
(2,-2,-2)
(-2,2,-2)
(2,2,-2)
(-2,-2,2)
(2,-2,2)
(-2,2,2)
(-3,-3,-3)
(3,-3,-3)
(-3,3,-3)
(3,3,-3)
(-3,-3,3)
(3,-3,3)
(-3,3,3)

(d)



(e)



(f)

FIGURE C.5 – Énergie comparée des coefficients d'ondelettes (en bleu pour le canal, en rouge pour l'environnement) pour chaque patient. Chaque ensemble est numéroté par $(\pm n, \pm n, \pm n)$, où n est le niveau de décomposition, et le signe indique le filtre utilisé dans cette direction (correspondant successivement à X, Y et Z) : $+$ pour le passe bas (donc l'approximation) et $-$ pour le passe-haut (donc les détails).

C.4 Conclusion

Ces premières expériences sur le canal mandibulaire ne nous ont pas permis d'identifier un moyen robuste sur lequel s'appuyer pour réaliser une approche de segmentation directe du canal mandibulaire. Il nous apparaît donc nécessaire de s'appuyer sur d'autres structures pour parvenir à une identification du trajet du canal. Plusieurs solutions sont possibles, notamment dans la littérature (voir chapitre 2) : en passant par les foramens mandibulaire et mentonnier, une approche par plus court chemin donne des résultats cohérents, mais il est alors nécessaire de procéder à une analyse transversale pour segmenter l'épaisseur du canal. Une approche par forme de type modèle actif permet aussi d'estimer approximativement la position canalaire, avec un risque d'erreur si le canal du sujet possède un placement inhabituel. À la suite de ces expériences, et dans la logique de s'appuyer sur notre modèle structurel et sur les segmentations antérieures, nous avons décidé de nous concentrer sur la segmentation dentaire qui, en plus de son intérêt intrinsèque, fournirait une information forte pour accéder au canal.

Annexe D

Description du modèle structurel

Nous décrivons ici le contenu de notre modèle, en passant en revue les différentes structures que nous avons intégrées. Notons que pour le moment, la modélisation reflète une anatomie dans le cas sain pour un adulte. L'enrichissement avec des connaissances sur l'enfant et dans le cadre de pathologies constituerait un développement important qui permettrait de traiter des problèmes plus spécifiques. Concernant les informations sur l'intensité, comme le CBCT ne permet pas d'utiliser d'échelle de référence (de type Hounsfield), la caractérisation est volontairement qualitative. Nous distinguons principalement quatre grandes catégories : faible, moyenne, forte, hyper-intense, qui correspondent respectivement aux intensités apparentes de l'air, des tissus mous, des tissus durs (os) et enfin de l'émail (de même que le métal). Toutefois les intensités étant stables dans nos images (voir partie 3.1), nous pouvons définir des valeurs pour ces classes (voir table 4.2).

Fosses nasales

Les fosses nasales, au sens large, représentent toutes les cavités supérieures de la voie aérienne, c'est-à-dire les narines, les cavités nasales (avec les cornets et la narine interne), les sinus paranasaux et la partie terminale du nasopharynx. Ces structures sont donc principalement des couloirs aériens, impliqués dans la ventilation, mais participent aussi à l'olfaction, la gustation ainsi qu'à la défense immunitaire. Les traitements dentaires amènent parfois les dentistes à devoir prendre en compte cet espace, notamment dans le cas de pose d'implants dans l'os maxillaire (l'implant peut parfois pénétrer dans les cavités en présence d'une résorption osseuse), ou encore dans le cas de traitements orthodontiques, pour s'assurer d'une bonne structuration globale. Voici les caractéristiques de cet élément :

— au-dessus de l'os maxillaire,

— adjacent à l'os maxillaire,

— moins intense que l'os maxillaire,

— apparence homogène, pas de texture,

— intensité faible, similaire à l'intensité de l'air,

— forme complexe, polylobée et variable (voir figure D.1).

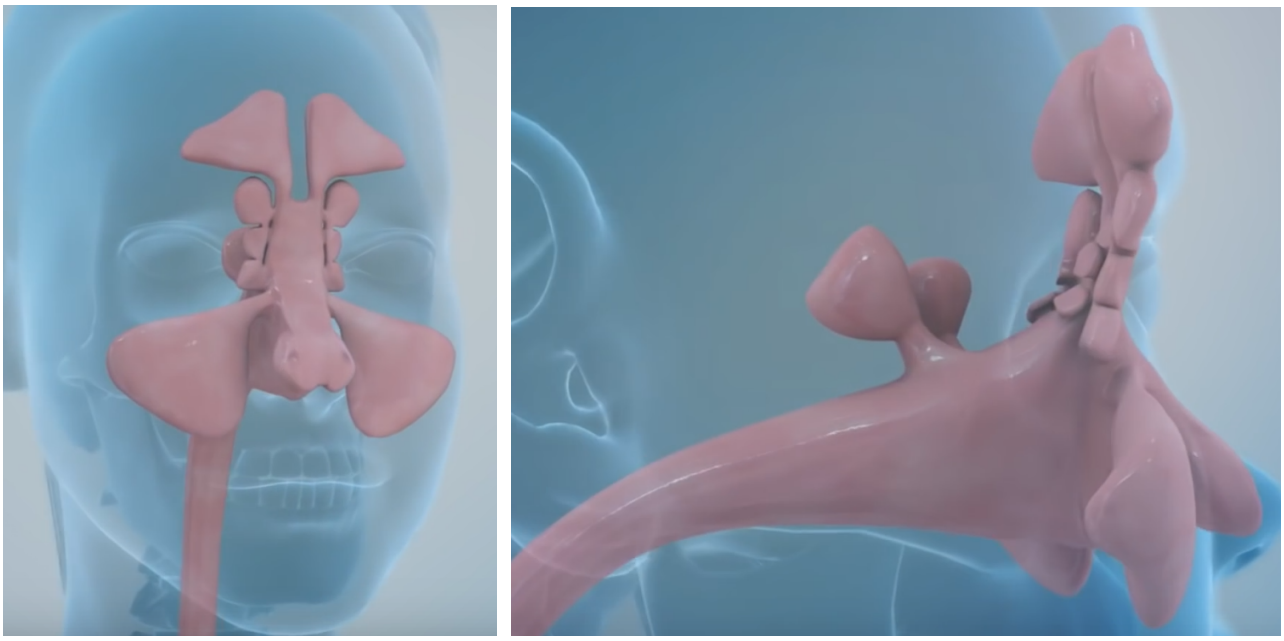
Os maxillaires et mandibule

Les os maxillaires, au nombre de deux (droit et gauche, de chaque côté de la ligne médiane), constituent le support de l'arcade dentaire supérieure (voir figure D.2 (a)), et intéressent à ce titre les praticiens, principalement pour juger de leur densité osseuse et de la place disponible pour la dentition. Caractéristiques :

— en-dessous des fosses nasales,

— au-dessus de l'arcade dentaire supérieure,

— adjacents aux fosses nasales, à l'arcade dentaire supérieure et à la cavité buccale,



(a) (b)

FIGURE D.1 – Représentation 3D de la forme des fosses nasales. (a) Vue de face. (b) Vue de profil. Images provenant de www.fauquierent.net, extraites d'une vidéo de chirurgie des sinus.

— constitués d'os cortical et d'os trabéculaire,

— l'os cortical entoure l'os trabéculaire,

— l'os cortical constitue le contour de l'os,

— l'os cortical est une zone d'aspect homogène,

— l'os trabéculaire est une zone texturée (dont les caractéristiques sont inconnues),

— les deux maxillaires sont symétriques et soudés (approximativement au niveau du philtrum), c'est-à-dire qu'il est parfois difficile de visualiser la jonction. Nous pouvons donc conceptualiser ces os comme un ensemble unique,

— l'os cortical possède une forte intensité, similaire à celle des autres structures osseuses. Les filaments osseux de l'os trabéculaire possèdent une intensité de même nature, mais le signal est parfois imperceptible en fonction du diamètre du filament, et de la résolution de l'image. La partie de tissu conjonctif est d'intensité moyenne, similaire à celle des autres tissus mous,

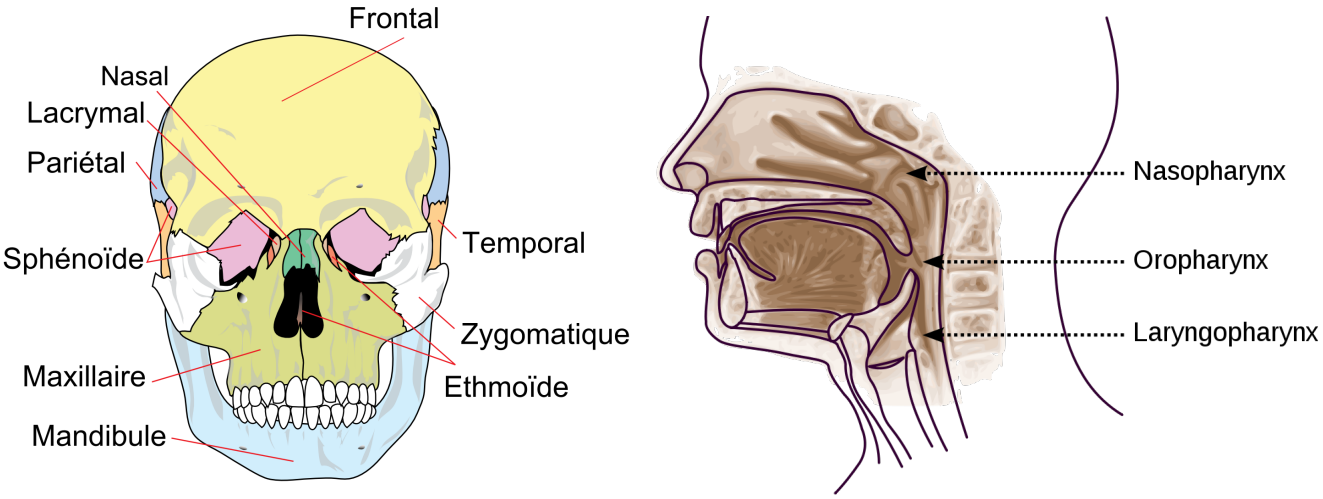
— forme globale de tétraèdre avec pour chaque maxillaire une face antérieure (sous-cutanée et palpable), une face postérieure tournée vers la tempe, une face orbitale constituant la partie basse de l'orbite et une base qui constitue une paroi des cavités nasales et buccales. Le sommet se trouve au niveau de la jonction avec l'os zygomatique.

La mandibule (ou os mandibulaire, figure D.3) est cliniquement importante dans toutes les spécialités dentaires : l'orthodontiste l'étudie pour connaître la place disponible pour la denture, pour son rôle dans l'occlusion et dans le prognathisme ; l'endodontiste et l'implantologue vérifient son état osseux ainsi que la position du canal mandibulaire qui le traverse ; le chirurgien planifie les interventions de résection osseuse. C'est aussi l'os qui porte l'arcade dentaire inférieure. Contrairement aux maxillaires, la mandibule finit par être un organe unique à la suite de la fusion de deux os initiaux durant la croissance, et la symphyse n'est pas visible à l'image. Caractéristiques :

— en-dessous de l'arcade dentaire inférieure et de la cavité buccale,

— devant le pharynx,

— constitué d'os cortical et d'os trabéculaire,



(a) (b)

FIGURE D.2 – Schéma des os du crâne humain en façade (a) et des voies aéro-digestives (b). Illustrations extraites de Wikipédia.

— l'os cortical entoure l'os trabéculaire,

— l'os cortical constitue le contour de l'os,

— l'os cortical est une zone d'aspect homogène,

— l'os trabéculaire est une zone hétérogène, de texture non modélisée,

— l'os cortical possède une forte intensité, similaire à celle des autres structures osseuses. Les filaments osseux de l'os trabéculaire possèdent une intensité de même nature, mais le signal est parfois imperceptible en fonction du diamètre du filament, et de la résolution de l'image. La partie de tissu conjonctif est d'intensité moyenne, similaire à celle des autres tissus mous,

— la forme de la mandibule est une arche dont les extrémités sont recourbées vers le sommet du crâne et le sommet est en position frontale. L'angle formé par les branches est variable. Plus il est important, plus le menton apparaît pointu (forme en V). À l'inverse, si les branches sont presque parallèles, le menton sera plat, tendant vers une forme en U.

Cavité buccale

La cavité buccale est simplement la partie aérienne de la bouche. Si son identification n'est pas primordiale pour les praticiens, son volume donne des indications sur l'espace global disponible en orthodontie. Caractéristiques :

— adjacente aux arcades dentaires et au pharynx,

— en-dessous des fosses nasales et de l'os maxillaire,

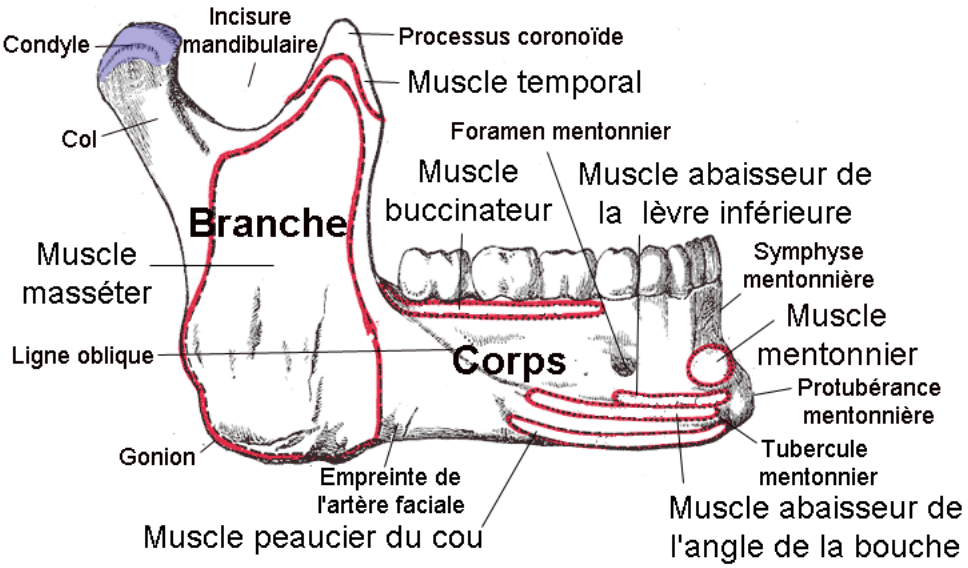
— devant le pharynx,

— au-dessus de l'os mandibulaire,

— intensité de l'air (faible),

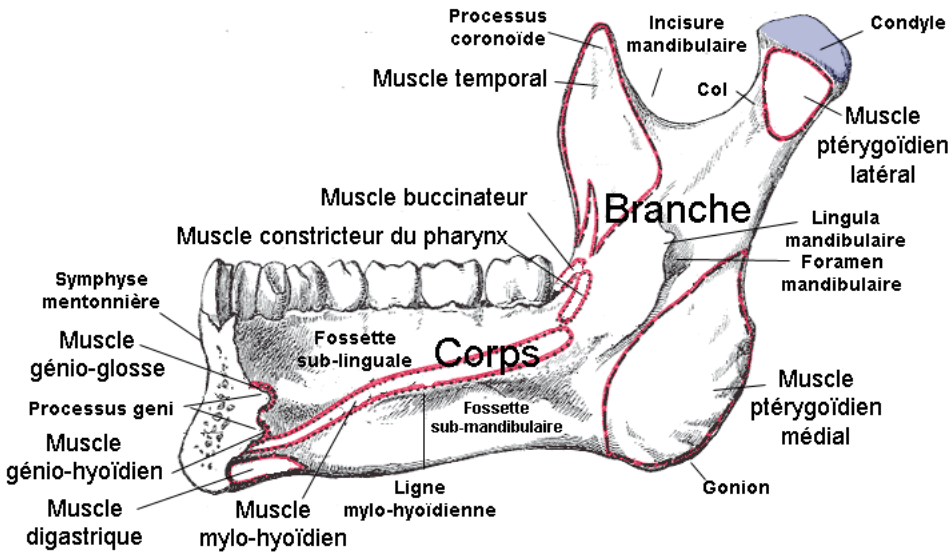
— zone homogène,

— forme variable selon le positionnement mandibulaire, l'ouverture buccale et le positionnement lingual. La partie centrale est globalement de forme hémi-cylindrique (la courbure étant donnée par le palais).



Mandibule
(vue latérale droite)

(a)



Mandibule
(vue médiale droite)

(b)

FIGURE D.3 – Anatomie de la mandibule. Illustrations de Wikipédia.

Pharynx

Le pharynx est l'espace aérien qui constitue le carrefour des voies respiratoire et digestive (figure D.2 (b)). Il se compose successivement du nasopharynx, de l'oropharynx et du laryngopharynx. Selon le champ de vue, il n'apparaît pas toujours dans les images, mais son étude est importante en orthondontie pour garantir que l'évolution de la dentition n'impacte pas le volume et le diamètre des voies. Dans le

cas contraire, des problématiques liées à la respiration ou la phonation peuvent apparaître (apnée du sommeil, ronflement, dysphonie). Caractéristiques :

— derrière la cavité buccale, l'os mandibulaire et l'os maxillaire,

— adjacent à la cavité buccale et aux fosses nasales,

— intensité faible,

— zone homogène,

— forme tubulaire tendant à se resserrer dans sa partie basse.

Canal mandibulaire

Le canal mandibulaire est une structure tubulaire qui commence au niveau du foramen mandibulaire derrière l'épine de Spix (ou lingula mandibulaire, voir figure D.3) pour ressortir au niveau du foramen mentonnier. Cette structure se retrouve de chaque côté et contient un faisceau de fibres nerveuses ainsi que des artères, qui irriguent la partie inférieure de la face. C'est une zone critique pour toute intervention ou contrôle dentaire et les praticiens doivent l'identifier afin de s'assurer qu'il ne sera pas endommagé. Dans le cas contraire, des hémorragies, des douleurs neurologiques ou même des anesthésies permanentes peuvent survenir. Caractéristiques :

— inclus dans la mandibule et traverse l'os trabéculaire,

— longe l'os cortical,

— connecte le foramen mentonnier au foramen mandibulaire,

— intensité moyenne,

— texture non différentiable de celle de l'os trabéculaire (voir annexe C),

— forme tubulaire.

Foramens mentonniers

Les foramens mentonniers sont deux ouvertures dans la façade frontale de la mandibule, situés approximativement sous les prémolaires. Ils sont cliniquement importants car ils constituent la sortie du canal mandibulaire de chaque côté. Ils fournissent donc un indice dans son identification. Caractéristiques :

— en-dessous des prémolaires (le plus souvent à l'aplomb de la seconde),

— inclus dans la mandibule, plus précisément dans l'os cortical,

— adjacents à l'os trabéculaire,

— intensité moyenne, d'aspect homogène,

— forme d'œillet de diamètre moyen de 2,7mm.

Arcade dentaire supérieure

L'arcade dentaire supérieure correspond à l'ensemble des dents maxillaires. Elle comprend donc seize dents (quatre incisives, deux canines, quatre prémolaires et six molaires, dont deux "de sagesse" qui ne sont pas toujours externalisées) correspondant aux numéros dentaires de 11 à 18 (côté droit) et de 21 à 28 (côté gauche) débutant aux incisives. L'arcade n'est donc pas une structure clairement délimitée en soit, mais un ensemble d'objets. Son identification revêt un intérêt pour l'obtention de mesures orthodontiques liées à la forme ou à son angulation. Caractéristiques :

— constituée des seize dents maxillaires,

— en-dessous des maxillaires, au-dessus de l'arche inférieure et devant le pharynx,

— adjacente aux maxillaires et à la cavité buccale,

— l'arcade constitue un alignement local des dents (voir partie 4.3.2),

— forme d'arche dépendant de celle des maxillaires. La courbure sera déterminée par le support osseux.

Arcade dentaire inférieure

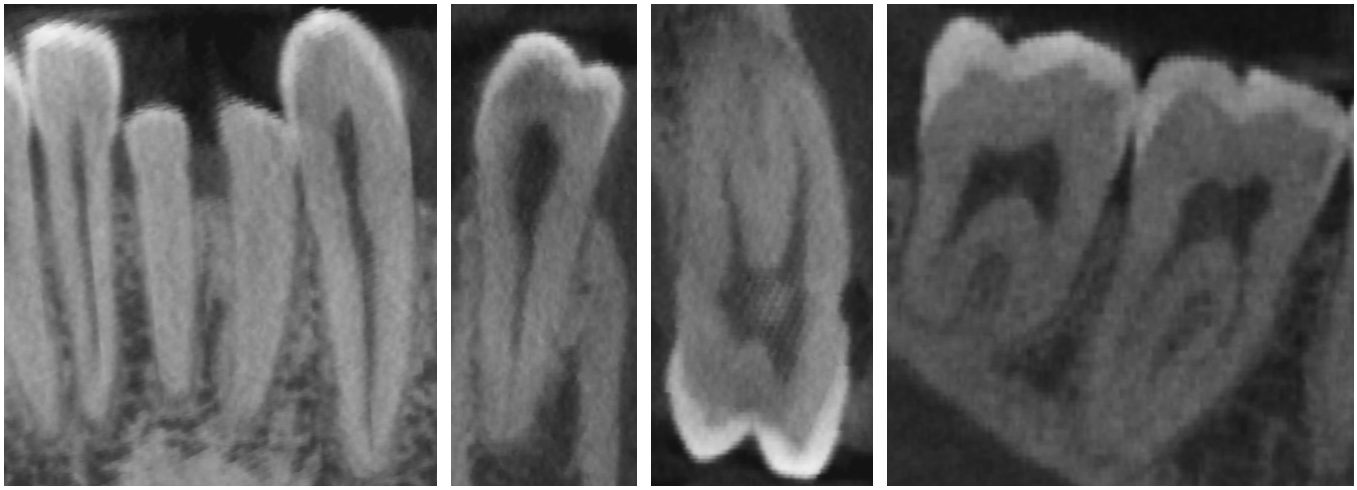
L'arcade dentaire inférieure est le pendant mandibulaire de l'arcade supérieure. Elle comprend donc elle aussi seize dents, distribuées de la même manière et correspondant aux numéros 41 à 48 (côté droit) et de 31 à 38 (côté gauche). Ici aussi son intérêt clinique se retrouve en orthodontie. Caractéristiques :

- constituée des seize dents mandibulaires,
- au-dessus de la mandibule, en-dessous de l'arche dentaire supérieure et devant le pharynx,
- adjacente à la mandibule et à la cavité buccale,
- l'arcade constitue un alignement local des dents,
- forme d'arche dépendant de celle de la mandibule. Usuellement plus facilement évasée que l'arche supérieure.

Dent

La dent est l'organe central en clinique (sa structure a été présentée en figure 1.3). Son identification et sa délimitation sont donc d'un intérêt primordial comme présenté dans les chapitres précédents. Caractéristiques :

- position proche d'autres dents, pouvant aller jusqu'à l'adjacence,
- ancrée (voir la définition dans la partie suivante) soit dans les maxillaires (n° 11 à 28), soit dans la mandibule (n° 31 à 48),
- adjacente à la cavité buccale,
- au-dessus de la mandibule et en-dessous des maxillaires,
- constituée d'émail, de dentine, d'une chambre pulpaire et de canaux radiculaires,
- l'émail entoure partiellement la dentine,
- la chambre pulpaire et les canaux radiculaires sont inclus dans la dentine,
- la chambre pulpaire est adjacente aux canaux radiculaires et constitue leur lieu de réunion,
- tous ces tissus sont d'aspect homogène, sans texture notable,
- l'émail est hyper-intense, la dentine a une intensité forte. La chambre pulpaire et les canaux sont d'intensité moyenne,
- chaque type de dent à une forme caractéristique (voir chapitre 3),
- les canaux radiculaires ont une forme tubulaire,
- la forme de la chambre pulpaire varie grandement selon le type de dent. Les dents antérieures (incisives et canines) ont généralement une chambre de forme tubulaire dans le prolongement des canaux et il est difficile de positionner la séparation. Les prémolaires mandibulaires ont une chambre globalement ovoïde tandis que celles maxillaires présentent une forme plus complexe en forme de "H". Enfin les molaires tendent à posséder une chambre de la forme d'un disque aux faces concaves (voir figure D.4).



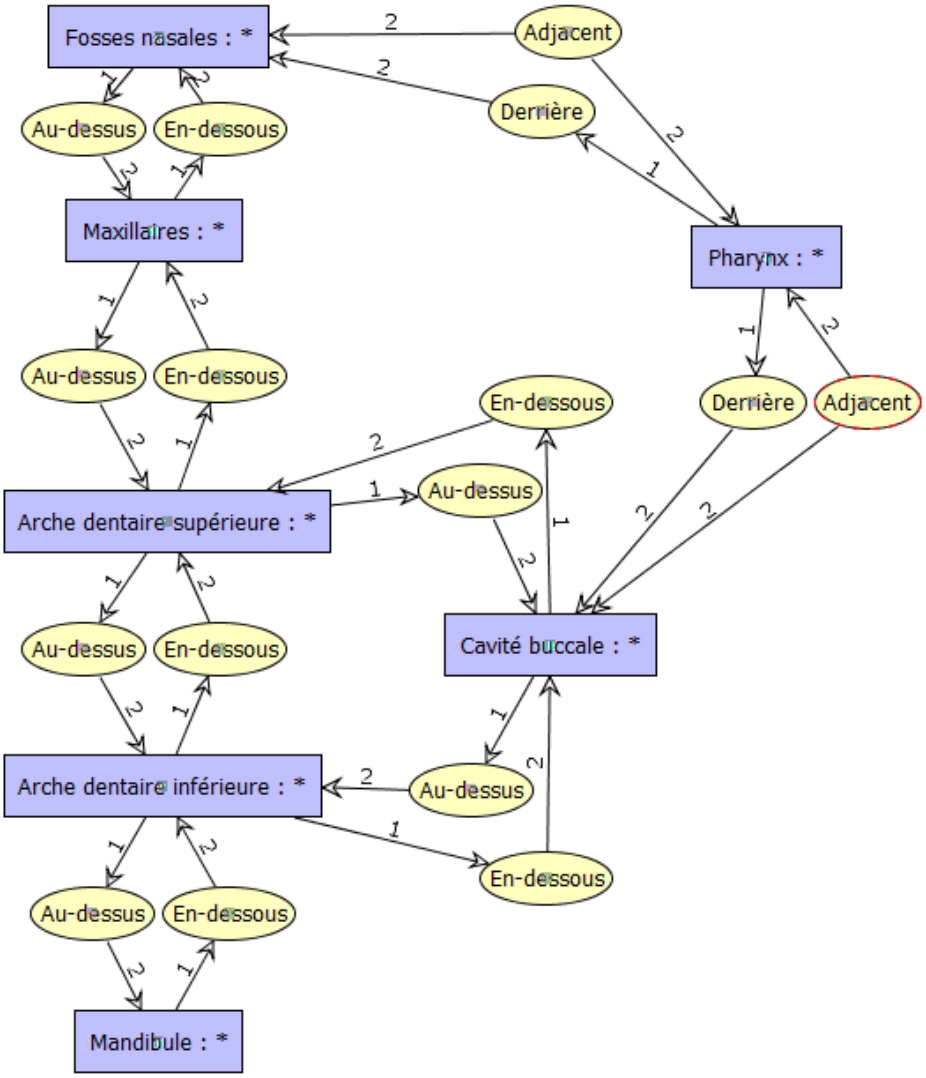
(a) (b) (c) (d)

FIGURE D.4 – Illustration des formes rencontrées pour la chambre pulpaire. (a) Incisives et canine (à droite). (b) Prémolaire mandibulaire. (c) Prémolaire maxillaire. (d) Molaires.

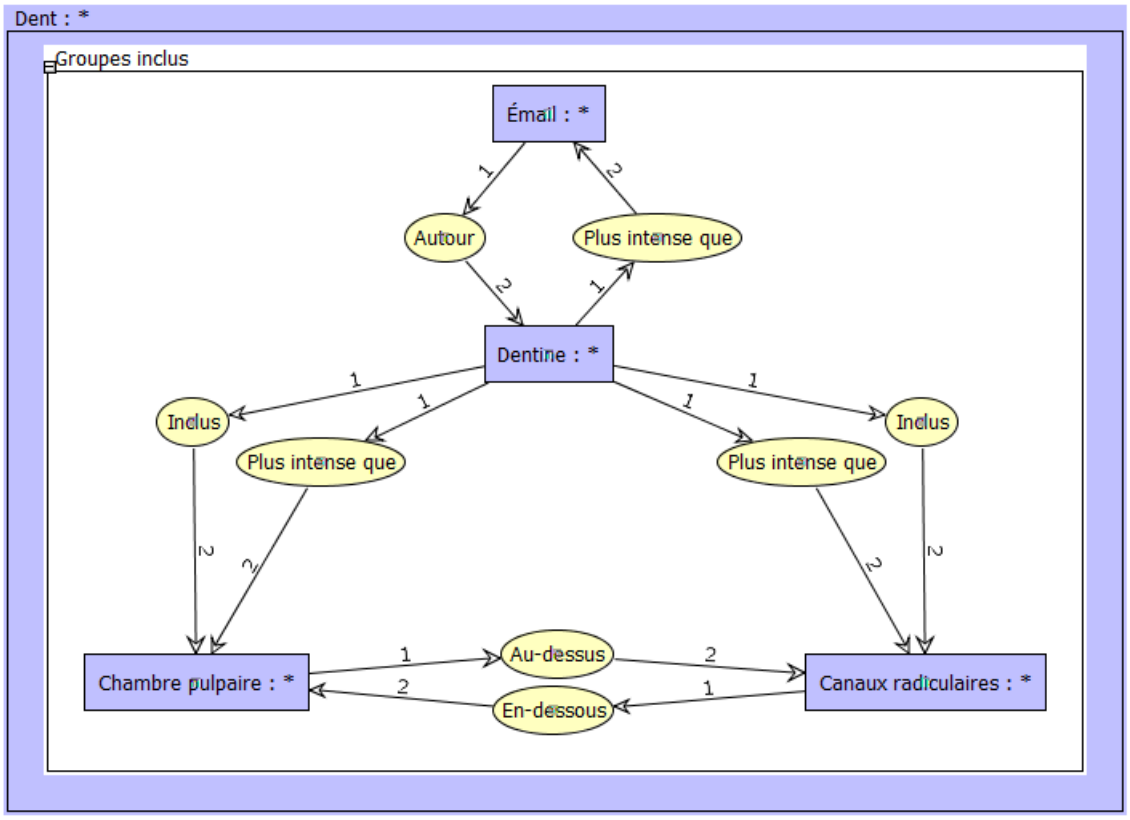
Annexe E

Représentation graphique du modèle structurel

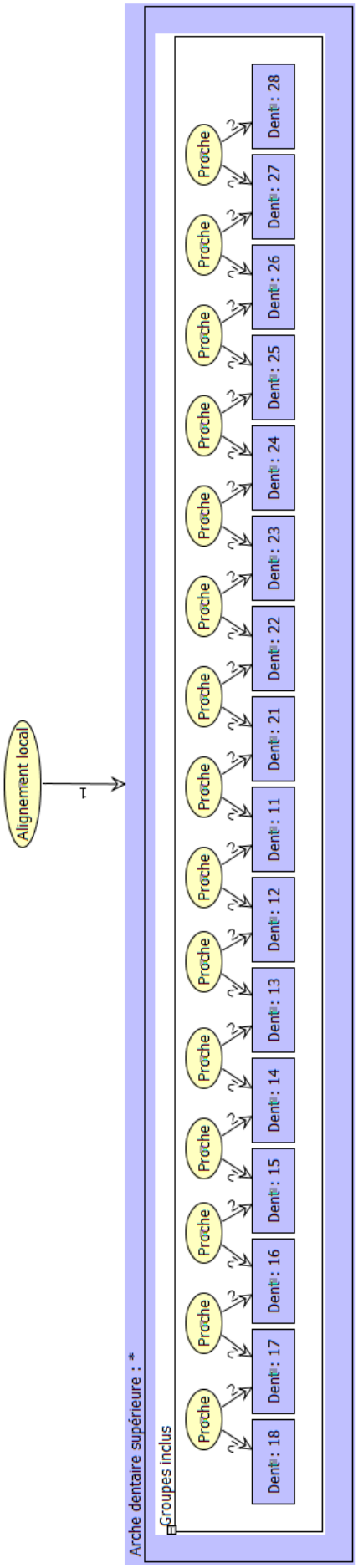
Ci-après est illustrée notre ontologie sous sa forme graphique, telle qu'elle peut être observée dans l'application CoGui (http ://www.lirmm.fr/cogui/).



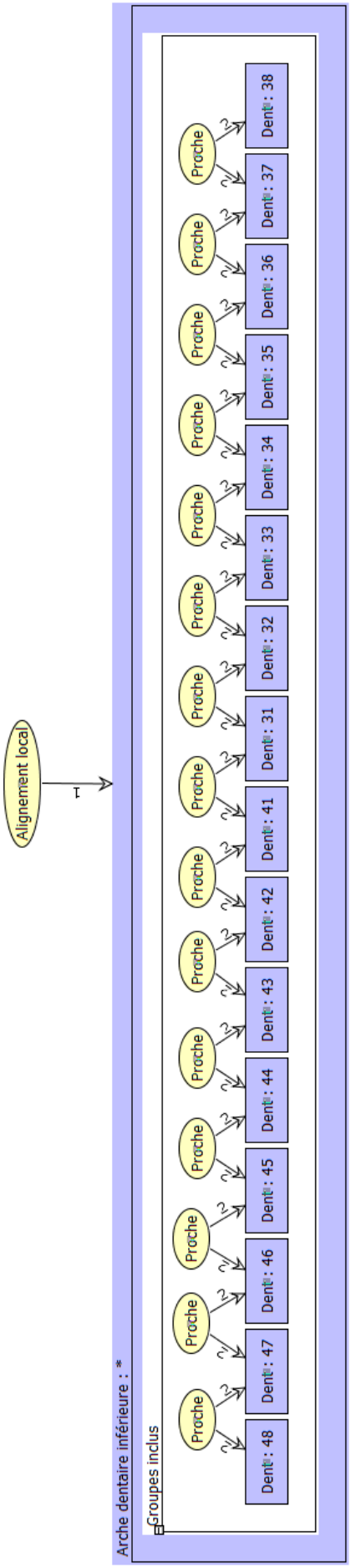
(a) Vue globale du modèle avec les principales structures de la zone maxillo-faciale



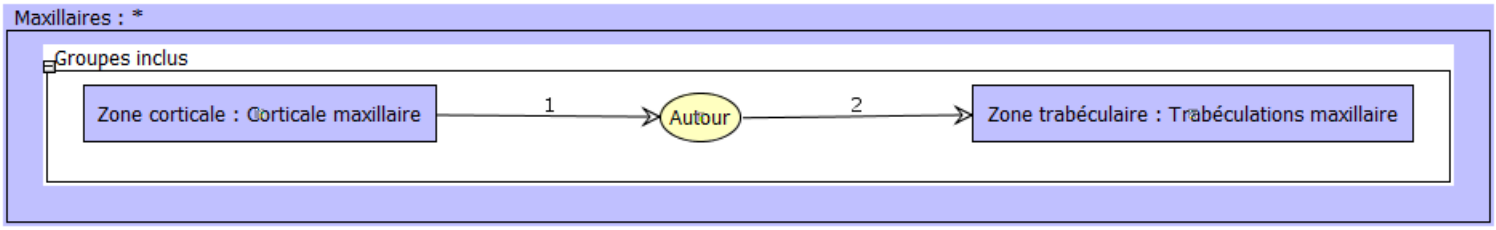
(b) Concept général d'une dent. Ce concept est spécialisé selon le type de dent, ce qui lui confère des attributs (de forme notamment, non représentés ici) différents.



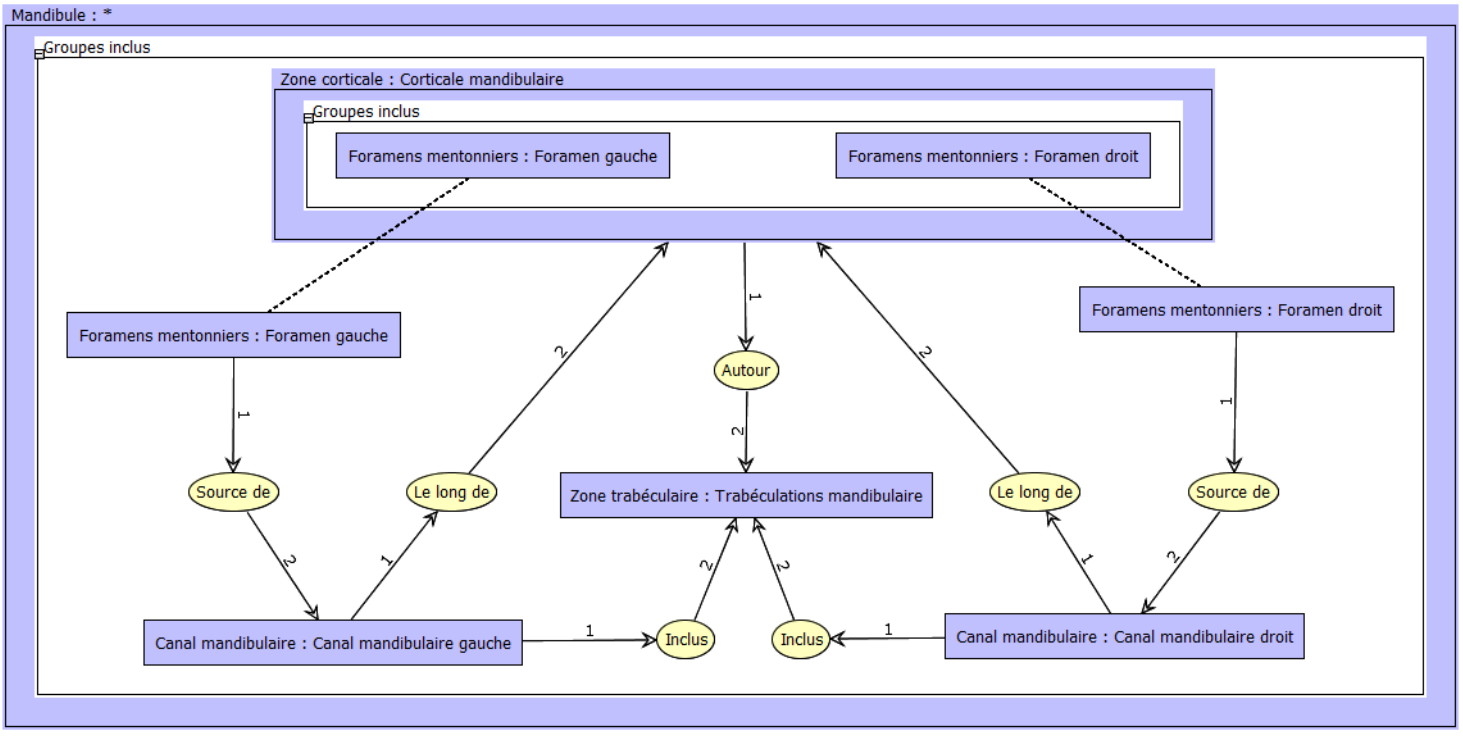
(c) Les arches dentaires sont un groupement de dents aligné localement. Ici l'arche maxillaire. Notons que chaque dent est individualisée par son numéro dentaire.



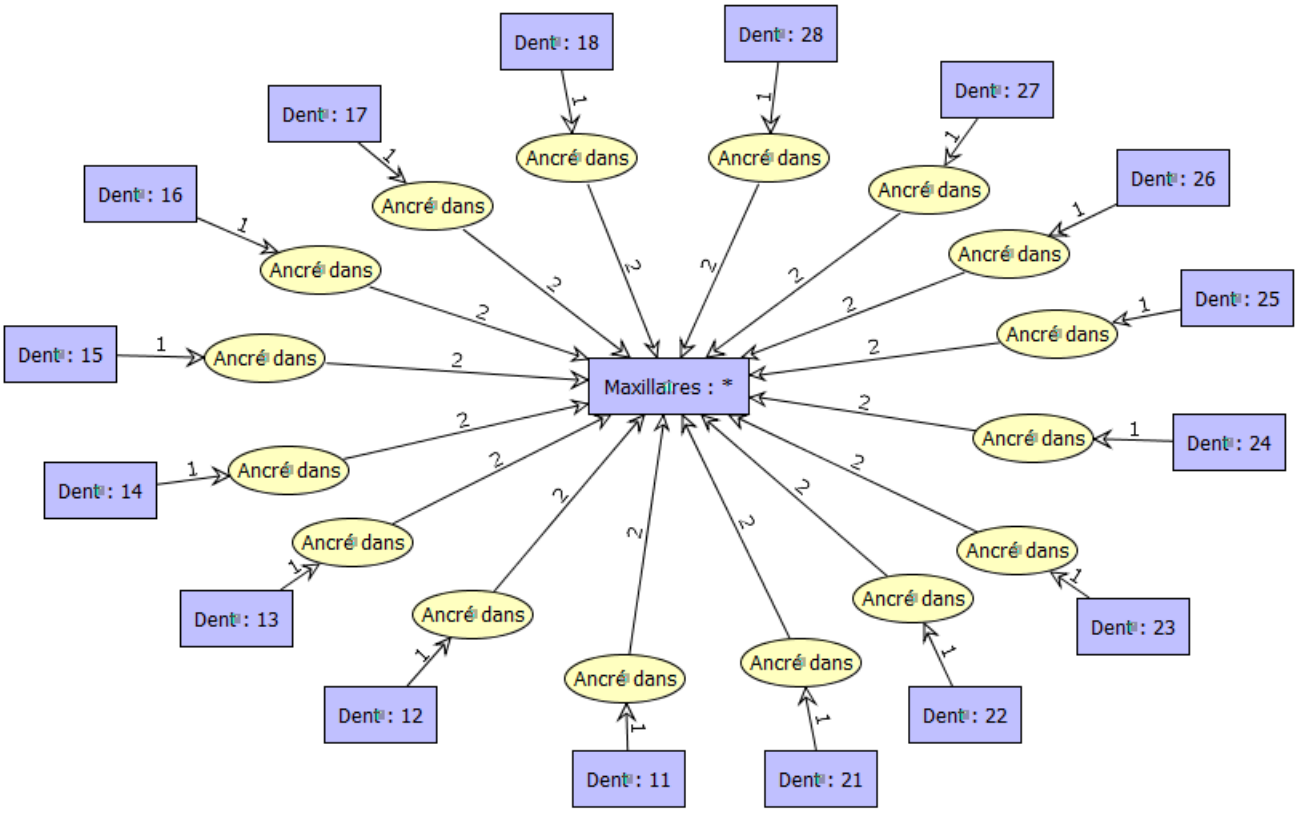
(d) Les arches dentaires sont un groupement de dents aligné localement. Ici l'arche mandibulaire. Notons que chaque dent est individualisée par son numéro dentaire.



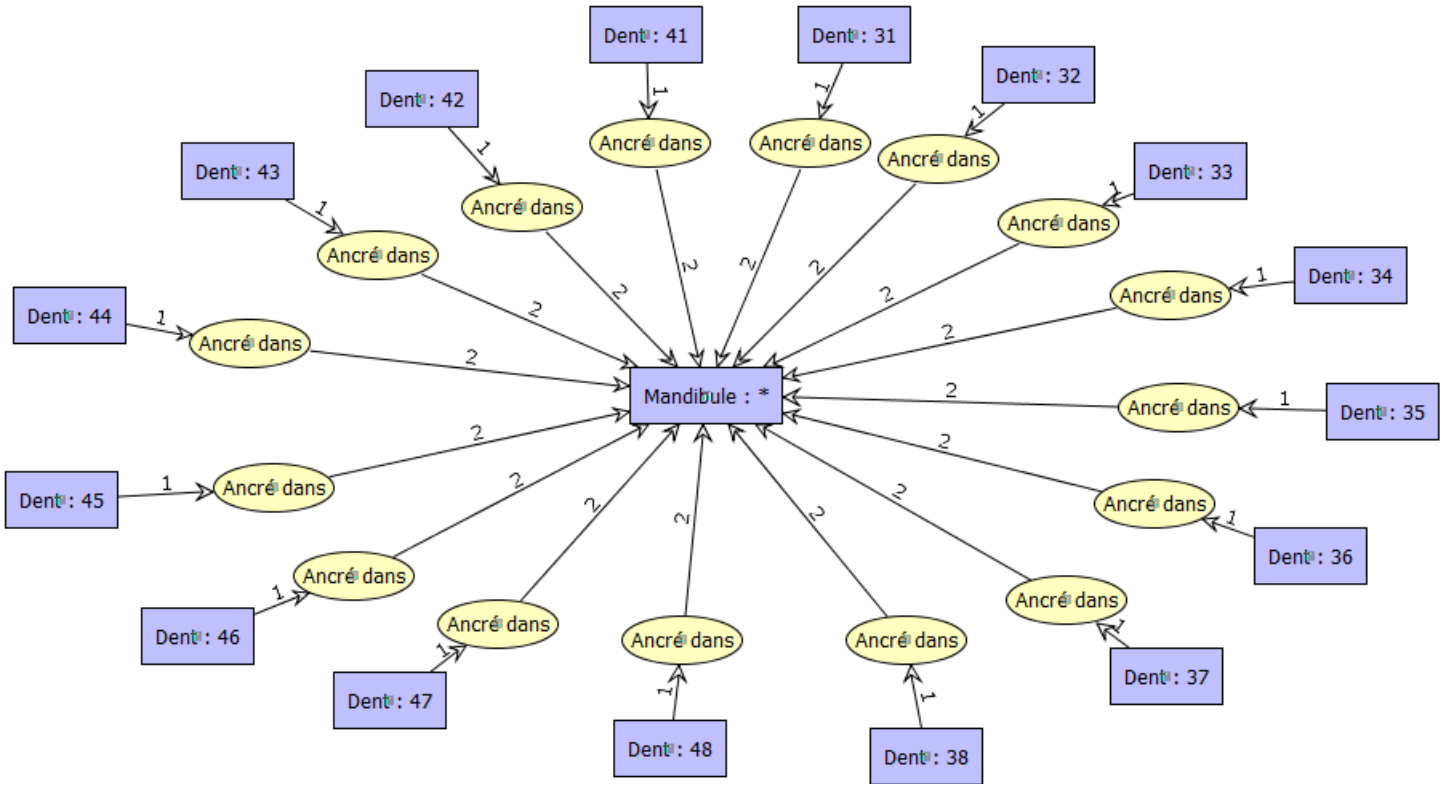
(e) Concept de l'ensemble uni des os maxillaires.



(f) Concept d'une mandibule.



(g) Chaque dent est individuellement liée à son support osseux par une relation d'ancrage. Ici les dents maxillaires.



(h) Chaque dent est individuellement liée à son support osseux par une relation d'ancrage. Ici les dents mandibulaires.

FIGURE E.1 – Modèle structurel de la zone maxillo-faciale.

Annexe F

Algorithme de localisation d'arche dentaire

En préambule de la thèse a été réalisée une méthode permettant de définir un modèle de la courbure de l'arche dentaire. Cet algorithme a pour but de permettre la réalisation de manière automatique d'une image pseudo-panoramique telle que présentée dans la figure C.1 dans le logiciel de visualisation de l'entreprise. Cette méthode repose sur des travaux pré-existants dans l'entreprise et ne constitue donc pas une contribution nouvelle.

Dans un premier temps nous réalisons le seuillage des données par la méthode d'Otsu [101] afin de ne récupérer que les structures denses comme l'os et les dents. Nous construisons ensuite une carte d'accumulation 2D grâce à ce volume binaire en sommant les voxels selon la direction axiale telle que :

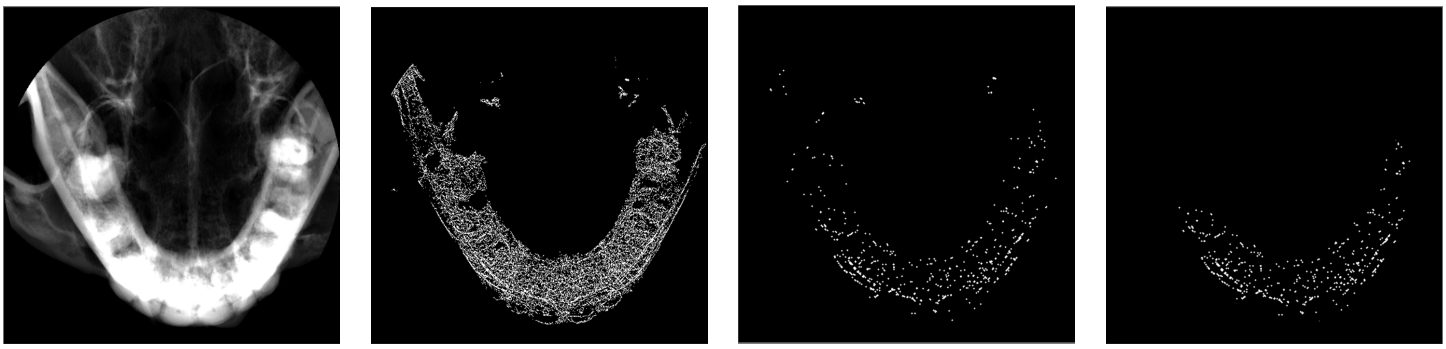
$$I_{acc}(x,y) = \sum_{z=0}^{Z-1} I_{binary}(x,y,z)$$

Nous recherchons ensuite les points saillants de cette carte, le but étant de garder l'information ne concernant que l'arche dentaire, par la méthode suivante :

1. Calcul du gradient de I_{acc} dans les directions x et y ∇I_{acc} : $\begin{pmatrix} \frac{\partial I_{acc}}{\partial x} \\ \frac{\partial I_{acc}}{\partial y} \end{pmatrix}$.
2. Normalisation du gradient pour avoir $[m_x, m_y]$ compris dans $[0,1]$.
3. Calcul des dérivées partielles de m_x et m_y ; $[s_x, s_y] = \nabla m_x$ et $[c_x, c_y] = \nabla m_y$.
4. L'ampleur des discontinuités D est obtenue comme :
   $D(x,y) = \sqrt{s_x^2(x,y) + s_y^2(x,y) + c_x^2(x,y) + c_y^2(x,y)}$.
5. Les crêtes sont séparées des vallées par le signe du laplacien dans l'image d'accumulation.
   $D1(x,y) = \begin{cases} D(x,y) & si \quad \text{sgn}(\nabla^2 I_{acc}(x,y)) < 0, \\ 0 & \text{sinon.} \end{cases}$
6. Élimination des points de $D1$ où les niveaux de gris des points correspondants de I_{acc} sont inférieurs à 20% de l'intensité maximale :$D2(x,y) = \begin{cases} D1(x,y) & si \quad I_{acc}(x,y) > 0.2 \times \max[I(x,y)], \\ 0 & \text{sinon.} \end{cases}$
7. Nous obtenons une carte de points saillants par seuillage de $D2$ avec la méthode d'Otsu.

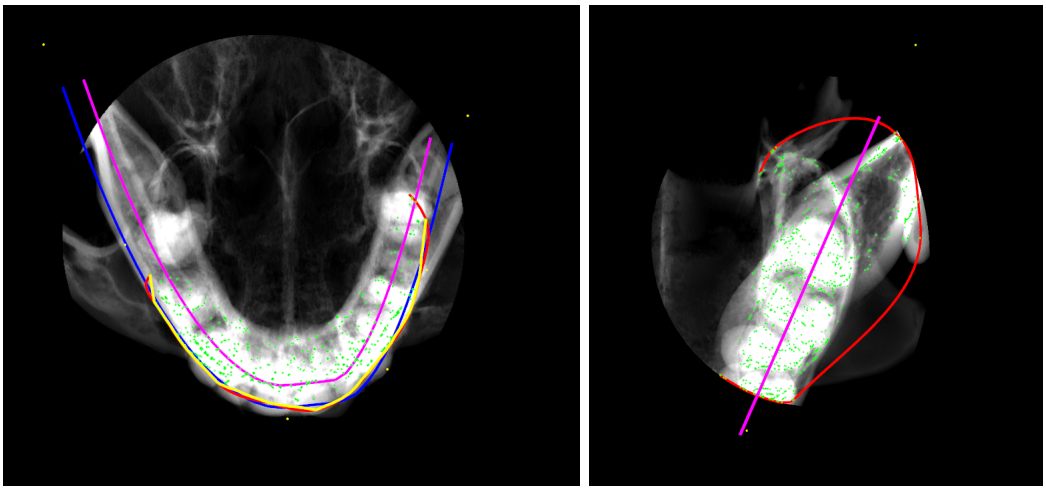
La carte de points saillants est filtrée par une ouverture morphologique puis les ensembles de points peu importants (moins de 80 voisins dans un rayon de 40 pixels) sont éliminés par seuillage avec une méthode par les plus proches voisins (kNN). Le processus est illustré en figure F.1.

Les points restants sont considérés comme étant représentatifs de l'arche dentaire. L'enveloppe convexe de ces points est calculée et la première approximation du contour externe de l'arche est définie en éliminant la section de l'enveloppe correspondant à l'arrière de l'arche (segment le plus long entre deux points de l'enveloppe ainsi que les segments adjacents jusqu'à inflexion de la courbe). À ce point il est important de souligner que le champ de vue initial peut contenir l'arche entière ou seulement un fragment, qui peut être courbe (fragment antérieur) ou assimilable à une droite



(a) Carte d'accumulation. (b) Carte de points saillants. (c) Points saillants après nettoyage. (d) Point saillants après filtrage des plus proches voisins.

FIGURE F.1 – Étapes de la sélection des points.



(a) Résultat sur un champ de vue de l'arche totale. (b) Résultat sur un petit champ de vue postérieur.

FIGURE F.2 – Résultats de l'approximation de l'arche. Sont représentés les points saillants (vert), l'enveloppe convexe initiale ouverte (rouge), l'enveloppe convexe suréchantillonnée (jaune), l'approximation parabolique externe (bleu) et le modèle final (violet).

(fragment postérieur). Nous analysons alors la forme du nuage de points saillants en calculant la hauteur du triangle formé par les deux extrémités de l'approximation et le point le plus éloigné de ces dernières. Si la hauteur est inférieure à un seuil de 15mm (distance anatomique), nous considérons que nous sommes en présence d'une section droite, donc postérieure de l'arche. Pour ces sections, une analyse en composantes principales est réalisée sur le nuage de point, et le premier vecteur propre nous permet de modéliser une droite à partir du barycentre des points. Cette droite est le modèle de l'arche. À l'inverse, si la section est détectée comme antérieure, la première approximation du contour externe est suréchantillonnée pour avoir une distribution uniforme, puis séparée en deux parties égales. Sur chaque partie, une parabole constitue l'approximation finale du contour externe de l'arche. Le modèle de courbure de l'arche est ensuite obtenu en faisant évoluer le contour externe de 4,5mm vers l'intérieur de l'arche (normale en chaque point). Un exemple de résultat est présenté en figure F.2.

Références

[1] D. Adalsteinsson and J. A. Sethian. "A Fast Level Set Method for Propagating Interfaces". *Journal of Computational Physics* 118.2 (1995), pp. 269–277 (cit. on p. 9).

[2] Y. Amit and D. Geman. *Randomized Inquiries about Shape; an Application to Handwritten Digit Recognition*. Technical report. Dept of Statistics, University of Chicago, 1994, p. 49 (cit. on p. 18).

[3] J. Atif, C. Hudelot, G. Fouquier, I. Bloch, and E. Angelini. "From Generic Knowledge to Specific Reasoning for Medical Image Interpretation Using Graph based Representations". *International Joint Conference on Artificial Intelligence - IJCAI*. 2007, pp. 224–229 (cit. on pp. 22, 104).

[4] J. Atif, C. Hudelot, O. Nempont, N. Richard, B. Batrancourt, E. Angelini, and I. Bloch. "GRAFIP: A Framework for the Representation of Healthy and Pathological Cerebral Information". *IEEE International Symposium on Biomedical Imaging - ISBI*. 2007, pp. 205–208 (cit. on pp. 22, 87).

[5] J. Atif, O. Nempont, O. Colliot, E. Angelini, and I. Bloch. "Level set Deformable Models Constrained by Fuzzy Spatial Relations". *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems - IPMU*. 2006, pp. 236–241 (cit. on pp. 11, 22, 86).

[6] I. Banerjee, G. Patané, and M. Spagnuolo. "Combination of Visual and Symbolic Knowledge: A Survey in Anatomy". *Computers in Biology and Medicine* 80 (2017), pp. 148–157 (cit. on p. 21).

[7] C. Berger, T. Géraud, R. Levillain, N. Widynski, A. Baillard, and E. Bertin. "Effective Component Tree Computation with Application to Pattern Recognition in Astronomical Imaging". *IEEE International Conference on Image Processing - ICIP*. Vol. 4. 2007, pp. 41–44 (cit. on p. 109).

[8] J. Besag. "Statistical Analysis of Dirty Pictures". *Journal of the Royal Statistical Society* 48.3 (1986), pp. 259–302 (cit. on p. 40).

[9] S. Beucher and F. Meyer. "The Morphological Approach to Segmentation: the Watershed Transformation". *Mathematical Morphology in Image Processing* (1993), pp. 433–481 (cit. on p. 13).

[10] I. Biederman. "Recognition-by-components: a Theory of Human Image Understanding." *Psychological Review* 94.2 (1987), pp. 115–147 (cit. on p. 22).

[11] A. M. Bjorndal, W. G. Henderson, A. E. Skidmore, and F. H. Kellner. "Anatomic Measurements of Human Teeth Extracted from Males between the ages of 17 and 21 years". *Oral Surgery, Oral Medicine, Oral Pathology* 38.5 (1974), pp. 791–803 (cit. on p. 33).

[12] R. Blanc, M. Reyes, C. Seiler, and G. Székely. "Conditional Variability of Statistical Shape Models based on Surrogate Variables". *Medical Image Computing and Computer-Assisted Intervention - MICCAI*. Vol. LNCS 5762. 2009, pp. 84–91 (cit. on pp. 65, 74, 105).

[13] R. Blanc, C. Seiler, G. Székely, L. P. Nolte, and M. Reyes. "Statistical Model based Shape Prediction from a Combination of Direct Observations and Various Surrogates: Application to Orthopaedic Research". *Medical Image Analysis* 16.6 (2012), pp. 1156–1166 (cit. on p. 105).

[14] I. Bloch. “Fuzzy Relative Position between Objects in Image Processing: a Morphological Approach”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21.7 (1999), pp. 657–664 (cit. on pp. 22, 84).

[15] I. Bloch. “Fuzzy Spatial Relationships for Image Processing and Interpretation: a Review”. *Image and Vision Computing* 23.2 (2005), pp. 89–110 (cit. on pp. 22, 75, 78).

[16] I. Bloch. “Fuzzy Spatial Relationships from Mathematical Morphology for Model-Based Pattern Recognition and Spatial Reasoning”. *International Conference on Discrete Geometry for Computer Imagery - DGCI*. Vol. LNCS 2886. 2003, pp. 16–33 (cit. on p. 22).

[17] I. Bloch, O. Colliot, O. Camara, and T. Géraud. “Fusion of Spatial Relationships for Guiding Recognition, Example of Brain Structure Recognition in 3D MRI”. *Pattern Recognition Letters* 26.4 (2005), pp. 449–457 (cit. on pp. 22, 104).

[18] I. Bloch, O. Colliot, and R. M. Cesar. “On the Ternary Spatial Relation "Between".” *IEEE Transactions on Systems, Man, and Cybernetics.* 36.2 (2006), pp. 312–27 (cit. on pp. 22, 76).

[19] I. Bloch and A. Ralescu. “Directional Relative Position between Objects in Image Processing: a Comparison between Fuzzy Approaches”. *Pattern Recognition* 36.7 (2003), pp. 1563–1582 (cit. on pp. 22, 75, 83).

[20] M. Bojarski, A. Choromanska, K. Choromanski, B. Firner, L. Jackel, U. Muller, and K. Zieba. “VisualBackProp: Efficient Visualization of CNNs”. *ArXiv Online* 1611.05418 (2016) (cit. on p. 23).

[21] Y. Boykov and G. Funka-Lea. “Graph Cuts and Efficient N-D Image Segmentation”. *International Journal of Computer Vision* 70.2 (2006), pp. 109–131 (cit. on p. 41).

[22] Y. Boykov and M.-P. Jolly. “Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images”. *IEEE International Conference on Computer Vision - ICCV*. 2001, pp. 105–112 (cit. on pp. 10, 37).

[23] Y. Boykov and V. Kolmogorov. “An Experimental Comparison of Min-cut/Max-flow Algorithms for Energy Minimization in Vision”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26.9 (2004), pp. 1124–1137 (cit. on pp. 10, 40, 89).

[24] Y. Boykov, O. Veksler, and R. Zabih. “Fast Approximate Energy Minimization via Graph Cuts”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23.11 (2001), pp. 1222–1239 (cit. on pp. 10, 67–69).

[25] L. Brahim, K. Okba, and L. Robert. “Mathematical Framework for Yopological Relationships between Ribbons and Regions”. *Journal of Visual Languages and Computing* 26 (2015), pp. 66–81 (cit. on p. 22).

[26] L. Breiman. “Random forests”. *Machine Learning* 45.1 (2001), pp. 5–32 (cit. on p. 18).

[27] B. Caldairou. “Arbre des Composantes Connexes : Méthodologie et Application à la Segmentation d’Images Médicales”. Thèse de Master. Ecole Supérieure de Chimie, Physique et Electronique de Lyon, 2008 (cit. on p. 13).

[28] E. Carlinet and T. Géraud. “A Comparison of Many Max-Tree Computation Algorithms”. *International Symposium on Mathematical Morphology - ISMM*. Vol. LNCS 7883. 2013, pp. 73–85 (cit. on p. 109).

[29] E. Carlinet and T. Géraud. “MToS: A Tree of Shapes for Multivariate Images”. *IEEE Transactions on Image Processing* 24.12 (2015), pp. 5330–5342 (cit. on p. 13).

[30] V. Caselles, R. Kimmel, and G. Sapiro. “Geodesic Active Contours”. *International Journal of Computer Vision* 22.1 (1997), pp. 61–79 (cit. on pp. 9–10, 17).

[31] T. F. Chan and L. A. Vese. “Active Contours Without Edges”. *IEEE Transactions on Image Processing* 10.2 (2001), pp. 266–277 (cit. on pp. 9–10).

[32] X. Chen, J. K. Udupa, A. Alavi, and D. A. Torigian. “GC-ASM: Synergistic Integration of Graph-Cut and Active Shape Model Strategies for Medical Image Segmentation”. *Computer Vision and Image Understanding* 117.5 (2013), pp. 513–524 (cit. on pp. 17, 37).

[33] P. Christiano, J. Leike, T. B. Brown, M. Martic, S. Legg, and D. Amodei. “Deep Reinforcement Learning from Human Preferences”. *ArXiv Online* 1706.03741 (2017) (cit. on p. 23).

[34] E. Clementini and P. Di Felice. “Qualitative Representation of Positional Information”. *Artificial Intelligence* 95 (1997), pp. 317–356 (cit. on p. 22).

[35] O. Colliot. “Représentation, Évaluation et Utilisation de Relations Spatiales pour l’Interprétation d’Images. Application à la Reconnaissance de Structures Anatomiques en Imagerie Médicale”. Thèse de Doctorat. Télécom ParisTech, 2003 (cit. on p. 22).

[36] O. Colliot, O. Camara, and I. Bloch. “Integration of Fuzzy Spatial Relations in Deformable Models—Application to Brain MRI Segmentation”. *Pattern Recognition* 39.8 (2006), pp. 1401–1414 (cit. on pp. 11, 22).

[37] R. W. Conners, M. M. Trivedi, and C. A. Harlow. “Segmentation of a High-Resolution Urban Scene Using Texture Operators”. *Computer Vision, Graphics, and Image Processing* 25.3 (1984), pp. 273–310 (cit. on p. 132).

[38] T. F. Cootes, G. J. Edwards, and C. J. Taylor. “Active Appearance Models”. *European Conference on Computer Vision - ECCV*. Vol. LNCS 1407. 1998, pp. 484–498 (cit. on p. 28).

[39] T. F. Cootes, C. J. Taylor, D. Cooper, and J. Graham. “Active Shape Models-Their Training and Application”. *Computer Vision and Image Understanding* 61 (1995), pp. 38–59 (cit. on pp. 17, 28).

[40] C. Couprie, L. Grady, L. Najman, and H. Talbot. “Power Watershed: A Unifying Graph-based Optimization Framework”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.7 (2011), pp. 1384–1399 (cit. on pp. 12, 14, 54–55, 65, 87).

[41] A. Criminisi and J. Shotton. *Decision Forests for Computer Vision and Medical Image Analysis*. Springer London, 2013 (cit. on p. 18).

[42] A. Criminisi, J. Shotton, D. Robertson, and E. Konukoglu. “Regression Forests for Efficient Anatomy Detection and Localization in CT Studies”. *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging.* 2011, pp. 106–117 (cit. on p. 18).

[43] W. De Vos, J. Casselman, and G. R. J. Swennen. “Cone-Beam Computerized Tomography (CBCT) Imaging of the Oral and Maxillofacial Region: A Systematic Review of the Literature”. *International Journal of Oral and Maxillofacial Surgery* 38 (2009), pp. 609–625 (cit. on p. 6).

[44] T. Deschamps and L. D. Cohen. “Fast Extraction of Tubular and Tree 3D Surfaces with Front Propagation Methods”. *International Conference on Pattern Recognition - ICPR*. Vol. 1. 2002, pp. 731–734 (cit. on p. 60).

[45] A. Desolneux, L. Moisan, and J.-M. Morel. “From Gestalt Theory to Image Analysis”. *Interdisciplinary Applied Mathematics*. Vol. 34. 2008, p. 285 (cit. on p. 22).

[46] Q. Dou, L. Yu, H. Chen, Y. Jin, X. Yang, J. Qin, and P. A. Heng. “3D Deeply Supervised Network for Automated Segmentation of Volumetric Medical Images”. *Medical Image Analysis* 41 (2017), pp. 40–54 (cit. on p. 20).

[47] G. J. Edwards, C. J. Taylor, and T. F. Cootes. “Learning to Identify and Track Faces in Image Sequences”. *IEEE International Conference on Automatic Face and Gesture Recognition - ICAFGR*. 1998, pp. 260–265 (cit. on p. 28).

[48] T. Evain, X. Ripoche, J. Atif, and I. Bloch. “Fuzzy “Along” Spatial Relation in 3D. Application to Anatomical Structures in Maxillofacial CBCT”. *International Conference on Image Analysis and Processing - ICIAP*. Vol. LNCS 9279. 2015, pp. 271–281 (cit. on pp. 75, 104).

[49] T. Evain, X. Ripoche, J. Atif, and I. Bloch. “Semi-automatic Teeth Segmentation in Cone-Beam Computed Tomography by Graph-Cut with Statistical Shape Priors”. *IEEE International Symposium on Biomedical Imaging - ISBI*. Vol. 2. 2017, pp. 1197–1200 (cit. on pp. 55, 64, 103).

[50] J. Faure, A. Oueiss, J. Treil, S. Chen, V. Wong, and J.-m. Inglese. “Céphalométrie 3D et Intelligence Artificielle”. *Revue d’Orthopédie Dento-Faciale* 50.3 (2016), pp. 315–334 (cit. on p. 98).

[51] L. R. Ford and D. R. Fulkerson. “Maximal Flow Through a Network”. *Canadian Journal of Mathematics* 8 (1956), pp. 399–404 (cit. on pp. 10, 40).

[52] G. Fouquier. “Optimisation de Séquences de Segmentation combinant Modèle Structurel et Focalisation de l’Attention Visuelle. Application à la Reconnaissance de Structures”. Thèse de Doctorat. Télécom ParisTech, 2010 (cit. on pp. 22, 87, 104, 106).

[53] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. “Multiscale Vessel Enhancement Filtering”. *International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI*. Vol. LNCS 1496. 1998, pp. 130–137 (cit. on p. 60).

[54] D. Freedman and T. Zhang. “Interactive Graph Cut based Segmentation with Shape Priors”. *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*. Vol. 1. 2005, pp. 755–762 (cit. on pp. 17–18, 38).

[55] J. Freeman. “The Modelling of Spatial Relations”. *Computer Graphics and Image Processing* 4 (1975), pp. 156–171 (cit. on p. 22).

[56] M. Gahegan. “Proximity Operators for Qualitative Spatial Reasoning”. *Spatial Information Theory : a theoretical basis for GIS*. Vol. LNCS 988. 1995, pp. 31–44 (cit. on p. 22).

[57] Y. Gan, Z. Xia, J. Xiong, Q. Zhao, Y. Hu, and J. Zhang. “Toward Accurate Tooth Segmentation from Computed Tomography Images using a Hybrid Level set Model.” *Medical physics* 42.1 (2015), p. 14 (cit. on p. 11).

[58] H. Gao and O. Chae. “Individual Tooth Segmentation from CT Images using Level set Method with Shape and Intensity Prior”. *Pattern Recognition* 43.7 (2010), pp. 2406–2417 (cit. on pp. 10–11, 62).

[59] M. Gargouri, J. Tierny, E. Jolivet, P. Petit, and E. D. Angelini. “Accurate and Robust Shape Descriptors for the Identification of Rib Cage Structures in CT-Images with Random Forests”. *IEEE International Symposium on Biomedical Imaging - ISBI*. 2013, pp. 65–68 (cit. on pp. 18, 54–55).

[60] R. Gauriau. “Méthodes Multi-Organes rapides avec A Priori de Forme pour la Localisation et la Segmentation en Imagerie Médicale 3D”. Thèse de Doctorat. Télécom ParisTech, 2015 (cit. on p. 18).

[61] S. Geman and D. Geman. “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6.6 (1984), pp. 721–741 (cit. on p. 40).

[62] T. Géraud, E. Carlinet, S. Crozet, and L. Najman. “A Quasi-linear Algorithm to Compute the Tree of Shapes of nD Images”. *International Symposium on Mathematical Morphology - ISMM*. Vol. LNCS 7883. 2013, pp. 98–110 (cit. on pp. 13, 109).

[63] P. Geurts, D. Ernst, and L. Wehenkel. “Extremely Randomized Trees”. *Machine Learning* 63.1 (2006), pp. 3–42 (cit. on p. 18).

[64] B. Glocker, O. Pauly, E. Konukoglu, and A. Criminisi. “Joint Classification-Regression Forests for Spatially Structured Multi-object Segmentation”. *European Conference on Computer Vision - ECCV*. Vol. LNCS 7575. 4. 2012, pp. 870–881 (cit. on p. 18).

[65] E. S. Gopi and P. Palanisamy. “Fast Computation of PCA Bases of Image Subspace using its Inner-product Subspace”. *Applied Mathematics and Computation* 219.12 (2013), pp. 6729–6732 (cit. on p. 29).

[66] L. Grady. “Random Walks for Image Segmentation”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28.11 (2006), pp. 1768–1783 (cit. on p. 10).

[67] N. M. Grande, G. Plotino, R. Pecci, R. Bedini, C. H. Pameijer, and F. Somma. “Microcomputerized Tomographic Analysis of Radicular and Canal Morphology of Premolars with Long Oval Canals”. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology* 106.3 (2008) (cit. on p. 60).

[68] D. M. Greig, B. T. Porteous, and A. H. Seheult. “Exact Maximum A Posteriori Estimation for Binary Images”. *Journal of the Royal Statistical Society* 51.2 (1989), pp. 271–279 (cit. on pp. 10, 40).

[69] D. Grosgeorge, C. Petitjean, J. N. Dacher, and S. Ruan. “Graph Cut Segmentation with a Statistical Shape Model in Cardiac MRI”. *Computer Vision and Image Understanding* 117.9 (2013), pp. 1027–1035 (cit. on pp. 18, 35).

[70] R. M. Haralick. “Statistical and Structural Approaches To Texture”. *Proceedings of the IEEE* 67.5 (1979), pp. 786–804 (cit. on p. 132).

[71] R. M. Haralick, K. Shanmugan, and I. Dinstein. “Textural Features for Image Classification”. *IEEE Transactions on Systems, Man, and Cybernetics.* 3.6 (1973), pp. 610–621 (cit. on p. 132).

[72] W. G. Hayward and M. J. Tarr. “Spatial Language and Spatial Representation.” *Cognition* 55 (1995), pp. 39–84 (cit. on p. 22).

[73] L. T. Hiew, S. H. Ong, and K. W. C. Foong. “Tooth Segmentation From Cone-Beam CT Using Graph Cut”. *Asia-Pacific Signal and Information Processing Association - APSIPA*. 2010, pp. 272–275 (cit. on p. 11).

[74] T. K. Ho. “Random Decision Forests”. *International Conference on Document Analysis and Recognition - ICDAR*. 1995, pp. 278–282 (cit. on p. 18).

[75] C. Hudelot, J. Atif, and I. Bloch. “Fuzzy Spatial Relation Ontology for Image Interpretation”. *Fuzzy Sets and Systems* 159.15 (2008), pp. 1929–1951 (cit. on p. 104).

[76] J. E. Iglesias, E. Konukoglu, A. Montillo, Z. Tu, and A. Criminisi. “Combining Generative and Discriminative Models for Semantic Segmentation of CT Scans via Active Learning”. *NeuroImage* 108 (2011), pp. 25–36 (cit. on p. 18).

[77] D. Kainmueller, H. Lamecker, H. Seim, M. Zinser, and S. Zachow. “Automatic Extraction of Mandibular Nerve and Bone from Cone-Beam CT Data”. *International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI*. Vol. LNCS 5762. 2009, pp. 76–83 (cit. on p. 105).

[78] K. Kamnitsas, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker. “Efficient Multi-scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation”. *Medical Image Analysis* 36 (2017), pp. 61–78 (cit. on pp. 19–20).

[79] M. Kass, A. Witkin, and D. Terzopoulos. “Snakes: Active Contour Models”. *International Journal of Computer Vision* 1.4 (1988), pp. 321–331 (cit. on p. 9).

[80] O. Koenig, L. P. Reiss, and S. M. Kosslyn. “The Development of Spatial Relation Representations: Evidence from Studies of Cerebral Lateralization”. *Journal of Experimental Child Psychology* 50 (1990), pp. 119–130 (cit. on p. 22).

[81] V. Kolmogorov and R. Zabih. “What Energy Functions can be Minimized via Graph Cuts?” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26.2 (2004), pp. 147–159 (cit. on pp. 38, 67–68).

[82] A. Krizhevsky, I. Sutskever, and G. E. Hinton. “ImageNet Classification with Deep Convolutional Neural Networks”. *Conference on Neural Information Processing Systems - NIPS*. Vol. 25. 2012, pp. 1097–1105 (cit. on p. 19).

[83] F. Le Ber, J. Lieber, and A. Napoli. “Les Systèmes à Base de Connaissances”. *Encyclopédie de l’informatique et des systèmes d’information*. 2006, pp. 1197–1208 (cit. on p. 21).

[84] J. L. Leopold, C. L. Sabharwal, and K. J. Ward. “Spatial Relations between 3D Objects: The Association between Natural Language, Topology, and Metrics”. *Journal of Visual Languages and Computing* 27.1 (2015), pp. 29–37 (cit. on p. 22).

[85] M. E. Leventon, W. E. L. Grimson, and O. Faugeras. “Statistical Shape Influence in Geodesic Active Contours”. *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*. Vol. 1. 2000, pp. 316–323 (cit. on pp. 17, 54–55).

[86] C. Li, C. Xu, C. Gui, and M. Fox. “Level Set Evolution without Re-Initialization: A New Variational Formulation”. *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*. Vol. 1. 2005, pp. 430–436 (cit. on p. 10).

[87] M. Li, A. Shekhovtsov, and D. Huber. “Complexity of Discrete Energy Minimization Problems”. *European Conference on Computer Vision - ECCV*. Vol. LNCS 9905. 2016, pp. 834–852 (cit. on p. 40).

[88] S. Livens, P. Scheunders, G. Van de Wouwer, and D. Van Dyck. “Wavelets for Texture Analysis”. *International Journal of Computer Science and Information Management* (1997), pp. 581–585 (cit. on p. 135).

[89] J. Long, E. Shelhamer, and T. Darrell. “Fully Convolutional Networks for Semantic Segmentation”. *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*. 2015, pp. 3431–3440 (cit. on p. 19).

[90] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. “PET-CT Image Registration in the Chest using Free-form Deformations”. *IEEE Transactions on Medical Imaging* 22.1 (2003), pp. 120–128 (cit. on p. 29).

[91] Y. Miki, C. Muramatsu, T. Hayashi, X. Zhou, T. Hara, A. Katsumata, and H. Fujita. “Classification of Teeth in Cone-Beam CT using Deep Convolutional Neural Network”. *Computers in Biology and Medicine* 80 (2017), pp. 24–29 (cit. on p. 23).

[92] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. “Human-level Control through Deep Reinforcement Learning”. *Nature* 518.7540 (2015), pp. 529–533 (cit. on p. 23).

[93] R. S. Montero and E. Bribiesca. “State of the Art of Compactness and Circularity Measures”. *International Mathematical Forum* 4.27 (2009), pp. 1305 –1335 (cit. on p. 76).

[94] A. Montillo, J. Shotton, J. Winn, J. E. Iglesias, D. Metaxas, and A. Criminisi. “Entangled Decision Forests and their Application for Semantic Segmentation of CT Images”. *International Conference on Information Processing in Medical Imaging - IPMI*. Vol. LNCS 6801. 2011, pp. 184–196 (cit. on p. 18).

[95] J. N. Morgan and J. A. Sonquist. “Problems in the Analysis of Survey Data, and a Proposal”. *Journal of the American Statistical Association* 58.302 (1963), p. 415 (cit. on p. 18).

[96] B. Naegel, N. Passat, N. Boch, and M. Kocher. “Segmentation using Vector-Attribute Filters : Methodology and Application to Dermatological Imaging”. *International Symposium on Mathematical Morphology - ISMM*. 2007, pp. 239–250 (cit. on p. 13).

[97] L. Najman and M. Couprie. “Building the Component Tree in Quasi-linear Time”. *IEEE Transactions on Image Processing* 15.11 (2006), pp. 3531–3539 (cit. on pp. 13, 109).

[98] O. Nempont. “Modèles Structurels Flous et Propagation de Contraintes pour la Segmentation et la Reconnaissance d’Objets dans les Images”. Thèse de Doctorat. Télécom ParisTech, 2009 (cit. on pp. 22, 87, 104, 106).

[99] H. Noh, S. Hong, and B. Han. “Learning Deconvolution Network for Semantic Segmentation”. *IEEE International Conference on Computer Vision - ICCV*. 2015, pp. 1520–1528 (cit. on p. 19).

[100] S. Osher and J. A. Sethian. “Fronts Propagating with Curvature-dependent Speed: Algorithms based on Hamilton-Jacobi Formulations”. *Journal of Computational Physics* 79.1 (1988), pp. 12–49 (cit. on p. 9).

[101] N. Otsu. “A Threshold Selection Method from Gray-level Histograms”. *Automatica* 20.1 (1975), pp. 62–66 (cit. on pp. 35, 155).

[102] Y. Pei, X. Ai, H. Zha, T. Xu, and G. Ma. “3D Exemplar-based Random Walks for Tooth Segmentation from Cone-Beam Computed Tomography Images”. *Medical Physics* 43.9 (2016), pp. 5040–5050 (cit. on pp. 15–16).

[103] O. A. B. Penatti, K. Nogueira, and J. A. Dos Santos. “Do Deep Features generalize from Everyday Objects to Remote Sensing and Aerial Scenes Domains?” *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*. 2015, pp. 44–51 (cit. on p. 20).

[104] P. Proff. “Malocclusion, Mastication and the Gastrointestinal System”. *Journal of Orofacial Orthopedics / Fortschritte der Kieferorthopädie* 71.2 (2010), pp. 96–107 (cit. on p. 2).

[105] A. Rosenfeld. “The Fuzzy Geometry of Image Subsets”. *Pattern Recognition Letters* 2 (1984), pp. 311–317 (cit. on pp. 22, 78).

[106] F. Saint-Pierre, G. Fanelli, L. Mosnegutu, and F. Devaux. *Tomographie Volumique à Faisceau Conique de la Face (Cone Beam)*. Rapport technique. Haute Autorité de Santé, 2009, pp. 1–10 (cit. on p. 5).

[107] P. Salembier, A. Oliveras, and L. Garrido. “Antiextensive Connected Operators for Image and Sequence Processing”. *IEEE Transactions on Image Processing* 7.4 (1998), pp. 555–570 (cit. on pp. 13, 109).

[108] W. C. Scarfe and A. G. Farman. “What is Cone-Beam CT and How Does it Work?” *Dental Clinics of North America* 52.4 (2008), pp. 707–730 (cit. on p. 5).

[109] N. Sebe and M. S. Lew. “Wavelet Based Texture Classification”. *International Conference on Pattern Recognition - ICPR*. Vol. 3. 2000, pp. 959–962 (cit. on p. 135).

[110] M. Sepehrian, A. M. Deylami, and R. A. Zoroofi. “Individual Teeth Segmentation in CBCT and MSCT Dental Images using Watershed”. *Iranian Conference on Biomedical Engineering - ICBME*. 2013, pp. 27–30 (cit. on pp. 14–15).

[111] J. A. Sethian. “Fast Marching Methods and Level Set Methods for Propagating Interfaces”. *Computational Fluid Mechanics from von Karman Institute Lecture Series*. von Karman Institute, 1998 (cit. on p. 9).

[112] A. Shariff, M. J. Egenhofer, and D. M. Mark. “Natural-Language Spatial Relations Between Linear and Areal Objects : The Topology and Metric of English”. *International Journal of Geographical Information Science* 12.3 (1998), pp. 215–246 (cit. on p. 22).

[113] K. Simonyan and A. Zisserman. “Very Deep Convolutional Networks for Large-Scale Image Recognition”. *International Conference on Learning Representations - ICRL*. 2015, pp. 1–14 (cit. on pp. 19–20).

[114] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. “Content-based Image Retrieval at the End of the Early Years”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.12 (2000), pp. 1349–1380 (cit. on p. 22).

[115] W. Sweldens. “The Lifting Scheme: A Construction of Second Generation Wavelets”. *SIAM Journal on Mathematical Analysis* 29.2 (1998), pp. 511–546 (cit. on p. 135).

[116] A. A. Taha and A. Hanbury. “Metrics for Evaluating 3D Medical Image Segmentation: Analysis, Selection, and Tool.” *BMC Medical Imaging* 15 (2015), p. 29 (cit. on p. 45).

[117] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang. “Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?” *IEEE Transactions on Medical Imaging* 35.5 (2016), pp. 1299–1312 (cit. on p. 20).

[118] C. M. Takemura, R. M. Cesar, and I. Bloch. “Modeling and Measuring the Spatial Relation “Along”: Regions, Contours and Fuzzy Sets”. *Pattern Recognition* 45.2 (2012), pp. 757–766 (cit. on pp. 22, 75–79).

[119] H. Urien, I. Buvat, N. Rougon, M. Soussan, and I. Bloch. “Brain Lesion Detection in 3D PET Images Using Max-Trees and a New Spatial Context Criterion”. *International Symposium on Mathematical Morphology - ISMM*. Vol. LNCS 10225. 2017, pp. 455–466 (cit. on p. 13).

[120] M. C. Vanegas. “Relations Spatiales et Raisonnement Spatial pour l’Interprétation des Images d’Observation de la Terre utilisant un Modèle Structurel.” Thèse de Doctorat. Télécom ParisTech, 2011 (cit. on pp. 22, 84–85, 87).

[121] M. C. Vanegas, I. Bloch, and J. Inglada. "A Fuzzy Definition of the Spatial Relation "Surround"-Application to Complex Shapes". *Conference of the European Society for Fuzzy Logic and Technology - EUSFLAT*. 2011, pp. 844–851 (cit. on pp. 22, 75).

[122] M. C. Vanegas, I. Bloch, and J. Inglada. "Alignment and Parallelism for the Description of High-Resolution Remote Sensing Images". *IEEE Transactions on Geoscience and Remote Sensing* 51.6 (2013), pp. 3542–3557 (cit. on pp. 22, 83).

[123] L. Wang, Y. Gao, F. Shi, G. Li, K.-c. Chen, Z. Tang, J. J. Xia, and D. Shen. "Automated Segmentation of CBCT Image with Prior-Guided Sequential Random Forest". *Medical Computer Vision: Algorithms for Big Data.* Vol. LNCS 9601. 2016, pp. 72–82 (cit. on p. 23).

[124] C. D. Wickens, J. Lee, Y. Liu, and S. E. Gordon Becker. *An Introduction to Human Factors Engineering*. 2nd Ed. Pearson, 2004, p. 608 (cit. on p. 107).

[125] Z. Xia, Y. Gan, L. Chang, J. Xiong, and Q. Zhao. "Individual Tooth Segmentation from CT Images scanned with Contacts of Maxillary and Mandible Teeth". *Computer Methods and Programs in Biomedicine* 138 (2017), pp. 1–12 (cit. on p. 11).

[126] Y. Xu. "Espaces des Formes basés sur des Arbres : Définition et Applications en Traitement d'Images et Vision par Ordinateur". PhD thesis. Université Paris Est, 2013 (cit. on p. 13).

[127] Y. Xu, T. Géraud, and I. Bloch. "From Neonatal to Adult Brain MR Image Segmentation in a Few Seconds Using 3D-like Fully Convolutional Network and Transfer Learning". *IEEE International Conference on Image Processing - ICIP*. 2017 (cit. on pp. 20, 56).

[128] L. A. Zadeh. "Fuzzy Logic = Computing with words". *IEEE Transactions on Fuzzy Systems* 4.2 (1996), pp. 103–111 (cit. on p. 22).

[129] L. A. Zadeh. "Fuzzy Sets". *Information and Control* 8.3 (1965), pp. 338–353 (cit. on p. 22).

[130] Z. Zou, S.-d. Luo, J.-s. Pan, S.-j. Liu, and S.-h. Liao. "A Semi-automatic Segmentation for Tooth on Cone Beam CT Volume Following the Anatomic Guidance". *Journal of Information Hiding and Multimedia Signal Processing* 8.4 (2017), pp. 774–783 (cit. on pp. 12–13).

[131] R. Zwick, E. Carlstein, and D. V. Budescu. "Measures of Similarity Among Fuzzy Concepts: A Comparative Analysis". *International Journal of Approximate Reasoning* 1 (1987), pp. 221–242 (cit. on p. 84).

Nouvelles méthodes de segmentation en imagerie tomographique volumique à faisceau conique dentaire

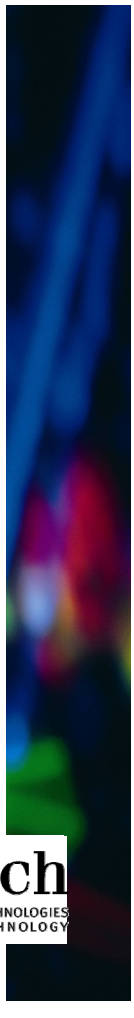
Timothée EVAIN

RESUMÉ : La tomographie à faisceau conique (CBCT) est devenue la modalité de référence pour les praticiens du domaine dentaire. Sa relative nouveauté et ses spécificités impliquent que le domaine du traitement de ces images est peu développé à l'heure actuelle. En partenariat industriel avec Carestream Dental, le premier volet de la thèse a conduit à développer une méthode de segmentation semi-automatique de chaque dent, reposant sur l'utilisation de contraintes de forme et d'intensité, et formulée comme un problème de minimisation d'énergie résolu par coupure de graphe. Les résultats montrent une bonne qualité de segmentation avec un coefficient de Dice moyen de $0,958$. Une application logicielle a été réalisée pour la mise en œuvre de la méthode de segmentation auprès des praticiens, en tenant compte des contraintes techniques et temporelles imposées par le contexte clinique, ainsi que du profil des utilisateurs. Un travail préliminaire d'extension de l'approche par coupure de graphe pour segmenter simultanément plusieurs dents à été réalisé, montrant la nécessité de rendre les contraintes de formes plus adaptées aux images et de modifier la méthode d'optimisation pour atteindre des temps de calcul compatibles avec la pratique clinique. Un second volet prospectif des travaux concerne la constitution d'un modèle structurel de la région maxillo-faciale pour formaliser les connaissances a priori sur les organes et leurs interactions. Ce modèle est un graphe conceptuel où sont représentés les concepts des structures et des relations. En particulier, les relations d'alignement et "le long de" ont été modélisées en 3D dans le cadre des ensembles flous.

MOTS-CLEFS : Tomographie à faisceau conique (CBCT) ; Imagerie dentaire ; Segmentation sous contraintes de forme ; Modèle structurel ; Relations spatiales

ABSTRACT : Cone-Beam computed tomography (CBCT) is the new standard imaging method for dental practitioners. The image processing field of CBCT data is still underdeveloped due to the novelty of the method and its specificities compared to traditional CT. With Carestream Dental as industrial partner, the first part of this work is a new semi-automatic segmentation protocol for teeth, based on shape and intensity constraints, through a graph-cut optimization of an energy formulation. Results show a good quality of segmentation with an average Dice coefficient of 0.958. A fully functional implementation of the algorithm has led to a software available for dentists, taking into account the clinical context leading to temporal and technical difficulties. A preliminary extension to multi-objects segmentation showed the necessity to get more stringent shape constraints as well as a better optimization algorithm to get acceptable computation times. The second part of this thesis, more prospective, is about the creation of a structural model of the maxillo-facial space, to formalize the a priori knowledge on organs and theirs spatial relations. This model is a conceptual graph where structures and relationships are seen as concepts. In particular, the spatial relations "Along" and "Aligned", modeled in a fuzzy set framework, have been extended to 3D objects.

KEY-WORDS : Cone-Beam computed tomography (CBCT) ; Dental imaging ; Shape constrained segmentation ; Structural model ; Spatial relations



TELECOM ParisTech

EDiTE

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY