N°d'ordre NNT : 2023ISAL0008

THESE de DOCTORAT DE L'INSA LYON,
membre de l'Université de Lyon

Ecole Doctorale N° 160
Électronique, Électrotechnique, Automatique

Spécialité/ discipline de doctorat :
Traitement du signal et de l'image

Soutenue publiquement le 07/02/2023, par :
Louise Friot--Giroux

Méthodes de reconstruction avancées en tomographie dentaire par faisceau conique

---

Devant le jury composé de :

Bousse, Alexandre	Chargé de recherche, LaTIM, INSERM	Rapporteur
Sidky, Emil	Professor, University of Chicago (USA)	Rapporteur
Chouzenoux, Émilie	Chargée de recherche, INRIA Saclay	Examinatrice
Desbat, Laurent	Professeur, Université de Grenoble	Examinateur, Président du jury
Maxim, Voichița	Maître de conférence, INSA Lyon	Directrice de thèse
Peyrin, Françoise	Directrice de recherche, INSERM	Directrice de thèse
Maury, Colombe	Ingénieur R&D, Carestream Dental	Invitée

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Département FEDORA – INSA Lyon – Ecoles Doctorales

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	**CHIMIE DE LYON** https://edchimie-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage secretariat@edchimie-lyon.fr	**M. Stéphane DANIELE** C2P2-CPE LYON-UMR 5265 Bâtiment F308, BP 2077 43 boulevard du 11 Novembre 1918 69616 Villeurbanne directeur@edchimie-lyon.fr
E.E.A.	**ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE** https://edeea.universite-lyon.fr Sec. : Stéphanie CAUVIN Bâtiment Direction INSA Lyon Tél : 04.72.43.71.70 secretariat.edeea@insa-lyon.fr	**M. Philippe DELACHARTRE** INSA Lyon Laboratoire CREATIS Bâtiment Blaise Pascal, 7 avenue Jean Capelle 69621 Villeurbanne CEDEX Tél : 04.72.43.88.63 philippe.delachartre@insa-lyon.fr
E2M2	**EVOLUTION, ECOSYSTEME, MICROBIOLOGIE, MODELISATION** http://e2m2.universite-lyon.fr Sec. : Bénédicte LANZA Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.88.62 secretariat.e2m2@univ-lyon1.fr	**Mme Sandrine CHARLES** Université Claude Bernard Lyon 1 UFR Biosciences Bâtiment Mendel 43 boulevard du 11 Novembre 1918 69622 Villeurbanne CEDEX sandrine.charles@univ-lyon1.fr
EDISS	**INTERDISCIPLINAIRE SCIENCES-SANTE** http://ediss.universite-lyon.fr Sec. : Bénédicte LANZA Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.88.62 secretariat.ediss@univ-lyon1.fr	**Mme Sylvie RICARD-BLUM** ICBMS – UMR 5246 CNRS – Université de Lyon Bâtiment Raulin – 2ème étage Nord 43 boulevard du 11 Novembre 1918 69622 Villeurbanne Cedex Tél : 04.72.44.82.32 sylvie.ricard-blum@univ-lyon1.fr
INFOMATHS	**INFORMATIQUE ET MATHEMATIQUES** http://edinfomaths.universite-lyon.fr Sec. : Renée EL MALHEM Bât. Blaise Pascal, 3e étage Tél : 04.72.43.80.46 infomaths@univ-lyon1.fr	**M. Hamamache KHEDDOUCI** Université Claude Bernard Lyon 1 Bâtiment Nautibus 43 boulevard du 11 Novembre 1918 69622 Villeurbanne Cedex Tél : 04.72.44.83.69 hamamache.kheddouci@univ-lyon1.fr
Matériaux	**MATERIAUX DE LYON** http://ed34.universite-lyon.fr Sec. : Yann DE ORDENANA Tél : 04.72.18.62.44 yann.de-ordenana@ec-lyon.fr	**M. Stéphane BENAYOUN** Ecole Centrale de Lyon, Laboratoire LTDS 36 avenue Guy de Collongue 69134 Ecully Cedex Tél : 04.72.18.64.37 stephane.benayoun@ec-lyon.fr
MEGA	**MECANIQUE, ENERGETIQUE, GENIE CIVIL, ACOUSTIQUE** http://edmega.universite-lyon.fr Sec. : Stéphanie CAUVIN Tél : 04.72.43.71.70 Bâtiment Direction INSA Lyon mega@insa-lyon.fr	**M. Jocelyn BONJOUR** INSA Lyon Laboratoire CETHIL Bâtiment Sadi-Carnot 9 rue de la Physique 69621 Villeurbanne Cedex jocelyn.bonjour@insa-lyon.fr
ScSo	**ScSo*** https://edsciencessociales.universite-lyon.fr Sec. : Mélina FAVETON INSA : J.Y. TOUSSAINT Tél : 04.78.69.77.79 melina.faveton@univ-lyon2.fr	**M. Christian MONTES** Université Lumière Lyon 2 86 rue Pasteur 69365 Lyon Cedex 07 christian.montes@univ-lyon2.fr

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

« Être scientifique, c'est parfois être un grand enfant. »

Valentin Gautier, le 18 octobre 2022

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Remerciements

Je tiens à exprimer en premier lieu ma gratitude aux membres du jury qui ont accepté de se pencher sur mon travail. Merci à Emilie Chouzenoux et Laurent Desbat pour leur retour en tant qu'examinateur. Je remercie plus particulièrement Alexandre Bousse et Emil Sidky pour leur relecture attentive de ce manuscrit, ainsi que pour leur présence à la soutenance malgré la distance à parcourir et les difficultés de transport actuelles.

Je suis particulièrement reconnaissante auprès de mes deux directrices de thèses, Voichita et Françoise, pour m'avoir fait confiance et m'avoir permis d'effectuer cette thèse. Je remercie chaleureusement Colombe pour sa grande disponibilité tout au long de ces trois années, et ses précieux conseils en reconstruction.

Un grand merci à ceux qui ont contribué à l'écriture de ce manuscrit. Les relecteurs : Pauline, Théo, Valentin, Tony et Odyssée. Et les sauveurs du chapitre 6 : PJ (stagiaire non rémunéré le plus qualifié de CREATIS) dont l'aide précieuse m'a donné le courage d'entraîner des U-Net 3D deux semaines avant de rendre la thèse ; et Matthis, sans qui les performances de ces réseaux seraient encore très moyennes.

Côté entreprise, même si la situation sanitaire ne m'aura pas permis de passer beaucoup de temps en votre compagnie, je garde d'excellents souvenirs de mon partage de bureau avec Gaël, Thibaud et Elyes. Merci également à Yoann, Flore, Claire et Mélanie pour leur accueil chaleureux au sein de l'équipe. Et enfin, je remercie Xavier pour sa disponibilité et son dévouement.

Ce manuscrit signe la fin d'une aventure riche en rencontres et en émotion. J'ai passé trois belles années aux côtés des autres doctorants du laboratoire. Je vous remercie tous pour les pauses interminables autour de pédantix, les nombreuses soirées Kfet, Toï-Toï ou autres. Une mention spéciale à l'équipe K-fet des premiers temps : Ludmilla, Florian, Charles, Maxime, Benoît, et à la relève, plus sage et mieux organisée : Nicolas, Pierre, Gaspard, Juliette, Émile, Romain, Maëlys, Morgane, Robin, Valentin. Et Matthis, mon gars sûr pour aller boire un verre le soir, malheureusement, cette grande qualité te sort de la catégorie "sage", et t'isole dans une phase à part.

Je pense également à Audrey (aka DreyO Les bons plans), Anne-Lise, Antonio, Sébastien, Zexian, Daria, Valentine, Frank, Nathan, Raphaël, Rémi, Julia, Nicolas, et Nolann (que je remercie pour l'ébauche de ces remerciements).

Suzanne, merci pour les précieux conseils en début de thèse et le soutien que tu as pu m'apporter.

J'exprime ma profonde gratitude à Pierre. Ton originalité et ton renouvellement vestimentaire quotidien ont été grandement appréciés. J'espère un jour être capable d'égaler tes clins d'oeil si élégants.

Une pensée nostalgique pour Gaspard, que je remercie pour les nombreuses (sûrement trop) heures passées à jouer à Geoguessr, le développement d'un véritable dépassement de soi-même pour des résultats rapides et efficaces. Tu resteras aussi le croissantage dont je suis le plus fière !

Merci à Mom qui a réponse à n'importe quelle question, et qui m'a accompagnée et supportée pendant nos vacances à l'autre bout du monde.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Je suis de plus reconnaissante auprès d'Hervé pour son SAV point thèse de qualité, et remercie plus généralement l'ensemble du laboratoire pour l'aide et le soutien que j'ai pu recevoir, en particulier de la part de Carole, Pierre, Marion, Anthony, Marie et Manu.

Un immense merci à mes co-bureau Cyril, Mehdi, Samaneh, Alexandre et Sophie ! Merci à Mehdi et Samaneh de nous avoir bien fait rire avec leur remise en question de la langue française et les nouveaux mots inventés qui manquaient à notre vocabulaire. Merci à Sophie, chef pâtissier du labo, qui nous aura régalé avec ses gâteaux et ses soirées autour d'un bon repas. Je te remercie pour toutes les fois où tu as mis beaucoup d'énergie à me remonter le moral, mais aussi pour les fois où tu as su me laisser exercer ma passion seule : tirer la gueule dans un coin. Enfin, Cyril le mâle alpha du bureau. Le rayon de soleil dont tout le monde a besoin, jamais sans restes d'idées ou de blagues improbables pour casser le quotidien. Le quart d'heure de 18h15 me manquera énormément.

J'esquisse également un sourire au souvenir du groupe 10 de l'an dernier, qui m'a donné un but dans la vie pendant quelques mois : leur faire diagonaliser des matrices.

Quelques mots pour mes deux frères de thèse, Théo et Valentin.
Un immense merci à Théo, sans qui le dixième de cette thèse n'aurait pas vu le jour. Je suis très heureuse d'avoir pu partager ces têtes à têtes en exclusivité avec toi en début de thèse, à t'observer engloutir des assiettes qui m'auraient nourrie une semaine. Je te suis infiniment reconnaissante du temps que tu as passé à m'écouter râler sur à peu près tout, à me réconforter et à sécher mes larmes. Tu as bien mérité ton titre de meilleur directeur de thèse.
Valentin, pour reprendre la métaphore de la famille, tu es le petit frère qu'on prend un véritable plaisir à embêter, mais qu'on aime beaucoup (mes oreilles se désolidarisent de cette fin de phrase). Une pensée spéciale à l'introduction à Geoguessr et aux nombreuses parties de babyfoot ! Bon courage pour la fin de thèse, tu gères.

Un remerciement particulier à ma colocataire Ludmilla, grâce à toi, je garde finalement de bons souvenirs du confinement. Merci de nous avoir fait découvrir la Guadeloupe, ses plages et le marché de St Anne ! Ton enthousiasme et ton optimisme ont été une grande source de réconfort et d'inspiration pendant cette thèse (même s'il faut reconnaître qu'il me reste du travail de ce côté là. . .)

Enfin, merci Tony pour tous les massages du cou, qui m'ont permis de rédiger plus détendue et avec la capacité de tourner la tête. Merci d'avoir pris soin de moi pendant cette période compliquée, d'avoir tout fait pour rendre ma vie plus simple et agréable. Hâte de vivre ma best life à tes côtés.

Je m'excuse auprès de tous ceux qui attendaient la version audio-descriptive de cette thèse, ce projet utopique n'a pas eu le temps de voir le jour.
Je propose à la place une playlist pouvant accompagner chaque doctorant dans son parcours vers le titre de Docteur. Elle est composée de sorte à représenter au mieux les différentes étapes rencontrées, des moments de joie aux moments de désespoirs intenses, avec des musiques spéciales rédaction et d'autres plus adaptés au débogage d'un code qui ne veut pas fonctionner (deezer et spotify).





Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Table des matières

Introduction **1**

1 Principe de la tomographie **3**
I Aspects physiques de la radiographie . . . 4
1 Tube de Coolidge . . . 5
2 Création des rayons X . . . 5
3 Interactions rayons X - matière . . . 7
4 Détection et formation de l'image . . . 9
II Géométries d'acquisition . . . 10
III Applications en imagerie médicale . . . 12
1 Système CBCT en imagerie dentaire . . . 12
2 En imagerie dentaire . . . 13

2 Méthodes de reconstruction tomographique **17**
I Formalisme mathématique . . . 18
1 Transformées usuelles en tomographie . . . 19
2 Formulation du problème inverse . . . 21
II Reconstructions analytiques . . . 22
1 Rétro-projection en géométrie 2D parallèle . . . 23
2 En géométrie divergente 2D . . . 24
3 En CBCT . . . 26
III Reconstructions itératives . . . 27
1 Méthodes algébriques . . . 27
2 Méthodes statistiques . . . 30
3 Accélération de Nesterov . . . 34
IV Artefacts de reconstruction . . . 34

3 Reconstructions itératives régularisées **45**
I Problème inverse et optimisation . . . 46
1 Formulation variationnelle . . . 46
2 Optimisation régulière . . . 46
3 Optimisation convexe . . . 48
4 Formulation bayésienne . . . 52
II Méthodes régularisées par variation totale . . . 53
1 Variation Totale . . . 53
2 SIRT régularisé . . . 55
3 MLEM régularisé . . . 56
4 Algorithme de Chambolle-Pock en tomographie . . . 57
III Application au fantôme de Shepp-Logan 2D . . . 58
1 Présentation des données . . . 58
2 Comparaison des résultats . . . 59

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

3 Application à des données expérimentales 60

4 Reconstructions itératives régularisées en CBCT faible-dose 67
I Reconstruction tomographique en imagerie dentaire 68
1 Reconstruction CBCT 68
2 Dose et bruit 69
3 Projections tronquées 69
II Algorithmes et métriques utilisés 70
1 Paramètres de reconstruction 70
2 Métriques d'évaluation 71
III Description des données 73
1 Géométrie d'acquisition 73
2 Acquisitions des données expérimentales 73
3 Fantômes numériques 74
IV Résultats sur fantômes 75
V Résultats sur données expérimentales 81
VI Analyse et conclusion de l'étude 86

5 Reconstruction tomographique et apprentissage profond 95
I Introduction à l'apprentissage profond 96
1 Apprentissage automatique 96
2 Apprentissage profond 98
3 Apprentissage et paramètres d'un réseau 101
II Réseaux de neurones convolutifs 103
1 Couches de convolution 104
2 Couches de pooling 105
3 Architecture des CNN 106
III Apprentissage profond en CT 107
1 En post-traitement 108
2 En pré-traitement 109
3 Méthodes end-to-end 110
4 Méthodes semi ou non-supervisées 112

6 Application d'un réseau de neurones en post-traitement 119
I Post-traitement avec un U-Net 120
1 Description des réseaux utilisés 120
2 Entraînement des réseaux 123
II Évaluation des U-Net 125
1 Évolution des métriques pendant l'entraînement 125
2 Évaluation sur les métriques 126
3 Évaluation visuelle 127
4 Amélioration du U-Net-3D 129
III Comparaison avec une régularisation TV 130
1 Détail des algorithmes 131
2 Comparaison 131

Conclusion 139

Liste des communications 143

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Table des figures

1.1 Spectre électromagnétique. 5
1.2 Radiographies d'une pivoine[1] et de la main de Frank Nicolet et sa fracture du cinquième métacarpe après une chute à vélo[2]. 5
1.3 Tube de Coolidge 6
1.4 Spectre de rayons X à la sortie du tube (en rose) et spectre filtré par une plaque d'aluminium d'épaisseur 4 mm (en orange). Simulations faites avec la librairie Spekpy[1] en Python (Bujila *et al.*; 2020; Poludniowski *et al.*; 2021). 7
1.5 Interactions photon - matière 8
1.6 (a) Pourcentage des trois principales interactions dans l'eau en fonction de l'énergie des rayons X. (b) Pourcentage d'énergie transmise en fonction de l'énergie des rayons X. Simulations effectuées à l'aide de la librairie XrayDB[1] en Python. 9
1.7 Coefficient d'atténuation linéaire pour l'iode, les os et l'eau en fonction de l'énergie. Simulations effectuées à l'aide de la librairie XrayDB en Python. 10
1.8 Différentes générations de scanner. 11
1.9 Illustration des différentes géométries d'acquisition. De gauche à droite : parallèle, en éventail, conique. 12
1.10 Schéma d'une acquisition CBCT. 13
1.11 Schéma d'une dent. 14

2.1 Transformée de Radon : intégrale de f sur la droite perpendiculaire à $\vec{\alpha_\theta}$ et qui passe par le point de coordonnées $(u\cos\theta, u\sin\theta)$ 20
2.2 Projection d'angle 0° avec rayons parallèles d'une mâchoire. 20
2.3 Exemple d'un sinogramme. 21
2.4 Illustration de la transformée rayons X en géométrie (a) en éventail et (b) conique pour une droite passant par l'origine. L'intégration se fait sur la droite représentée en rouge. 22
2.5 Différents filtres utilisés pour la rétro-projection filtrée. 24
2.6 Rétro-projection d'une coupe d'une mâchoire, 180 projections allant de 0 à 180°. A gauche : rétro-projection simple, à droite : rétro-projection filtrée. 25
2.7 Couverture angulaire nécessaire en géométrie (a) parallèle et (b) en éventail. Le rayon violet en bas de la projection à 0° est symétrique à celui de la projection à 180° en parallèle. En éventail, la projection à 180° ne donne pas les mêmes rayons, alors que celle à 180° + angle de cône a le même premier rayon violet de celle à 0°. 25
2.8 Introduction des notations utilisées dans la formule de FDK. 26
2.9 Représentation géométrique de l'algorithme ART avec $M = N = 2$. 29
2.10 Illustration de l'importance d'inclure la sensibilité dans la reconstruction. Les rayons jaunes qui arrivent sur le détecteur à 0° ont rencontré 12 points de la grille, alors que les mêmes rayons roses à 40° en rencontrent 6 pour le premier et environ 16 pour le second. 30

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

2.11 Reconstructions MLEM avec différents nombres d'itérations : 1, 10, 100 et 500. Sur la ligne du haut, reconstructions à partir de projections non bruitées, et sur la ligne du bas, à partir de projections bruitées par du bruit poissonnien. Pour ces reconstructions, 180 projections parallèles régulièrement échantillonnées entre 0 et 180° ont été utilisées. Un bruit de Poisson a été ajouté aux projections. 33
2.12 Illustration d'un mouvement de tête d'un patient créant un artefact de bougé-patient (a) sans correction, (b) avec correction (Hsieh; 2009). L'artefact se présente sous forme de rayures obliques dans le crâne du patient. 36
2.13 (a) Sinogramme avec un pixel du détecteur défectueux. (b) Exemple d'artefact en anneau (Buzug; 2011). 37
2.14 Différents artéfacts, avec effets sur le sinogramme et l'image reconstruite. (a) Reconstruction normale, 300 projections en géométrie parallèle, régulièrement espacées avec des angles allant de 0 à π et un détecteur de 300 pixels (b) Reconstruction avec projections bruitées par un bruit gaussien de moyenne nulle et de variance $\sigma = 0.02 \times \max(\text{sinogramme})$, (c) Reconstruction avec 200 projections et des angles allant de 0 à $\frac{2\pi}{3}$ seulement, (d) Reconstruction avec projections tronquées : taille du détecteur = 200 pixels, (e) Reconstruction avec peu de projections, seulement 1 sur 5, et (f) Reconstruction avec une dent en métal. 38
2.15 Illustration de l'artefact de cone-beam sur le fantôme *Defrise*, créé avec la librairie TomoPhantom (Kazantsev *et al.*; 2018). (a) Coupe du fantôme de face, (b) projection cone-beam, (c) coupe de face de la reconstruction, (d) coupe de profil de la reconstruction. 39

3.1 Débruitage d'une photographie[1] d'une personne souhaitant rester anonyme, à l'aide de la variation totale. a : Image nette, b : image bruitée par un bruit gaussien avec $\sigma = 0.11 \times v_{\max}$, c : image débruitée, 500 itérations, $\alpha = 30$. 54
3.2 Fantôme de Shepp-Logan utilisé pour les simulations. 59
3.3 Reconstructions du fantôme de Shepp-Logan. 60
3.4 Profil du fantôme original et des trois reconstructions itératives. 60
3.5 Evolution de l'erreur quadratique moyenne selon le nombre d'itérations pour les algorithmes SIRT-TV, KL-TV et MLEM-TV. 61
3.6 Valeurs de la fonction coût pour MLEM-TV et KL-TV en fonction (a) des itérations, (b) du temps en secondes. 61
3.7 (a) Une projection de l'échantillon de CoOCNT. (b) Coupe de la reconstruction MLEM-TV du CoOCNT. 62

4.1 Coupes axiales et sagittales de la sensibilité, pour le volume normal et pour le volume augmenté. 71
4.2 Schéma de la géométrie d'acquisition, à l'échelle. 73
4.3 (a) Coupes axiales et sagittales des fantômes utilisés. En haut, le fantôme de Shepp-Logan et en bas, le fantôme de mâchoire. (b) Schéma de coupes axiale, coronale et sagittale sur un patient. 74
4.4 Evolution des fonctions coûts lors de la reconstruction du fantôme de mâchoire en très faible dose. A gauche, les fonctions coûts en fonction des itérations, et à droite, en fonction du temps. 76
4.5 Evolution de la MSE lors de la reconstruction du fantôme de mâchoire en très faible dose, en fonction du temps. 76
4.6 A gauche : le profil étudié correspond à la ligne rose. A droite : profil des différentes reconstructions. 77
4.7 Coupes axiales du fantôme de mâchoire. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV et (f) KL-TV. 78

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

4.8 Une coupe coronale et son agrandissement pour le fantôme de mâchoire. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV et (f) KL-TV. 79
4.9 Une coupe coronale et son agrandissement pour le fantôme de Shepp-Logan. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV et (f) KL-TV. 80
4.10 Différence entre les reconstructions et la vérité terrain. (a) SIRT-TV, (b) MLEM-TV et (c) KL-TV. 81
4.11 Coupes axiales des reconstructions dentaires, en faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV, (f) KL-TV. 84
4.12 Coupes axiales contenant du métal, extraites de reconstructions faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV, (f) KL-TV. Les flèches rouges indiquent les molaires gauches. 84
4.13 Agrandissement d'une dent sur une coupe coronale, en faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV, (f) KL-TV. Les flèches bleues, vertes et bleu clair pointent respectivement vers le canal de la dent, l'os trabéculaire et le canal mandibulaire. 85
4.14 Coupes axiales des reconstructions dentaires, en très faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV, (f) KL-TV. Les canaux radiculaires de la première molaire droite sont identifiés par les flèches jaunes. 85
4.15 Coupes axiales contenant du métal, en très faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV, (f) KL-TV. La flèche orange pointe les canaux radiculaires supprimés par MLEM et SIRT-TV, et la flèche blanche un canal qui n'est reconstruit par aucune méthode. 86
4.16 Agrandissement d'une dent dans une coupe coronale, en très faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV (f) KL-TV. Les flèches rose, bleue et bleu clair indiquent respectivement la pulpe, le canal de la dent et le canal mandibulaire. 87

5.1 Illustration de sous-apprentissage et de sur-apprentissage dans le cas d'une régression polynomiale et d'une classification sur les données d'entraînement. 99
5.2 Illustration des opérations logiques effectuées à l'aide de neurones artificiels. Figure inspirée de (Géron; 2022). Chaque flèche représente une connexion, et il faut au moins deux entrées positives pour activer le neurone. 99
5.3 Représentation des étapes suivies par une entrée x dans un neurone de McCulloch-Pitts. Le neurone multiplie x par le poids w et lui ajoute le biais b. La fonction d'activation ρ est ensuite appliquée à ce résultat pour donner y, la sortie. 100
5.4 Schéma d'un perceptron à deux entrées et trois sorties. 100
5.5 Architecture d'un perceptron multi-couche. 101
5.6 Représentation graphique des fonctions d'activation sigmoïde, tangente hyperbolique et LeakyReLU pour un paramètre λ égal à 0.1. 103
5.7 Différents filtres appliqués à la photographie[1] d'un graffiti d'une souris (a). (b) Filtre vertical, (c) filtre horizontal, (d) flou uniforme, (e) gradient et (f) contours. 105
5.8 Architecture typique d'un CNN utilisé en classification. 106
5.9 Architecture du U-Net original (Ronneberger *et al.*; 2015). La taille des images/feature maps est notée à gauche des blocs bleus représentant les feature map multi-canaux. Le nombre de canaux est écrit au-dessus de chaque bloc. Les blocs blancs sont les copies des blocs bleus avec pointillés qui sont concaténées au bloc bleu en sortie de la déconvolution. conv $d \times s$ correspond à une convolution avec un filtre de taille $d \times d$. 107

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

6.1 Exemple d'artefacts présents dans les images de sortie des réseaux. A gauche : on observe un halo flou autour de la couronne en métal lorsque la valeur du gradient est seuillée à 10^{-2}. A droite : l'artefact en damier, caractéristique des couches de déconvolution dans les zones de haute intensité. 121
6.2 Architecture du U-Net utilisé pour le post-traitement. Inspiré de (Jin *et al.*; 2017) avec modification de la fonction d'activation et des couches de upsampling. 121
6.3 Schéma du fonctionnement du U-Net-MP. Trois réseaux Γ_A, Γ_S et Γ_C sont respectivement entraînés sur l'ensemble des coupes axiales, sagittales et coronales. Pour l'inférence, le volume passe, coupe par coupe, par chacun des trois réseaux, puis les trois volumes obtenus sont moyennés pour obtenir le volume final. 122
6.4 (a) Illustration des rayures verticales présentes dans une coupe coronale d'une reconstruction FDK faible dose. (b) et (c) Exemple d'une coupe axiale (et son agrandissement) d'un volume auquel on a appliqué un U-Net entraîné sur les coupes sagittales. Image légèrement saturée pour permettre une meilleure visualisation en impression. On peut remarquer les rayures verticales correspondant à la direction d'entraînement. 123
6.5 (a-c) Coupes de patients de la base de données d'entraînement. (d) Coupe du patient utilisé pour les tests de la section III. 124
6.6 Moyennes et écart-type (σ) des courbes de MSE et MAE évaluées sur les données d'entraînement et de validation, en fonction des époques, des 32 entraînements du réseau Γ_A et du réseau U-Net-3D. 126
6.7 Visualisation des valeurs des métriques sur chacun des volumes. Plus la couleur d'un volume est rose, moins la métrique associée est bonne. 127
6.8 Exemples de coupes de volumes en sortie des réseaux. Les deux premières colonnes correspondent aux meilleures valeurs des métriques, et les deux dernières aux moins bonnes. Chaque ligne correspond à un réseau. Première ligne de gauche à droite : coupes axiales des patients 19, 5, 24 et 8 en sortie du réseau Γ_A. Seconde ligne : patients 19, 5, 24 et 26 en sortie de Γ_C. Troisième ligne : coupes sagittales en sortie de Γ_S des patients 19, 5, 24 et 29. Dernière ligne : coupes axiales des patients 19, 6, 17 et 23 en sortie du réseau U-Net-3D. 128
6.9 Courbes de la fonction coût (MAE train) sur les 100 époques, et des MAE et MSE évaluées sur les données de validation (val) pour le réseau U-Net-3D+. 129
6.10 Coupe axiale et agrandissement d'une coupe coronale des volumes en sorties des réseaux U-Net-3D et U-Net-3D+. 130
6.11 Reconstructions d'une coupe axiale contenant une couronne en métal. 132
6.12 Agrandissement de la coupe axiale présentée dans la figure 6.11, provenant des volumes en sortie des réseaux. 132
6.13 Agrandissements de reconstructions d'une coupe axiale ne contenant pas de métal. La flèche orange pointe vers le ligament alvéolaire dentaire de la deuxième prémolaire droite. 133
6.14 Agrandissements de reconstructions d'une coupe sagittale, sur la deuxième prémolaire mandibulaire gauche. 134
6.15 Agrandissement de reconstructions d'une coupe coronale, sur une zone de la mâchoire contenant le canal mandibulaire. 134

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Liste des tableaux

4.1 Paramètres utilisés pour les reconstructions du fantôme de Shepp-Logan. N est le nombre d'itérations de reconstruction, NTV est le nombre d'itérations TV et α est le paramètre TV utilisé pour déterminer l'importance de la régularisation. 75
4.2 Paramètres utilisés pour les reconstructions du fantôme numérique de la mâchoire. 75
4.3 NRMSE, PSNR et SSIM pour les reconstructions du fantôme de la mâchoire pour les différents algorithmes. 77
4.4 Temps de reconstruction pour les fantômes de Shepp-Logan et de mâchoire pour les différents algorithmes. Le nombre d'itérations utilisé pour la reconstruction et le débruitage TV est rappelé entre parenthèse. 81
4.5 Paramètres utilisés pour la reconstruction des données expérimentales. 82
4.6 CNR et corrélation des reconstructions des données expérimentales, en faible dose et très faible dose. Corr A est la valeur de la corrélation sur le volume entier, alors que Corr B est calculée sur une région 2D contenant des dents et Corr C sur une zone de l'arrière-plan. 83
4.7 Résumé des cinq algorithmes. Nous rappelons sur quelles méthodes ils sont basés, la fonctionnelle qu'ils minimisent, leur temps de reconstruction, leur capacité à réduire le bruit, les artefacts et à restaurer correctement les détails. 87

6.1 Moyennes et écart-types des valeurs des métriques sur les volumes d'évaluation pour les 32 entraînements de la validation croisée. 126
6.2 Valeurs des métriques évaluées sur le patient 32, pour les volumes en sortie du U-Net-3D, du U-Net-3D+, ainsi que le U-Net-MP pour comparer. 130
6.3 Valeurs des métriques sur volume entier et les trois ROIs. 135

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Introduction

Contextes et objectifs

Cette thèse a été réalisée au sein du laboratoire d'imagerie médicale CREATIS en collaboration avec l'entreprise Carestream Dental. L'entreprise Carestream Dental développe des dispositifs à base de rayons X intra et extraoraux, des solutions CAD/CAM [1] et des logiciels pour l'imagerie. Ces systèmes sont destinés à l'étude de la région maxillo-faciale et s'adressent à des dentistes, des spécialistes et des radiologues. Leur utilisation se popularise et demande la mise en place de nouvelles technologies de traitement d'image et d'analyse.

La dentisterie a beaucoup évolué au cours du siècle dernier, en grande partie grâce aux progrès de la technologie d'imagerie. Les rayons X ont été découverts en 1895, et depuis lors, l'imagerie médicale par rayons X a joué un rôle de plus en plus important dans tous les domaines, du diagnostic à la planification du traitement. Aujourd'hui, l'imagerie dentaire représente une part essentielle des soins dentaires de routine, permettant aux dentistes de détecter les problèmes à un stade précoce et de fournir un traitement plus efficace. L'imagerie dentaire peut être utilisée pour examiner les dents, les mâchoires et les structures environnantes, et peut aider à diagnostiquer un large éventail de problèmes.

Parmi les équipements proposés par l'entreprise, une gamme de scanners de tomographie à faisceau conique (Cone-Beam Computed Tomography : CBCT) est commercialisée. Ces scanners permettent de mettre à portée des praticiens l'acquisition de données 3D pour un coût financier et temporel raisonnable, ainsi qu'un impact radiologique faible pour les patients. Le CBCT est largement utilisé en imagerie dentaire, il fournit des images tridimensionnelles du patient à partir de mesures extérieures de celui-ci. Le rayonnement X transmis par la matière est mesuré et les tissus sont reconstruits à partir de ces mesures.

Cependant, les rayons X sont ionisants, ils peuvent endommager l'ADN des cellules et être la source de cancers. Ainsi, l'un des défis de l'imagerie médicale est de réduire la dose d'exposition aux radiations, tout en obtenant des images de haute qualité. Cela est particulièrement important pour la tomographie, qui utilise une dose de rayonnement plus élevée que les radiographies.

La reconstruction des images en faible dose pose cependant des problèmes, car elle peut provoquer des artefacts qui rendent l'interprétation des images difficile. De nouveaux algorithmes de reconstruction sont développés pour améliorer la qualité des images. La reconstruction d'images 3D en CBCT avec des algorithmes analytiques proposée jusqu'à présent par l'entreprise a l'avantage d'être rapide, mais les artefacts liés à l'angle limité, aux projections manquantes, à la tomographie locale d'une région d'intérêt ou encore la présence de bruit, peuvent diminuer la qualité de l'image.

Les méthodes de reconstruction itératives avec prise en compte d'informations a priori semblent mieux adaptées dans une perspective de réduction du nombre de projections et de la dose au patient. Le premier objectif de cette thèse est de définir des méthodes de reconstruction

1. Conception et fabrication assistées par ordinateur.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

itératives basées sur des techniques de régularisation avec des critères adaptés à l'imagerie orale et d'identifier les algorithmes de résolution du problème contraint les plus rapides et les plus fiables.

Depuis quelques années, les méthodes d'apprentissage profond émergent en imagerie médicale au vu des excellents résultats obtenus en traitement d'image classique. Ces méthodes surpassent les méthodes traditionnelles pour le débruitage, la segmentation et la reconnaissance d'images. Leur application à la tomographie pourrait repousser les limites de l'imagerie médicale, permettant un meilleur diagnostic et une réduction de la dose. Le second objectif de cette thèse est d'étudier le potentiel de ce type de méthode et de les tester dans le cadre de la reconstruction tomographique en imagerie dentaire faible-dose.

Contributions et organisation du manuscrit

Le chapitre 1 présente en détail le principe de la tomographie. La chaîne d'acquisition d'une radiographie est rapidement décrite, puis les caractéristiques importantes de la tomographie, nécessaires à la compréhension de la suite de la thèse, sont présentées. Dans le chapitre 2, une liste non exhaustive d'algorithmes de reconstruction est donnée. Nous décrivons les algorithmes analytiques et quelques algorithmes itératifs.

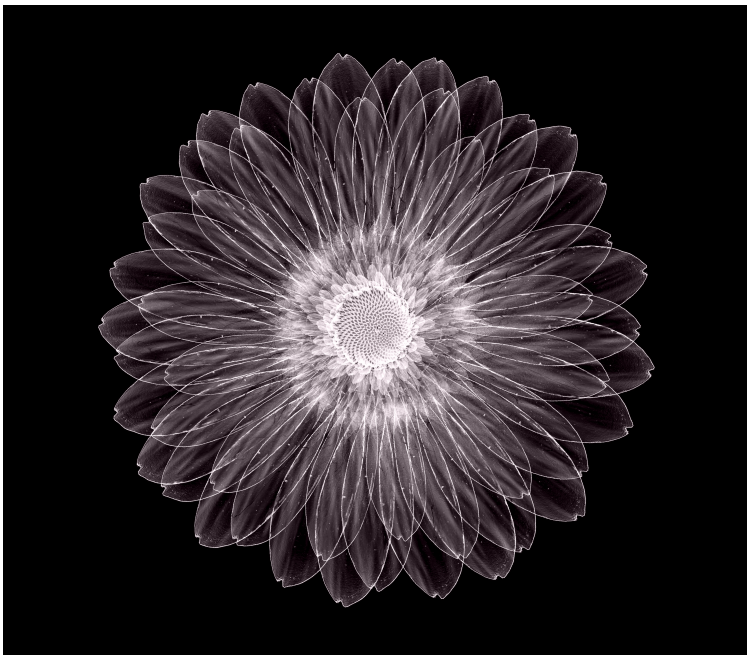
Les méthodes de reconstruction itératives régularisées, qui incluent de la régularisation pendant le processus de reconstructions, sont introduites dans le chapitre 3. Ces algorithmes ont été appliqués au fantôme de Shepp-Logan 2D dont les projections sont corrompues par un bruit gaussien-Poisson. Ils ont ensuite été adaptés à la géométrie conique dans le chapitre 4, tout d'abord sur des fantômes, puis sur des données expérimentales d'un fantôme anthropomorphique acquises avec un scanner de Carestream Dental. Lors de cette dernière étape, les algorithmes ont dû s'adapter à la présence de métal et aux projections tronquées rencontrées en imagerie dentaire.

Le chapitre 5 introduit les méthodes d'apprentissage profond et synthétise différentes manières de les utiliser en reconstruction tomographique en passant en revue l'état de l'art actuel sur ce sujet.

Une étude d'application des méthodes d'apprentissage profond au CBCT dentaire est donnée dans le chapitre 6. Des U-Net ont été entraînés de manière supervisée à améliorer la qualité de reconstructions faites avec un algorithme analytique à partir de données faibles doses. Les méthodes utilisées sont étudiées puis comparées aux méthodes analytiques et itératives.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

1 | Principe de la tomographie



I Aspects physiques de la radiographie 4
1 Tube de Coolidge 5
2 Création des rayons X 5
3 Interactions rayons X - matière 7
4 Détection et formation de l'image 9
II Géométries d'acquisition 10
III Applications en imagerie médicale 12
1 Système CBCT en imagerie dentaire 12
2 En imagerie dentaire 13

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Le but de ce chapitre est de donner le contexte physique de la reconstruction tomographique en imagerie dentaire. Nous allons tout d'abord expliquer les fondements de la radiographie, qui constitue la base de la tomographie par rayons X. Pour cela, nous allons suivre le cheminement des rayons X, de leur création jusqu'à leur arrivée sur le détecteur en passant par la traversée du patient. Nous expliquerons comment les interactions entre les rayons et la matière permettent d'obtenir une image de l'intérieur du patient. Nous verrons ensuite comment ont évolué les scanners au fil du temps et les différentes géométries qu'ils utilisent. Enfin, nous parlerons des diverses applications en imagerie médicale, et plus particulièrement en imagerie dentaire. Les informations données dans ce chapitre s'appuient principalement sur les livres (Buzug; 2011) et (Hsieh; 2009).

I Aspects physiques de la radiographie

La radiographie est une technique d'imagerie qui utilise des rayons X (en imagerie médicale), ou gamma (pour les applications industrielles). Ces rayons, extrêmement pénétrants, traversent le patient ou l'objet, et les rayons transmis permettent d'obtenir une image de l'intérieur du volume considéré.

Les rayons X sont un rayonnement électromagnétique similaire à la lumière visible mais de longueur d'onde bien plus faible, allant de quelques picomètres à quelques nanomètres. Contrairement aux rayons lumineux qui sont réfléchis par les objets, les rayons X passent au travers de la plupart des matériaux et sont invisibles à l'œil nu. En 1895, le physicien Wilhelm Röntgen découvre par hasard des rayons émanant d'un tube de verre qui apparaissent sur du papier photographique. Par leur caractère inconnu, il nomme ces rayons les rayons X. La figure 1.1 illustre la position des rayons X dans le spectre électromagnétique. Entre 0.01 nm et 0.1 nm ce sont les rayons X dits "durs", utilisés en radiographie médicale ; au dessus de 0.1 nm, ce sont des rayons "mous" moins pénétrants.

Le lien entre longueur d'onde et énergie est donné par la formule :

$$E = h\nu = \frac{hc}{\lambda}$$

avec $h = 6.63 \times 10^{-34}$ J.s la constante de Planck, $c = 3 \times 10^{8}$ m.s^{-1} la vitesse de la lumière dans le vide et λ la longueur d'onde. Ainsi, les rayons X ont une énergie comprise entre une centaine d'eV à plusieurs MeV.

La radiographie utilise le fait que les rayons X peuvent traverser la matière. Ils sont absorbés de manière différente selon le matériau rencontré. Ainsi, les matériaux à forte densité absorbent plus les rayons X que ceux à faible densité. C'est cette propriété qui est utilisée en médecine et qui permet, par exemple, de mettre en évidence les os par rapport au reste du corps humain lors d'une radiographie. La figure 1.2 montre les radiographies d'une fleur et d'une main.

Les rayons X sont utilisés par les radiologues mais également par les dentistes car les os et les dents sont bien contrastés sur la radiographie. On les retrouve également dans les aéroports pour vérifier les bagages, et en archéologie pour révéler des couches dissimulées dans les peintures ou les manuscrits. Les astrophysiciens les utilisent pour étudier la composition des étoiles.

En médecine, l'arrivée des rayons X a été une grande avancée pour l'imagerie médicale. Grâce à eux, les médecins ont pu observer l'intérieur des patients de manière rapide, et sans intervention chirurgicale.

Image de couverture : radiographie d'un gerbera, licence standard obtenue sur Adobe Stock.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

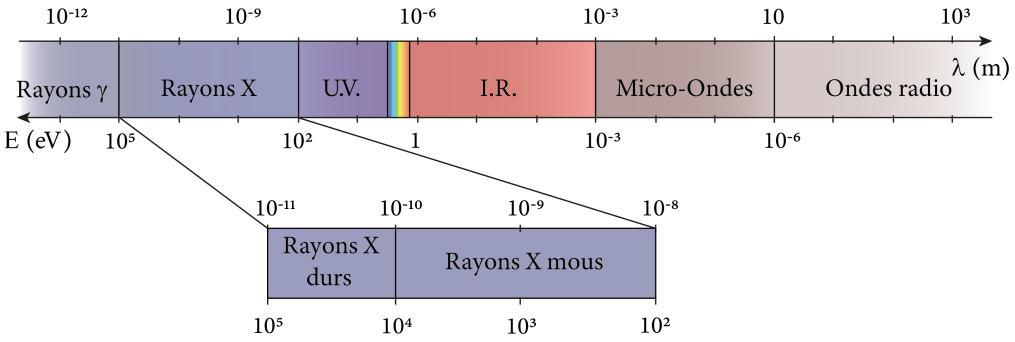


FIGURE 1.1 – Spectre électromagnétique.



FIGURE 1.2 – Radiographies d'une pivoine[1] et de la main de Frank Nicolet et sa fracture du cinquième métacarpe après une chute à vélo[2].
[1] Avec l'autorisation de l'auteur © Nick Veasey et [2] autorisation de Frank Nicolet.

1 Tube de Coolidge

Les rayons X sont principalement produits dans un tube de Coolidge. Il s'agit d'un tube en verre dans lequel se trouvent deux électrodes : une cathode et une anode.

La cathode produit un champ d'électrons qui sont envoyés sur l'anode afin de créer le faisceau de rayons X. Ces deux électrodes sont placées dans une chambre vide en verre pour limiter les interactions avec l'air. Les rayons X étant ionisants, il est important d'ajouter une protection autour de la cage en verre pour protéger les opérateurs et les patients des radiations. On utilise une protection en plomb : grâce à son numéro atomique élevé, elle empêche les rayons X de quitter le tube. Enfin, à la sortie du tube se trouve une plaque en aluminium dont le but est de filtrer les rayons X de basses énergies. Un schéma de ce tube est visible dans la figure 1.3.

2 Création des rayons X

La création des rayons X résulte d'interactions entre particules relevant de la physique quantique. Elle s'effectue en trois étapes. La première consiste en la création d'un nuage électronique par la cathode. Cette dernière est composée d'un filament de métal que l'on fait chauffer par ef-

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

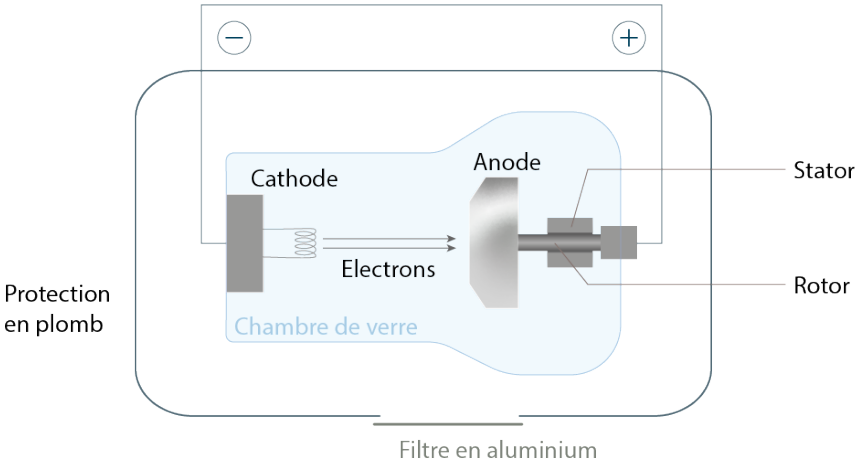


FIGURE 1.3 – Tube de Coolidge

fet Joule. Le filament se réchauffe et la chaleur provoque une agitation des électrons superficiels dont une partie se détache pour constituer un nuage électronique, c'est le processus d'émission thermo-ionique. L'intensité du faisceau d'électrons est proportionnelle à l'intensité du courant de chauffe de la cathode et à la durée du chauffage.

Pendant la deuxième étape, les électrons arrachés des atomes du filament sont accélérés et envoyés sur l'anode. L'anode est portée à très haute tension, la différence de potentiel entre l'anode et la cathode crée un champ électrique qui accélère les électrons en direction de l'anode. Plus la différence de potentiel est grande, plus l'énergie cinétique des électrons est grande. La cathode et l'anode sont placées dans une chambre vide en verre afin d'éviter que les électrons ne soient déviés par des molécules d'air.

La troisième étape se produit lorsque les électrons interagissent avec l'anode, c'est à ce moment-là que se créent les rayons X. Une fois arrivés sur l'anode, les électrons sont déviés ou entrent en collision avec les atomes de celle-ci.

1. Dans le premier cas, on parle de rayonnement de freinage (*Bremsstrahlung*) : l'électron s'approche d'un atome et à cause de sa charge négative, il est attirée par le noyau d'après la loi électrostatique de Coulomb. Il est dévié de sa trajectoire et subit une décélération. La perte d'énergie cinétique de l'électron conduit à l'émission d'un rayonnement. Plus l'écart d'énergie cinétique de l'électron est grand, plus l'énergie du photon émis sera grande. L'effet Bremsstrahlung crée un spectre continu, aussi appelée "radiation blanche".
2. Dans le second cas, l'électron incident entre en collision avec un des électrons d'une couche profonde d'un atome et l'éjecte de sa couche, l'atome se retrouve ionisé. Pour rétablir l'équilibre, un électron d'une couche supérieure descend remplacer l'électron éjecté. En changeant de couche, cet électron perd de l'énergie, ce qui aboutit à l'émission d'une radiation caractéristique dont l'énergie dépend des couches de l'atome mises en jeu lors de cette interaction. Les différences d'énergie entre les différentes couches d'un atome sont fixées et caractéristiques de cet atome. Ainsi, ce type d'interaction crée des raies caractéristiques dans le spectre d'émission.
3. Enfin, l'électron peut directement rentrer en collision avec le noyau d'un atome. Dans cette configuration, l'intégralité de l'énergie de l'électron se transforme en photon selon l'effet Bremsstrahlung. L'énergie du rayon X produit par cette interaction est ainsi maximale, cependant, la probabilité d'une telle collision est très faible.

Durant ce processus, seule une faible partie de l'énergie incidente est transformée en rayons

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

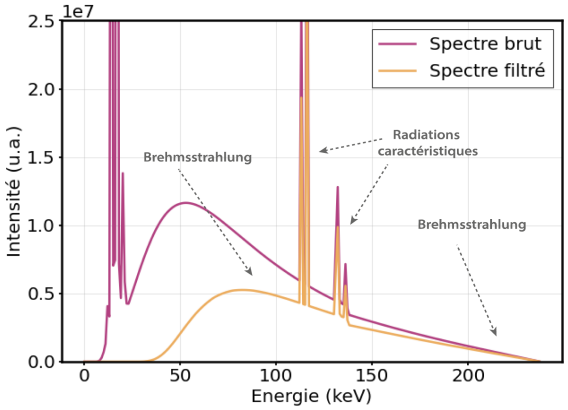


FIGURE 1.4 – Spectre de rayons X à la sortie du tube (en rose) et spectre filtré par une plaque d'aluminium d'épaisseur 4 mm (en orange). Simulations faites avec la librairie Spekpy[1] en Python (Bujila *et al.*; 2020; Poludniowski *et al.*; 2021).

[1] `https://bitbucket.org/spekpy/spekpy_release/wiki/Home`

X, 99% de cette énergie se dissipe en chaleur. Le rendement énergétique est très faible et proportionnel au numéro atomique de l'anode et à la différence de potentiel entre l'anode et la cathode. Pour maximiser le rendement, il est donc préférable de choisir un matériau de numéro atomique élevé et qui soit capable de résister à la chaleur ; en général l'anode est en tungstène, molybdène ou cuivre.

Le spectre de rayons X est obtenu en additionnant le spectre continu créé par l'effet Bremsstrahlung et les radiations caractéristiques qui forment des pics. Un exemple de spectre est montré dans la figure 1.4. Dans ce spectre continu, on retrouve des photons de basse énergie qui sont absorbés par les tissus mous. Ces rayons augmentent la dose de rayons X envoyée au patient et n'apportent aucune contribution utile (Webb et Kagadis; 2003). Ils sont donc filtrés à l'aide d'une plaque en aluminium qui absorbe uniquement les rayons de faible énergie.

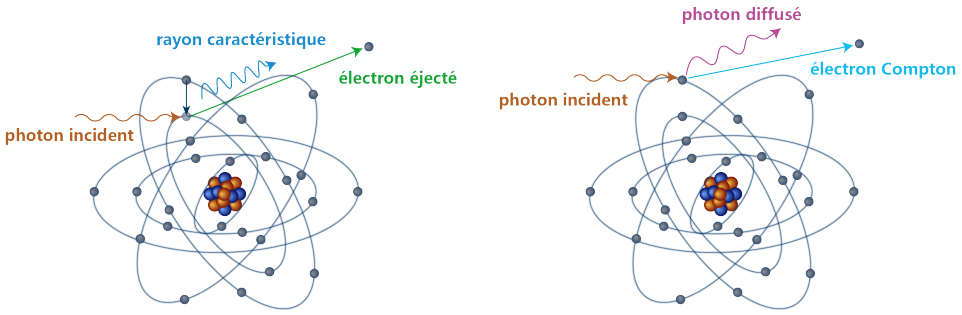
Le spectre obtenu est un spectre polychromatique, c'est-à-dire qu'il contient plusieurs longueurs d'onde différentes. En tomographie, nous verrons dans la suite qu'il est préférable de travailler avec des spectres monochromatiques. Un spectre monochromatique pourrait théoriquement être obtenu par filtrage mais au détriment du flux de photons, ce qui compromet son utilisation en médecine. Des faisceaux monochromatiques à fort flux de photons peuvent être obtenus sur des synchrotrons, ce qui permet d'obtenir des images tomographiques quantitatives à fort rapport signal sur bruit, mais cette technique n'est toutefois pas utilisable en imagerie dentaire.

3 Interactions rayons X - matière

Après sa sortie du tube de Coolidge, le faisceau de rayons X traverse le patient, et les rayons non absorbés par le patient forment l'image de radiographie. Durant leur traversée les photons des rayons X interagissent de différentes manières avec la matière. Le type d'interaction dépend de l'énergie des rayons X et des numéros atomiques des atomes rencontrés, on en trouve cinq différents, dont seulement trois dans le domaine d'énergie utilisé en imagerie par rayons X.

Effet photo-électrique Le photon incident éjecte un électron de l'atome et est complètement absorbé. L'atome se retrouve ionisé et crée un rayonnement caractéristique secondaire pour se stabiliser. C'est l'interaction la plus utile pour la radiographie médicale, elle apparaît pour des énergies entre 10 et 300 keV et dépend du numéro atomique de l'atome.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés



(a) Effet photo-électrique (b) Effet Compton

FIGURE 1.5 – Interactions photon - matière

Cette dernière propriété est importante : les os, principalement composés de calcium dont le numéro atomique $Z = 20$ est supérieur à celui de l'oxygène présent dans l'eau ($Z = 8$), absorberont plus les rayons X par effet photo-électrique que les tissus mou. C'est ce qui permettra de distinguer les deux matières sur la radiographie.

Effet Compton Un photon incident de grande énergie entre en collision avec un électron de l'atome mais n'est pas absorbé. L'électron est éjecté et le photon incident, qui a perdu une partie de son énergie, est diffusé dans une direction aléatoire faisant un angle dépendant de l'énergie initiale et de l'énergie perdue. La diffusion Compton est donc aussi appelée diffusion incohérente et elle intervient dans la même plage d'énergie que l'effet photo-électrique. Sa probabilité ne dépend que de la densité électronique du matériau - le nombre d'électrons sur la couche de valence - et non pas du numéro atomique de l'atome. L'absence de lien avec le numéro atomique ne permet pas d'obtenir un bon contraste entre les différents matériaux de densité similaire, et l'angle aléatoire du photon diffusé crée du bruit. Par conséquent, les appareils de tomographie essaient de minimiser l'impact de l'effet Compton, avec une grille par exemple.

Diffusion de Rayleigh - diffusion cohérente Le photon excite l'atome sans collision avec un électron. Le retour à l'état stable de l'atome émet un nouveau photon de même énergie. Cette interaction n'apparaît que lorsque l'énergie du photon est relativement basse et est peu utile en imagerie médicale.

Les deux autres interactions, la création de paires ainsi que la photo-désintégration, n'ont lieu que dans des tranches d'énergie très hautes non utilisées en radiographie médicale. La création de paires se produit lorsque le photon interagit au voisinage du noyau de l'atome et qu'un électron et un positon sont créés. Si le photon a une très grande énergie, l'impact de celui-ci avec un atome peut causer la destruction d'une partie du noyau : c'est la photo-désintégration.

La figure 1.5 représente les deux principales interactions en radiographie : l'effet photo-électrique et l'effet Compton. Afin de mieux comprendre l'importance relative de chaque interaction, la figure 1.6a montre le pourcentage d'interaction dans l'eau selon l'énergie du photon. L'effet photo-électrique est prédominant en dessous de 30 keV ; au-delà, l'effet Compton prédomine. La diffusion de Rayleigh a une faible probabilité de se produire pour toutes les énergies. La figure 1.6b complète la figure précédente et montre que pour les faibles énergies, la majorité de l'énergie transférée est due à l'effet photo-électrique.

D'après ce qui précède, en traversant le patient, les rayons X peuvent être transmis, diffusés ou absorbés. En effectuant une radiographie, nous mesurons l'atténuation que subit le faisceau

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

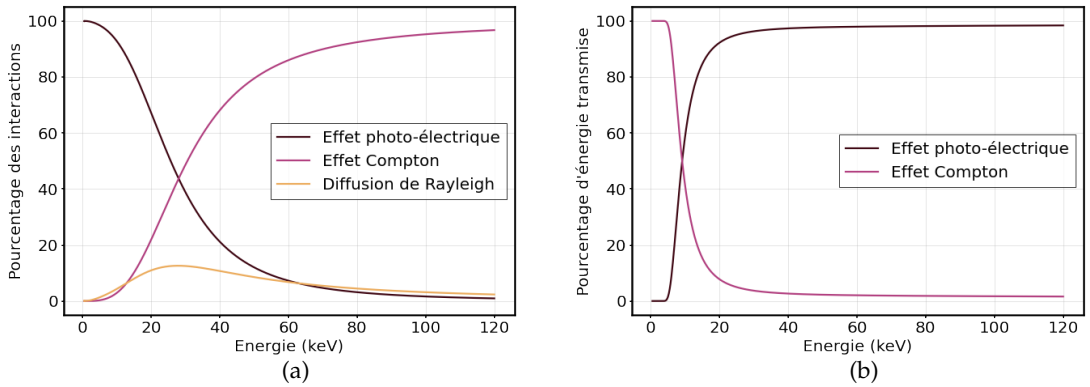


FIGURE 1.6 – (a) Pourcentage des trois principales interactions dans l'eau en fonction de l'énergie des rayons X. (b) Pourcentage d'énergie transmise en fonction de l'énergie des rayons X. Simulations effectuées à l'aide de la librairie XrayDB[1] en Python.

[1] `https://xraypy.github.io/XrayDB/python.html`

de rayons X. Le faisceau étant polychromatique (cf figure 1.4), son atténuation vérifie l'équation suivante pour un matériau homogène :

$$I = \int_0^{E_{\max}} I_0(E) \mathrm{e}^{-\mu(E)L} dE \tag{1.1}$$

avec I_0 l'intensité du faisceau X incident, L l'épaisseur du matériau traversé (cm) et μ le coefficient d'absorption linéaire du matériau (cm^{-1}).

En pratique, on suppose le rayonnement monochromatique, l'équation (2.32) devient :

$$I = I_0 e^{-\mu L} \tag{1.2}$$

aussi appelée loi d'absorption de Beer-Lambert.

Ainsi, l'intensité reçue par le détecteur dépend du coefficient d'absorption du matériau. La figure 1.7 représente l'évolution du coefficient selon l'énergie pour trois matériaux différents. A partir de 30 keV, l'iode a un coefficient d'absorption beaucoup plus élevé que l'os et l'eau, ce qui en fait un très bon agent de contraste. Il est utilisé pour observer les organes et le sang en angiographie. Le coefficient de l'os est supérieur à celui de l'eau, que l'on identifie aux tissus mous, d'où la nette différence entre le squelette et les tissus sur une radiographie.

4 Détection et formation de l'image

Une fois que les rayons X ont traversé le corps du patient, il faut les quantifier. Au début de la radiographie, on utilisait des films photosensibles avec un fonctionnement similaire à celui de la photographie argentique classique. Des cristaux d'argent présents sur le film sont ionisés lorsqu'ils reçoivent le rayonnement, ils créent des amas d'atomes d'argent qui forment l'image visible. Ainsi, l'image est plus sombre là où elle a reçu plus de rayons, ce qui explique que les tissus mous qui laissent passer plus de rayons ressortent en noir et les os en blanc sur une radiographie.

Lors de l'apparition des premiers capteurs plans, ces films ont été délaissés au profit de leur version numérique qui permet de faire le traitement des images obtenues par ordinateur. On

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

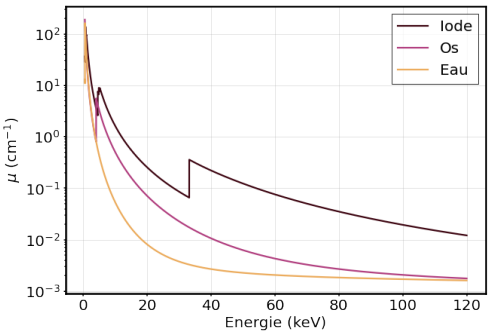


FIGURE 1.7 – Coefficient d'atténuation linéaire pour l'iode, les os et l'eau en fonction de l'énergie. Simulations effectuées à l'aide de la librairie XrayDB en Python.

retrouve deux types de détecteurs, les détecteurs indirects et les détecteurs directs (Rowlands; 2002).

Les *détecteurs indirects* sont constitués d'un scintillateur et de photo-diodes. Le scintillateur convertit les rayons X en rayons lumineux. Avec un matériau tel que l'iodure de césium, des photons de fluorescence sont émis par effet photo-électrique dans le domaine du visible. Les photons lumineux sont ensuite transformés en signaux électriques par les photo-diodes. Cependant, lors des interactions dans le scintillateur, certains photons X sont perdus, et les photons lumineux produits n'ont pas nécessairement la même direction que le photon X qui les a produit.

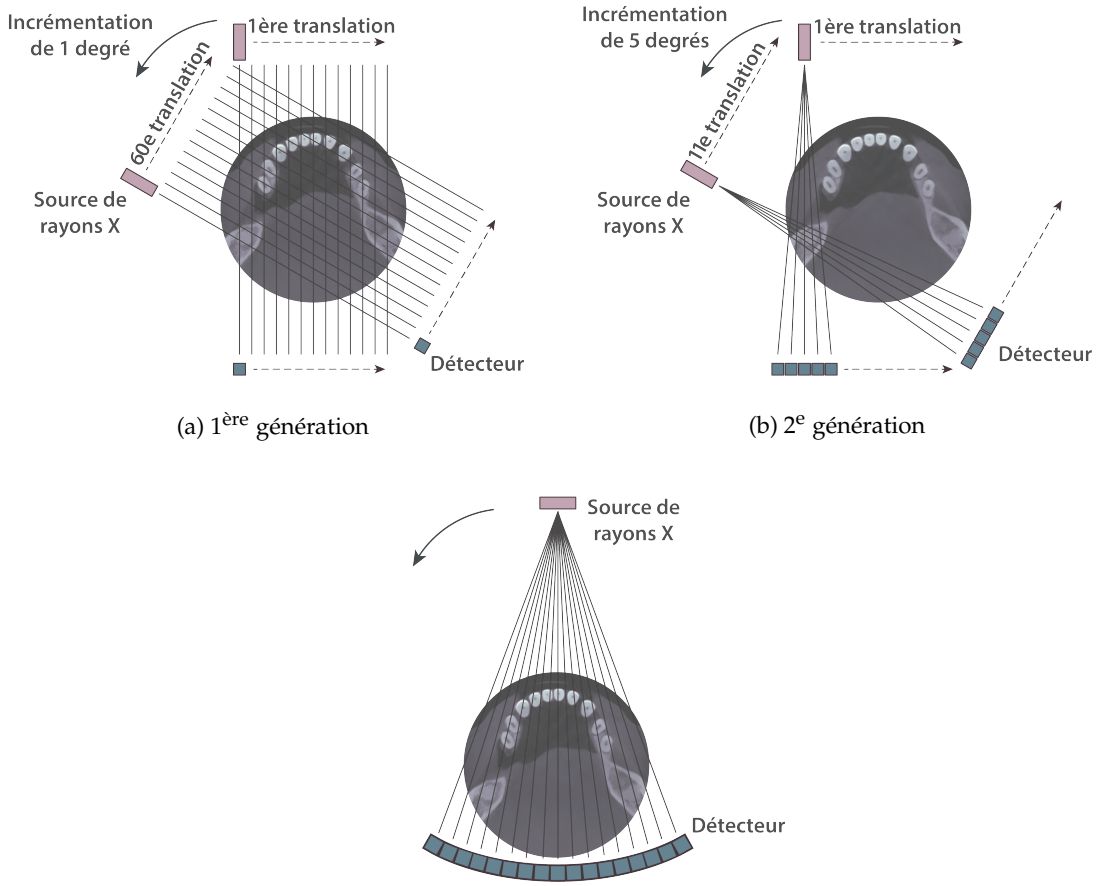
Les *capteurs directs* détectent les photons X sans transformation préalable et les convertissent directement en signaux électriques. Les photons X sont convertis en paires d'électron-trou dans un semi-conducteur tel que le sélénium amorphe, et les électrons sont collectés pour quantifier le nombre de photon X présents. Par rapport aux détecteurs indirects, ceux-ci ont l'avantage de pouvoir également déterminer l'énergie des rayons X et sont donc utilisés en imagerie spectrale. De plus, ces capteurs ont une résolution supérieure à celle des capteurs indirects (Ristić; 2013).

II Géométries d'acquisition

Comme exposé en introduction, le but de la tomographie est d'observer l'intérieur d'un objet ou d'un patient de manière non invasive à partir de mesures extérieures. En tomographie par rayons X, ces mesures extérieures sont des mesures d'atténuation, prises selon différents angles de vue autour du patient. Au fil des années et des différentes avancées technologiques, les scanners et les techniques d'acquisition ont évolué. Nous présentons ici différentes générations de scanners, que l'on retrouve figure 1.8 et géométries d'acquisition, figure 1.9.

La première géométrie d'acquisition utilisait des rayons X parallèles entre eux : c'est la géométrie parallèle utilisée par les scanners de première génération. Leur but était de reconstruire une coupe, c'est-à-dire une image 2D, de l'intérieur d'un patient. Ils utilisaient un détecteur mono-pixel qui se déplaçait linéairement, en même temps que la source, pour scanner le patient selon un certain angle, ce qui permettait d'obtenir une mesure 1D. Puis l'opération était répétée autant de fois que nécessaire afin d'obtenir les acquisitions sous plusieurs angles (figure 1.8a). Avec ce type de scanner, le temps d'acquisition était très long (environ 5 minutes) et la qualité des reconstructions diminuée par les mouvements du patient. Cette géométrie n'est aujourd'hui plus utilisée en imagerie médicale.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés



(a) 1ère génération

(b) 2e génération

(c) 3e génération

FIGURE 1.8 – Différentes générations de scanner.

Pour la géométrie en éventail [1], la source est ponctuelle et le faisceau de rayons X a une forme d'éventail. Le détecteur est plat ou incurvé. On retrouve cette géométrie dans les scanners de seconde et troisième génération. Les scanners de seconde génération avaient pour principal but de réduire considérablement le temps d'acquisition par rapport à la première génération. Le faisceau de rayons X a une forme d'éventail et le détecteur multi-pixels permet de réduire le nombre d'angles d'acquisition (figure 1.8b). Un scan pouvait être effectué en moins d'une minute avec cette génération de scanner. La troisième génération de scanner a vu disparaître la translation nécessaire pour chaque angle. Le détecteur est devenu suffisamment grand pour contenir l'intégralité de l'objet dans son champ de vue (figure 1.8c). Ainsi, seule une rotation autour du patient était suffisante, ce qui a permis de réduire le temps d'acquisition en dessous de 3 secondes.

Aujourd'hui, les détecteurs utilisés peuvent être plans ou incurvés, indépendamment de la forme du faisceau, et le scanner n'effectue qu'une seule rotation autour du patient, sans translation et en un temps très court.

Les scanners présentés jusqu'à présent permettent de reconstruire une image 2D. Si l'on veut reconstruire un volume 3D du patient, on peut répéter plusieurs fois un scan 2D pour reconstruire plusieurs coupes qui formeront le volume 3D. Aujourd'hui, le couple source-

1. *fan beam* pour faisceau en éventail dans la littérature

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

détecteur suit une trajectoire hélicoïdale : ils tournent plusieurs fois autour du patient en effectuant une légère translation pour obtenir les mesures sur plusieurs coupes. Une dernière technique est d'utiliser un détecteur rectangulaire composé de plusieurs lignes de pixels.

La géométrie conique [1] utilise un faisceau en forme de cône. La source reste ponctuelle, mais plusieurs plans de faisceau en éventail sont collectés simultanément pour couvrir un volume. Cependant, la source étant ponctuelle, un seul plan est perpendiculaire à l'axe de rotation, les autres sont inclinés. Pour cette géométrie, l'acquisition se fait en n'effectuant qu'un seul tour autour du patient. Le processus de reconstruction tomographique est noté dans ce cas CBCT (*Cone Beam Computed Tomography*). Les données obtenues en géométrie conique correspondent à des radiographies.

Dans la suite de cette thèse, c'est cette dernière géométrie qui nous intéressera. Cependant, la formulation géométrique étant complexe, les premières explications sur la reconstruction tomographique resteront dans le cadre simplifié de la reconstruction 2D et de la géométrie parallèle. On retrouve les trois différentes géométries dans la figure 1.9.

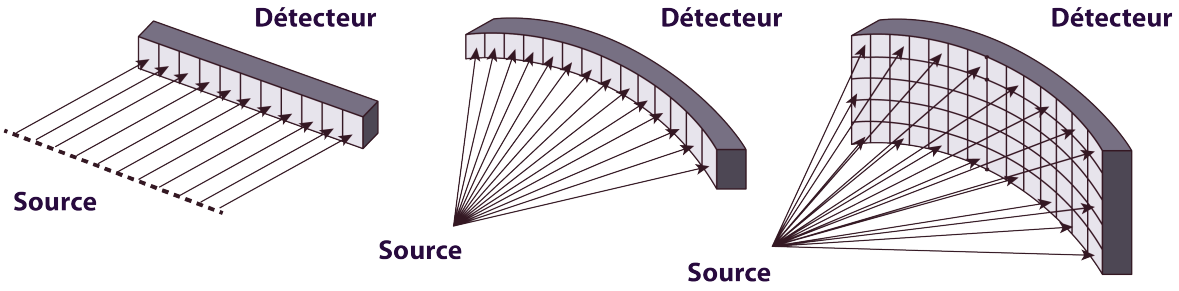


FIGURE 1.9 – Illustration des différentes géométries d'acquisition. De gauche à droite : parallèle, en éventail, conique.

III Applications en imagerie médicale

Les utilisations de la tomographie en imagerie médicale sont nombreuses. Puisqu'elle utilise la transmission des rayons X, on comprend qu'elle bénéficie des avantages de celle-ci et qu'elle sera utile pour observer les os et par exemple, détecter des fractures. Mais elle permet également de s'affranchir de la superposition des différents matériaux traversés sur la radiographie, et donc de déterminer leur position relative. Elle améliore aussi le contraste dans les tissus et des anomalies dans les poumons peuvent par exemple être détectées grâce à une image de scanner. Enfin, avec l'apparition des produits de contraste, elle trouve des applications en angiographie et peut imager les vaisseaux sanguins ou encore le cerveau.

Nous allons, dans la suite de cette partie, parler plus précisément de la tomographie par faisceau conique (CBCT) et de ses applications en imagerie dentaire.

1 Système CBCT en imagerie dentaire

Lors d'un examen de scanner CBCT, le patient est installé assis ou debout et garde sa tête immobile. La source de rayons X et le détecteur vont effectuer une rotation simultanée autour du patient et acquérir les images radiographiques nécessaires à la reconstruction. La figure 1.10 illustre le fonctionnement du CBCT et sa géométrie d'acquisition. Une fois les projections

1. *cone beam* dans la littérature

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

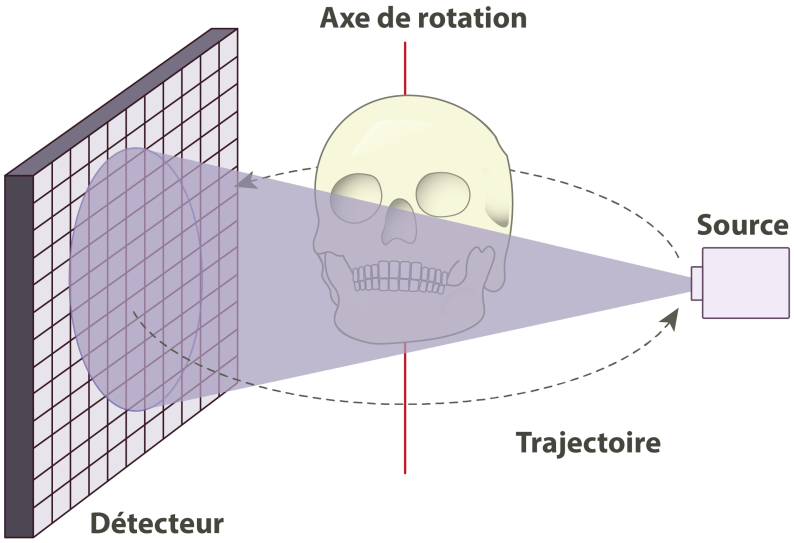


Figure 1.10 – Schéma d'une acquisition CBCT.

récupérées par l'ordinateur, elles sont pré-traitées avant qu'un algorithme de reconstruction n'effectue la reconstruction 3D. Les images obtenues peuvent ensuite subir encore quelques post-traitements, en particulier lors d'acquisition à faible dose ou de présence de métal, et sont finalement prêtes pour le diagnostic et/ou la modélisation 3D.

Comparé aux scanners conventionnels, le CBCT offre plusieurs avantages. Il permet de diminuer la dose envoyée au patient tout en offrant une meilleure résolution spatiale grâce à ses petits voxels isotropiques (Grondin; 2013; Dartus *et al.*; 2021). Par la position assise ou debout du patient, les scanners CBCT sont de plus petite taille que les scanners conventionnels. Ils peuvent donc être installés dans les cabinets dentaires. Ils sont également beaucoup moins coûteux, ce qui les rend facilement accessibles pour les dentistes qui exercent dans des cabinets en dehors d'un milieu hospitalier.

2 En imagerie dentaire

En imagerie dentaire, les scanners CBCT sont utilisés pour produire des images 3D des structures dentaires, des tissus mous, des canaux contenant les nerfs et des os présents dans la région maxillo-faciale. Ils fournissent des images détaillées de l'os et permettent de détecter des anomalies dans cette région : la denture, les structures osseuses du visage, la cavité nasale et les sinus. Nous retrouvons dans la figure 1.11 un schéma d'une dent avec les différentes structures anatomiques importantes.

Les points importants à examiner dans un scanner de la mâchoire sont :

- la zone ligamentaire autour des dents
- l'os alvéolaire, pour la pose d'implant
- les potentielles lésions autour des racines des dents
- les canaux à l'intérieur des dents, leur longueur et leur largeur
- la position du canal mandibulaire, particulièrement importante pour ne pas avoir de contact avec le nerf à l'intérieur pendant une intervention chirurgicale.

La tomographie en imagerie dentaire est une modalité d'imagerie particulièrement difficile, les structures importantes recherchées par le praticien sont souvent très fines et nécessitent une excellente qualité de reconstruction. Cependant, de nombreux facteurs physiques dégradent la

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

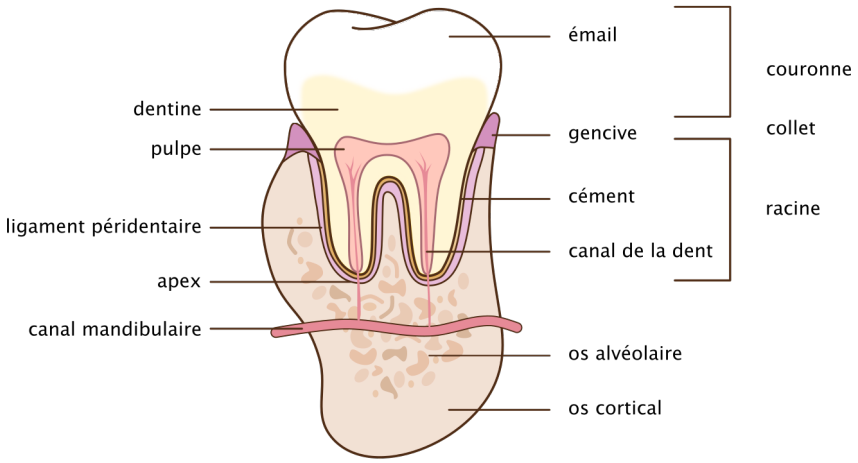


FIGURE 1.11 – Schéma d'une dent.

qualité de l'image. La position assise du patient le rend moins stable que lorsqu'il est allongé, et ses mouvements lors de l'acquisition créent des artefacts dans l'image reconstruite s'ils ne sont pas corrigés. De plus, la zone maxillo-faciale contient souvent du métal (implants, bagues orthodontiques, etc.) à l'origine d'artefacts métalliques dans les images. Le métal absorbant bien plus les rayons X que les os ou les tissus mous, les projections manquent d'informations pour représenter correctement le volume à reconstruire.

De part son aspect cancérigène, il est nécessaire de réduire la dose au maximum afin de limiter les risques pour le patient. Cependant, en diminuant la dose envoyée, l'importance relative du bruit statistique naturellement présent dans les acquisitions augmente et le rapport signal sur bruit [1] diminue, la qualité de la reconstruction est alors dégradée. En imagerie de la mâchoire, afin de ne pas irradier le patient avec des rayons inutiles, le scanner envoie des rayons seulement dans la région d'intérêt qui est souvent plus petite que la taille de la tête du patient. On obtient alors ce que l'on appelle des projections tronquées, qui ne contiennent pas l'intégralité de l'objet sur chaque angle, et qui génèrent des incohérences lors du processus de reconstruction. Les différents artefacts, ainsi que leurs causes et leurs conséquences, seront détaillés dans le chapitre 2.

Dans ce premier chapitre, nous avons expliqué comment les rayons X sont produits et comment ils interagissent avec la matière pour donner les images de radiographie utilisées en imagerie médicale. Grâce aux propriétés d'absorption propres à chacun des matériaux, il est possible de distinguer les différents matériaux qui composent l'intérieur d'un objet ou d'un patient, sans intervention chirurgicale.

La tomographie à rayons X utilise des données d'atténuation, aussi appelées projections, de l'objet considéré pour reconstruire des images en coupe de celui-ci. Cela permet de séparer des structures qui se superposent dans une radiographie. C'est une technique d'imagerie utilisée dans divers domaines, mais dont les applications sont surtout connues du grand public en imagerie médicale.

En imagerie dentaire, les scanners CBCT ont permis de rendre accessible la tomographie

1. SNR pour *Signal to Noise Ratio*

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

3D à un grand nombre de praticiens. Elle est utilisée dans de nombreuses interventions et domaines : odontologie, parodontie, chirurgie, implantologie, orthodontie ou encore pédodontie.

Le prochain chapitre traitera de la partie reconstruction, de ses bases mathématiques et des premiers algorithmes qui ont été utilisés.

Bibliographie

Bujila, R., Omar, A. et Poludniowski, G. (2020). A validation of SpekPy : A software toolkit for modelling X-ray tube spectra. *Physica Medica*, 75:44–54. xi, 7

Buzug, T. M. (2011). Computed tomography. *In Springer handbook of medical technology*, pages 311–342. Springer. 4

Dartus, J., Jacques, T., Martinot, P., Pasquier, G., Cotten, A., Migaud, H., Morel, V. et Putman, S. (2021). Intérêt du Cone-Beam CT en comparaison du scanner multibarettes conventionnel dans le bilan étiologique de douleurs sur prothèse totale de genou. *Revue de Chirurgie Orthopédique et Traumatologique*, 107(3):358–365. 13

Grondin, O. (2013). Intérêt du Cone Beam en cabinet d'omnipratique libéral. *Bordeaux : Université Bordeaux*, 2. 13

Hsieh, J. (2009). *Computed tomography : principles, design, artifacts, and recent advances*. SPIE press. 4

Poludniowski, G., Omar, A., Bujila, R. et Andreo, P. (2021). SpekPy v2. 0—a software toolkit for modeling x-ray tube spectra. *Medical Physics*, 48(7):3630–3637. xi, 7

Ristić, G. S. (2013). The digital flat-panel X-ray detectors. *In Proceedings of the third conference on medical physics and biomedical engineering, Association for medical physics and biomedical engineering in cooperation with the European federation of organisations for medical physics*, pages 65–71. 10

Rowlands, J. (2002). The physics of computed radiography. *Physics in medicine & biology*, 47(23):R123. 10

Webb, A. et Kagadis, G. (2003). Introduction to biomedical imaging. *Med. Phys*, 30(107153):8. 7

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

2 Méthodes de reconstruction tomographique



I	**Formalisme mathématique**		**18**
	1	Transformées usuelles en tomographie	19
	2	Formulation du problème inverse	21
II	**Reconstructions analytiques**		**22**
	1	Rétro-projection en géométrie 2D parallèle	23
	2	En géométrie divergente 2D	24
	3	En CBCT	26
III	**Reconstructions itératives**		**27**
	1	Méthodes algébriques	27
	2	Méthodes statistiques	30
	3	Accélération de Nesterov	34
IV	**Artefacts de reconstruction**		**34**

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Dans ce chapitre nous introduisons le principe de la reconstruction tomographique ainsi que les méthodes de l'état de l'art. Le but de ces méthodes est de cartographier l'intérieur de l'objet à partir des projections par rayons X ; c'est-à-dire déterminer les coefficients d'atténuation linéaire des matériaux traversés.

Nous commencerons par introduire les bases mathématiques de la tomographie, notamment la transformée de Radon. Elle nous permettra de modéliser les projections par rayons X. Nous parlerons ensuite des méthodes de reconstruction que l'on divise en deux groupes.

Le premier groupe comprend les méthodes analytiques basées sur l'inversion de la transformée de Radon. Les méthodes itératives, quant à elles, considèrent le problème de reconstruction comme une résolution d'équations linéaires et reconstruisent le volume en plusieurs étapes.

Mises à part les parties clairement explicitées, ce chapitre théorique restera dans le cas de reconstruction d'images 2D et donc de projections 1D, avec une géométrie parallèle, afin de simplifier le cadre d'étude et d'en améliorer la compréhension. Le détecteur sera toujours considéré plat et la trajectoire circulaire. Indépendamment de la dimension, nous appellerons *voxels* les pixels ou voxels de l'objet à reconstruire, et *pixels* les pixels du détecteur.

I Formalisme mathématique

Nous avons vu dans le chapitre précédent que les rayons X sont atténués lors de leur parcours par les matériaux traversés. Lorsque le matériau est homogène et le faisceau monochromatique, en notant I_0 l'intensité du faisceau incident, la loi d'absorption de Beer-Lambert donne :

$$I = I_0 \mathrm{e}^{-\mu L}$$

avec I l'intensité mesurée, L l'épaisseur traversée en cm et μ le coefficient d'absorption linéaire du matériau. Lorsqu'on effectue la reconstruction tomographique, l'intensité initiale du faisceau, la mesure de l'intensité transmise sont connues et nous cherchons à déterminer le coefficient d'atténuation μ.

Si le matériau n'est pas homogène, dans le cas discret, l'équation de Beer-Lambert s'écrit en prenant la somme des coefficients des n différents matériaux traversés par le faisceau :

$$I = I_0 \exp\left(-\sum_{i=1}^{n} \mu_n L_n\right)$$

où μ_i est le coefficient d'atténuation du matériau i et L_i la distance parcourue par les photons dans ce matériau.

Pour le corps humain, et les objets de manière générale, le coefficient d'atténuation linéaire est mieux représenté comme une fonction continue par morceaux :

$$I = I_0 \exp\left(-\int_{\Delta} \mu(l)dl\right) \tag{2.1}$$

avec Δ la ligne parcourue et $\mu(l)$ la valeur du coefficient linéaire d'atténuation au point de coordonnée l sur la droite. Nous pouvons réécrire l'équation (2.1) de la manière suivante :

$$\ln\left(\frac{I_0}{I}\right) = \int_{\Delta} \mu(l)dl. \tag{2.2}$$

Ainsi, puisque l'on connaît I et I_0, nous connaissons l'intégrale du coefficient linéaire d'atténuation de l'objet le long de la ligne traversée. Nous obtenons une équation de cette forme

Image de couverture : radiographie de champignons par cadop de Pixabay.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

pour chaque ligne de rayons X que peut mesurer un pixel du détecteur, et pour chaque inclinaison de détecteur effectuée autour du patient. La résolution du problème de reconstruction tomographique repose donc sur la résolution de ces équations intégrales.

L'équation (2.2) est également valable en 3D si l'on considère un coefficient d'atténuation dépendant de l'espace :

$$\ln\left(\frac{I_0}{I}\right) = \int_\Delta \mu(x, y, z)dl. \tag{2.3}$$

1 Transformées usuelles en tomographie

L'intégration de la fonction représentant les coefficients d'atténuation peut s'exprimer de deux manières différentes que nous présentons dans la suite.

Transformée de Radon

En géométrie parallèle, l'équation (2.2) peut être écrite à l'aide de la transformée de Radon (Radon; 1917). L'objet à reconstruire est modélisé par une fonction f qui à tout point de l'espace x associe le coefficient d'atténuation de l'objet en ce point.

Soit $f \in \mathcal{S}(\mathbb{R}^n)$, avec $\mathcal{S}(\mathbb{R}^n)$ l'espace de Schwartz, i.e. f est indéfiniment dérivable et à décroissance rapide, ainsi que toutes ses dérivées. La transformée de Radon $\mathcal{R}$ est définie par les intégrales de f sur les hyperplans de $\mathbb{R}^n$. Soit $\vec{\xi} \in \mathbb{S}_1^{n-1}$ un vecteur unitaire de $\mathbb{R}^n$. La transformée de Radon de f, $\mathcal{R}f : \mathbb{S}_1^{n-1} \times \mathbb{R} \mapsto \mathbb{R}$, est l'intégrale de f sur l'hyperplan perpendiculaire à $\vec{\xi}$ de distance signée à l'origine u et s'écrit :

$$\mathcal{R}f(\vec{\xi}, u) = \int_{\mathbb{R}^n} f(x)\delta(u - \vec{\xi} \cdot x)dx_1 \cdots dx_n. \tag{2.4}$$

Dans la suite, nous ne considérerons que le cas $n = 2$ dans lequel les hyperplans sont des droites et la transformée de Radon de f est l'intégrale de f le long d'une ligne. Dans $\mathbb{R}^2$, une ligne est caractérisée par deux paramètres, l'angle qu'elle fait avec l'axe des ordonnées, $\theta \in [0, 2\pi[$, et sa distance signée à l'origine, $u \in \mathbb{R}$:

$$\mathcal{L}(\theta, u) = \left\{\vec{v} \in \mathbb{R}^2 \text{ tels que } \vec{v} \cdot \vec{\alpha_\theta} = u\right\}$$

avec $\vec{\alpha_\theta} = (\cos\theta, \sin\theta)$, le vecteur unitaire qui fait un angle θ avec l'axe des abscisses. Nous définissons également le vecteur $\vec{\beta_\theta} = (-\sin\theta, \cos\theta)$, perpendiculaire à $\vec{\alpha_\theta}$. Les notations sont illustrées dans la figure 2.1. La transformée de Radon bidimensionnelle est alors :

$$\mathcal{R}f(\theta, u) = \int_{\mathcal{L}(\theta,u)} f(\vec{v})d\vec{v} = \int_{\mathbb{R}} f(u\vec{\alpha_\theta} + t\vec{\beta_\theta})dt. \tag{2.5}$$

On retrouve ainsi la forme de l'équation (2.2) avec la fonction f qui représente les coefficients d'atténuation linéaire μ. Le détecteur mesure des intégrales selon des lignes de f. L'ensemble de ces intégrales le long du détecteur pour un angle donné - en faisant varier u - est appelé *projection*. Dans une acquisition réelle, on échantillonne à la fois les angles de projections et les projections. Une projection contient donc autant de mesures intégrales qu'il y a de pixels sur le détecteur. La figure 2.2 représente la projection pour un angle d'une coupe de mâchoire. On remarque bien les "vagues" au centre du graphique qui correspondent aux dents qui absorbent plus les rayons X que le reste de la mâchoire.

Un scanner effectue des acquisitions sur plusieurs angles tout autour du patient de manière à récupérer de nombreuses projections et permettre de reconstruire une image de l'intérieur du

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

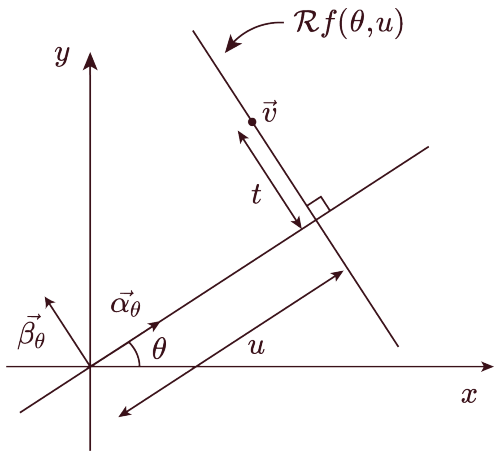


FIGURE 2.1 – Transformée de Radon : intégrale de f sur la droite perpendiculaire à $\vec{\alpha}_\theta$ et qui passe par le point de coordonnées $(u\cos\theta, u\sin\theta)$

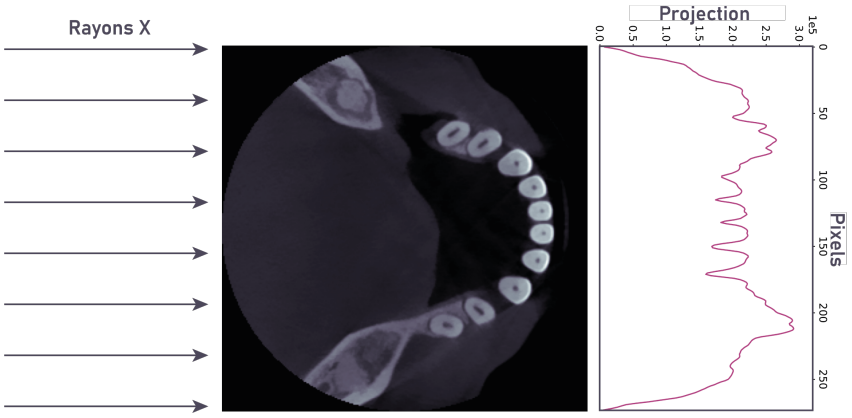


FIGURE 2.2 – Projection d'angle 0° avec rayons parallèles d'une mâchoire.

patient. L'ensemble de ces projections peut être visualisé dans une image appelée *sinogramme* : les projections sont ajoutées les unes à la suite des autres dans une image, chaque colonne correspond à une projection à un angle donné et sur chaque ligne on retrouve les valeurs mesurées par un pixel aux différents angles. Un sinogramme est donc une image de taille *nombre de pixels du détecteur* × *nombre d'angles*. La figure 2.3 donne un exemple de sinogramme. Le nom *sinogramme* vient du fait qu'un point de l'objet laisse une empreinte sinusoïdale sur le sinogramme puisque le détecteur tourne autour de celui-ci. Ainsi, les données acquises par le scanner correspondent à un échantillonnage de la transformée de Radon.

En dimension 3, la transformée de Radon intègre la fonction f sur des plans et non plus sur des lignes. Ceci ne correspond pas au phénomène physique que nous voulons modéliser, d'où la nécessité d'introduire une autre transformée : la transformée rayons X.

Transformée rayons X

En géométrie divergente, c'est-à-dire lorsque les rayons X ne sont pas émis avec des rayons parallèles, comme c'est le cas en géométrie en éventail ou en cône, la transformée de Radon ne s'applique plus. On utilise la transformée rayons X qui effectue les intégrales sur des demi-

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

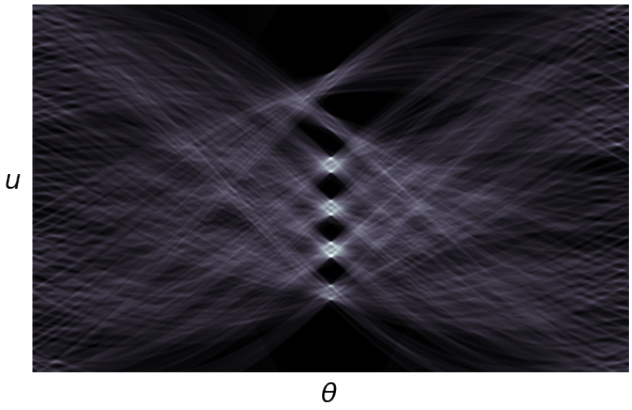


FIGURE 2.3 – Exemple d'un sinogramme.

droites plutôt que sur des hyperplans.

En dimension n, la transformée rayons X calcule l'intégrale de f sur les demi-droites issues de la source qui traversent f. Soit $\vec{\xi}$ un vecteur unitaire de $\mathbb{R}^n$, et $\vec{\xi}^{\perp}$ l'hyperplan passant par l'origine, orthogonal à $\vec{\xi}$. On paramétrise la droite sur laquelle on intègre avec son vecteur directeur $\vec{\xi}$ et son point d'intersection avec l'hyperplan $\vec{\xi}^{\perp}$, a. La transformée rayons X de f est donnée par l'intégrale sur $\mathbb{R}^+$ suivante :

$$Xf\left(\vec{\xi}, a\right) = \int_{\mathbb{R}^+} f\left(a + \lambda\vec{\xi}\right) d\lambda,\ a \in \vec{\xi}^{\perp} \tag{2.6}$$

Ainsi, en dimension 2 et en géométrie parallèle, la transformée de Radon et la transformée rayons X coïncident, supposant l'origine en dehors de l'objet.

La figure 2.4a illustre la transformée rayons X dans le cas de la géométrie en éventail. Comme dans le cas général, $\vec{\xi}$ est le vecteur directeur de la droite. Il faut que $\vec{\xi}$ soit de sorte à ce que la droite soit dans le faisceau délimité par l'*angle d'éventail* φ. L'hyperplan $\vec{\xi}^{\perp}$ n'est que la droite perpendiculaire à $\vec{\xi}$ qui passe par l'origine, confondue avec a ici. La géométrie conique est représentée dans la figure 2.4b. On définit ici ψ, l'*angle de cône* qui représente la "hauteur" du cône. Ici, $\vec{\xi}^{\perp}$ n'est pas représenté mais c'est le plan perpendiculaire à $\vec{\xi}$ qui passe par l'origine.

2 Formulation du problème inverse

Le problème direct en tomographie consiste à estimer les projections p à partir du volume f. Connaissant f, il suffit alors d'appliquer l'opérateur de Radon ou de transformée rayons X pour obtenir p. Ce problème direct s'écrit :

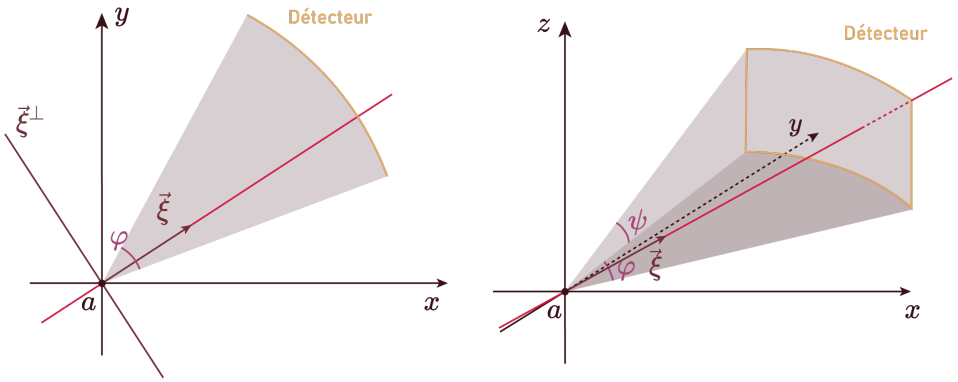
$$Af = p \tag{2.7}$$

avec A l'opérateur linéaire de la transformée de Radon ou de la transformée rayons X.

Le problème inverse, quant à lui, vise à retrouver le volume à partir des projections. Ce problème inverse satisfait les conditions d'Hadamard si :

- il y a existence d'une solution
- il y a unicité de la solution
- la solution dépend continument des données

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés



(a) (b)

FIGURE 2.4 – Illustration de la transformée rayons X en géométrie (a) en éventail et (b) conique pour une droite passant par l'origine. L'intégration se fait sur la droite représentée en rouge.

Lorsqu'au moins l'une de ces trois conditions n'est pas vérifiée, on dit que le problème est mal-posé.

En réalité, la formulation de ce problème n'est pas continue, on a un nombre fini de projection, un nombre fini de pixels sur le détecteur ou encore un nombre fini de voxels dans le volume. Il faut donc discrétiser ce problème à un moment. On peut utiliser des formules intégrales basées sur la transformée de Radon et de Fourier, puis discrétiser pour obtenir la reconstruction ; ou, discrétiser dès le départ en considérant A comme une matrice, f et p comme des vecteurs et utiliser des techniques d'algèbre linéaire pour inverser le problème. Nous présentons, dans ce qui suit, des algorithmes pour ces deux cas de figure.

En pratique, de nombreux paramètres physiques font que la reconstruction tomographique est un problème mal-posé : nombre de mesures insuffisant, mesures corrompues par du bruit, etc. Nous allons donc présenter dans la suite des algorithmes qui permettent d'effectuer cette reconstruction, en commençant par le cas de données idéales, puis en prenant en compte les imperfections que peuvent contenir les mesures.

Avant toute reconstruction, les images acquises par le scanner sont d'abord transformées pour que leurs valeurs soient exprimées en unité Hounsfield, définie par :

$$\mathrm{HU}(\mu) = 1000 \times \frac{\mu - \mu_{\mathrm{eau}}}{\mu_{\mathrm{eau}} - \mu_{\mathrm{air}}}.$$

II Reconstructions analytiques

Dans cette section, nous présentons les algorithmes analytiques utilisés en tomographie. Nous détaillerons le cas de la rétro-projection filtrée qui effectue la reconstruction dans le cadre de la géométrie parallèle 2D, puis nous parlerons de sa généralisation en CBCT.

Les méthodes analytiques s'appuient sur une description continue du problème : f et les opérateurs sont des fonctions. La discrétisation du problème s'effectue dans un second temps pour s'adapter à l'échantillonnage des données.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

1 Rétro-projection en géométrie 2D parallèle

Reprenons la formule (2.5) de la transformée de Radon :

$$\mathcal{R}f(\theta,u) = \int_{\mathcal{L}(\theta,u)} f(\vec{v})d\vec{v} = \int_{\mathbb{R}} f\left(u\vec{\alpha_\theta} + t\vec{\beta_\theta}\right) dt.$$

Il faut retrouver $f : \mathbb{R}^2 \mapsto \mathbb{R}^+$ à partir des valeurs de sa transformée de Radon.

Rétro-projection simple

La rétro-projection simple est l'opérateur adjoint de la transformée de Radon, elle consiste à attribuer la valeur de $Rf(\theta,u)$ à tout point sur la ligne définie par $u\vec{\alpha_\theta} + t\vec{\beta_\theta}$, puis à faire la somme sur toutes les contributions issues des différentes projections. En notant $\mathcal{R}^*$ l'opérateur de rétro-projection on a :

$$\mathcal{R}^* f(x,y) = \int_0^\pi \mathcal{R}f\ (\theta, x\cos\theta + y\sin\theta)\ d\theta \tag{2.8}$$

La rétro-projection simple ne correspond pas à l'inversion de la transformée de Radon. En pratique, l'image obtenue est floue (voir figure 2.6).

Rétro-projection filtrée

En utilisant le théorème de coupe-projection, nous pouvons montrer que la transformée de Radon est inversible et que le problème (2.7) admet une unique solution.

> **Théorème coupe-projection** : La transformée de Fourier d'une projection est égale à une ligne de la transformée de Fourier de l'image, ligne qui passe par l'origine et fait un angle θ avec l'axe des abscisses. Mathématiquement :
>
> $$\widehat{\mathcal{R}f}\ (\theta,\nu) = \widehat{f}\ (\nu\cos\theta, \nu\sin\theta)\ .$$

En effet, la transformée de Fourier de la projection $\mathcal{R}f(\theta,\cdot)$ est :

$$\begin{aligned}\widehat{\mathcal{R}f}\ (\theta,\nu) &= \int_{\mathbb{R}} \mathcal{R}f\ (\theta,u)\, \mathrm{e}^{-2i\pi\nu u} du \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f\left(u\vec{\alpha} + t\vec{\beta}\right) dt \right) \mathrm{e}^{-2i\pi\nu u} du\end{aligned}$$

En notant (x,y) les coordonnées du vecteur $u\vec{\alpha} + t\vec{\beta}$ dans le repère $x0y$, on a

$$\begin{aligned}u &= \vec{\alpha}\cdot\left(u\vec{\alpha} + t\vec{\beta}\right) \\ &= \vec{\alpha}\cdot(x,y) \\ &= (\cos\theta,\sin\theta)\cdot(x,y) \\ &= x\cos\theta + y\sin\theta\end{aligned}$$

Ainsi, $\nu u = \nu\,(x\cos\theta + y\sin\theta) = x\times\nu\cos\theta + y\times\nu\sin\theta = x\nu_x + y\nu_y$. On effectue alors un changement de variable correspondant à une rotation d'angle θ :

$$\widehat{\mathcal{R}f}\ (\theta,\nu) = \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\mathrm{e}^{-2i\pi\left(x\nu_x + y\nu_y\right)} dxdy$$

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

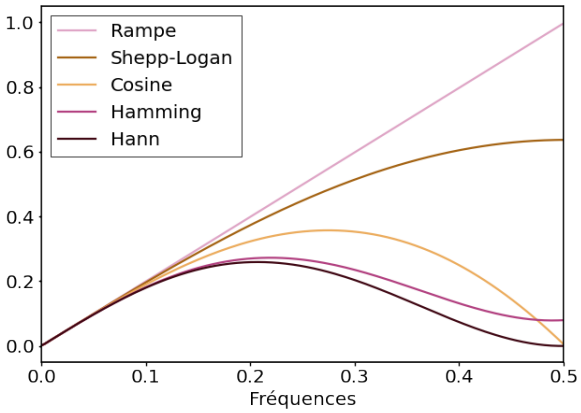


FIGURE 2.5 – Différents filtres utilisés pour la rétro-projection filtrée.

ce qui correspond à la transformée de Fourier 2D de f.

Ce théorème nous dit que si l'on dispose de toutes les projections, on peut retrouver la transformée de Fourier de f. On peut ensuite déterminer f en utilisant la transformée de Fourier inverse :

$$\begin{aligned} f(x,y) &= \int_{\mathbb{R}}\int_{\mathbb{R}} \widehat{f}(\nu_x,\nu_y) \mathrm{e}^{2i\pi(x\nu_x+y\nu_y)} d\nu_x d\nu_y \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \widehat{\mathcal{R}f}\,(\theta,\nu)\, \mathrm{e}^{2i\pi(x\nu_x+y\nu_y)} d\nu d\theta \\ &= \int_{0}^{\pi}\int_{\mathbb{R}} \widehat{\mathcal{R}f}\,(\theta,\nu)\, \mathrm{e}^{2i\pi\nu(x\cos\theta+y\sin\theta)} |\nu| d\nu d\theta \end{aligned}$$

en effectuant un changement de variables associant les coordonnées catésiennes aux coordonnées polaires : $g(\nu_x,\nu_y) = (\theta,u)$ avec $\theta \in [0,\pi[$ et $\nu \in \mathbb{R}$. Pour un angle $\varphi \in [\pi, 2\pi[$, on prendra $\theta = \varphi - \pi$ et pour ν, l'opposé du rayon.

Cette dernière ligne est la formule de la rétro-projection filtrée. On peut la voir comme une convolution de $\mathcal{R}f$ avec le filtre $|\nu|$.

Le terme $|\nu|$ qui correspond au déterminant de la jacobienne du changement de variable g est un filtre passe-haut appelé *filtre rampe* ou filtre de Ram-Lak. Ce filtre met la moyenne du signal à 0, ce qui va introduire des valeurs négatives qui atténueront les artefacts en étoile et gommer l'information qui a été rétro-projetée en dehors du volume par la rétro-projection simple. Ce filtre permet aussi d'augmenter la résolution spatiale mais il amplifie les haute-fréquences, et donc le bruit et cela perturbe la qualité de la reconstruction. Pour corriger cela, on peut utiliser d'autres filtres qui atténuent ces hautes fréquences. On retrouve dans la figure 2.5 différents filtres que l'on peut comparer avec le filtre rampe.

Ainsi, la rétro-projection filtrée consiste en deux étapes illustrées par la double intégrale : on filtre les projections avec le filtre rampe, puis on effectue la rétro-projection de ces projections filtrées.

Afin d'illustrer leur différence, une coupe de mâchoire reconstruite avec une rétro-projection simple et une rétro-projection filtrée est montrée dans la figure 2.6.

2 En géométrie divergente 2D

En géométrie divergente, plusieurs techniques sont utilisées pour reconstruire le volume à partir des projections.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

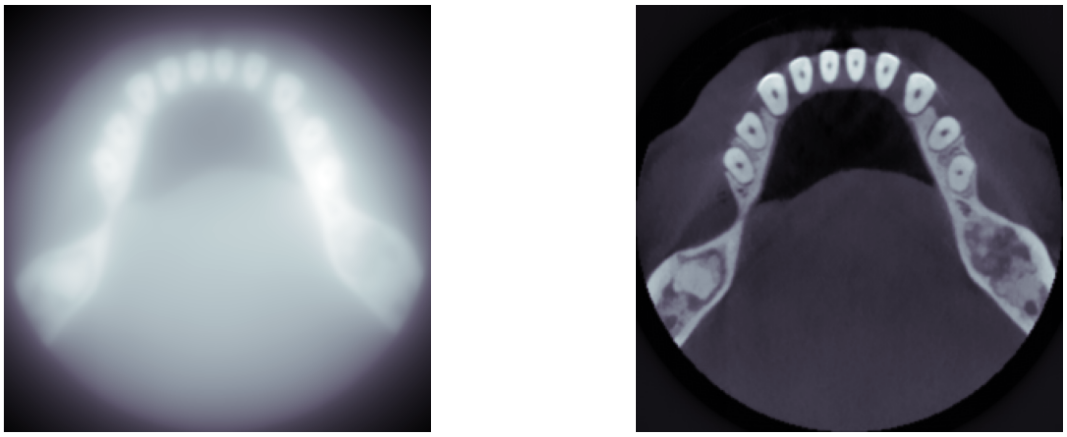


FIGURE 2.6 – Rétro-projection d'une coupe d'une mâchoire, 180 projections allant de 0 à 180°. A gauche : rétro-projection simple, à droite : rétro-projection filtrée.

Dans un premier cas, lorsque la source est suffisamment éloignée de l'objet et du détecteur, on peut considérer que les données correspondent à des données acquises en géométrie parallèle. On utilise alors les formules données précédemment. Cette méthode fonctionne correctement jusqu'à un angle d'éventail de 15° (Bonnet; 2000). L'angle d'éventail a été défini dans la figure 2.4a.

La seconde méthode consiste à réarranger les projections divergentes en un ensemble de projections parallèles en effectuant une interpolation. On peut retrouver les formules de rétro-projection filtrée dans le cas de faisceau en éventail avec détecteur plat ou incurvé dans (Kak et Slaney; 2001).

Enfin, on peut calculer une formule d'inversion directe qui soit adaptée aux données divergentes. On trouve des exemples de telles formules dans (Lakshminarayanan; 1975; Horn; 1979; Webb; 1982).

La reconstruction en géométrie divergente nécessite une couverture angulaire plus large qu'en géométrie parallèle où 180° suffisent. Avec un angle de faisceau α, il faut couvrir l'intervalle $[0, 180 + \alpha]$ pour obtenir toutes les données nécessaires à la reconstruction. Cependant, il y a également des données redondantes qui apparaissent en augmentant la taille de l'intervalle, il faut alors pondérer leur contribution lors de la rétro-projection. La figure 2.7 explique la différence entre la géométrie parallèle et en éventail par rapport à la couverture angulaire.

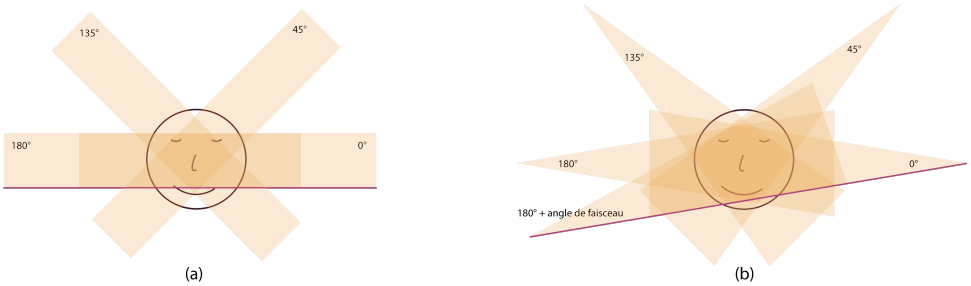


FIGURE 2.7 – Couverture angulaire nécessaire en géométrie (a) parallèle et (b) en éventail. Le rayon violet en bas de la projection à 0° est symétrique à celui de la projection à 180° en parallèle. En éventail, la projection à 180° ne donne pas les mêmes rayons, alors que celle à 180° + angle de cône a le même premier rayon violet de celle à 0°.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

3 En CBCT

Le problème de reconstruction consiste à retrouver une fonction de $\mathbb{R}^3$ à partir de projections bidimensionnelles. L'algorithme de Feldkamp, Davis et Kress (Feldkamp *et al.*; 1984), noté dans la suite FDK, permet d'obtenir une approximation de la fonction f dans le cas d'une trajectoire circulaire du système d'acquisition.

Soient (x_d, y_d) les coordonnées d'un pixel sur le détecteur, on note $P(x_d, y_d)$ la valeur mesurée par le détecteur pour ce pixel. On note θ, comme précédemment, l'angle qui détermine la position de la source et du détecteur dans le plan de rotation.

La formule d'inversion de l'algorithme de FDK est donnée par :

$$P'(x_d, y_d) = \frac{D_{SD}}{\sqrt{D_{SD}^2 + x_d^2 + y_d^2}} P(x_d, y_d) \tag{2.9}$$

$$f(x, y, z) = \frac{1}{2} \int_{\theta \in \mathbb{S}^1} \frac{D_{SO}^2}{(D_{SO} - u)^2} \int_{\mathbb{R}} P'\left(\tau, \frac{zD_{SD}}{D_{SO} - u}\right) h\left(\frac{tD_{SD}}{D_{SO} - u} - \tau\right) d\tau d\theta \tag{2.10}$$

avec D_{SD} et D_{SO} respectivement la distance source-détecteur et source-origine, et (t, u) les coordonnées du point (x, y) dans le repère $(t0u)$ qui est une rotation d'angle θ du repère $(x0y)$. Les notations sont montrées dans la figure 2.8. h correspond à la réponse impulsionnelle du filtre rampe. Ainsi, pour chaque projection trois étapes sont effectuées :

1. on pondère la valeur du pixel pour prendre en compte la divergence, c'est l'équation (2.9),
2. on filtre à l'aide du filtre rampe h,
3. et on fait une rétro-projection pondérée 3D.

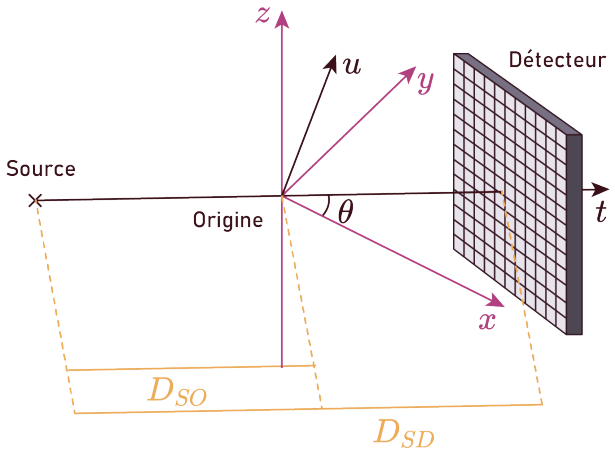


FIGURE 2.8 – Introduction des notations utilisées dans la formule de FDK.

L'algorithme de FDK s'applique sur une trajectoire d'acquisition circulaire, pour laquelle la transformée de Radon n'est complète que dans le plan de la source (le plan horizontal qui contient la source). Ainsi, la reconstruction FDK n'est exacte que dans ce plan et correspond dans ce cas à la reconstruction en géométrie divergente 2D. En dehors de ce plan, la qualité de la reconstruction se dégrade avec l'augmentation de la distance avec le plan de source. Cependant, il est important de noter que ce n'est pas tant l'algorithme FDK en lui-même, mais la trajectoire d'acquisition circulaire qui fait que la reconstruction ne peut être exacte partout.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

III Reconstructions itératives

Contrairement aux méthodes analytiques qui s'appuient une description continue du problème et qui reconstruisent le volume en une étape, les algorithmes de reconstruction itérative utilisent dès le départ une forme discrétisée des objets et des opérateurs. Le problème inverse est exprimé matriciellement, et la reconstruction s'appuie sur les méthode itératives de l'algèbre linéaire. Dans le cadre d'acquisitions non idéales (présence de bruit, couverture angulaire inférieure à 180°, nombre de projections faible, etc.), ces méthodes sont plus robustes que les méthodes analytiques.

Ces méthodes itératives consistent à estimer le volume de manière itérative, en minimisant une fonctionnelle. Cette fonctionnelle permet de prendre en compte la physique des acquisitions, la distribution du bruit, et d'ajouter un a priori sur la solution recherchée. Nous présentons dans cette partie les premiers algorithmes itératifs présents dans l'état de l'art de la reconstruction tomographique. Nous les classons en deux catégories : les algorithmes algébriques qui sont basés sur une interprétation géométrique du problème, et les algorithmes bayésiens qui étaient initialement utilisés en tomographie à émission de positons (TEP).

1 Méthodes algébriques

Pour les méthodes algébriques, la reconstruction tomographique consiste à déterminer les coefficients d'absorption en résolvant un système d'équations linéaires.

En 2D, on définit un objet f discrétisé en $N = N_x \times N_y$ voxels carrés. On note M_θ le nombre de projections et M_d le nombre de pixels sur le détecteur. Pour le pixel du détecteur i, la transformée de Radon se discrétise selon la formule suivante :

$$p_i = \sum_{j \in J} a_{ij} f_j \tag{2.11}$$

avec p_i la valeur mesurée par le pixel i, a_{ij} la longueur d'un segment, obtenu comme intersection de la droite $\Delta_{i,\theta}$ avec le voxel j, $\Delta_{i,\theta}$ étant la droite qui relie la source au centre du pixel i sur le détecteur, et f_j est la valeur du coefficient d'atténuation du voxel j.

Nous pouvons réécrire cette équation matriciellement :

$$p = Af \tag{2.12}$$

avec $p \in \mathbb{R}^M$ le vecteur des projections, $M = M_\theta \times M_d$ étant le nombre de mesures. La matrice $A = (a_{ij}) \in \mathcal{M}_{M \times N}(\mathbb{R})$ contient les contributions de chaque voxel dans les mesures des pixels. La valeur a_{ij} est la contribution du voxel j dans la mesure du pixel détecteur i.

La matrice A n'est pas carrée lorsque le nombre de voxels est différent du nombre de mesures. Cependant, même lorsqu'elle est carrée, elle est de très grande taille, creuse et non nécessairement inversible (en particulier en présence de bruit ou de projections manquantes). Ainsi le problème ne peut être résolu en calculant l'inverse de A pour obtenir $f = A^{-1}p$. De plus, lorsque A n'est pas carrée, le système n'a pas une unique solution.

Solution au sens des moindres carrés Une solution à ce problème serait de considérer le problème de minimisation quadratique suivant :

$$\arg\min_f \|Af - p\|_2^2 \tag{2.13}$$

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

On peut alors calculer la pseudo-inverse de Moore-Penrose (Moore; 1920) de A : $A^{\dagger} = (A\,{}^tA)^{-1}\,{}^tA$. La solution du problème équation (2.13) est alors donnée par :

$$\hat{f}_{\mathrm{MC}} = A^{\dagger}p = \left(A\,{}^tA\right)^{-1}\,{}^tA \tag{2.14}$$

Cependant, lorsque le problème est mal posé, cette solution est particulièrement sensible au bruit et dépend de l'inversibilité de $A\,{}^tA$. En effet, si la matrice $(A\,{}^tA)$ a des valeurs propres proches de 0, son inversion amplifiera les hautes fréquences et donc le bruit.

Approche itérative Les méthodes algébriques calculent itérativement une solution approchée de f. Elles utilisent la méthode de Kaczmarz (Kaczmarz; 1937; Natterer; 2001), à l'étape $k+1$, la valeur du voxel j est calculée de la manière suivante :

$$f_j^{(k+1)} = f_j^{(k)} + \frac{p_i - \sum_{l=1}^{N} f_l^{(k)} a_{il}}{\sqrt{\sum_{l=1}^{N} a_{il}^2}} a_{ij} \tag{2.15}$$

où $i = k \bmod (m) + 1$. L'équation (2.15) correspond à la mise à jour du voxel j à partir de la projection i, il faut effectuer cette mise à jour par rapport à toutes les projections avant de passer à l'itération $k+1$. Le terme $\sum_{l=1}^{N} f_l^{(k)} a_{il}$ correspond à la projection du volume à l'étape k. Les méthodes algébriques peuvent donc être résumées ainsi : on part d'une image d'initialisation (0, 1 ou une simple rétro-projection) et on la met à jour à chaque étape de l'itération, jusqu'à convergence :

- Projection de la reconstruction en cours,
- Différence entre projections de la reconstruction et projections acquises,
- Rétro-projection de la différence et mise à jour de la reconstruction.

L'équation (2.15) vient de l'écriture sous forme de système de l'équation (2.12) :

$$\begin{cases} a_1 \cdot f = a_{11}f_1 + a_{12}f_2 + \cdots + a_{1N}f_N = p_1 \\ a_2 \cdot f = a_{21}f_1 + a_{22}f_2 + \cdots + a_{2N}f_N = p_2 \\ \qquad \vdots \\ a_M \cdot f = a_{M1}f_1 + a_{M2}f_2 + \cdots + a_{MN}f_N = p_M \end{cases} \tag{2.16}$$

où a_i est le vecteur $(a_{i1}, a_{i2}, \cdots, a_{iN})$ et $a_i \cdot f$ le produit scalaire entre a_i et f.

Chaque équation de ce système est l'équation d'un hyperplan dans $\mathbb{R}^N$. La méthode de Kaczmarz suppose que si la solution existe, elle se trouve à l'intersection de ces hyperplans. La projection orthogonale successive de la solution courante sur chacun des hyperplans permet ainsi de se rapprocher de manière itérative de la solution recherchée. La projection orthogonale sur l'hyperplan i s'écrit :

$$\Pi_i(f) = f - \frac{a_i \cdot f - p_i}{\|a_i\|^2} a_i \tag{2.17}$$

avec a_i le vecteur d'une ligne de la matrice A. On retrouve l'équation (2.15).

La figure 2.9 donne une représentation graphique de l'algorithme dans le cas où on a seulement deux hyperplans qui sont des droites.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

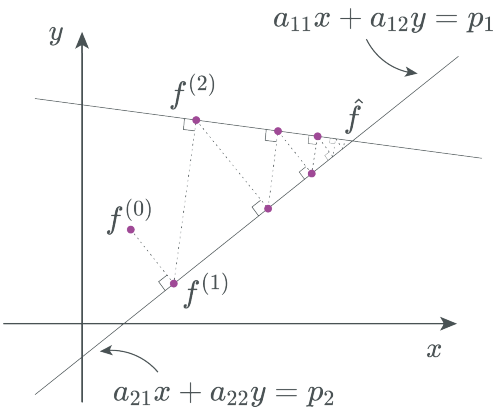


FIGURE 2.9 – Représentation géométrique de l'algorithme ART avec $M = N = 2$.

ART

La première méthode de reconstruction itérative, ART (*Algebraic Reconstruction Techniques)* (Gordon *et al.*; 1970) est une application directe de la méthode de Kaczmarz avec l'ajout de la contrainte de positivité des coefficients d'atténuation. On obtient l'étape de mise à jour suivante :

$$f_j^{(k+1)} = \Pi_i \left(f^{(k)} \right) \tag{2.18}$$

Gordon *et al.* (1970) ont également proposé une version multiplicative de leur algorithme : MART (*Multiplicative Algebraic Reconstruction Technique*) dont le terme de mise à jour est donné par :

$$f_j^{(k+1)} = \frac{p_i}{\sum_{l=1}^{N} f_l^{(k)} a_{il}} f_j^{(k)}. \tag{2.19}$$

La vitesse de convergence de ces algorithmes dépend de l'angle entre les hyperplans : plus l'angle se rapproche de 90° et plus la convergence est rapide. De plus, lorsque M est inférieur à N, l'intersection des hyperplans ne donne pas un unique point, ou lorsque le système est incompatible, la solution peut ne pas exister. Dans ce cas, ART calcule plutôt $\arg\min_f \|Af - p\|_2^2$.

Enfin, cette méthode est particulièrement sensible au bruit et ses reconstructions comportent beaucoup d'erreurs (Gilbert; 1972; Andersen et Kak; 1984). Une version avec relaxation proposée par Gordon (1974) permet de réduire le bruit dans les reconstructions, mais au détriment de la vitesse de convergence.

SIRT

L'algorithme SIRT (*Simultaneous Iterative Reconstruction Technique*) proposé dans (Gilbert; 1972) est une modification de ART. Son nom provient du fait qu'il utilise simultanément toutes les projections pour mettre à jour un pixel, contrairement à ART qui n'utilise les projections qu'une par une. Cette différence rend l'algorithme de SIRT beaucoup plus stable qu'ART, même en présence d'erreurs de mesure dans les projections.

La formule de reconstruction est donnée par une itération de Landweber (Landweber; 1951) :

$$f_j^{(k+1)} = f_j^{(k)} - \lambda \frac{\sum_{i=1}^{M} \frac{\sum_{n=1}^{N} a_{in} f_n^{(k)} - p_i}{\sum_{n=1}^{N} a_{in}} \cdot a_{ij}}{\sum_{i=1}^{M} a_{ij}}$$

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

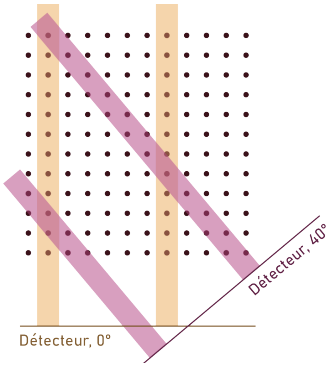


FIGURE 2.10 – Illustration de l'importance d'inclure la sensibilité dans la reconstruction. Les rayons jaunes qui arrivent sur le détecteur à 0° ont rencontré 12 points de la grille, alors que les mêmes rayons roses à 40° en rencontrent 6 pour le premier et environ 16 pour le second.

avec λ le paramètre SIRT, un paramètre de relaxation permettant de contrôler la vitesse de convergence et la stabilité de l'algorithme. Matriciellement, cette équation devient :

$$f^{(k+1)} = SIRT\left(f^{(k)}\right) = f^{(k)} - \lambda \frac{1}{{}^tA\mathbf{1}} {}^tA\left[\frac{Af^{(k)} - p}{A\mathbf{1}}\right] \tag{2.20}$$

où $\mathbf{1}$ désigne le vecteur composé de 1, ${}^tA$ est la transposée de la matrice A, et les opérations vectorielles sont effectuées termes à termes.

La reconstruction avec SIRT converge moins rapidement qu'avec ART au début des itérations, mais elle est plus robuste. De plus, SIRT prend en compte la sensibilité du système, en introduisant le terme ${}^tA\mathbf{1}$. En effet, selon l'angle de la projection considéré, le nombre de voxels du volume traversés par un rayon n'est pas le même (voir figure 2.10). La normalisation effectuée lors de la division par ${}^tA\mathbf{1}$ permet d'accorder la juste importance relative pour chaque rayon.

Malgré sa ressemblance avec ART, SIRT peut également être considéré comme un algorithme bayésien qui minimiserait la fonctionnelle $\|Af - p\|_2^2$. Nous parlerons de ces algorithmes dans la section 2. D'un point de vue statistique, minimiser cette fonctionnelle est équivalent à maximiser la log-vraisemblance si l'on considère que les données suivent une distribution gaussienne, i.e quand $p = Af + \varepsilon$ avec $\varepsilon = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ et ε_i indépendants et identiquement distribués selon une loi normale $\mathcal{N}(0, 1)$.

Les méthodes algébriques résolvent le problème de reconstruction tomographique en cherchant la solution d'une équation linéaire. Elles peuvent reconstruire correctement le volume même avec un nombre limité de projections ou lorsque les projections contiennent des incohérences. Elles permettent ainsi de pallier certains défauts des méthodes analytiques. Cependant, elles ne prennent pas en compte la statistique des données contrairement aux méthodes suivantes.

2 Méthodes statistiques

En tomographie par émission (TEP), le rayonnement n'est pas envoyés par une source extérieure, mais par un traceur radioactif injecté dans le patient. On compte alors, dans ce cas, les photons provenant de l'intérieur du patient. Les méthodes statistiques sont utilisées pour prendre en compte la nature stochastique du comptage des photons par le détecteur. Elles sont adaptées à des cas de faibles statistiques photoniques, et sont donc indiquées pour la reconstruction en TEP, mais également dans le cas de la tomographie par transmission à faible dose.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

MLEM

L'algorithme MLEM (*Maximum Likelihood Expectation Maximization*), proposé par Shepp et Vardi (1982) et Lange *et al.* (1984), est très utilisé en reconstruction TEP. Cet algorithme maximise la log-vraisemblance de données que l'on suppose suivre une distribution de Poisson à l'aide de l'algorithme EM (*Expectation Maximization*) (Dempster *et al.*; 1977).

Dans le cas de la reconstruction tomographique, les données disponibles sont les projections p et on considère qu'elles suivent une distribution de Poisson de moyenne Af. On cherche à estimer le volume, représenté par le vecteur f. Ceci revient à maximiser la vraisemblance d'obtenir p sachant f donnée par :

$$\pi\left(p|f\right) = \prod_i \left(\sum_j a_{ij} f_j\right)^{p_i} \frac{e^{-\sum_j a_{ij} f_j}}{p_i!} \tag{2.21}$$

et par croissance du logarithme, cela revient à maximiser la log-vraisemblance :

$$L(f) = \log\left(\pi\left(p|f\right)\right) = \sum_i \left(p_i \log\left(\sum_j a_{ij} f_j\right) - \sum_j a_{ij} f_j - \log\left(p_i!\right)\right). \tag{2.22}$$

On reconnaît une forme approchée de la divergence de Kullback-Leibler dans les termes de droite dépendant de f. Titterington (1987) ont démontré que la minimisation de la divergence de Kullback-Leibler équivaut à la maximisation de la fonction de vraisemblance correspondant à des données suivant une distribution de Poisson.

Puisque l'on ne connaît pas f, l'idée de l'algorithme EM est d'introduire une variable latente Z qui correspond à la contribution du voxel j du volume dans la valeur du pixel i du détecteur. On a alors $p_i = \sum_j z_{ij}$ et les z_{ij} suivent une loi de Poisson de moyenne $a_{ij} f_j$ pour tout i et j. f est ainsi devenu un paramètre de la variable aléatoire Z et on cherche donc la valeur de ce paramètre qui maximise la probabilité d'obtenir les données p.

L'algorithme EM maximise la vraisemblance Q d'une variable aléatoire dépendant d'un paramètre et de données non observables, en deux étapes. On cherche premièrement à estimer la log-vraisemblance pour un certain paramètre f, c'est l'étape E (*Expectation)* :

$$Q\left(f|f^{(k)}\right) = \mathbb{E}_Z\left(\log \pi\left(p, f\right)|f^{(k)}\right) \tag{E}$$

avec p les projections observées, $f^{(k)}$ les paramètres actuels, f le paramètre à trouver et π la densité de la distribution de probabilité des données. Cette étape compense la connaissance incomplète de la variable Z en estimant l'espérance de sa vraisemblance avec un paramètre f quelconque au début. Une fois Q estimée, on maximise sa valeur par rapport à f : étape M (*Maximization*). Soit,

$$f^{(k+1)} = \arg\max_f Q\left(f|f^{(k)}\right). \tag{M}$$

Ces deux étapes sont ensuite répétées jusqu'à la convergence. Les conditions de convergence sont généralement satisfaites en tomographie et la démonstration est donnée dans (Dempster *et al.*; 1977).

Après application de l'algorithme EM à la log-vraisemblance donnée dans l'équation (2.22), on obtient l'équation de mise à jour de MLEM suivante :

$$f_j^{(k+1)} = f_j^{(k)} \frac{1}{\sum_i a_{ij}} \sum_i \frac{a_{ij} p_i}{\sum_l a_{il} f_l^{(k)}} \tag{2.23}$$

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

où f_j est la j-ième composante du vecteur représentant le volume. On peut également l'écrire sous forme matricielle :

$$f^{(k+1)} = \frac{f^{(k)}}{{}^tA\mathbf{1}} \, {}^tA \left(\frac{p}{A f^{(k)}} \right). \tag{2.24}$$

L'algorithme MLEM étant un algorithme multiplicatif, il doit être initialisé par des valeurs strictement positives. Il assure ainsi la positivité de la reconstruction à chaque étape. Cependant on peut noter quelques inconvénients ayant poussé le développement de nouvelles méthodes :

- L'algorithme est très sensible au bruit. Au fur et à mesure des itérations, un effet "ciel étoilé" apparaît. MLEM n'est donc jamais utilisé jusqu'à sa convergence, les itérations devant être stoppées assez rapidement pour limiter cet effet. Un nombre d'itérations optimal doit être trouvé, afin que la qualité ne soit pas dégradée. La figure 2.11 présente une reconstruction MLEM avec projections bruitées et non bruitées pour différents nombres d'itérations. Dans le cas non bruité, la reconstruction à l'itération 500 est mieux résolue alors que dans le cas de projections bruitées, le bruit introduit par MLEM dégrade la qualité de l'image et c'est finalement à l'itération 100 que l'on voit le plus de détails.
- Le nombre d'itérations nécessaires pour obtenir une reconstruction correcte est assez élevé.
- MLEM ne converge pas à la même vitesse dans toutes les zones de l'image : lorsque les zones ont de faibles valeurs (zones froides), l'algorithme converge plus lentement. En effet, en notant $L(f)$ la log-vraisemblance négative de l'équation (2.24) devient :

$$f_j^{(k+1)} = f_j^{(k)} + \frac{f_j^{(k)}}{\sum_l a_{jl}} \frac{\partial L}{\partial f_j} \left(f_j^{(k)} \right) \tag{2.25}$$

  Donc le pas de la descente de gradient est proportionnel à l'activité du voxel f_j.
- La contrainte de positivité apporte également un biais positif dans l'image reconstruite, important dans les zones froides et particulièrement lorsque les projections sont bruitées et en faible nombre. Il est compensé par un biais négatif dans les zones chaudes proches des zones froides.

Variantes de MLEM

Proposé par Hudson et Larkin (1994), l'algorithme OSEM (*Ordered Subset Expectation Maximization*) est une accélération de MLEM. L'idée est d'appliquer successivement l'algorithme MLEM à de petits groupes de projections plutôt qu'à l'intégralité des projections. On subdivise les projections en m groupes, on applique l'étape MLEM de l'équation (2.24) à chaque groupe, puis on réitère sur l'ensemble des groupes jusqu'à la convergence.

Nous pouvons remarquer que le cas $m = M$ correspond au cas limite où il y a seulement une projection par groupe. On obtient alors l'équivalent de ART pour SIRT avec MLEM. Cet algorithme donne un intermédiaire entre les deux configurations : considérer l'intégralité des projections qui peut aboutir à des opérations assez lourdes, et une seule projection qui risque d'entraîner une convergence plus lente et une plus grande sensibilité au bruit.

OSEM ne converge pas, mais entre plutôt dans un cycle limite. Cependant, les itérations état de toute façon arrêtées prématurément dans MLEM, l'absence de convergence n'est pas un problème. De plus, il est toujours possible d'effectuer quelques itérations MLEM à la fin afin d'affiner la convergence si on le souhaite. Cependant, nous pouvons mentionner qu'Ahn et Fessler (2003) ont proposé deux versions de OSEM globalement convergentes.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

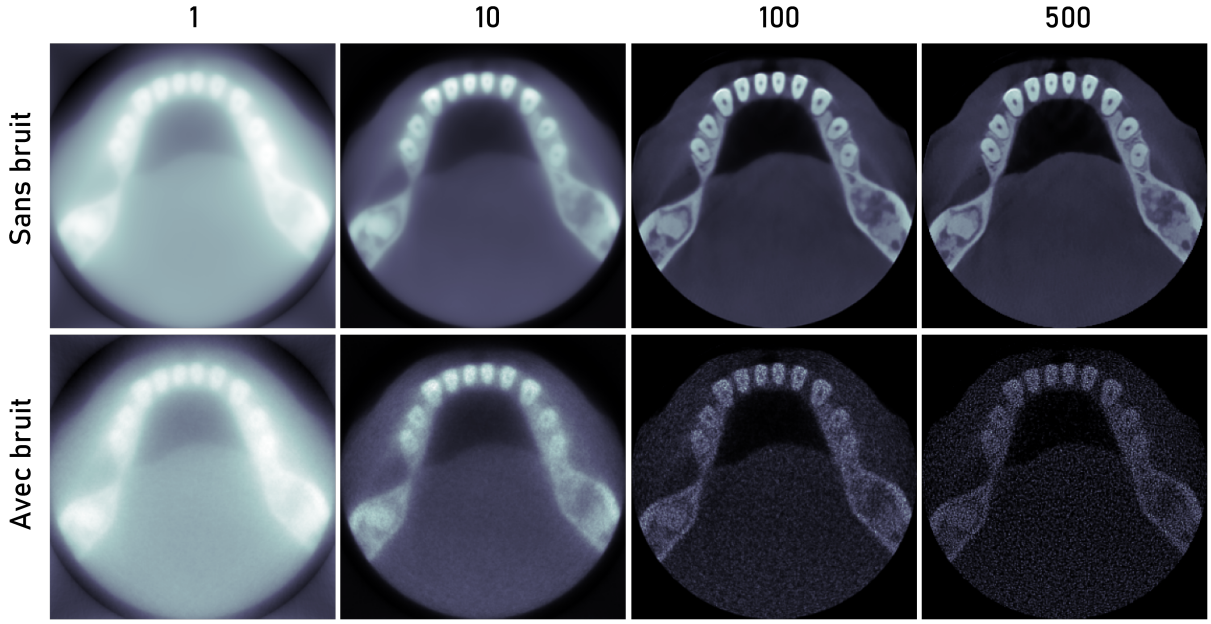


FIGURE 2.11 – Reconstructions MLEM avec différents nombres d'itérations : 1, 10, 100 et 500. Sur la ligne du haut, reconstructions à partir de projections non bruitées, et sur la ligne du bas, à partir de projections bruitées par du bruit poissonnien. Pour ces reconstructions, 180 projections parallèles régulièrement échantillonnées entre 0 et 180° ont été utilisées. Un bruit de Poisson a été ajouté aux projections.

D'autres adaptions de MLEM ont été proposées, parmi celles-ci nous pouvons citer ABEMML (Byrne; 1998) qui généralise MLEM en contraignant la solution à être à valeurs dans un intervalle $[\alpha, \beta]$ avec α pouvant être négatif. Cette valeur de α négative permet de diminuer le biais positif des reconstructions MLEM.

Méthodes statistiques en CT

En transmission, le modèle de bruit Poisson, qui découle de la loi de Beer-Lambert (2.2), est donné par la formule suivante (Anthoine *et al.*; 2012) :

$$p_i \sim \mathcal{P}\left(I_{0,i} \exp\left(-\left[Af\right]_i\right)\right) \tag{2.26}$$

avec $\mathcal{P}(\lambda)$ qui décrit une loi de Poisson de paramètre λ, et $I_{0,i}$ le nombre de photons envoyés depuis la source dans la direction du pixel i.

La log-vraisemblance négative obtenue dans ce cas là est :

$$L_{CT}(f) = -\sum_i p_i \log\left(I_{0,i} \exp\left(-\left[Af\right]_i\right)\right) - \left(I_{0,i} \exp\left(-\left[Af\right]_i\right)\right). \tag{2.27}$$

Ce qui revient à minimiser :

$$L_{CT}(f) = \sum_i p_i \left[Af\right]_i + I_{0,i} \exp\left(-\left[Af\right]_i\right). \tag{2.28}$$

Maximum de vraisemblance Lange *et al.* (1984) ont développé un algorithme MLEM pour ce problème. Cependant, l'étape de maximisation (M) conduit à une équation transcendante, difficile à résoudre. Nuyts *et al.* (1998) ont proposé une méthode qui maximise directement la

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

fonction de vraisemblance à l'aide d'une montée de gradient. L'étape de mise à jour du volume est donnée par :

$$f^{(k+1)} = f^{(k)} + \frac{\alpha}{N}\left(1 - \frac{{}^tAp}{{}^tA\left(I_0 \exp\left(-Af^{(k)}\right)\right)}\right), \tag{2.29}$$

avec α un paramètre de relaxation et N le nombre de voxels dans le volume.

Moindres carrés pondérés Le problème de reconstruction en tomographie par transmission peut être résolu avec la méthodes des moindres carrés pondérés (WLS [1]). Celle-ci calcule une solution approchée du problème (2.28) en cherchant une estimation du problème suivant :

$$\hat{f} = \arg\min_{f \geq 0} \frac{1}{2}\|Af - p\|_W^2 \tag{2.30}$$

où W est une matrice diagonale de poids représentant la variance des rayons, et $\|\cdot\|_W^2$ défini par $\|x\|_W^2 = {}^txWx$. La matrice W permet de prendre en compte l'intensité des rayons X et le bruit électronique (Kandarpa; 2022), cependant les images reconstruites restent sensibles au bruit et aux artefacts.

3 Accélération de Nesterov

Pour améliorer la convergence des descentes de gradient, Nesterov (1983) ont introduit une étape supplémentaire dans les algorithmes. Cette technique d'accélération permet d'augmenter le taux de convergence de méthodes de premier ordre pour les problèmes convexes. Les algorithmes MLEM et SIRT étant tous les deux des algorithmes de descente de gradient, peuvent bénéficier de cette accélération.

La reconstruction actuelle est remplacée par une combinaison linéaire de celle-ci avec la reconstruction de l'itération précédente, il se déroule de la manière suivante : on calcule une suite de réels $\left(t^{(n)}\right)$ telle que $t^{(0)} = 1$ et pour tout $k \in \mathbb{N}^*$:

$$t^{(k)} = \frac{1}{2}\left(1 + \sqrt{1 + 4\left(t^{(k-1)}\right)^2}\right)$$

et la k-ième itération du volume, pour SIRT ou MLEM, est mise à jour avec :

$$f^{(k)} \leftarrow f^{(k)} + \frac{t^{(k+1)} - 1}{t^{(k)}}\left(f^{(k)} - f^{(k-1)}\right) \tag{2.31}$$

L'algorithme FISTA (*Fast Iterative Shrinkage-Thresholding Algorithm)* (Beck et Teboulle; 2009) qui consiste à utiliser l'accélération de Nesterov avec un algorithme de gradient proximal est un exemple très connu d'application de cette accélération.

IV Artefacts de reconstruction

Les méthodes analytiques peuvent reconstruire correctement un volume en des temps très rapides : 1 à 2 secondes pour une coupe et quelques dizaines de secondes pour un volume 3D (avec parallélisation sur GPU). Leur formulation basée sur une description continue du problème fait qu'en théorie, avec un nombre infini de projections et beaucoup de temps, on peut reconstruire le volume. En pratique, on ne dispose que d'un nombre fini de projections et

1. *Weighted Least Squares*

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

même assez faible pour diminuer l'irradiation que subit le patient. Les méthodes itératives sont plus robustes face aux incohérences dans les projections, ou lorsque le nombre de projections est faible, mais elles restent sensibles à bien d'autres phénomènes physiques. La présence de bruit, de métal dans les projections, les imprévus physiques ou encore les erreurs de modélisation sont autant de sources d'artefacts dans les reconstructions. Ces erreurs de reconstruction peuvent conduire à de mauvais diagnostics médicaux et ne sont donc pas souhaitables. Nous présentons dans cette partie une liste non exhaustive d'artefacts de reconstruction, leurs causes et les techniques utilisées pour les réduire.

Durcissement de faisceau

Lorsque l'on a utilisé la formule de Beer-Lambert pour modéliser l'atténuation du faisceau à travers le patient, nous avons supposé que le faisceau était monochromatique. Cependant l'atténuation des rayons ne dépend pas uniquement du chemin parcouru, mais également des interactions rayons X - matière détaillées dans la partie 3 du chapitre 1, qui elles-mêmes dépendent de la longueur d'onde. Or, nous avons vu que le spectre créé par un scanner médical était polychromatique. L'approximation faite par l'équation (2.1) conduit à des artefacts de durcissement de faisceau [1]. Le spectre du faisceau de rayons X est modifié le long de son chemin à travers le patient : les rayons de faible énergie (rayons mous) sont plus facilement absorbés que les rayons de haute énergie (rayons durs), d'où le terme durcissement de faisceau. Pour prendre en compte les différentes longueurs d'onde, il faudrait intégrer sur l'énergie et remplacer cette équation par :

$$I = \int_0^{E_{\max}} I_0(E) e^{-\int_L \mu(x,E)dx} dE. \tag{2.32}$$

Cette équation conduit à une relation non linéaire entre les projections et les coefficients d'atténuation, incohérente avec la modélisation utilisée et laissant apparaître des zones d'ombre dans la reconstruction. Pour corriger ces artefacts, il existe plusieurs solutions (Hsieh; 2009) :

- Filtration du faisceau à la sortie du tube pour absorber les rayons de faible énergie (Barrett et Keat; 2004).
- Utilisation d'un fantôme d'eau pour calculer la correction à effectuer numériquement et application de la correction aux tissus mous qui ne diffèrent pas beaucoup de l'eau (Brooks et Di Chiro; 1976; Kijewski et Bjärngard; 1978). On utilise un matériau dont on connaît les propriétés pour faire la calibration, cependant, cette technique n'est pas utile pour les zones d'os.
- Scan du patient avec deux valeurs de voltage différentes, et donc des énergies différentes (Coleman et Sinclair; 1985; Goodsitt; 1995). Les deux images obtenues peuvent ensuite être combinées pour corriger l'artefact de durcissement de faisceau. Mais cette technique nécessite deux scans, et est donc peu utilisée en pratique.

Bougé-patient

Les artefacts de bougé-patient apparaissent lorsque le patient bouge pendant l'acquisition. Cela peut être dû à un mouvement de la part du patient ou, plus difficile à contrôler, le mouvement respiratoire ou cardiaque. Dans ces conditions, les projections mesurées contiennent des incohérences : certaines projections récupéreront l'atténuation en un point à un certain endroit et d'autres projections le verront ailleurs. L'application des algorithmes de reconstruction crée donc un flou sur le volume reconstruit et des stries près des zones de grande intensité.

1. *Beam hardening* dans la littérature

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

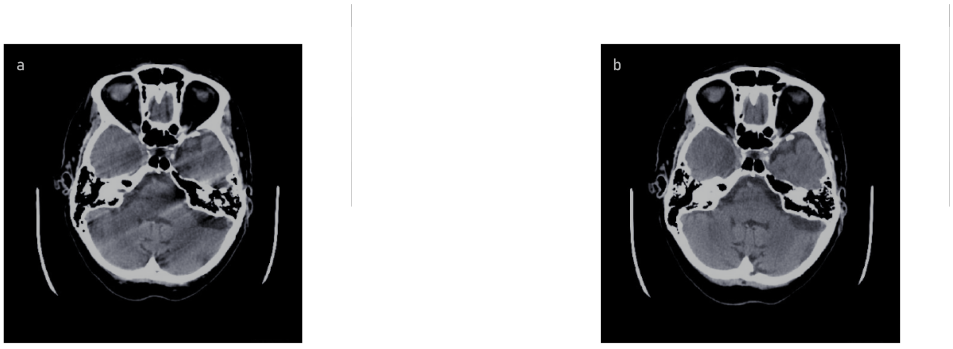


FIGURE 2.12 – Illustration d'un mouvement de tête d'un patient créant un artefact de bougé-patient (a) sans correction, (b) avec correction (Hsieh; 2009). L'artefact se présente sous forme de rayures obliques dans le crâne du patient.

Pour limiter les effets de cet artefact, l'immobilisation du patient et la réduction du temps d'acquisition sont des solutions évidentes. Cependant, demander au patient de retenir sa respiration le temps du scan ou à un enfant d'être parfaitement immobile n'est pas toujours possible. Dans la mesure où elle oblige à diminuer la dose (mAs) et/ou le nombre de projections, la réduction du temps d'acquisition diminue le rapport signal sur bruit des projections et la qualité de la reconstruction. On peut alors pondérer l'importance des projections dans la reconstruction en déterminant les projections qui auraient subi les plus grands mouvements (souvent les premières et les dernières) (Pelc et Glover; 1986). La figure 2.12 montre l'impact des mouvements du patient sur la qualité de l'image ainsi que sa correction.

Artefacts dus au détecteur

Un pixel défectueux sur le détecteur peut considérablement détériorer la qualité de l'image. Sur le sinogramme, cela se traduira en une ligne de pixel noire horizontale qui crée une discontinuité (cf figure 2.13a). Lorsque l'on effectue la reconstruction, des artefacts en forme d'anneaux blancs apparaissent sur l'image.

La réduction de ces artefacts peut se faire en supprimant directement la ligne discontinue sur le sinogramme (Hsieh; 1994) ou en corrigeant l'image reconstruite en cherchant les anneaux (Freundlich; 1987; Hsieh; 1996).

La figure 2.13b montre un exemple d'un tel artefact dans une image d'un torse.

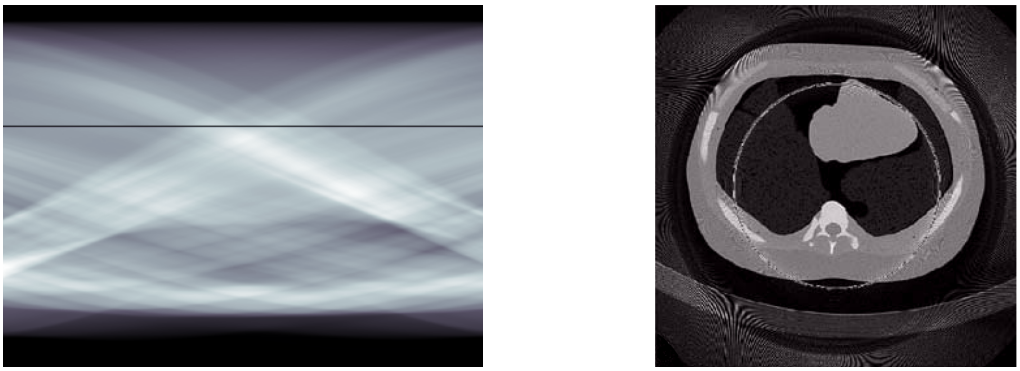
Plus généralement, un mauvais étalonnage des pixels qui n'ont pas le même gain entre eux conduit également à des artefacts en forme d'anneaux.

Artefacts métalliques

Un problème bien connu en reconstruction tomographique est la présence de métal à l'intérieur du patient, et ceci est particulièrement fréquent en CBCT dentaire. A cause du numéro atomique élevé des métaux, ils absorbent la quasi-totalité, voire la totalité, des rayons X qui sont censés les traverser. Cela produit des inconsistances de même type que celles du durcissement de faisceau et crée d'importantes raies noires et blanches en étoile autour de la zone contenant du métal ainsi qu'une baisse du contraste.

Pour remédier à ce problème, il existe de nombreuses techniques dont la plus utilisée consiste à faire de l'*inpainting* sur les projections. On cherche sur les projections les endroits où le métal apparaît, et on essaie de remplacer les valeurs par une interpolation des données

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés



(a) (b)

FIGURE 2.13 – (a) Sinogramme avec un pixel du détecteur défectueux. (b) Exemple d'artefact en anneau (Buzug; 2011).

correctes (Glover et Pelc; 1981).

Angles manquants

Nous avons vu précédemment qu'un certain intervalle angulaire devait être respecté pour pouvoir appliquer les algorithmes de reconstruction correctement : 180° en géométrie parallèle et 180° + angle de faisceau en géométrie divergente. Lorsque l'intervalle nécessaire ne peut pas être scanné intégralement, les méthodes de reconstruction ne disposent pas de l'information suffisante pour effectuer une reconstruction correcte. On observe alors des artefacts dans les directions des angles manquants.

Bruit

La réduction de la dose envoyée au patient diminue le SNR des projections. Celles-ci sont bruitées et les reconstructions également. Un bruit important dans une image dégrade de manière drastique la capacité à distinguer les détails dans l'image. La tomographie à faible dose est un sujet actuel important, qui fait d'ailleurs l'objet de cette thèse et de nombreuses techniques de réduction de bruit ont été développées. On peut par exemple commencer par traiter les projections avant la reconstruction (Manduca *et al.*; 2009) ou faire du post-traitement sur l'image reconstruite (Giraldo *et al.*; 2009) ou encore combiner ces deux techniques. D'autres méthodes reconstruisent le volume en utilisant une régularisation pour le débruiter en même temps, nous parlerons de ces méthodes dans le chapitre 3.

Faible nombre de projections

Une méthode pour diminuer le bruit lors d'acquisitions continues à faible dose est de diminuer le nombre de projections. Une projection peut être vue comme l'intégration de la dose reçue sur un certain pas angulaire. Dans le cas de l'imagerie faible dose, on choisit un pas angulaire plus grand pour récupérer plus de photons et augmenter le SNR. Dans ce cas, les projections sont moins nombreuses mais moins bruitées. Une autre manière de diminuer la dose d'irradiation est de limiter le nombre de projections en effectuant une acquisition pulsée : la source de rayons X émet de manière discontinue. Dans les deux cas, le faible nombre de projections crée des artefacts étoilés qui sont d'autant plus importants que le nombre de

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

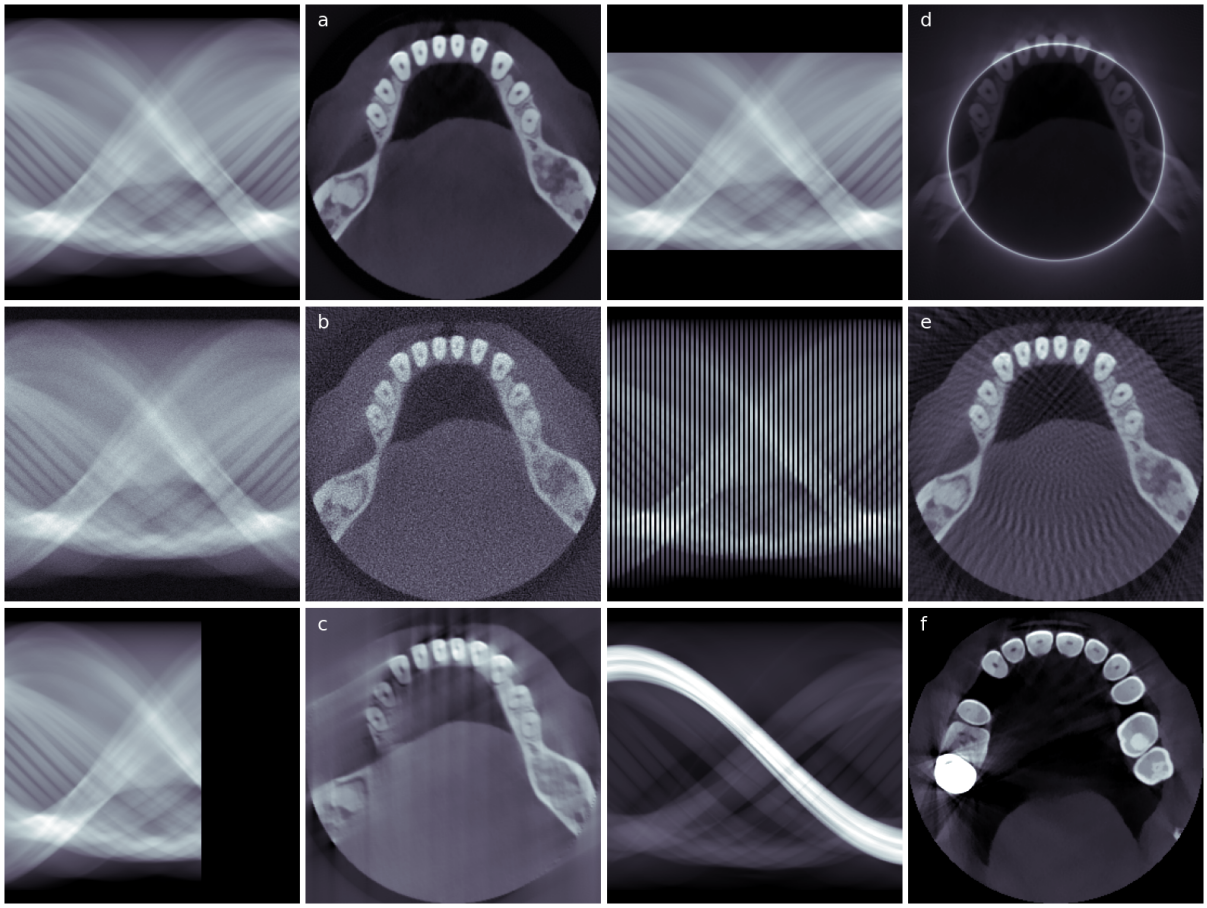


FIGURE 2.14 – Différents artéfacts, avec effets sur le sinogramme et l'image reconstruite. (a) Reconstruction normale, 300 projections en géométrie parallèle, régulièrement espacées avec des angles allant de 0 à π et un détecteur de 300 pixels (b) Reconstruction avec projections bruitées par un bruit gaussien de moyenne nulle et de variance $\sigma = 0.02 \times \max(\text{sinogramme})$, (c) Reconstruction avec 200 projections et des angles allant de 0 à $\frac{2\pi}{3}$ seulement, (d) Reconstruction avec projections tronquées : taille du détecteur = 200 pixels, (e) Reconstruction avec peu de projections, seulement 1 sur 5, et (f) Reconstruction avec une dent en métal.

projections est faible. Nous verrons également dans la suite comment réduire cet artefact. Comme pour le bruit, il est possible de travailler à différentes étapes de la reconstruction : interpolation des projections (Bertram *et al.*; 2004), régularisation pendant la reconstruction (Hu *et al.*; 2017) et post-traitement de la reconstruction (Tipnis *et al.*; 2010).

Projections tronquées et tomographie intérieure

On parle de tomographie intérieure lorsque le champ de vue (FoV [1]) est plus petit que l'objet. Dans ce cas, l'objet n'est pas complètement inclus dans chacune des projections, on parle alors de projections tronquées. Deux raisons peuvent aboutir à cette configuration : 1) une région d'intérêt assez petite, auquel cas il n'est pas utile d'irradier l'intégralité du patient et 2), un détecteur moins large que l'objet. Si l'on reconstruit directement le volume, un cercle de forte intensité apparaît autour du FOV. Une des corrections possibles est d'approximer le patient avec un cylindre d'eau, et d'étendre les projections tronquées avec les projections de

1. *Field of View*

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

ce cylindre. Nous parlerons plus en détail des différentes méthodes pour corriger cet artefact dans la suite.

La figure 2.14 montre l'impact des artefacts dus au bruit, aux angles manquants, aux projections tronquées, au faible nombre de projections et à la présence de métal. Le sinogramme est également présenté dans chacun des cas pour visualiser le lien entre l'information manquante sur le sinogramme et la dégradation de l'image.

Artefact de cone beam

La géométrie conique du faisceau en CBCT entraîne des erreurs dans les reconstructions dans les zones éloignées du plan de la source. Nous en avons déjà parlé dans la partie reconstruction avec l'algorithme FDK, mais la déformation et le manque d'information dans ces zones induit des artefacts avec les reconstructions itératives également.

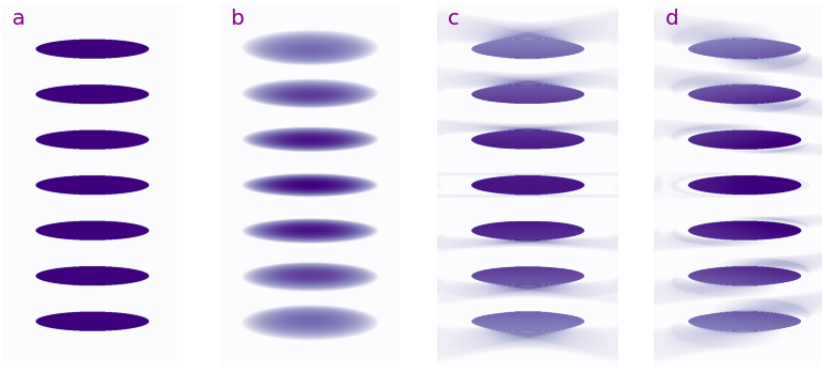


FIGURE 2.15 – Illustration de l'artefact de cone-beam sur le fantôme *Defrise*, créé avec la librairie TomoPhantom (Kazantsev *et al.*; 2018). (a) Coupe du fantôme de face, (b) projection cone-beam, (c) coupe de face de la reconstruction, (d) coupe de profil de la reconstruction.

La figure 2.15 illustre très bien cet artefact. Le fantôme *Defrise*, de taille 1024×512×512 voxels de côté 0.1 mm, a été projeté avec une géométrie conique. La source était centrée sur le volume et située à 200 mm du centre du volume. La figure 2.15b montre l'étalement progressif des projections en fonction de l'éloignement par rapport au plan de la source. L'artefact de cone beam est bien visible sur les figures 2.15c et 2.15d. La forme du faisceau est bien visible dans cette dernière.

Dans ce chapitre, nous avons présenté les bases mathématiques de la reconstruction tomographique avec la transformée de Radon. Nous avons décrit deux types de reconstructions : les reconstructions analytiques et les reconstructions itératives.

Les reconstructions analytiques utilisent une formulation continue du problème et la transformée de Radon inverse. Ces méthodes ont l'avantage d'être très rapides mais elles sont aussi particulièrement sensibles aux perturbations dans les données. La présence de bruit ou d'incohérences dans les données crée des artefacts qui diminuent la qualité de l'image et peut conduire à un diagnostic erroné de la part du médecin.

Les méthodes itératives permettent de prendre en compte la physique du système de détection et même la distribution probabiliste des données pour les méthodes statistiques. Elles

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

sont davantage robustes face aux incohérences mais restent assez sensibles au bruit présent dans les projections, voire augmentent le bruit après un certain nombre d'itérations.

De nombreux phénomènes physiques, liés au modèle physique de la tomographie, au patient ou au scanner utilisé, compromettent la qualité de l'image en créant des artefacts dans les reconstructions, en particulier en CBCT dentaire (Jaju *et al.*; 2013). La présence de métal et les projections tronquées sont les principales causes d'artefacts dans les images. Puis la réduction de la dose implique l'apparition de bruit et d'artefacts étoilés causés par le faible nombre de projections.

Le chapitre suivant décrit quelques méthodes de reconstruction itérative régularisée dont le principe est de débruiter le volume pendant la reconstruction. Elles résolvent le problème de reconstruction en le transformant en un problème de minimisation de fonctionnelle. En ajoutant un terme de régularisation à la fonctionnelle, elles permettent l'introduction d'informations a priori sur la solution.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Bibliographie

Ahn, S. et Fessler, J. A. (2003). Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms. *IEEE transactions on medical imaging*, 22(5):613–626. 32

Andersen, A. H. et Kak, A. C. (1984). Simultaneous algebraic reconstruction technique (SART) : a superior implementation of the ART algorithm. *Ultrasonic imaging*, 6(1):81–94. 29

Anthoine, S., Aujol, J.-F., Boursier, Y. et Melot, C. (2012). Some proximal methods for Poisson intensity CBCT and PET. *Inverse Problems & Imaging*, 6(4):565–598. 33

Barrett, J. F. et Keat, N. (2004). Artifacts in CT : recognition and avoidance. *Radiographics*, 24(6):1679–1691. 35

Beck, A. et Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1):183–202. 34

Bertram, M., Rose, G., Schafer, D., Wiegert, J. et Aach, T. (2004). Directional interpolation of sparsely sampled cone-beam CT sinogram data. *In 2004 2nd IEEE International Symposium on Biomedical Imaging : Nano to Macro (IEEE Cat No. 04EX821)*, pages 928–931. IEEE. 38

Bonnet, S. (2000). *Approches multi résolution en reconstruction tomographique 3D : Application à l'angiographie cérébrale*. Thèse de doctorat, Lyon, INSA. 25

Brooks, R. A. et Di Chiro, G. (1976). Beam hardening in X-ray reconstructive tomography. *Physics in medicine & biology*, 21(3):390. 35

Buzug, T. M. (2011). Computed tomography. *In Springer handbook of medical technology*, pages 311–342. Springer. xii, 37

Byrne, C. (1998). Iterative algorithms for deblurring and deconvolution with constraints. *Inverse Problems*, 14(6):1455. 33

Coleman, A. et Sinclair, M. (1985). A beam-hardening correction using dual-energy computed tomography. *Physics in medicine & biology*, 30(11):1251. 35

Dempster, A. P., Laird, N. M. et Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society : Series B (Methodological)*, 39(1):1–22. 31

Feldkamp, L. A., Davis, L. C. et Kress, J. W. (1984). Practical cone-beam algorithm. *Josa a*, 1(6):612–619. 26

Freundlich, D. A. (1987). Ring artifact correction for computerized tomography. US Patent 4,670,840. 36

Gilbert, P. (1972). Iterative methods for the three-dimensional reconstruction of an object from projections. *Journal of theoretical biology*, 36(1):105–117. 29

Giraldo, J. C. R., Kelm, Z. S., Guimaraes, L. S., Yu, L., Fletcher, J. G., Erickson, B. J. et McCollough, C. H. (2009). Comparative study of two image space noise reduction methods for computed tomography : bilateral filter and nonlocal means. *In 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3529–3532. IEEE. 37

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Glover, G. H. et Pelc, N. J. (1981). An algorithm for the reduction of metal clip artifacts in CT reconstructions. *Medical physics*, 8(6):799–807. 37

Goodsitt, M. M. (1995). Beam hardening errors in post-processing dual energy quantitative computed tomography. *Medical physics*, 22(7):1039–1047. 35

Gordon, R. (1974). A tutorial on ART (algebraic reconstruction techniques). *IEEE Transactions on Nuclear Science*, 21(3):78–93. 29

Gordon, R., Bender, R. et Herman, G. T. (1970). Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of theoretical Biology*, 29(3):471–481. 29

Horn, B. K. (1979). Fan-beam reconstruction methods. *Proceedings of the IEEE*, 67(12):1616–1623. 25

Hsieh, J. (1994). Radiation detector offset and afterglow compensation technique. US Patent 5,331,682. 36

Hsieh, J. (1996). Guided ringfix algorithm for image reconstruction. US Patent 5,533,081. 36

Hsieh, J. (2009). *Computed tomography : principles, design, artifacts, and recent advances*. SPIE press. xii, 35, 36

Hu, Z., Gao, J., Zhang, N., Yang, Y., Liu, X., Zheng, H. et Liang, D. (2017). An improved statistical iterative algorithm for sparse-view and limited-angle CT image reconstruction. *Scientific reports*, 7(1):1–9. 38

Hudson, H. et Larkin, R. (1994). Accelerated image reconstruction using ordered subsets of projection data. *IEEE Transactions on Medical Imaging*, 13(4):601–609. 32

Jaju, P. P., Jain, M., Singh, A., Gupta, A. *et al.* (2013). Artefacts in cone beam CT. *Open Journal of Stomatology*, 3(05):292. 40

Kaczmarz, M. (1937). Approximate solution of systems of linear equations. Rapport technique. 28

Kak, A. C. et Slaney, M. (2001). *Principles of computerized tomographic imaging*. SIAM. 25

Kandarpa, V. S. S. (2022). *Tomographic image reconstruction with direct neural network approaches*. Thèse de doctorat, Université de Bretagne occidentale-Brest. 34

Kazantsev, D., Pickalov, V., Nagella, S., Pasca, E. et Withers, P. J. (2018). TomoPhantom, a software package to generate 2D-4D analytical phantoms for CT image reconstruction algorithm benchmarks. *SoftwareX*, 7:150–155. xii, 39

Kijewski, P. K. et Bjärngard, B. E. (1978). Correction for beam hardening in computed tomography. *Medical physics*, 5(3):209–214. 35

Lakshminarayanan, A. (1975). *Reconstruction from divergent ray data*. State University of New York (Buffalo). Department of Computer Science. 25

Landweber, L. (1951). An iteration formula for Fredholm integral equations of the first kind. *American journal of mathematics*, 73(3):615–624. 29

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Lange, K., Carson, R. *et al.* (1984). EM reconstruction algorithms for emission and transmission tomography. *J Comput Assist Tomogr*, 8(2):306–16. 31, 33

Manduca, A., Yu, L., Trzasko, J. D., Khaylova, N., Kofler, J. M., McCollough, C. M. et Fletcher, J. G. (2009). Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT. *Medical physics*, 36(11):4911–4919. 37

Moore, E. H. (1920). On the reciprocal of the general algebraic matrix. *Bull. Am. Math. Soc.*, 26:394–395. 28

Natterer, F. (2001). The Mathematics of Computerized Tomography (Philadelphia, PA : SIAM). 28

Nesterov, Y. E. (1983). A method for solving the convex programming problem with convergence rate $O(1/k^2)$. *In Dokl. Akad. Nauk SSSR,*, volume 269, pages 543–547. 34

Nuyts, J., De Man, B., Dupont, P., Defrise, M., Suetens, P. et Mortelmans, L. (1998). Iterative reconstruction for helical ct : a simulation study. *Physics in Medicine & Biology*, 43(4):729. 33

Pelc, N. J. et Glover, G. H. (1986). Method for reducing image artifacts due to projection measurement inconsistencies. US Patent 4,580,219. 36

Radon, J. (1917). On the determination of functions from their integrals along certain manifolds. *Ber. Verh, Sachs Akad Wiss.*, 69:262–277. 19

Shepp, L. et Vardi, Y. (1982). Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, 1(2):113–122. 31

Tipnis, S., Ramachandra, A., Huda, W., Hardie, A., Schoepf, J., Costello, P., Flohr, T. et Sedlmair, M. (2010). Iterative reconstruction in image space (IRIS) and lesion detection in abdominal CT. *In Medical Imaging 2010 : Physics of Medical Imaging*, volume 7622, pages 846–857. SPIE. 38

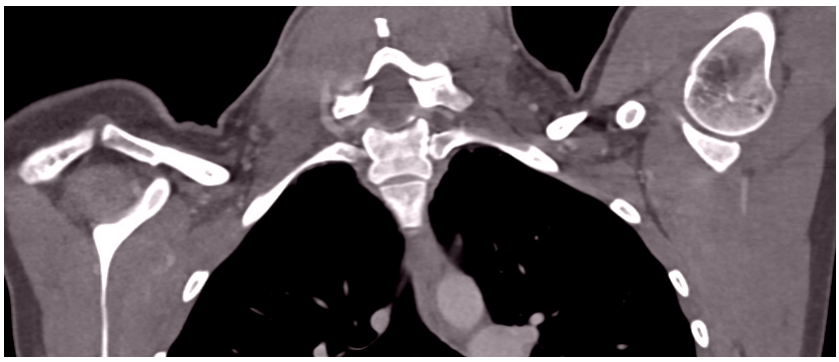
Titterington, D. (1987). On the iterative image space reconstruction algorthm for ECT. *IEEE transactions on medical imaging*, 6(1):52–56. 31

Webb, S. (1982). A modified convolution reconstruction technique for divergent beams. *Physics in Medicine and Biology*, 27(3):419–423. 25

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

3 Reconstructions itératives régularisées



I Problème inverse et optimisation 46
1 Formulation variationnelle 46
2 Optimisation régulière 46
3 Optimisation convexe 48
4 Formulation bayésienne 52
II Méthodes régularisées par variation totale 53
1 Variation Totale 53
2 SIRT régularisé 55
3 MLEM régularisé 56
4 Algorithme de Chambolle-Pock en tomographie 57
III Application au fantôme de Shepp-Logan 2D 58
1 Présentation des données 58
2 Comparaison des résultats 59
3 Application à des données expérimentales 60

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Dans ce chapitre nous allons introduire les méthodes itératives régularisées pour la résolution de problèmes inverses, qui pourront être utilisées pour la tomographie. Ces méthodes permettent de régulariser la solution pendant le processus de reconstruction itérative et ainsi de réduire le bruit et les artefacts présents dans les reconstructions du chapitre précédent. Nous repartirons de la formulation problème inverse $Af = p$ et investiguerons les différentes méthodes permettant de résoudre ce problème et les appliquerons en 2D et en géométrie parallèle. Nous étudierons ensuite le comportement et la comparaison de ces algorithmes sur le fantôme de Shepp-Logan 2D et des données expérimentales 3D parallèles.

I Problème inverse et optimisation

Exprimé de façon mathématique, le problème de reconstruction tomographique est un problème inverse linéaire. Nous avons vu dans le chapitre précédent que la matrice A n'est pas toujours inversible et même lorsque c'est le cas, la dimension de A est suffisamment grande pour que le calcul de son inverse soit très compliqué. Les méthodes présentées dans le chapitre 2 cherchent à résoudre ce problème de différentes manières. Les méthodes analytiques visent à inverser directement la formule pour obtenir f, nous avons vu que ces algorithmes étaient très sensibles aux conditions d'acquisition et au bruit. Ensuite, nous avons introduit les algorithmes itératifs classiquement utilisés qui sont également encore assez sensibles au nombre de vues, et qui s'éloignent de la solution désirée au fil des itérations, au profit d'une image de plus en plus bruitée. Toutefois, la formulation variationnelle du problème permet d'exploiter des informations a priori sur la solution, et la résolution du problème peut être mise en œuvre par divers algorithmes d'optimisation. Dans cette partie, nous présentons les méthodes classiques de minimisation de fonctionnelles. Les fonctionnelles convexes constituent un cas particulier important, en effet, elles ont tendance à avoir des minima globaux.

1 Formulation variationnelle

Lorsque le problème inverse est mal-posé, les méthodes variationnelles régularisées sont largement utilisées. Elles considèrent un problème de minimisation où l'on cherche f^* qui minimise une fonctionnelle de type :

$$J(f) = d\left(Af, p\right) + \alpha R(f) \tag{3.1}$$

avec $d(Af, p)$ le terme d'attache aux données qui assure la correspondance entre la solution $f \in \mathbb{R}^N$ et les projections $p \in \mathbb{R}^M$, $R(f)$ est le terme de régularisation qui permet d'ajouter un a priori sur la solution et α est le paramètre de régularisation dont la valeur est choisie en fonction de l'importance que l'on veut donner à la régularisation. Nous présentons de façon non exhaustive des principes de base sur les algorithmes d'optimisation et les différentes approches permettant de chercher une solution approchée de f^*.

2 Optimisation régulière

On suppose ici que J est régulière avec les conditions suivantes :

- $J \in C^2\left(\mathbb{R}^N, \mathbb{R}\right)$
- ∇J est lipschitzienne de constante L sur $\mathbb{R}^N$.

Image de couverture : coupe d'un angioscanner des membres supérieurs de moi-même.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Les conditions d'existence d'un minimum local en f^* sont (Ciarlet; 1990) :

$$\begin{cases} \nabla J(f^*) &= 0 \\ {}^t g H_J(f) g &\geq 0 \quad \forall g \in \mathbb{R}^N, \ \forall f \text{ dans un voisinage de } f^* \end{cases} \tag{3.2}$$

avec ∇J et H_J respectivement le gradient et la Hessienne de J. Lorsque ces conditions sont vérifiées, deux méthodes sont principalement utilisées pour approcher la solution du problème d'optimisation : la méthode de Newton et la descente de gradient qui est une variante de la première.

Méthode de Newton

Nous venons de voir que lorsque J admet un minimum local en f^*, nous avons $\nabla J(f^*) = 0$. La méthode de Newton s'intéresse à la résolution de l'équation $\nabla J(f) = 0$, également appelée *équation d'Euler*, i.e. à la recherche des zéros de l'application dérivée ∇J. On initialise la suite $\left(f^{(k)}\right)_k$ avec $f^{(0)}$ puis on suit le schéma itératif présenté dans la suite.

On approche ∇J au voisinage de $f^{(0)}$ en considérant sa tangente en f :

$$\nabla J(f) \approx \nabla J\left(f^{(0)}\right) + H_J\left(f^{(0)}\right)\left(f - f^{(0)}\right) \tag{3.3}$$

d'après la formule de Taylor. Pour trouver un zéro de cette fonction, il suffit résoudre :

$$0 = \nabla J\left(f^{(0)}\right) + H_J\left(f^{(0)}\right)\left(f - f^{(0)}\right). \tag{3.4}$$

En résolvant cette équation, nous obtenons un $f^{(1)}$ plus proche de f^* que $f^{(0)}$. On réitère le processus jusqu'à ce que la différence entre deux itérations successives soit inférieur à un seuil fixé, indiquant que la convergence de l'algorithme a été atteinte. Le schéma itératif de la méthode de Newton est donné par :

$$f^{(k+1)} = f^{(k)} - \left(H_j\left(f^{(k)}\right)\right)^{-1} \nabla J\left(f^{(k)}\right) \tag{3.5}$$

Ce schéma a une convergence quadratique, les conditions et la démonstration de cette convergence peuvent être trouvées dans (Ciarlet; 1990).

Descente de gradient

Le calcul de la Hessienne à chaque étape, dans la méthode de Newton, peut rapidement se retrouver compliqué ou coûteux. Une simplification possible de cet algorithme peut consister à considérer la Hessienne comme la matrice identité multipliée par une constante le long des itérations. C'est l'idée de la descente de gradient. À partir d'une initialisation $f^{(0)} \in \mathbb{R}^N$, le pas de descente de gradient s'écrit :

$$f^{(k+1)} = f^{(k)} - \alpha \nabla J\left(f^{(k)}\right) \tag{3.6}$$

où α est le pas de la descente. La convergence de cet algorithme est assurée dès que $\alpha \leq \frac{2}{L}$ (Wright *et al.*; 1999), cependant elle reste assez lente et dépendante des propriétés de la fonctionnelles. D'autre méthodes jouant sur le pas descente ou les directions ont permis d'accélérer cet algorithme. Nous pouvons par exemple citer la méthode du gradient conjugué qui utilise des vecteurs de direction conjugués, ou la méthode du gradient à pas optimal qui calcule le pas optimal $\alpha^{(k)}$ à chaque étape (Wright *et al.*; 1999).

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Nous avons vu des conditions nécessaires et suffisantes à l'existence d'un minimum pour les fonctionnelles C^2 dans le système (3.2). Cependant, la condition de régularité sur J est une condition assez forte, qui n'est pas toujours vérifiée. De plus, la deuxième condition peut être difficile à vérifier si le calcul de la Hessienne n'est pas aisé. Les fonctionnelles convexes sont très utilisées en optimisation car les conditions de régularité peuvent être réduites. Celles-ci présentent des propriétés spécifiques sur leur minimum :

- un minimum local est un minimum global
- la première condition de (3.2) est une condition nécessaire et suffisante de minimalité. Si J est de classe C^1 sur $\mathbb{R}^N$:

$$f \text{ est un minimum de } J \Leftrightarrow \nabla J(f) = 0 \tag{3.7}$$

Ainsi, ces fonctionnelles ont leurs propres méthodes d'optimisation, et la partie suivante en présente quelques unes rapidement.

3 Optimisation convexe

Il existe de nombreuses raisons pour lesquelles les fonctions convexes sont intéressantes en optimisation. Par exemple, elles peuvent être facilement minimisées à l'aide de la descente de gradient. De plus, les fonctions convexes ont tendance à avoir des minima globaux, ce qui signifie qu'elles peuvent être optimisées plus largement que les fonctions non convexes. Enfin, les fonctions convexes sont souvent présentes dans la nature, ce qui en fait de bons modèles pour les problèmes d'optimisation du monde réel.

Les plus représentatives sont la norme L^2, l'exponentielle et le logarithme. Les fonctions quadratiques sont souvent utilisées pour modéliser des phénomènes physiques, tels que le mouvement d'un pendule ou la trajectoire d'un projectile. La fonction exponentielle est souvent utilisée pour modéliser la croissance ou la décroissance. Par exemple, pour modéliser la croissance d'une population, on peut utiliser une fonction exponentielle. La fonction logarithme est souvent utilisée pour modéliser le comportement de systèmes qui évoluent lentement dans le temps. Par exemple, le logarithme du temps nécessaire à la désintégration d'une substance radioactive peut être utilisé pour modéliser la désintégration elle-même. De plus, le logarithme de la distance entre deux objets peut être utilisé pour modéliser la force de gravité entre eux.

Nous commençons cette partie par introduire quelques définitions qui seront utilisées dans la suite.

Lorsque J est simplement considérée convexe, et non nécessairement différentiable, nous introduisons la notion de sous-gradient qui généralise le gradient lorsque ce dernier n'existe pas. Un vecteur $g \in \mathbb{R}^N$ est un *sous-gradient* de J en f lorsque

$$J(v) \geq J(f) + \langle g, v - f \rangle, \ \forall\, v \in \mathbb{R}^N. \tag{3.8}$$

L'ensemble des vecteurs g qui satisfont (3.8) est appelé *sous-différentiel* de J en f et est noté $\partial J(f)$. La relation (3.7) s'écrit également avec le sous-gradient :

$$f \text{ est un minimum de } J \Leftrightarrow 0 \in \partial J(f) \tag{3.9}$$

Une fonctionnelle J est dite *semi-continue inférieurement* si pour toute suite (f_k) convergeant vers f, on a

$$J(f) \leq \liminf_{k \to +\infty} J(f_k)$$

et elle est dite *propre* lorsque pour toute suite (f_k) telle que

$$\sup_{k,l} \|f_k - f_l\|_2 < +\infty \ \text{ et } \ \sup_k J(f_k) < +\infty$$

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

il existe une sous-suite convergeant dans $\mathbb{R}^N$.

Soit une fonctionnelle convexe, semi-continue inférieurement propre H, on définit son *opérateur proximal* par :

$$\text{prox}_\sigma[H](z) = \arg\min_{z'} \left\{ H(z') + \frac{\|z - z'\|_2^2}{2\sigma} \right\} \tag{3.10}$$

L'opérateur proximal permet de généraliser la descente de gradient lorsque les fonctionnelles ne sont pas différentiables.

Descente de sous-gradient

On peut généraliser la descente de gradient et l'équation (3.6) avec la notion de sous-gradient :

$$f^{(k+1)} = f^{(k)} - \alpha_k g_k, \quad g_k \in \partial J(f) \tag{3.11}$$

avec α_k les pas de descente. La convergence est assurée dans les cas suivants (Wright *et al.*; 1999) :

- $\sum_k \alpha_k = +\infty$
- $\sum_k (\alpha_k)^2 < +\infty$

avec un taux de convergence $O\left(\frac{1}{\sqrt{k}}\right)$. Cette méthode converge donc assez lentement et n'a pas de très bonnes performances (Polyak; 1987; Nesterov; 2003). Elle reste utilisée lorsque l'opérateur proximal de J est difficile à calculer, mais que ∂J se calcule facilement.

Descente de gradient implicite et algorithme proximal

Lorsque J est différentiable, mais non lipschitzienne, la descente de gradient 3.6 présentée précédemment ne vérifie plus les conditions de convergence. Cependant, on peut considérer le schéma de descente de gradient implicite suivant :

$$f^{(k+1)} = f^{(k)} - \alpha \nabla J\left(f^{(k+1)}\right) \tag{3.12}$$

La solution $f^{(k+1)}$ de l'équation (3.12) est ainsi un point critique de la fonction $f \mapsto J(f) + \frac{1}{2\alpha}\|f - f^{(k)}\|_2^2$. Puisque J est convexe et semi-continue inférieurement, ce point critique est un minimiseur de cette fonctionnelle et ce minimiseur est unique. On définit la régularisée de Moreau-Yosida J_α de J par :

$$J_\alpha(f) = \min_g J(g) + \frac{\|g - f\|_2^2}{2\alpha}. \tag{3.13}$$

∇J_α est $\frac{1}{\alpha}$-Lipschitzienne et on a la relation suivante (Chambolle et Pock; 2016) :

$$\text{prox}_\alpha[J](f) = f - \alpha \nabla J_\alpha(f). \tag{3.14}$$

Cette équation montre qu'une itération de gradient implicite sur J est équivalente à une itération explicite sur J_α :

$$f^{(k+1)} = \text{prox}_\alpha[J]\left(f^{(k)}\right) = f^{(k)} - \alpha \nabla J_\alpha\left(f^{(k)}\right). \tag{3.15}$$

Le schéma implicite (3.12) est alors facilement implémentable et converge en $O\left(\frac{1}{k}\right)$.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Forward-Backward Splitting

Le schéma de Forward-Backward (FB) splitting minimise le problème suivant :

$$\min_f F(f) + G(f), \tag{3.16}$$

avec F une fonctionnelle convexe, avec un gradient Lipschitzien, et G une fonctionnelle convexe, semi-continue inférieure propre dont l'opérateur proximal est connu.

L'idée de ce schéma est de résoudre ce problème de minimisation en effectuant un pas de descente de gradient explicite sur F qui est régulière, et un pas de descente implicite, en utilisant l'opérateur proximal, sur G. Le schéma converge en $O\left(\frac{1}{k}\right)$ et s'écrit :

$$f^{(k+1)} = \text{prox}_\alpha[G]\left(f^{(k)} - \alpha\nabla F\left(f^{(k)}\right)\right) \tag{3.17}$$

En utilisant l'accélération de Nesterov sur cette méthode, on obtient l'algorithme FISTA (Beck et Teboulle; 2009) (cf chapitre 2), qui a un taux de convergence de $O\left(\frac{1}{k^2}\right)$.

Méthodes primales-duales

Les méthodes primales-duales peuvent être utilisées pour résoudre des problèmes d'optimisation qui s'écrivent sous la forme :

$$\min_f F(Af) + G(f) \tag{3.18}$$

avec $A : X \to Y$ un opérateur linéaire continu, $F : Y \to \mathbb{R}$ et $G : X \to \mathbb{R}^+$ deux fonctionnelles convexes semi-continues inférieures propres, et X et Y des espaces de Hilbert.

Ces méthodes considèrent le problème primal (3.18) et le problème dual qui lui est associé (3.21), et ces deux formulations aboutissent à un problème de point selle. L'algorithme de Chambolle-Pock résout ce problème en effectuant une descente sur la variable primale et une montée sur la variable duale en utilisant le schéma de splitting FB.

On définit la conjuguée d'une fonctionnelle convexe H à l'aide de la transformée de Legendre (Rockafellar; 1970) :

$$H^*(z) = \max_{z'} \{\langle z, z'\rangle - H(z')\}. \tag{3.19}$$

D'après le théorème de Moreau-Fenchel, la fonctionnelle H peut être retrouvée en appliquant la conjugaison à H^* :

$$H(z) = \max_{z'} \{\langle z, z'\rangle - H^*(z')\}. \tag{3.20}$$

Le problème primal (3.18) est équivalent au problème dual suivant :

$$\max_p \left\{-F^*(p) - G^*\left(-{}^tAp\right)\right\}. \tag{3.21}$$

Ces deux problèmes sont connectés à travers le problème de point selle :

$$\min_f \max_p \{\langle Af, p\rangle + G(f) - F^*(p)\}. \tag{3.22}$$

La maximisation de (3.22) sur p, en utilisant (3.20), donne le problème primal et sa minimisation sur f, le problème dual.

La résolution du problème dual conjointement au problème primal permet d'assurer la convergence de l'algorithme. L'écart dual est la différence entre l'itération du problème primal et celle du problème dual. On dit que l'algorithme a convergé lorsque cet écart est inférieur à un seuil fixé. Nous détaillons dans la suite l'algorithme de Chambolle-Pock (Chambolle et Pock; 2011) qui est une méthode de référence dans le cadre des méthodes primales-duales.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Algorithme de Chambolle-Pock L'algorithme de Chambolle-Pock, aussi connu sous le nom de *primal-dual hybrid gradient method* (PDHG), résout simultanément les problèmes (3.18) et (3.21). Il est utilisé dans de nombreux problèmes d'optimisation avec des applications très différentes. Le principe est de séparer chaque itération en deux :

- Une étape qui maximise (3.22) sur p, avec f fixé
- Et une seconde étape qui minimise (3.22) sur f en gardant p fixé.

L'algorithme de Chambolle-Pock utilise le schéma de splitting FB pour résoudre ces deux étapes. Il est décrit dans l'algorithme 1, les conditions et démonstrations de convergence de cet algorithme sont données dans (Chambolle et Pock; 2011).

Algorithme 1: Pseudo-code pour N itérations de l'algorithme de Chambolle-Pock

Initialisation de $\tau,\ \sigma,\ \theta$
$f^{(0)}, p^{(0)}, n = 0$
$\bar{f}^{(0)} \leftarrow f^{(0)}$
while $n \leq N$ **do**
  $p^{(k+1)} \leftarrow \text{prox}_{\sigma}[F^*]\left(p^{(k)} + \sigma A \bar{f}^{(k)}\right)$
  $f^{(k+1)} \leftarrow \text{prox}_{\tau}[G]\left(f^{(k)} - \tau^t A p^{(k+1)}\right)$
  $\bar{f}^{(k+1)} \leftarrow f^{(k+1)} + \theta\left(f^{(k+1)} - f^{(k)}\right)$
  $k \leftarrow k + 1$

ADMM

L'algorithme ADMM proposé par Glowinski et Marroco (1975) pour résoudre les problèmes de la forme :

$$\underset{Kf+Kp'=b}{\arg\min}\ F(p) + G(f) \tag{3.23}$$

est un algorithme connu et très utilisé pour optimiser des problèmes convexes non-réguliers. Lorsque $b = 0$, $K = A$ et $K' = -Id$, on retrouve le problème (3.18).

Le principe de l'algorithme est d'introduire un multiplicateur de Lagrange z pour la contrainte, et de réécrire le problème comme un problème d'optimisation de point selle pour un Lagrangien augmenté :

$$\min_{f,p} \sup_{z} F(p) + G(f) + \langle z, p - Af\rangle + \frac{\gamma}{2}\|p - Af\|^2, \tag{3.24}$$

avec γ un paramètre.

Une façon de résoudre ce problème est d'alterner les étapes de descente en f et p, puis de mettre à jour z (Gabay et Mercier; 1976). Les différentes étapes sont résumées dans l'algorithme 2.

Algorithme 2: Pseudo-code pour N itérations de l'algorithme ADMM

Initialisation de $\gamma > 0$, $f^{(0)}, z^{(0)}$
while $n \leq N$ **do**
  $p^{(k+1)} \leftarrow \arg\min_p F(p) + \langle z^{(k)}, p\rangle + \frac{\gamma}{2}\|p - Af^{(k)}\|^2$
  $f^{(k+1)} \leftarrow \arg\min_f G(f) - \langle z^{(k)}, Af\rangle + \frac{\gamma}{2}\|p^{(k+1)} - Af\|^2$
  $z^{(k+1)} \leftarrow z^{(k)} + \gamma\left(p^{(k+1)} - Af^{(k+1)}\right)$

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Généralement, on réécrit l'algorithme ADMM sous une forme réduite (Boyd *et al.*; 2011) en posant $u = \frac{1}{\gamma} z$, le problème (3.24) est alors équivalent à :

$$\min_{f,p} \sup_{u} F(p) + G(f) + \frac{\gamma}{2}\|p - Af + u\|_2^2 - \frac{\gamma}{2}\|u\|_2^2. \tag{3.25}$$

La forme réduite est souvent plus courte que (3.24) et plus simple à implémenter.

La convergence de cet algorithme a été étudiée sous plusieurs conditions sur F et G, notamment sous les hypothèses de convexité et non-linéarité (Boyd *et al.*; 2011). Les principaux inconvénients de cette méthode sont le choix du paramètre γ et sa vitesse de convergence qui peut s'avérer assez lente. Cependant, sa structure fait qu'il est possible de paralléliser les opérations, et ainsi diminuer les temps de calcul.

4 Formulation bayésienne

Nous reprenons la formulation statistique de la section 2 du chapitre 2 où le volume f et les projections p sont considérées comme des variables aléatoires. Nous cherchons dans ce cas une estimation de f, avec un a priori sur la solution, qui soit la solution la plus probable d'avoir donné les mesures. On appelle cette estimation, l'estimation de maximum a posteriori (MAP), $\hat{f}_{\text{MAP}}$.

$$\hat{f}_{\text{MAP}} = \arg\max_{f} \mathbb{P}\left(f|p\right) \tag{3.26}$$

La formule de Bayes nous donne

$$\mathbb{P}\left(f|p\right) = \frac{\mathbb{P}\left(p|f\right)\mathbb{P}(f)}{\mathbb{P}(p)}, \tag{3.27}$$

en ne gardant que les termes dépendant de f, on a :

$$\hat{f}_{\text{MAP}} = \arg\max_{f} \mathbb{P}\left(p|f\right)\mathbb{P}(f).$$

Par croissance du logarithme, et en prenant le signe opposé, l'estimation du MAP équivaut à minimiser le problème suivant :

$$\hat{f}_{\text{MAP}} = \arg\min_{f} \left\{-\log \mathbb{P}\left(p|f\right) - \log \mathbb{P}(f)\right\} \tag{3.28}$$

Exemple avec bruit gaussien Nous montrons ici la formulation bayésienne dans le cas d'un bruit gaussien. On considère alors le modèle suivant :

$$p = Af + \varepsilon \tag{3.29}$$

avec $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma^2)$ iid. La probabilité de p sachant f est la probabilité de $\varepsilon = p - Af$ qui est

$$\mathbb{P}(p|f) = \exp\left(-\frac{1}{2\sigma^2}\|Af - p\|_2^2\right) \tag{3.30}$$

En reprenant l'équation (3.28), le problème devient :

$$\min_{f} \left\{-\log \mathbb{P}(f) + \frac{1}{2\sigma^2}\|Af - p\|_2^2\right\}. \tag{3.31}$$

On peut trouver plus de détails dans (Chambolle *et al.*; 2010).

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Sans considérer d'a priori sur la solution, nous retrouvons le problème minimisé par SIRT. Si l'on choisit un a priori gaussien pour f, de moyenne nulle et de variance égale à 1, on obtient la régularisation de Tikhonov (Tihonov; 1963) :

$$\min_f \frac{1}{2} \|Af - p\|_2^2 + \sigma^2 \|f\|_2^2. \tag{3.32}$$

Cette régularisation est utilisée car on connaît une solution explicite de ce problème : $\tilde{f} = \left({}^tAA + \sigma^2 I\right)^{-1} {}^tAp$.

Bruit de Poisson On considère désormais un bruit de Poisson sur les données et un a priori uniforme. Les projections suivent une loi de Poisson de moyenne Af : $p \sim \mathcal{P}(Af)$. La probabilité de f sachant p est alors :

$$\mathbb{P}(f|p) = \prod_i \frac{(Af)_i^{p_i}}{p_i!} e^{-(Af)_i}, \tag{3.33}$$

et en appliquant le logarithme, on obtient la maximisation suivante :

$$\max_f \sum_i [p \ln Af - Af]_i - \ln p_i!, \tag{3.34}$$

ce qui peut être lié à la minimisation la divergence de Kullback-Leibler en ne considérant que les termes dépendant de f. On retrouve la maximisation à effectuer dans (2.22) si A est l'opérateur de projection.

II Méthodes régularisées par variation totale

Nous rappelons que les méthodes régularisées considèrent un a priori sur la solution en ajoutant un terme de régularisation à la fonction coût à minimiser. On obtient alors la fonctionnelle présentée dans (3.1) :

$$J(f) = d\left(Af, p\right) + \alpha R(f)$$

Dans la suite, nous considérerons deux termes d'attaches aux données : la norme ℓ_2 lorsque les données seront supposées suivre une distribution gaussienne, et la divergence de Kullback-Leibler pour une distribution de Poisson.

Le terme de régularisation sera la Variation Totale (TV) qui préserve les contours, contrairement à la régularisation quadratique qui les lisse.

1 Variation Totale

La variation totale est une semi-norme définie pour toutes images $f \in \mathbb{R}^N$ par :

$$TV(f) = \sum_i |\nabla f| = \||\nabla f|\|_1 \tag{3.35}$$

où ∇f est le gradient de f et $\|| \cdot |\|_1$ correspond à la somme des modules des gradients sur chaque voxel de f. Elle n'est pas différentiable sur $\mathbb{R}^N$.

Initialement introduite pour débruiter des images corrompues par un bruit gaussien (Rudin *et al.*; 1992), la variation totale peut également être utilisée en reconstruction tomographique pour régulariser l'image pendant la reconstruction, en particulier en acquisitions à faible dose (Persson *et al.*; 2001; Sidky *et al.*; 2006; Sawatzky *et al.*; 2008; Anthoine *et al.*; 2012; Yan *et al.*; 2011).

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés



FIGURE 3.1 – Débruitage d'une photographie[1] d'une personne souhaitant rester anonyme, à l'aide de la variation totale. a : Image nette, b : image bruitée par un bruit gaussien avec $\sigma = 0.11 \times v_{\max}$, c : image débruitée, 500 itérations, $\alpha = 30$.

[1] Cette photographie a été prise le 1er Mars 2022 et s'intitule : "Avril, Mai, Juin, Juillet. Quatre mois de rédaction de thèse, je suis large. Je peux même ajouter des petits bonus."

Elle favorise l'apparition de zones constantes par morceaux tout en préservant les contours. Le problème de débruitage avec TV se formule ainsi :

$$\min_f \|f - g\|_2^2 + \alpha TV(f) \tag{3.36}$$

avec g l'image bruitée et α le paramètre TV.

Dans (Chambolle; 2004), un algorithme qui minimise rapidement la variation totale est proposé, avec des applications en débruitage et en super-résolution. La méthode proposée est basée sur la formulation duale de (3.36). La solution s'exprime sous la forme $u = g - \alpha \text{div}\varphi$, div étant l'opérateur de divergence et φ calculé itérativement en partant de $\varphi^{(0)} = 0$ puis

$$\varphi^{(k+1)} = \frac{\varphi^{(k)} + \tau\left(\text{div}\varphi^{(k)} - g/\alpha\right)}{1 + \tau|\text{div}\varphi^{(k)} - g/\alpha|} \tag{3.37}$$

avec $\tau \leq 1/8$ en 2D ($\tau \leq 1/12$ en 3D) le pas de descente de gradient, les opérations vectorielles effectuées termes à termes et $|M|$ la matrice composée des valeurs absolues des coefficients de M.

La figure 3.1 illustre les effets d'un débruitage TV sur une photographie en noir et blanc. Un bruit gaussien de moyenne nulle et d'écart-type $\sigma = 0.11 \times v_{\max}$, avec $v_{\max}$ la valeur maximale de l'image, a été ajoutée à l'image non bruitée. L'algorithme de débruitage TV a été ensuite appliqué sur 500 itérations avec le paramètre de régularisation $\alpha = 30$. Le bruit a été considérablement réduit et les contours sont bien préservés. La favorisation des zones constantes par morceaux donne cependant un effet "patch" à l'image débruitée.

La norme TV est utilisée depuis quelques années en reconstruction tomographique comme terme de régularisation dans la fonctionnelle (Sidky *et al.*; 2006). Elle a montré son efficacité dans les problèmes d'angles manquants (Persson *et al.*; 2001; Banjak *et al.*; 2018) et en tomographie par émission avec peu de photons (Sawatzky *et al.*; 2008).

Variation totale en 3D Nous donnons ici les détails de calcul de la variation totale en 3D qui sera utilisée dans le chapitre 4.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Pour une fonction f représentant une image tri-dimensionnelle de taille $J = M \times N \times N$, sa variation totale est donnée par

$$TV(f) = \sum_{i,j,k} \left| (\nabla f)_{i,j,k} \right|. \tag{3.38}$$

avec le gradient discret de f défini par le vecteur $(\nabla f)_{i,j,k} = \left((\nabla f)^1_{i,j,k}, (\nabla f)^2_{i,j,k}, (\nabla f)^3_{i,j,k} \right)$, où

$$(\nabla f)^1_{i,j,k} = \begin{cases} f_{i+1,j,k} - f_{i,j,k} & \text{si } i < M \\ 0 & \text{si } i = M \end{cases}, \quad (\nabla f)^2_{i,j,k} = \begin{cases} f_{i,j+1,k} - f_{i,j,k} & \text{si } j < N \\ 0 & \text{si } j = N \end{cases},$$

et $(\nabla f)^3_{i,j,k} = \begin{cases} f_{i,j,k+1} - f_{i,j,k} & \text{si } k < N, \\ 0 & \text{si } k = N \end{cases}$.

2 SIRT régularisé

La fonctionnelle à minimiser dans le cas de la régularisation TV de l'algorithme SIRT est :

$$\min_f \|Af - p\|_2^2 + \alpha TV(f) \tag{3.39}$$

avec α le paramètre TV.

Cette fonctionnelle est utilisée pour reconstruire des volumes à partir de projections corrompues par un bruit additif gaussien. Banjak *et al.* (2018) ont publié un algorithme minimisant cette équation dont le pseudo-code est donné dans l'algorithme 3. La méthode peut être résumée en trois étapes répétées jusqu'à la convergence :

1. un pas de reconstruction SIRT, donné par la formule (2.20)
2. régularisation TV de la solution courante
3. mise à jour du volume avec FISTA pour l'accélération.

Dans la suite, nous appellerons cet algorithme SIRT-TV. Cet algorithme contient deux boucles imbriquées : la boucle extérieure qui répète les trois étapes, et la boucle intérieure correspondant à la régularisation TV itérative de la deuxième étape.

Algorithme 3: SIRT-TV

$f^{(0)} \leftarrow 0,\ Niter \leftarrow 300,\ NTViter \leftarrow 100$
$\lambda \leftarrow 0.2,\ \alpha \leftarrow 0.2,\ t^{(0)} \leftarrow 1$
 # SIRT-FISTA-TV iteration
for $k \in [1, 2, \cdots, Niter]$ **do**
    # SIRT update
    $f^{(k)} = \texttt{SIRT}\left(f^{(k-1)}\right)$
    # TV minimization
    $f^{(k)} = \texttt{TVDenoise}(f^{(k)}, \alpha, NTViter)$
    # FISTA acceleration
    $t^{(k)} = \frac{1+\sqrt{1+4\left(t^{(k-1)}\right)^2}}{2}$
    $f^{(k)} = \texttt{FISTA}\left(f^{(k)}, f^{(k-1)}, t^{(k)}, t^{(k-1)}\right)$

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Nous pouvons remarquer que bien que SIRT soit classé parmi les méthodes algébriques, il peut également être vu comme un algorithme EM maximisant la vraisemblance pour des données gaussiennes (Yan *et al.*; 2011). De manière équivalente, cela correspond à la minimisation d'un terme quadratique de fidélité des données, comme vu dans le paragraphe 4.

3 MLEM régularisé

De la même façon que pour SIRT, l'algorithme MLEM peut être régularisé à l'aide de la variation totale :

$$\min_{f \geq 0} KL(Af, p) + \alpha TV(f) \tag{3.40}$$

avec α le paramètre TV et KL la divergence de Kullback-Leibler définie par :

$$KL(Q|P) = \sum_i Q(i) \ln\left(\frac{Q(i)}{P(i)}\right) \tag{3.41}$$

pour deux distributions de probabilités discrètes P et Q données. Ce problème effectue la reconstruction tomographique régularisée de données suivant une distribution de Poisson.

Sawatzky *et al.* (2008) ont proposé un schéma itératif implicite où la reconstruction est effectuée avec un pas MLEM et le volume est débruité grâce à un pas TV :

$$\begin{aligned} f^{\left(k+\frac{1}{2}\right)} &= \frac{f^{(k)}}{{}^tA\mathbf{1}} {}^tA\left(\frac{p}{Af^{(k)}}\right) \\ f^{(k+1)} &= f^{\left(k+\frac{1}{2}\right)} - \alpha \frac{f^{(k)}}{{}^tA\mathbf{1}} h^{(k+1)} \end{aligned} \tag{3.42}$$

avec $h^{(k+1)} \in \partial TV\left(f^{(k+1)}\right)$. Le vecteur $\mathbf{1}$ est le vecteur composé de 1 et de dimension donnée en indice.

Nous pouvons aussi citer le travail de Yan *et al.* (2011) qui effectue avec la même idée, une reconstruction du même type que SIRT-TV avec des données distribuées selon une loi de Poisson.

Le débruitage est un problème de minimisation que Maxim *et al.* (2018) ont proposé de réécrire de la manière suivante :

$$f^{(k+1)} = \arg\min_f \{H(f) + \alpha TV(f)\} \tag{3.43}$$

avec $H(f) = \left\langle f - f^{\left(k+\frac{1}{2}\right)} \log f, s \right\rangle$ et $s = {}^tA\mathbf{1}$ la sensibilité.

En utilisant le théorème de Fenchel-Rockafeller, nous obtenons l'itération :

$$f^{(k+1)} = \frac{s f^{\left(k+\frac{1}{2}\right)}}{s + \alpha \mathrm{div}\varphi^*} \tag{3.44}$$

avec φ^* obtenu itérativement avec le schéma suivant :

$$\varphi^{(i+1)} = \frac{\varphi^{(i)} - \tau z^{(k)}}{1 + \tau |z^{(k)}|} \tag{3.45}$$

à partir d'une initialisation $\varphi^{(0)}$ et avec la notation $z^{(k)} = \nabla\left(\frac{s f^{\left(k+\frac{1}{2}\right)}}{s + \mathrm{div}\varphi^{(k)}}\right)$.

La convergence de cet algorithme est prouvée dans (Maxim *et al.*; 2018) sous les conditions suivantes :

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

- $\alpha \leq \frac{s_{\min}}{4}$ et $\tau \leq \frac{(s_{\min} - 4\alpha)^2}{8\alpha \|sf^{(k+\frac{1}{2})}\|_\infty}$ en 2D
- $\alpha \leq \frac{s_{\min}}{6}$ et $\tau \leq \frac{(s_{\min} - 6\alpha)^2}{12\alpha \|sf^{(k+\frac{1}{2})}\|_\infty}$ en 3D

$s_{\min}$ étant la valeur la plus faible de s.

Nous utiliserons une version modifiée de cet algorithme dans laquelle nous avons ajouté une accélération FISTA dans le débruitage TV. Cet algorithme sera appelé MLEM-TV dans la suite.

4 Algorithme de Chambolle-Pock en tomographie

L'algorithme de Chambolle-Pock (Chambolle et Pock; 2011) s'applique au problème convexe (3.1). Sidky *et al.* (2012) ont appliqué cet algorithme dans plusieurs problèmes d'optimisation pour la reconstruction tomographique : en prenant la norme l_1, la norme quadratique et la divergence de Kullback-Leibler pour le terme d'attache aux données, avec et sans régularisation TV. Nous nous intéressons à la minimisation dans le cas suivant :

$$\arg\min_{f \geq 0} KL(Af, p) + \alpha TV(f) \tag{3.46}$$
$$\Leftrightarrow \arg\min_{f \geq 0} \sum_i \left[Af - p + p \log p - p \log\left(pos(Af)\right)\right]_i + \alpha TV(f)$$

où $pos(x) = \max(0, x)$ et avec toutes les opérations vectorielles effectuées termes à termes.

L'algorithme de Chambolle-Pock est appliqué sur les fonctions suivantes :

$$\begin{aligned} F(y, z) &= F_1(y) + F_2(z) \\ F_1(y) &= \sum_i \left[y - p + p \log p - p \log(pos(y))\right]_i \\ F_2(z) &= \alpha TV(z) \end{aligned}$$

ce qui revient à remplacer A dans l'équation primale (3.18) par (A, ∇) avec $G(f) = 0$. Le pseudo-code de l'algorithme obtenu est donné dans l'algorithme 4. Les indices P et I correspondent respectivement aux domaines des projections et des images. $|M|$ est la matrice formée en prenant la valeur absolue de chaque élément de M. Exceptés les multiplications matrices-vecteurs, les opérations sont effectuées terme à terme.

Algorithme 4: KL-TV non pré-conditionné

$L \leftarrow \|(A, \nabla)\|_2$; $\tau \leftarrow 1/L$; $\sigma \leftarrow 1/L$; $\theta \leftarrow 1$; $k \leftarrow 0$
$f^{(0)}, g^{(0)}, q^{(0)} = \mathbf{0}$
$\bar{f}^{(0)} \leftarrow f^{(0)}$
while $k \leq N$ **do**
  $g^{(k+1)} \leftarrow \frac{1}{2}\left(\mathbf{1}_P + g^{(k)} + \sigma A \bar{f}^{(k)} - \sqrt{(g^{(k)} + \sigma A \bar{f}^{(k)} - \mathbf{1}_P)^2 + 4\sigma p}\right)$
  $q^{(k+1)} \leftarrow \lambda\left(q^{(k)} + \sigma \nabla \bar{f}^{(k)}\right) / \max\left(\alpha \mathbf{1}_I, |q^{(k)} + \sigma \nabla \bar{f}^{(k)}|\right)$
  $f^{(k+1)} \leftarrow f^{(k)} - \tau^t A g^{(k+1)} + \tau div\, q^{(k+1)}$
  $\bar{f}^{(k+1)} \leftarrow f^{(k+1)} + \theta\left(f^{(k+1)} - f^{(k)}\right)$
  $k \leftarrow k + 1$

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cet algorithme est efficace mais sa convergence est assez lente. Pock et Chambolle (2011) ont proposé une version pré-conditionnée de l'algorithme de Chambolle-Pock qui consiste à remplacer les paramètres σ et τ par les matrices suivantes :

$$\begin{aligned} \sigma \leftarrow \Sigma &= \frac{\mathbf{1}_P}{|(A, \nabla)|\mathbf{1}_I} \\ \tau \leftarrow T &= \frac{\mathbf{1}_I}{|{}^t(A, \nabla)|\mathbf{1}_P} \end{aligned} \tag{3.47}$$

Nous utiliserons cette version de l'algorithme pour nos expérimentations et la noterons KL-TV.

Sidky *et al.* affirment que pour de faibles valeurs de α, l'algorithme pré-conditionné converge plus rapidement que l'algorithme 4.

Il existe d'autres façons d'accélérer l'algorithme de Chambolle-Pock en modifiant les paramètres d'initialisation, nous pouvons citer par exemple (Chambolle et Pock; 2016; Qiao *et al.*; 2019).

III Application au fantôme de Shepp-Logan 2D

Nous appliquons les algorithmes régularisés présentés précédemment : SIRT-TV (Banjak *et al.*; 2018), une version accélérée de MLEM-TV (Maxim *et al.*; 2018) avec FISTA et la version pré-conditionnée de KL-TV (Sidky *et al.*; 2012). Nous les comparons avec la rétro-projection filtrée et la reconstruction itérative avec MLEM sans régularisation. Les algorithmes sont appliqués sur le fantôme 2D de Shepp-Logan, sur des simulations corrompues par un bruit mixte gaussien et Poisson. Nous les comparons en terme de qualité de reconstruction et de temps de calcul.

Les algorithmes utilisés ont donc pour but de reconstruire des volumes à partir de projections bruitées par un bruit gaussien ou Poisson. Peu d'études se concentrent sur un bruit mixte gaussien-Poisson (Benvenuto *et al.*; 2008; Jezierska *et al.*; 2012; Calatroni *et al.*; 2017), qui est pourtant assez représentatif du bruit présent dans les données réelles. La prise en compte simultanée des deux types de bruit aboutit à des fonctionnelles complexes et des paramètres supplémentaires à optimiser (Calatroni *et al.*; 2017). Cette thèse s'intéresse à la reconstruction tomographique en imagerie dentaire en faible dose, et ce type de bruit peut être présent. Nous espérons avec cette étude simplifiée, investiguer la performance d'algorithmes plus simples, spécifiques à un unique type de bruit, sur des données mixtes.

Les calculs ont été effectués sur un GPU Quadro P2000 et les résultats de cette étude ont été publiés dans (Leuliet *et al.*; 2021).

1 Présentation des données

Pour les simulations, nous avons utilisé le fantôme de Shepp-Logan, avec une taille de 256×256 pixels, représenté dans la figure 3.2. Nous considérons une géométrie parallèle 2D et 180 projections entre 0 et 180°. Les projections bruitées ont été simulées à l'aide d'une loi de Poisson de moyenne les projections exactes. Nous avons ensuite ajouté un bruit gaussien additif avec pour écart-type $\sigma = 2\%$ de la valeur maximale des projections.

Après des expériences préliminaires, le nombre d'itérations de la boucle interne a été fixé à 100 pour SIRT-TV et 30 pour MLEM-TV. Afin de comparer la convergence des algorithmes, 700 itérations sont effectuées pour chacun d'entre eux. Pour le paramètre TV, nous avons choisi $\alpha = 1.5$ pour MLEM-TV et KL-TV qui partagent la même fonction coût, et $\alpha = 0.005$ pour SIRT-TV. Le paramètre SIRT, λ, vaut 0.2.

Les opérations de projection et rétro-projection ont été effectuées sur GPU à l'aide de la librairie ASTRA (Van Aarle *et al.*; 2015) en python. Les opérations matricielles, et en particulier

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

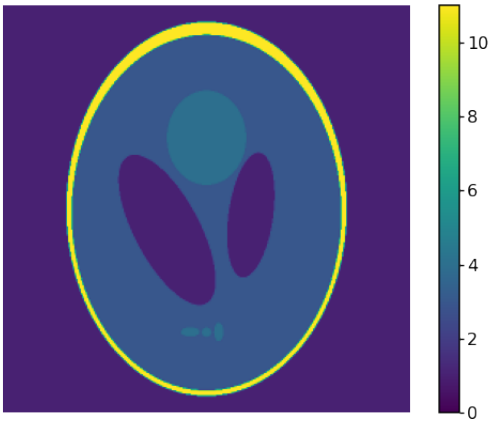


FIGURE 3.2 – Fantôme de Shepp-Logan utilisé pour les simulations.

la régularisation TV, sont également implémentées sur GPU grâce à la librairie CuPy (Okuta *et al.*; 2017).

Nous mesurons la qualité des reconstructions avec l'erreur quadratique moyenne (MSE pour *Mean Squared Error*) définie de la manière suivante pour deux images f et $g \in \mathbb{R}^N$:

$$\mathrm{MSE}(f,g) = \frac{1}{N}\sum_{i=1}^{N}(f_i - g_i)^2 . \tag{3.48}$$

L'évolution de la MSE et de la fonction coût a été tracée par rapport au nombre d'itérations et par rapport au temps d'exécution, dans l'objectif de comparer les vitesses de convergence en termes d'itérations et de temps.

2 Comparaison des résultats

Les reconstructions sont montrées dans la figure 3.3 : la rétro-projection filtrée (FBP) avec le filtre de Hamming, SIRT-TV, MLEM-TV et KL-TV. Le profil de la ligne centrale horizontale du fantôme original et des reconstructions est visible dans la figure 3.4.

A la convergence, la valeur de la MSE est respectivement 0.13 pour MLEM-TV et KL-TV, et 0.16 pour SIRT-TV. Les images montrées ici sont les reconstructions obtenues après convergence des algorithmes.

La figure 3.5 montre que la décroissance de la MSE est plus rapide pour MLEM-TV en termes d'itérations. Cependant, chaque itération de SIRT-TV et MLEM-TV contient non seulement une opération de projection et de rétro-projection, mais également la résolution du problème de minimisation TV de la solution courante. KL-TV n'est fait que d'une unique boucle, chaque itération est donc beaucoup plus rapide. La comparaison entre ces algorithmes par rapport aux itérations est ainsi assez difficile, le temps de calcul dépendant de l'implémentation des opérations de projections et rétro-projections et du nombre d'itérations TV pour MLEM-TV et SIRT-TV.

La ressemblance d'implémentation entre SIRT-TV et MLEM-TV rend la dépendance de temps de calcul par rapport à l'implémentation moindre, mais celle-ci reste très sensible au nombre d'itérations TV choisi. Dans nos expérimentations, nous avons choisi ces nombres de sorte qu'aucune différence ne soit visible dans les images reconstruites si nous choisissions un nombre d'itérations plus élevé. Dans cette configuration, nos résultats et la figure 3.5 suggèrent que MLEM-TV converge plus rapidement que SIRT-TV.

MLEM-TV et KL-TV minimisant la même fonction coût, $J(f) = KL(Af,p) + \alpha TV(f)$, on peut comparer la décroissance de celle-ci pour ces deux algorithmes. La figure 3.6 montre son

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

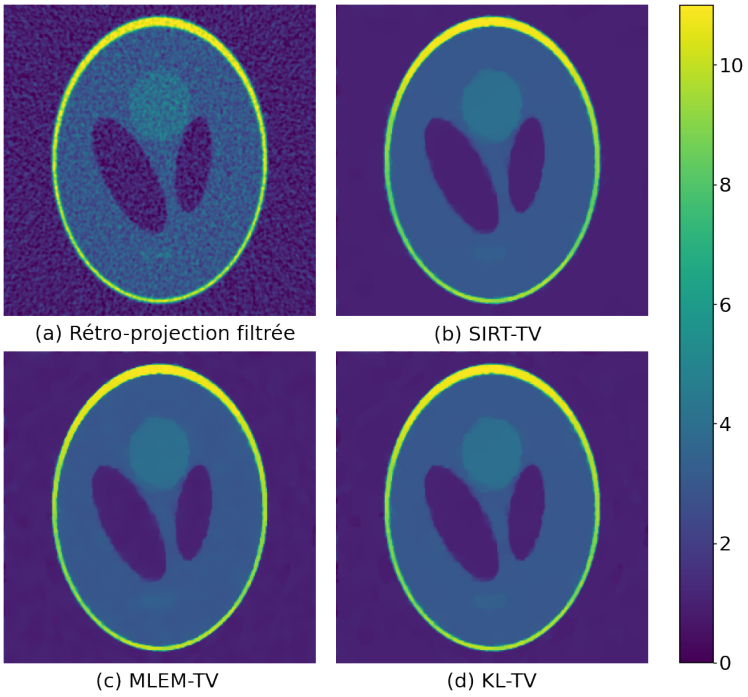


FIGURE 3.3 – Reconstructions du fantôme de Shepp-Logan.

évolution en fonction des itérations et du temps. Selon les itérations, on remarque la même tendance que dans la figure 3.5 : MLEM-TV converge plus rapidement que KL-TV. Mais lorsque l'on regarde le temps, la convergence de MLEM-TV n'est que très légèrement supérieure à celle de KL-TV. Cette comparaison par rapport au temps est plus parlante, surtout si l'on s'intéresse à l'utilisation de ces méthodes dans un cas clinique.

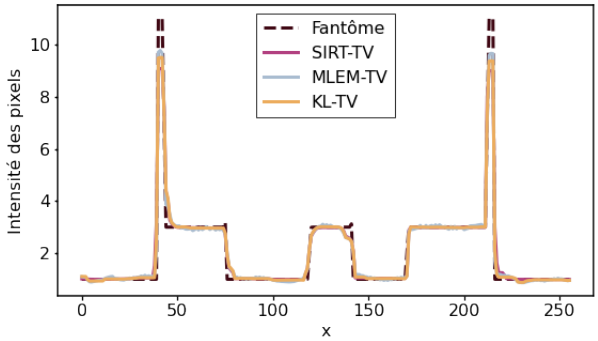


FIGURE 3.4 – Profil du fantôme original et des trois reconstructions itératives.

3 Application à des données expérimentales

Nous avons appliqué ces algorithmes à des données réelles issues de la reconstruction tomographique en microscopie électronique par transmission (TEM). Les données ont été acquises dans le laboratoire IPCMS à Strasbourg et sont les projections 3D parallèles d'un échantillon de CoOCNT. Pour ces expériences, des nanoparticules d'oxyde de cobalt ont été coulées à l'intérieur de canaux de nanotubes de carbone multi-parois par la décomposition

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

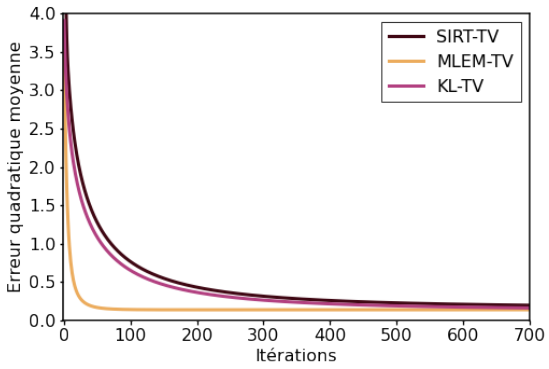
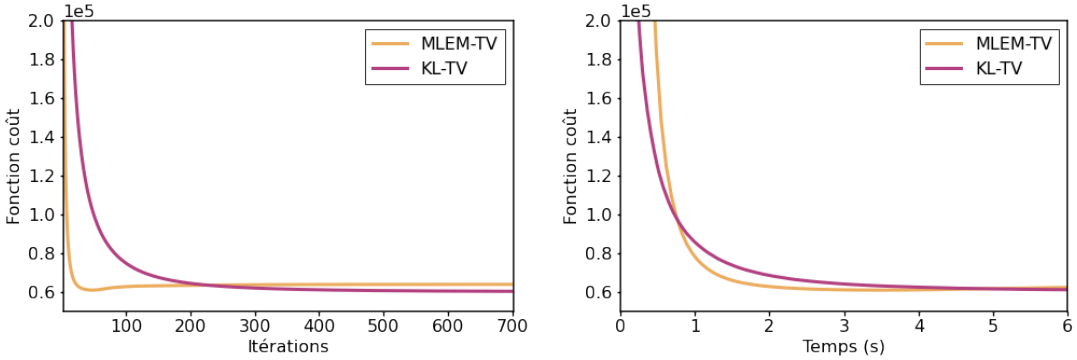


FIGURE 3.5 – Evolution de l'erreur quadratique moyenne selon le nombre d'itérations pour les algorithmes SIRT-TV, KL-TV et MLEM-TV.



(a) (b)

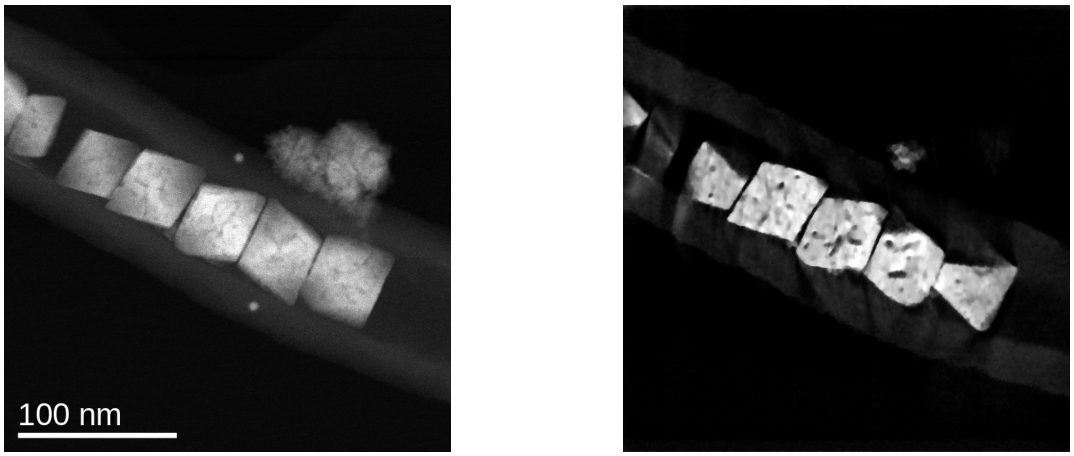
FIGURE 3.6 – Valeurs de la fonction coût pour MLEM-TV et KL-TV en fonction (a) des itérations, (b) du temps en secondes.

thermique contrôlée du stéarate de cobalt en présence d'acide oléique.

Le modèle de bruit mixte gaussien-Poisson est adapté à la modalité d'imagerie TEM (Öktem; 2008). Les reconstructions ont été effectuées à partir de 59 projections de l'échantillon, allant de -70 à 75 degrés. Un exemple d'une projection est visible dans la figure 3.7a. Nous avons reconstruit le volume 3D, composé de 512^3 voxels, en utilisant l'algorithme MLEM-TV, avec 700 itérations et $\alpha = 0.8$. La figure 3.7b représente une coupe de la reconstruction obtenue. Nous avons également reconstruit ce volume avec les algorithmes SIRT-TV et KL-TV, les résultats étaient semblables à celui de MLEM-TV avec une convergence un peu plus rapide pour MLEM-TV.

Nous avons appliqué les trois algorithmes itératifs considérés dans cette étude à des données corrompues par un bruit mixte gaussien et Poisson. Nos expérimentations sur données simulées et données réelles suggèrent que la convergence de SIRT-TV est plus lente que celle de MLEM-TV et KL-TV. Si l'on se réfère aux valeurs de la MSE, SIRT-TV donne des reconstructions de qualité inférieure. Cependant, dans les deux cas considérés, le bruit de Poisson était plus important

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés



(a) (b)

FIGURE 3.7 – (a) Une projection de l'échantillon de CoOCNT. (b) Coupe de la reconstruction MLEM-TV du CoOCNT.

que le bruit gaussien, donc ce résultat peut paraître normal. Des tests de reconstructions sur le fantôme de Shepp-Logan avec des contributions différentes pour le bruit Poisson et gaussien ont montré qu'effectivement, lorsque le bruit gaussien est plus important que le bruit Poisson, l'algorithme SIRT-TV donne de meilleurs résultats que MLEM-TV et KL-TV.

Dans les expérimentations présentées ici, l'algorithme MLEM-TV requiert moins d'itérations que KL-TV pour atteindre la convergence. Cependant, ces résultats dépendant fortement de plusieurs facteurs et peuvent changer d'une application à l'autre. Premièrement, le nombre d'itérations TV a été choisi de sorte que la convergence à l'intérieur de la boucle soit assurée, mais il est possible qu'un nombre plus bas d'itérations TV donne quand même un résultat similaire au regard du nombre d'itérations de la boucle extérieure. De même, afin de réduire le temps de MLEM-TV, nous pourrions effectuer le débruitage TV à partir d'un certain nombre d'itérations et non dès le début, où la reconstruction est encore très floue. Deuxièmement, les résultats dépendent de l'implémentation des algorithmes. Par exemple, l'utilisation d'un GPU au lieu d'un CPU donne des observations différentes du point de vue du temps de calcul. Enfin, le ratio du coût des projections/rétro-projections et de la boucle du débruitage est aussi un élément important.

Une remarque peut être faite sur la mémoire numérique nécessaire aux différents algorithmes. L'algorithme SIRT-TV nécessite au maximum le stockage simultané d'une fois les projections et 9 fois la taille du volume. L'algorithme MLEM a besoin d'une fois les projections et 11 fois la taille du volume. Cependant, si on enlève l'accélération FISTA dans la minimisation TV, on obtient la même chose que pour SIRT-TV. KL-TV diffère complètement des deux autres algorithmes dans son implémentation et nécessite au maximum 4 copies de la taille des projections et 7 copies de la taille du volume. Ainsi, dans le cas d'un volume important, il pourra être intéressant d'utiliser l'algorithme KL-TV ou MLEM-TV sans l'accélération FISTA du TV. Et au contraire, si les projections sont nombreuses, l'algorithme KL-TV demandera plus de mémoire.

Le choix du paramètre TV est décisif dans la reconstruction finale. Si le paramètre est trop petit, la reconstruction reste bruitée, et au contraire, s'il est trop grand, des détails importants peuvent être perdus. Des études se sont intéressées au calcul d'un paramètre optimal (Ito *et al.*; 2011; Bertero *et al.*; 2010). Nous avons appliqué l'algorithme proposé dans (Ito *et al.*; 2011) qui estime le paramètre optimal itérativement en utilisant les valeurs des reconstructions. Cette

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

méthode nécessite ainsi plusieurs reconstructions complètes successives, ce qui peut rapidement s'avérer assez coûteux en temps, surtout en 3D. Ici, pour la reconstruction du fantôme de Shepp-Logan, les formes à reconstruire étant assez simples et complètement adaptées à la régularisation TV, nous avons choisi manuellement les paramètres pour chaque méthode.

Dans ce travail préliminaire, nous avons évalué les performances de quelques algorithmes itératifs sur des problèmes de reconstructions à partir de projections corrompues par un bruit mixte gaussien-Poisson, acquises en géométrie parallèle. Le chapitre suivant adaptera ces algorithmes à la reconstruction 3D CBCT.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Bibliographie

Anthoine, S., Aujol, J.-F., Boursier, Y. et Melot, C. (2012). Some proximal methods for Poisson intensity CBCT and PET. *Inverse Problems & Imaging*, 6(4):565–598. 53

Banjak, H., Grenier, T., Epicier, T., Koneti, S., Roiban, L., Gay, A.-S., Magnin, I., Peyrin, F. et Maxim, V. (2018). Evaluation of noise and blur effects with SIRT-FISTA-TV reconstruction algorithm : Application to fast environmental transmission electron tomography. *Ultramicroscopy*, 189:109–123. 54, 55, 58

Beck, A. et Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1):183–202. 50

Benvenuto, F., La Camera, A., Theys, C., Ferrari, A., Lanteri, H. et Bertero, M. (2008). The study of an iterative method for the reconstruction of images corrupted by Poisson and Gaussian noise. *Inverse Problems*, 24(3):035016. 58

Bertero, M., Boccacci, P., Talenti, G., Zanella, R. et Zanni, L. (2010). A discrepancy principle for poisson data. *Inverse problems*, 26(10):105004. 62

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J. *et al.* (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122. 52

Calatroni, L., De Los Reyes, J. C. et Schönlieb, C.-B. (2017). Infimal convolution of data discrepancies for mixed noise removal. *SIAM Journal on Imaging Sciences*, 10(3):1196–1233. 58

Chambolle, A. (2004). An algorithm for total variation minimization and applications. *Journal of Mathematical imaging and vision*, 20(1-2):89–97. 54

Chambolle, A., Caselles, V., Cremers, D., Novaga, M. et Pock, T. (2010). An introduction to total variation for image analysis. *Theoretical foundations and numerical methods for sparse recovery*, 9(263-340):227. 52

Chambolle, A. et Pock, T. (2011). A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, 40(1):120–145. 50, 51, 57

Chambolle, A. et Pock, T. (2016). An introduction to continuous optimization for imaging. *Acta Numerica*, 25:161–319. 49, 58

Ciarlet, P. G. (1990). *Introduction a l'analyse numerique matricielle et a l'optimisation*. masson. 47

Gabay, D. et Mercier, B. (1976). A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & mathematics with applications*, 2(1):17–40. 51

Glowinski, R. et Marroco, A. (1975). Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. *Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique*, 9(R2):41–76. 51

Ito, K., Jin, B. et Takeuchi, T. (2011). A regularization parameter for nonsmooth Tikhonov regularization. *SIAM Journal on Scientific Computing*, 33(3):1415–1438. 62

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Jezierska, A., Chouzenoux, E., Pesquet, J.-C. et Talbot, H. (2012). A primal-dual proximal splitting approach for restoring data corrupted with Poisson-Gaussian noise. *In 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1085–1088. IEEE. 58

Leuliet, T., Friot--Giroux, L., Baaziz, W., Bretin, E., Ersen, O., Peyrin, F., Sixou, B. et Maxim, V. (2021). Efficiency of TV-regularized algorithms in computed tomography with Poisson-Gaussian noise. *In 2020 28th European Signal Processing Conference (EUSIPCO)*, pages 1294–1298. 58

Maxim, V., Feng, Y., Banjak, H. et Bretin, E. (2018). Tomographic reconstruction from Poisson distributed data : a fast and convergent EM-TV dual approach. working paper or preprint. 56, 58

Nesterov, Y. (2003). *Introductory lectures on convex optimization : A basic course*, volume 87. Springer Science & Business Media. 49

Öktem, O. (2008). Reconstruction methods in electron tomography. *Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT)*, pages 289–320. 61

Okuta, R., Unno, Y., Nishino, D., Hido, S. et Loomis, C. (2017). CuPy : A NumPy-Compatible Library for NVIDIA GPU Calculations. *In Proceedings of Workshop on Machine Learning Systems (LearningSys) in The Thirty-first Annual Conference on Neural Information Processing Systems (NIPS)*. 59

Persson, M., Bone, D. et Elmqvist, H. (2001). Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography. *Physics in Medicine & Biology*, 46(3):853–866. 53, 54

Pock, T. et Chambolle, A. (2011). Diagonal preconditioning for first order primal-dual algorithms in convex optimization. *In 2011 International Conference on Computer Vision*, pages 1762–1769. 58

Polyak, B. (1987). *Introduction to optimization (translated from the russian)*. Translations Series Publications Division, New York. 49

Qiao, Z., Zhu, Y., Redler, G. et Tang, S. (2019). The integrated acceleration of the Chambolle-Pock algorithm applied to constrained TV minimization in CT image reconstruction. *Inverse Problems in Science and Engineering*, 27(2):237–254. 58

Rockafellar, R. T. (1970). *Convex analysis*, volume 18. Princeton university press. 50

Rudin, L. I., Osher, S. et Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D : nonlinear phenomena*, 60(1-4):259–268. 53

Sawatzky, A., Brune, C., Wubbeling, F., Kosters, T., Schafers, K. et Burger, M. (2008). Accurate EM-TV algorithm in PET with low SNR. *In 2008 IEEE nuclear science symposium conference record*, pages 5133–5137. IEEE. 53, 54, 56

Sidky, E. Y., Jørgensen, J. H. et Pan, X. (2012). Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm. *Physics in Medicine & Biology*, 57(10):3065–3091. 57, 58

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Sidky, E. Y., Kao, C.-M. et Pan, X. (2006). Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. *Journal of X-ray Science and Technology*, 14(2):119–139. 53, 54

Tihonov, A. N. (1963). Solution of incorrectly formulated problems and the regularization method. *Soviet Math.*, 4:1035–1038. 53

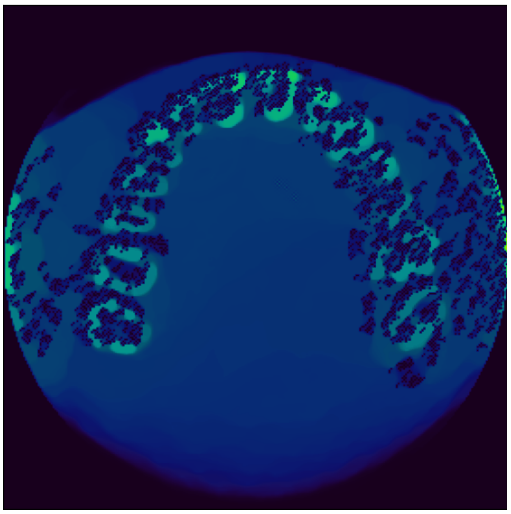
Van Aarle, W., Palenstijn, W. J., De Beenhouwer, J., Altantzis, T., Bals, S., Batenburg, K. J. et Sijbers, J. (2015). The ASTRA Toolbox : A platform for advanced algorithm development in electron tomography. *Ultramicroscopy*, 157:35–47. 58

Wright, S., Nocedal, J. *et al.* (1999). Numerical optimization. *Springer Science*, 35(67-68):7. 47, 49

Yan, M., Chen, J., Vese, L. A., Villasenor, J., Bui, A. et Cong, J. (2011). EM+TV Based Reconstruction for Cone-Beam CT with Reduced Radiation. *In Advances in Visual Computing*, pages 1–10. Springer Berlin Heidelberg. 53, 56

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

4 Reconstructions itératives régularisées en CBCT dentaire faible-dose



I Reconstruction tomographique en imagerie dentaire 68
1 Reconstruction CBCT 68
2 Dose et bruit 69
3 Projections tronquées 69
II Algorithmes et métriques utilisés 70
1 Paramètres de reconstruction 70
2 Métriques d'évaluation 71
III Description des données 73
1 Géométrie d'acquisition 73
2 Acquisitions des données expérimentales 73
3 Fantômes numériques 74
IV Résultats sur fantômes 75
V Résultats sur données expérimentales 81
VI Analyse et conclusion de l'étude 86

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Nous avons déjà mentionné la nécessité de réduire la dose envoyée au patient afin de limiter les risques causés par l'irradiation aux rayons X sur celui-ci. Cependant, plus l'intensité du faisceau de rayons X initial est faible - et de façon générale, la quantité de signal reçue par le capteur pendant la durée d'intégration des rayons X - plus le bruit est important dans les projections et donc dans les reconstructions. De nombreux algorithmes de débruitage d'images ont déjà été développés (Fan *et al.*; 2019), que l'on peut utiliser pour débruiter les reconstructions effectuées avec l'algorithme FDK ou les algorithmes itératifs sans régularisation (Diwakar et Kumar; 2018). Parmi ces algorithmes nous pouvons citer *Non-Local Means* (Buades *et al.*; 2005), BM3D (Dabov *et al.*; 2007) ou encore avec une régularisation TV (Rudin *et al.*; 1992; Chambolle; 2004). De notre côté, nous nous intéressons aux algorithmes itératifs régularisés qui effectuent la régularisation pendant la reconstruction en débruitant la reconstruction actuelle à chaque itération du processus de reconstruction itérative.

Bien qu'un large panel d'algorithmes itératifs soit disponible, ils ont la plupart du temps été appliqués en 2D seulement et plus rarement en CBCT. La géométrie spécifique du CBCT, couplée avec les particularités de l'imagerie dentaire apporte des difficultés dans le processus de reconstruction qui ne sont pas présentes en 2D. Nous pouvons, d'une part, souligner les ressources numériques importantes nécessaires à la reconstruction 3D. Ces ressources sont encore un problème pour les méthodes itératives actuellement malgré l'amélioration continue de la performance des GPUs. D'autre part, il faut également faire face aux problèmes induits par la troncature des projections lorsque la tête du patient est plus petite que la taille du détecteur.

Dans ce chapitre, nous comparons différents algorithmes itératifs avec régularisation TV dans le but de reconstruire des volumes à partir de projections tronquées, acquises à faible et très faible dose. Bien que notre comparaison ne soit pas exhaustive, elle devrait donner un bon aperçu du potentiel de ces méthodes. Nous donnerons les détails de leur adaptation à la géométrie conique et aux projections tronquées. Nos objectifs sont :

- évaluer la qualité des images reconstruites,
- évaluer les temps de calcul et les ressources numériques nécessaires,
- déterminer les améliorations potentielles.

Pour atteindre ces objectifs, nous avons appliqué ces algorithmes à des fantômes et des données expérimentales.

La plupart des résultats présentés ici ont été publiés dans l'article (Friot--Giroux *et al.*; 2022).

I Reconstruction tomographique en imagerie dentaire

1 Reconstruction CBCT

La tomographie à faisceau conique (CBCT) est une technique d'imagerie qui reconstruit un volume 3D à partir de projections 2D de celui-ci. En imagerie dentaire, le CBCT permet de réduire la dose envoyée au patient et de diminuer le temps d'acquisition comparé aux scanners conventionnels, rendant l'utilisation de ces scanners beaucoup plus fréquente (Tsiklakis *et al.*; 2005; Liang *et al.*; 2010).

La reconstruction en CBCT est encore largement effectuée avec les méthodes analytiques, permettant au praticien d'obtenir les images reconstruites en des temps très courts. L'algorithme FDK (Feldkamp *et al.*; 1984) est le plus utilisé malgré sa sensibilité au bruit et au nombre de projections. Les reconstructions avec cette méthode ont donc tendance à être bruitées et dégradées par de nombreux artefacts dans le cas d'acquisitions à faible dose. Il est utilisé seul

Image de couverture : Échec difficilement contestable d'une reconstruction itérative néanmoins très artistique.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

en haute dose (Baba *et al.*; 2004) ou associé à des pré et/ou post-traitements afin de réduire les artefacts et améliorer la qualité des images. Par exemple, les artefacts métalliques dans (Ibraheem *et al.*; 2012), ou encore le durcissement de faisceau et le bruit dans (Bayaraa *et al.*; 2020).

2 Dose et bruit

La dose de rayons X envoyée par le scanner est le premier facteur à jouer sur la qualité de l'image en termes de bruit. La baisse de la dose entraîne l'augmentation du bruit dans les mesures et ensuite dans les reconstructions si aucune correction n'est effectuée. Un points de recherche crucial en tomographie est d'obtenir des reconstructions de très bonne qualité, à partir de projections acquises en faible dose.

En reconstruction tomographique, les projections sont corrompues par du bruit ayant différentes origines, les plus importantes étant : le bruit électronique modélisé par un bruit gaussien, et le bruit photonique, supposé de Poisson. Le bruit photonique, également appelé bruit quantique, provient de la fluctuation statistique lors du comptage des photons X par le détecteur. Plus il y a de photons, et plus la mesure sera précise. Ainsi, le moyen le plus évident de réduire le bruit photonique est d'augmenter le nombre de photons X transmis et donc d'augmenter la dose (Diwakar et Kumar; 2018). Quant au bruit électronique, il provient des circuits électriques entre le détecteur et l'ordinateur, lors de la conversion signal analogique / signal électrique.

En tomographie par transmission, lorsque la dose est élevée, le bruit est souvent supposé uniquement gaussien par simplicité. Dans le cas d'acquisitions à faible dose, la nature poissonienne du comptage des photons doit être prise en compte. Le bruit n'est plus seulement additif et sa modélisation est une tâche complexe (Yu *et al.*; 2012; Leuschner *et al.*; 2021). Afin de prendre tout cela en compte, nos simulations incluent un bruit mixte gaussien-Poisson représentatif à la fois du mode d'acquisition faible dose et du bruit électronique.

La plupart des scanners CBCT proposent des modes de scan faible dose, utilisés lorsqu'il n'est pas nécessaire d'avoir des reconstructions d'excellente qualité. Les reconstructions obtenues doivent cependant encore permettre au médecin d'effectuer un diagnostic correct. Les expériences de cette partie consistent à appliquer les algorithmes itératifs présentés dans le chapitre 3 en CBCT, sur des données faibles doses mais également très faibles doses. L'obtention des données dans ces deux cas sera détaillée dans la suite.

3 Projections tronquées

Un des challenges de la reconstruction CBCT dentaire est la reconstruction à partir de projections tronquées. Ces troncatures ont lieu lorsque l'on ne veut reconstruire qu'un petit volume de la mâchoire, mais également dès que la tête du patient est plus grande que la taille du détecteur, ce qui est presque tout le temps le cas. Les projections tronquées non corrigées entraînent des artefacts dans l'image reconstruite, créent des contours clairs/sombres mais également des erreurs basses fréquences dans l'image qui gênent l'identification des tissus par leur niveau de gris et peuvent empêcher la segmentation de l'image par les programmes qui la manipulent.

Afin de réduire cet effet, plusieurs méthodes ont été développées. Des méthodes extrapolant les projections ont été proposées pour les reconstructions analytiques. Nous pouvons citer (Ohnesorge *et al.*; 2000) qui proposent un agrandissement des projections avec une symétrie miroir, ou (Hsieh *et al.*; 2004) où les projections sont étendues à l'aide de projections d'un cylindre d'eau supposé représenter la tête du patient.

Pour les reconstructions itératives, la stratégie la plus commune consiste à agrandir le champ de vue en reconstruisant un volume plus large (Dang *et al.*; 2016; Aootaphao *et al.*; 2022).

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

De cette manière, les informations issues des parties tronquées sont envoyées dans l'espace supplémentaire.

II Algorithmes et métriques utilisés

Le cadre de la reconstruction tomographique étant posé, les détails des méthodes utilisées pour résoudre ce problème sont décrits dans cette section.

1 Paramètres de reconstruction

Nous rappelons que les algorithmes utilisés visent à résoudre le problème suivant :

$$f^* = \arg\min_f d(p, Af) + \alpha TV(f) \tag{4.1}$$

avec $\alpha \in \mathbb{R}^+$ le paramètre de régularisation. La distance d est soit la norme 2 dans le cas de SIRT-TV, soit la divergence de Kullback-Leibler pour MLEM-TV et KL-TV. Nous appliquons ces trois algorithmes que nous comparons à MLEM sans régularisation et à l'algorithme analytique FDK, sur fantôme et données réelles, en faible dose et très faible dose. Les quelques modifications ajoutées aux algorithmes sont décrites dans ce qui suit.

Projections tronquées

Afin d'atténuer l'effet de la troncature et suivant l'idée de Maltz *et al.* (2007), nous avons étendu les projections en considérant le FoV inclus dans un cylindre pour la reconstruction FDK.

Dans nos expériences, cette technique n'a pas donné de résultats satisfaisants avec les algorithmes itératifs. Nous avons donc simplement reconstruit des volumes étendus comme dans (Dang *et al.*; 2016). La taille initiale des volumes était de $350 \times 275 \times 275$ voxels, nous avons ajouté 150 voxels sur chaque côté, reconstruisant ainsi des volumes de taille $650 \times 575 \times 575$ voxels. Le nombre de voxels ajoutés a été choisi de sorte à ce que l'artefact de troncature soit le moins visible possible, tout en gardant une taille de volume suffisamment petite pour que les volumes rentrent en mémoire et que les temps de calcul soient limités. Après la reconstruction, le volume est tronqué pour retrouver sa taille initiale.

Sensibilité

Contrairement à la géométrie parallèle, en CBCT, le nombre de rayons traversant un voxel varie à travers le volume. Une dose plus importante est reçue par la région d'intérêt (ROI). Mathématiquement, ceci se reflète dans la sensibilité définie par :

$$s = {}^tA\mathbf{1}. \tag{4.2}$$

Le terme "sensibilité" est plus connu en tomographie par émission où il représente la probabilité pour un photon émis dans le patient d'être détecté sur le détecteur.

Ainsi, en CBCT, la sensibilité est loin d'être constante (voire figure 4.1) et on observe même un écart considérable entre les valeurs à l'intérieur de la ROI et celles dans les coins du volume. L'écart devient encore plus important lorsque les volumes sont étendus pour diminuer les artefacts de projections tronquées.

Dans l'algorithme MLEM-TV, comme le pas de minimisation τ dépend de $s_{\min}$, la valeur minimale de la sensibilité, la convergence est ralentie de manière significative sur l'intégralité

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

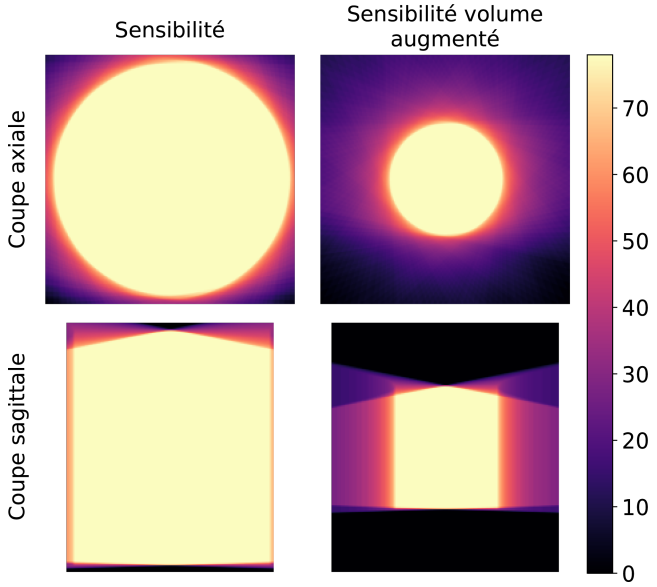


FIGURE 4.1 – Coupes axiales et sagittales de la sensibilité, pour le volume normal et pour le volume augmenté.

du volume, même si les valeurs faibles de la sensibilité se trouvent dans des régions tronquées après la reconstruction. En suivant l'idée de l'algorithme de Chambolle-Pock pré-conditionné (Pock et Chambolle; 2011), nous avons essayé une nouvelle version de l'algorithme donné dans (3.45) où τ est remplacé par la matrice suivante, avec opérations vectorielles effectuées termes à termes :

$$T = 0.9 \times \frac{(s - 6\alpha)^2}{12\alpha s f^{(k+\frac{1}{2})}}. \tag{4.3}$$

Cette formulation permet d'accélérer la convergence dans le centre du volume et donc dans la ROI.

Implémentation

Les algorithmes ont été implémentés en Python, et comme dans le chapitre précédent, les opérations de projection et rétro-projection sont effectuées sur GPU à l'aide la librairie ASTRA, et le reste des opérations matricielles sur GPU à l'aide de la librairie CuPy. Les reconstructions ont été faites sur un processeur Intel Xeon Gold 6226R avec un GPU Tesla V100.

2 Métriques d'évaluation

Pour les reconstructions des fantômes, puisque nous avons accès au volume initial f_{ref}, nous évaluons nos méthodes à l'aide des métriques suivantes :

- La racine de l'erreur quadratique moyenne normalisée (NRMSE [1]) définie par :

$$\text{NRMSE}(f, f_{\text{ref}}) = \frac{\|f - f_{\text{ref}}\|_2}{\|f_{\text{ref}}\|_2}$$

  La NRMSE calcule l'erreur quadratique cumulée entre le volume reconstruit et la vérité terrain. Une faible valeur de NRMSE indique une image reconstruite proche de la solution idéale.

1. *Normalized Root Mean Squared Error*

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

- Le PSNR [1] reflète la différence du niveau de bruit, en dB, entre deux images :

$$\mathrm{PSNR}(f, f_{\mathrm{ref}}) = 10 \log_{10} \left(\frac{\Delta^2}{\frac{1}{J} \sum_{j=1}^{J} (f_j - f_{\mathrm{ref},j})^2} \right),$$

  avec Δ la dynamique de donnée de f_{ref} (ie. la valeur maximale que peut prendre un voxel), et J le nombre total de voxels. Une grande valeur de PSNR indique une meilleure qualité de reconstruction.

- La SSIM [2] est utilisée pour mesurer la similarité entre deux images, en comparant la luminance, le contraste et les structures sur des patchs extraits des deux images (Wang *et al.*; 2004) :

$$\mathrm{SSIM}(f, f_{\mathrm{ref}}) = \frac{(2\mu_f \mu_{f_{\mathrm{ref}}} + c_1)(2\sigma_{f f_{\mathrm{ref}}} + c_2)}{(\mu_f^2 + \mu_{f_{\mathrm{ref}}}^2 + c_1)(\sigma_f^2 + \sigma_{f_{\mathrm{ref}}}^2 + c_2)},$$

  avec μ_i la moyenne de i, σ_{ij} la covariance entre i et j, $c_1 = 0.01 \times \Delta$ et $c_2 = 0.03 \times \Delta$, Δ défini comme pour le PSNR.

Pour les données expérimentales, la vérité terrain est une reconstruction FDK avec une dose normale. Cependant, la plage des valeurs de cette reconstruction étant différente de celles des nôtres, nous ne pouvons pas utiliser les mêmes métriques que pour les fantômes. Widmann *et al.* (2016) ont montré que le choix de la méthode de reconstruction avait un impact sur les valeurs HU des reconstructions, en particulier en faible-dose, et que les reconstructions obtenues avec les méthodes itératives montraient des valeurs différentes. Ils utilisent alors le CNR [3] et la corrélation pour comparer les reconstructions avec l'image de référence. Nous utiliserons donc également ces deux métriques pour nos évaluations.

Pour le calcul du CNR, nous comparons les valeurs moyennes d'un patch à l'intérieur d'une dent (μ_{obj}) avec celle d'un patch du fond de l'image (μ_{BG}), en normalisant par la variance (σ_{BG}) du bruit dans le fond :

$$\mathrm{CNR}(\mathrm{obj}, \mathrm{BG}) = 20 \log_{10} \left(\frac{|\mu_{\mathrm{obj}} - \mu_{\mathrm{BG}}|}{\sigma_{\mathrm{BG}}} \right).$$

Pour évaluer la corrélation, nous calculons les coefficients de corrélation de Pearson entre les images f et f_{ref} de même taille :

$$\mathrm{Corr}(f, f_{\mathrm{ref}}) = \frac{\mathrm{cov}(f, f_{\mathrm{ref}})}{\sigma_f \sigma_{f_{\mathrm{ref}}}},$$

avec μ_f et $\mu_{f_{\mathrm{ref}}}$ respectivement les moyennes de f et f_{ref},

$$\mathrm{cov}(f, f_{\mathrm{ref}}) = \frac{1}{J-1} \sum_{j=1}^{J} \left[(f_j - \mu_f)(f_{\mathrm{ref},j} - \mu_{f_{\mathrm{ref}}}) \right].$$

Nous comparons également les vitesses de convergence des différents algorithmes, en termes de nombre d'itérations et de temps, en observant l'évolution de la fonction coût.

1. *Peak Signal to Noise Ratio*
2. *Structural SIMilarity*
3. *Contrast to Noise Ratio*

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

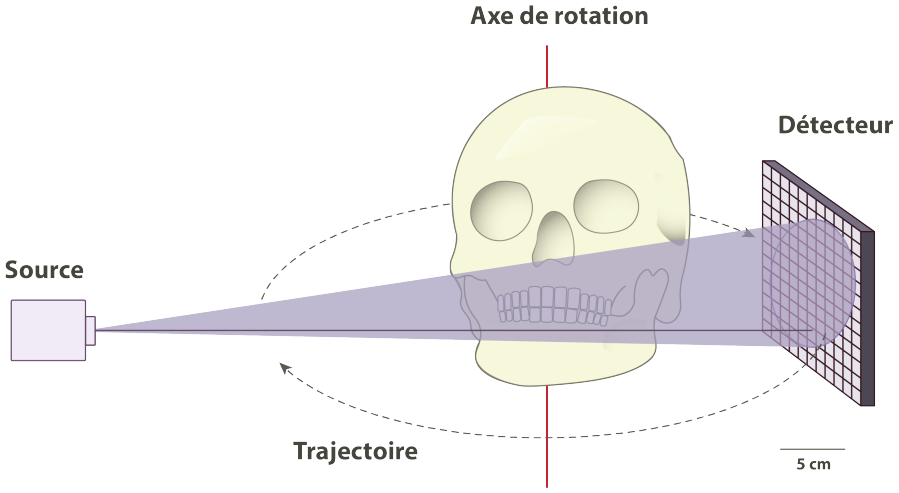


FIGURE 4.2 – Schéma de la géométrie d'acquisition, à l'échelle.

III Description des données

1 Géométrie d'acquisition

La géométrie d'acquisition du scanner CBCT utilisé est illustrée dans la figure 4.2. La source et le détecteur tournent autour de la tête du patient en suivant une trajectoire circulaire. L'axe de rotation est situé à 401.07 mm de la source et la distance source-détecteur est de 564.30 mm. La taille du détecteur est de 12×14 cm, avec 600×700 pixels carrés de 200 µm de côté. La source n'est pas centrée sur le détecteur, elle est décalée en hauteur. Les projections sont acquises dans des conditions de short-scan, c'est-à-dire sur un intervalle incluant [0 ; 180 + angle d'éventail].

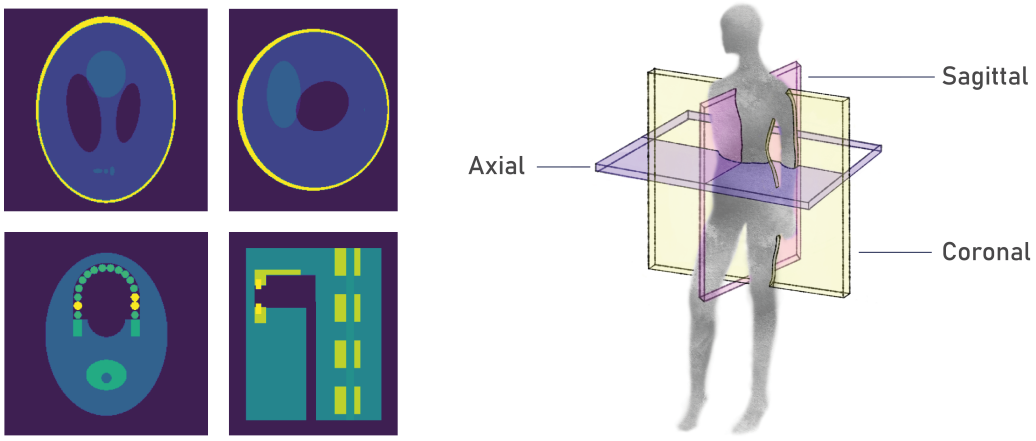
2 Acquisitions des données expérimentales

Des acquisitions CBCT d'un fantôme de tête anthropomorphique livré par W. Loy Gmbh [1] ont été réalisées sur un scanner CS 8200 3D de Carestream Dental. Un ensemble de 155 projections a été acquis avec un protocole faible-dose : 80 kV et 2 mA sur 3.1 s [2]. Afin de réduire la mémoire numérique nécessaire lors de la reconstruction, les projections ont été sous-échantillonnées en 300×350 pixels. Le volume reconstruit est composé de $350 \times 275 \times 275$ voxels cubiques de côté 300 µm, donnant ainsi un volume mesurant $105 \times 82.5 \times 82.5$ mm^3.

Au lieu d'ajouter un bruit simulé sur les projections réelles, nous avons choisi de simuler les acquisitions très faibles doses en diminuant le nombre de projections. Ceci correspond à une baisse de la dose sur des scanners ayant la capacité d'envoyer les rayons X de manière pulsée. Les deux techniques sont utilisées en imagerie médicale pour réduire la dose reçue par le patient (Liu *et al.*; 2015; Humphries *et al.*; 2019). Ainsi, nous considérons une version très faible dose de ces données, où seulement 78 projections sont régulièrement échantillonnées à partir des données faibles doses.

1. https://www.loy-gmbh.de/produkt/dental-phantom/
2. A titre de comparaison, les mêmes valeurs pour une acquisition normale dose peuvent être : 90 kV et 3.2 mA sur 15 s.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés



(a) (b)

FIGURE 4.3 – (a) Coupes axiales et sagittales des fantômes utilisés. En haut, le fantôme de Shepp-Logan et en bas, le fantôme de mâchoire. (b) Schéma de coupes axiale, coronale et sagittale sur un patient.

3 Fantômes numériques

Avant d'être appliqués aux données réelles, les algorithmes ont été testés sur des données simulées. Cela nous a permis de vérifier leur fiabilité en CBCT et d'ajuster les différents paramètres. Nous avons utilisé deux fantômes : le fantôme de Shepp-Logan tri-dimensionnel et un fantôme de mâchoire que nous avons créé en nous inspirant du fantôme FORBILD [1]. Ce fantôme de mâchoire est composé d'un cylindre représentant la tête du patient, différentes formes géométriques qui reproduisent la colonne vertébrale, les mâchoires ainsi que 31 dents. Trois dents de la mâchoire inférieure ont une densité similaire au métal. La figure 4.3a montre des coupes axiales et sagittales des fantômes utilisés et la figure 4.3b illustre dans quel plan se trouvent les coupes axiales, coronales et sagittales en imagerie médicale.

Dans le but d'appliquer par la suite facilement les algorithmes aux données réelles, nous avons choisi de prendre des fantômes de même taille que les volumes des données réelles et de considérer la même géométrie que celle du scanner utilisé pour les données expérimentales. Les volumes simulés ont donc une taille de $105 \times 82.5 \times 82.5$ mm^3 échantillonnés en $350 \times 275 \times 275$ voxels. Un total de 78 projections avec les mêmes angles que pour les données expérimentales (correspondant au cas très faible dose) ont été simulées à l'aide de la librairie ASTRA (Van Aarle *et al.*; 2016) en Python.

Dans le but de représenter au mieux le problème de reconstruction dentaire correspondant aux données expérimentales à notre disposition, les projections des fantômes numériques ont été simulées avec la même géométrie que celle décrite dans la section précédente.

Le niveau de bruit ajouté aux projections a été déterminé à partir des projections des données réelles. Pour simuler le bruit photonique, un bruit de Poisson a été ajouté, et un bruit gaussien a été ajouté pour simuler le bruit électronique. Le niveau de bruit gaussien a été estimé en utilisant la relation statistique entre la variance du bruit et les valeurs propres de la matrice de covariance des patchs extraits des images (Chen *et al.*; 2015). Les valeurs négatives ont été remplacées par des zéros.

Les fantômes ayant des structures assez simples, les reconstructions dans le cas correspon-

1. http://www.imp.uni-erlangen.de/phantoms/

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Méthode	**SIRT-TV**	**MLEM**	**MLEM-TV**	**KL-TV**
N	400	200	300	400
NTV	20		20	
α	9×10^{-6}		0.1	0.1

Tableau 4.1 – Paramètres utilisés pour les reconstructions du fantôme de Shepp-Logan. N est le nombre d'itérations de reconstruction, NTV est le nombre d'itérations TV et α est le paramètre TV utilisé pour déterminer l'importance de la régularisation.

Méthode	**SIRT-TV**	**MLEM**	**MLEM-TV**	**KL-TV**
N	400	200	400	500
NTV	20		20	
α	5×10^{-5}		0.1	0.1

Tableau 4.2 – Paramètres utilisés pour les reconstructions du fantôme numérique de la mâchoire.

dant au faible dose des données réelles étaient suffisamment de bonne qualité pour ne pas faire ressortir d'élément de comparaison intéressant. Nous avons donc choisi de ne présenter que les résultats correspondant au cas très faible dose.

IV Résultats sur fantômes

Dans cette partie, nous analysons les performances des algorithmes FDK, MLEM, SIRT-TV, MLEM-TV et KL-TV sur les données simulées. La plupart des résultats sont redondants entre les deux fantômes et ne seront donnés que pour le fantôme de la mâchoire. Seules les informations pertinentes tirées des résultats sur le fantôme de Shepp-Logan seront présentées.

Les méthodes itératives nécessitent de définir certains paramètres : le nombre d'itérations (N), le nombre d'itérations dans l'étape de débruitage TV (NTV) et le paramètre de régularisation α. Pour SIRT-TV, nous avons également dû déterminer la valeur du paramètre SIRT λ que nous avons fixé à 0.8. Des valeurs plus élevées ne permettaient pas de conserver la convergence de l'algorithme. Tous les paramètres ont, dans un premier temps, été choisis manuellement, en faisant varier leur valeur jusqu'à l'obtention d'un compromis visuellement acceptable entre suppression du bruit et conservation des détails de l'image. Dans un second temps, nous avons essayé de diminuer le nombre d'itérations et la valeurs des paramètres TV, tout en gardant une qualité d'image raisonnable et une valeur de MSE comparable à celle obtenue avec de nombreuses itérations.

La présence de dents en métal dans le fantôme de la mâchoire nécessite de prendre des paramètres différents de ceux pour le fantôme de Shepp-Logan. Les tableaux 4.1 et 4.2 résument les valeurs finales des paramètres pour les différents algorithmes, respectivement pour le fantôme de Shepp-Logan et de mâchoire.

La figure 4.4 montre l'évolution des fonctions coût en fonction du nombre d'itérations et du temps en secondes pour les différents algorithmes. Nous remarquons que l'algorithme SIRT-TV est le plus lent à converger pour cette implémentation. Ce résultat est en accord avec les observations faites dans le cas de la géométrie parallèle 2D du chapitre précédent.

Les courbes représentant la fonction coût semblent indiquer que la convergence est atteinte à partir de 100 itérations pour MLEM-TV et KL-TV, et 300 pour SIRT-TV. Cependant, la figure 4.5

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

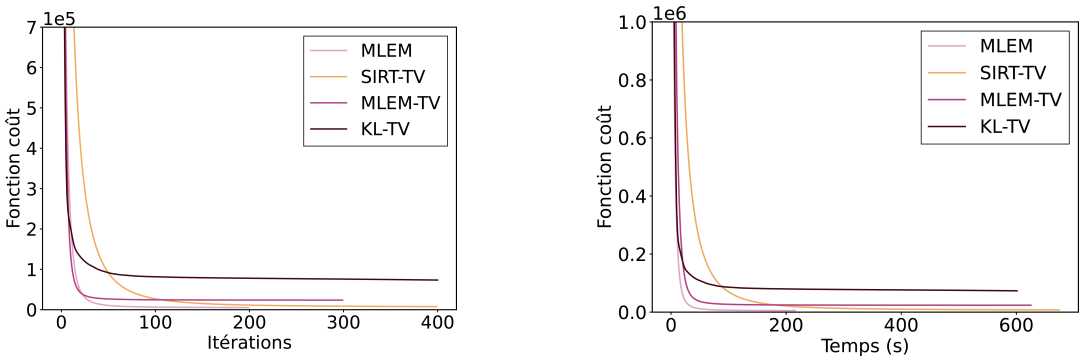


FIGURE 4.4 – Evolution des fonctions coûts lors de la reconstruction du fantôme de mâchoire en très faible dose. A gauche, les fonctions coûts en fonction des itérations, et à droite, en fonction du temps.

qui affiche l'évolution de la MSE en fonction des itérations, montre que la convergence de celle-ci est tresè différente de celle de fonction coût. Pour KL-TV, il faut attendre la 120e itération avant que la MSE ne se stabilise, et pour MLEM-TV et SIRT-TV, elle semble décroissante jusqu'à la fin des itérations déterminées et ne semble ainsi pas avoir convergé.

Visuellement, nous observons une amélioration de la qualité de reconstruction pour un nombre d'itérations allant jusqu'à 400 pour SIRT-TV et KL-TV et 300 pour MLEM-TV. La convergence en termes de fonction coût ne semble donc pas correspondre à un arrêt de l'amélioration de la qualité d'images reconstruite. Aussi, la MSE seule ne semble pas à même de refléter l'impression visuelle qui s'améliore en augmentant les itérations. Nous faisons ainsi le choix de déterminer manuellement le nombre d'itérations à partir de ces observations : nous fixons un critère d'arrêt pour les itérations tels que l'impression visuelle se stabilise.

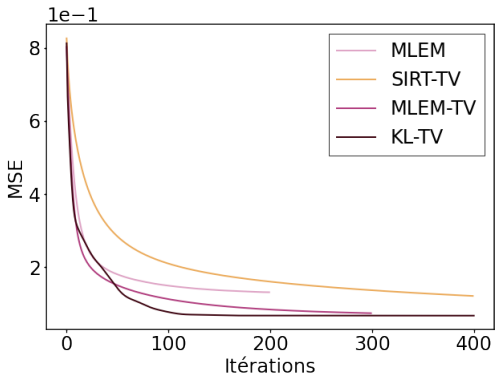


FIGURE 4.5 – Evolution de la MSE lors de la reconstruction du fantôme de mâchoire en très faible dose, en fonction du temps.

Les valeurs des NRMSE, PSNR et SSIM pour les différentes reconstructions sont données dans le tableau 4.3. Toutes les métriques reflètent la qualité inférieure des images fournies par FDK pour un faible nombre de projections. MLEM est moins affecté par les artefacts dus au faible nombre de vues, ce qui explique les bien meilleures valeurs des métriques comparé à FDK. Les métriques pour les trois autres reconstructions sont encore meilleures. Les méthodes statistiques régularisées basé sur une distribution de Poisson des données, obtiennent de meilleurs résultats que SIRT-TV. Ceci suggère que ces algorithmes sont plus appropriés pour le bruit mixte gaussien-Poisson considéré dans cette étude. Ils sont à peu près similaires, avec

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

une performance légèrement supérieure pour KL-TV.

Métriques	FDK	MLEM	SIRT-TV	MLEM-TV	KL-TV
NRMSE	0.248	0.229	0.046	0.031	**0.030**
PSNR	39.684	41.354	50.504	54.328	**57.216**
SSIM	0.841	0.976	0.996	0.998	**0.999**

TABLEAU 4.3 – NRMSE, PSNR et SSIM pour les reconstructions du fantôme de la mâchoire pour les différents algorithmes.

Un profil extrait du centre du volume est visible dans la figure 4.6. Nous pouvons remarquer le bruit présent dans les reconstructions FDK et MLEM et l'influence positive de la régularisation dans les autres méthodes.

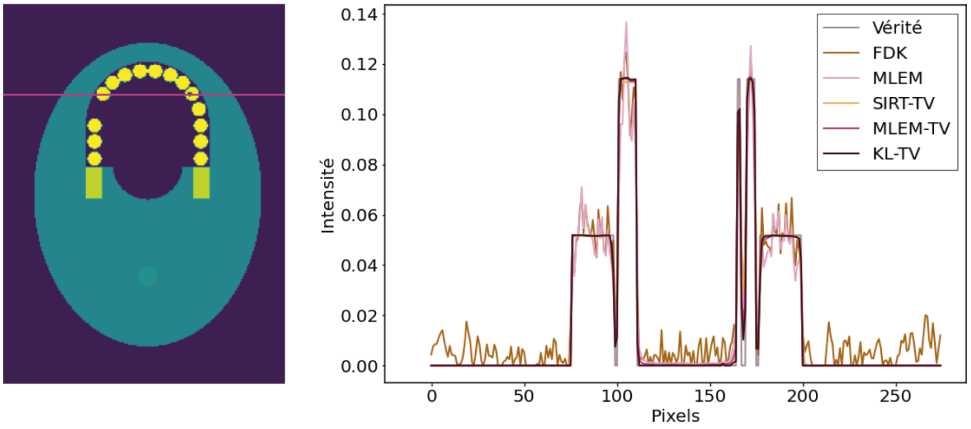


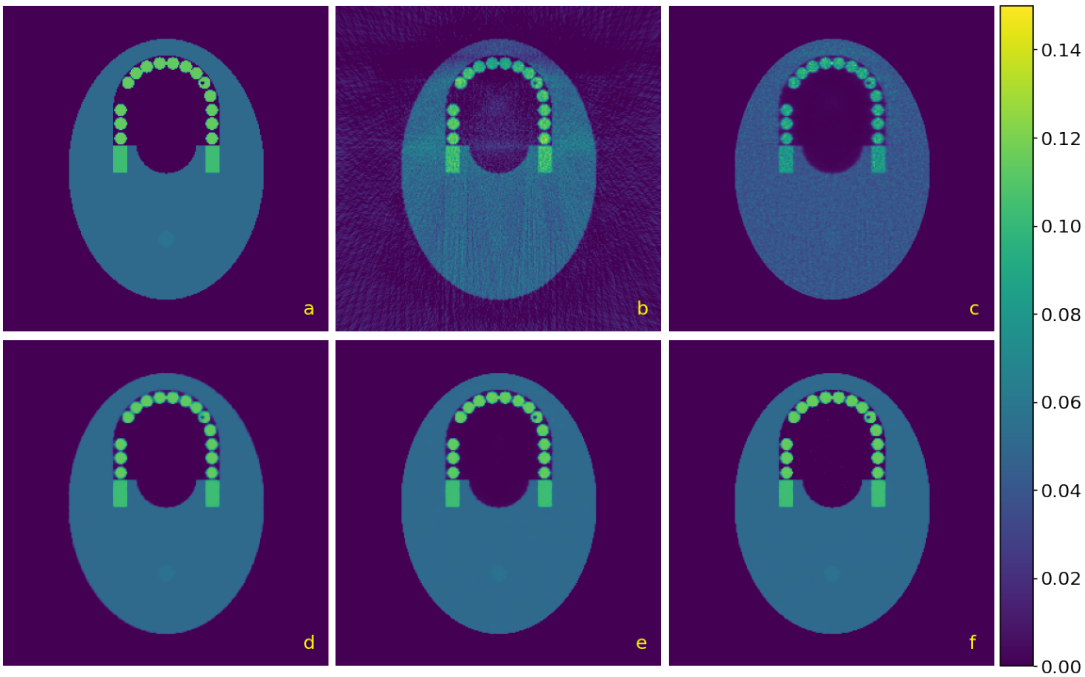
FIGURE 4.6 – A gauche : le profil étudié correspond à la ligne rose. A droite : profil des différentes reconstructions.

La figure 4.7 montre des coupes axiales des volumes reconstruits. Pour la coupe de la figure 4.7a, on peut voir que toutes les méthodes réussissent à reconstruire même les petits détails tels que la carie dans la première prémolaire mandibulaire droite. Les artefacts en étoile dus au faible nombre de projections peuvent être observés dans la reconstruction FDK. La présence de métal les fait ressortir davantage dans la figure 4.7b. Nous n'observons pas ces artefacts dans les reconstructions itératives. En particulier, nous pouvons souligner la capacité de MLEM à les supprimer, même sans utiliser de régularisation.

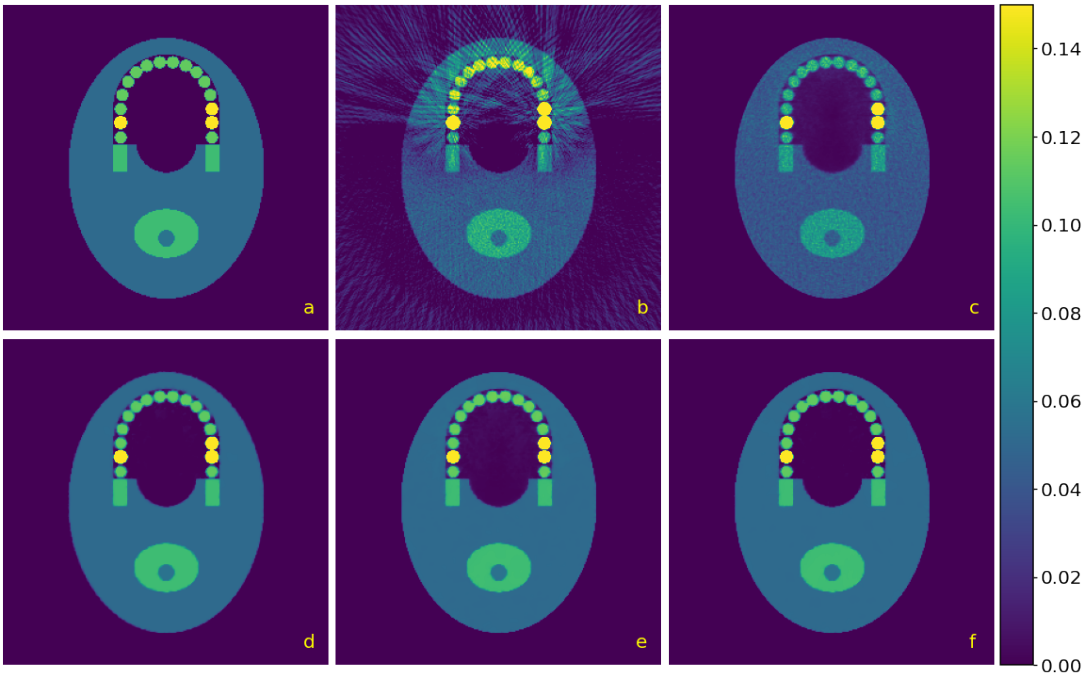
Une coupe coronale est visible dans la figure 4.8a. Nous pouvons voir que l'artefact de cône en haut du volume est particulièrement présent dans l'image FDK. L'algorithme MLEM réduit cet artefact en dehors de l'objet, mais la ligne horizontale qui délimite le fantôme est floue. La figure 4.8b montre un agrandissement de cette zone dans lequel il apparaît que les lignes verticales sont plutôt bien conservées avec les trois méthodes régularisées, alors que les lignes horizontales sont mieux conservées par KL-TV. Les trois méthodes effectuent une reconstruction précise et débruitée du volume, même si SIRT-TV présente des bords moins nets que MLEM-TV et KL-TV.

La figure 4.9 montre les reconstructions des différents algorithmes sur le fantôme de Shepp-Logan. Les mêmes observations que pour les fantômes de la mâchoire peuvent être faites par rapport à FDK et MLEM. Les reconstructions itératives sont assez similaires et ne permettent pas de faire de remarques aussi tranchées que pour le fantôme de mâchoire.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés



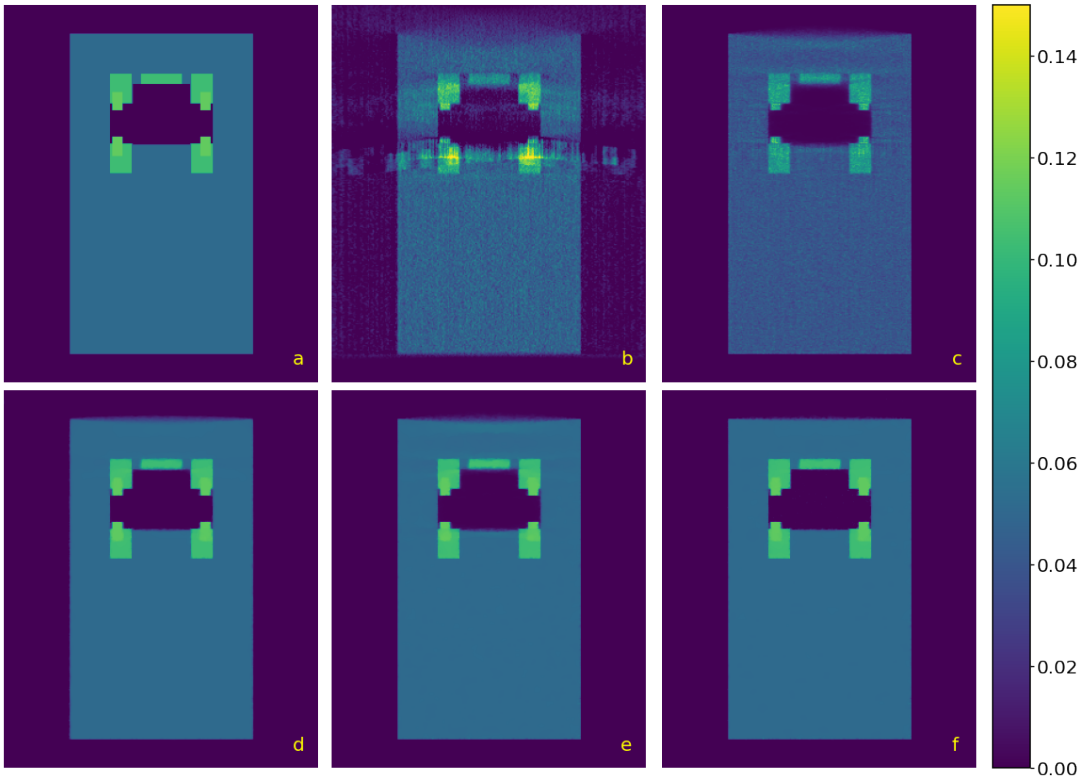
(a) Mâchoire supérieure



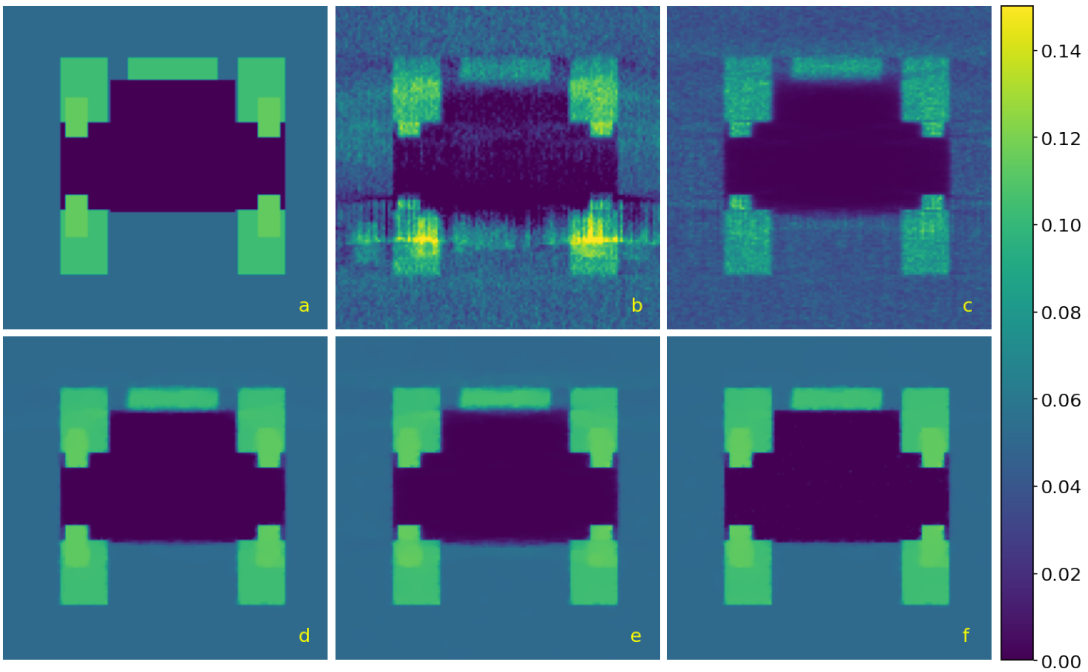
(b) Mâchoire inférieure

FIGURE 4.7 – Coupes axiales du fantôme de mâchoire. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV et (f) KL-TV.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés



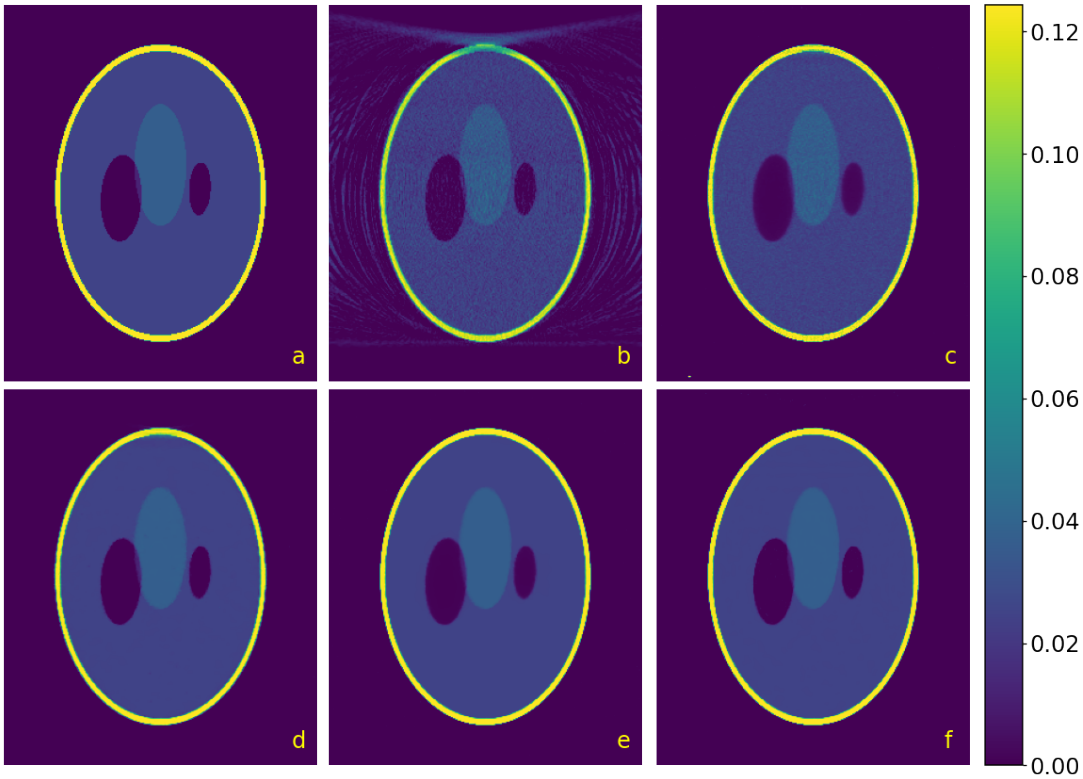
(a) Coupe coronale des volumes reconstruits.



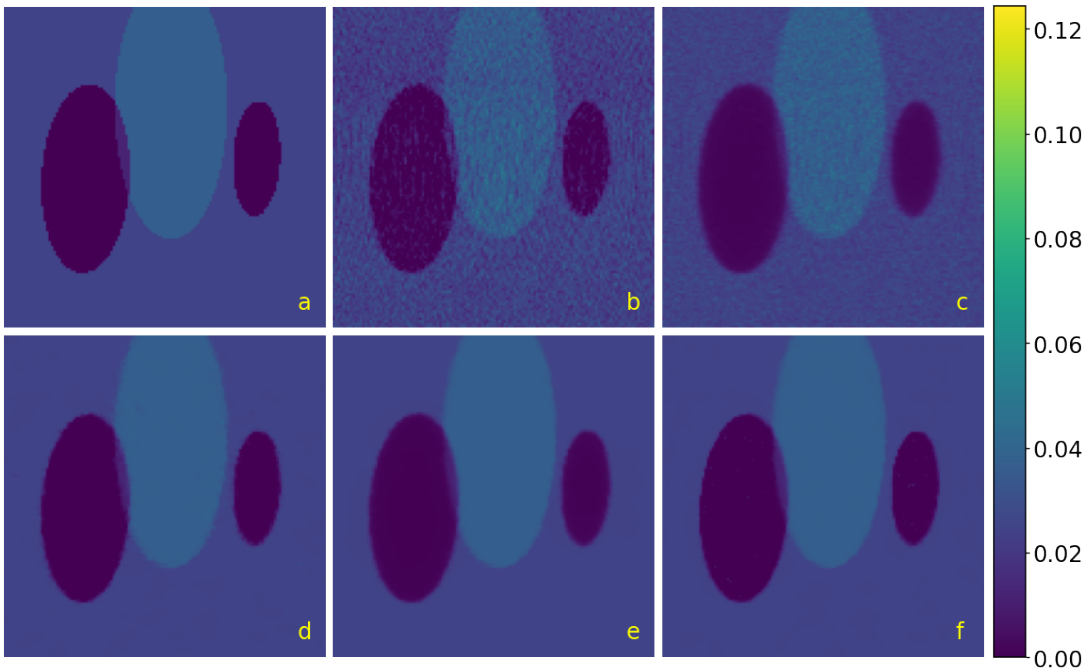
(b) Agrandissement de la coupe coronale précédente.

FIGURE 4.8 – Une coupe coronale et son agrandissement pour le fantôme de mâchoire. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV et (f) KL-TV.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés



(a) Coupe coronale des volumes reconstruits.



(b) Agrandissement de la coupe coronale précédente.

FIGURE 4.9 – Une coupe coronale et son agrandissement pour le fantôme de Shepp-Logan. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV et (f) KL-TV.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

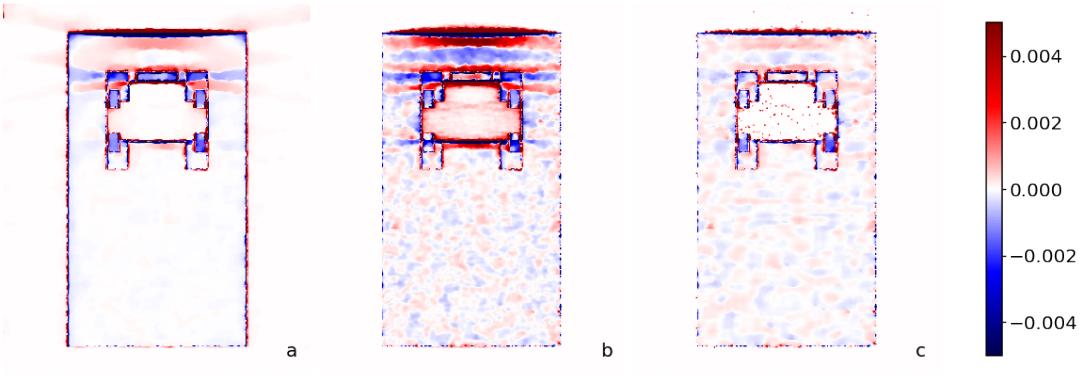


FIGURE 4.10 – Différence entre les reconstructions et la vérité terrain. (a) SIRT-TV, (b) MLEM-TV et (c) KL-TV.

Nous avons comparé la différence entre les reconstructions des méthodes itératives régularisées et le vrai volume. Des coupes coronales de cette différence sont visibles dans la figure 4.10. Les images sont volontairement saturées pour faire ressortir les défauts. Les mêmes observations sur les lignes horizontales que pour la figure 4.8 peuvent être faites, mais l'artefact de cône se voit mieux sur cette nouvelle figure. MLEM-TV et KL-TV ont tous les deux des difficultés à reconstruire correctement le vide à l'intérieur de la bouche, avec des aspects très différents. Les écarts correspondent à des biais positifs, ceux entre la reconstruction MLEM-TV et la vérité terrain sont a peu près continus, alors que ceux pour KL-TV se présentent sous la forme d'un bruit impulsionnel. Nous pouvons finalement remarquer sur cette figure que, globalement, SIRT-TV reconstruit plus fidèlement les zones constantes.

Les temps de reconstruction sont donnés dans le tableau 4.4. L'algorithme MLEM est évidemment beaucoup plus rapide que les autres algorithmes itératifs puisqu'il n'effectue pas de régularisation et que son nombre d'itérations est faible. Les méthodes régularisées montrent des temps comparables, avec un temps légèrement plus long pour MLEM-TV et SIRT-TV. KL-TV converge plus rapidement que MLEM-TV et SIRT-TV malgré son nombre d'itérations important. Comme pour les résultats en 2D, cela provient de la boucle interne dans MLEM-TV et SIRT-TV qui fait la minimisation TV.

FDK	**MLEM**	**SIRT-TV**	**MLEM-TV**	**KL-TV**
Shepp-Logan	(200)	(400-20)	(300-20)	(400)
00:00:02	00:04:19	00:09:05	00:09:31	00:08:35
mâchoire	(200)	(400-20)	(400-20)	(500)
00:00:02	00:04:19	00:12:55	00:12:40	00:10:47

TABLEAU 4.4 – Temps de reconstruction pour les fantômes de Shepp-Logan et de mâchoire pour les différents algorithmes. Le nombre d'itérations utilisé pour la reconstruction et le débruitage TV est rappelé entre parenthèse.

V Résultats sur données expérimentales

Nous appliquons désormais les algorithmes sur des données expérimentales faibles-doses acquises avec un scanner, et des données très faibles-doses obtenues en sous-échantillonnant

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

le nombre de projections faibles-doses.

Le tableau 4.5 résume les paramètres choisis pour les algorithmes, dans le cas faible dose et très faible dose. Nous avons choisi $\lambda = 0.9$ pour le paramètre SIRT. Comparé à la reconstruction du fantôme, les détails sont plus précis dans les données réelles et nous avons pu augmenter ce paramètre pour accélérer la convergence. Comme pour le fantôme, nous avons reconstruit le volume avec des nombre d'itérations et des paramètres différents, et choisi les meilleurs en terme de visualisation et métriques. Le nombre d'itérations est le même pour le cas faible dose et très faible dose. Le paramètre TV de SIRT a été augmenté pour la reconstruction très faible dose. L'augmentation de ce paramètre dans les reconstructions MLEM-TV et KL-TV résultait en une trop grande perte de détails. De même, l'augmentation du nombre d'itérations n'apportait aucune amélioration.

	Méthode	**SIRT-TV**	**MLEM**	**MLEM-TV**	**KL-TV**
Faible dose	N	400	400	400	700
	NTV	20		20	
	α	1×10^{-6}		0.05	0.05
Très faible dose	N	400	400	400	700
	NTV	20		20	
	α	2×10^{-6}		0.05	0.05

Tableau 4.5 – Paramètres utilisés pour la reconstruction des données expérimentales.

Comparé au fantôme, plus d'itérations ont été nécessaires pour les algorithmes, et en particulier pour KL-TV. Le nombre d'itérations TV n'a pas été changé, son augmentation n'améliorant pas le résultat. Les valeurs du paramètre TV sont un peu plus élevées, mais restent dans le même ordre de grandeur que celles utilisées dans les reconstructions du fantôme.

Les valeurs des métriques utilisées pour évaluer les reconstructions sont présentées dans le tableau 4.6. Les méthodes régularisées obtiennent les meilleurs résultats. La reconstruction MLEM-TV a le meilleur CNR, suivie de KL-TV. Nous pouvons remarquer que le CNR des reconstructions très faible dose est supérieur à celui des reconstructions faibles doses. Le CNR mesure le contraste dans une image. Puisque les reconstructions très faibles doses sont davantage bruitées, le paramètre de régularisation doit être plus élevé. Cependant, le débruitage TV favorisant les zones constantes par morceaux, l'augmentation de ce paramètre TV améliore le contraste de l'image au coût de la perte de certains détails. Ainsi, le paramètre TV a bien été augmenté pour SIRT-TV. Pour MLEM-TV et KL-TV, cela n'a pas été nécessaire : étant donné que les projections sont deux fois moins nombreuses, la valeur de $KL(Af, p)$ dans la fonction coût est plus faible. Donc, le même paramètre TV donne plus d'importance à la régularisation dans le cas très faible dose qu'en faible dose.

La corrélation a été calculée sur trois régions différentes : A est le volume entier, B est la région 2D que l'on voit dans la figure 4.13 et C est le fond de la coupe axiale visible dans la figure 4.11. Lorsque l'on considère l'intégralité du volume, SIRT-TV obtient la meilleure corrélation (Corr A), à égalité avec MLEM-TV en très faible dose. Ces résultats semblent en contradiction avec les observations visuelles que l'on peut faire de nos expériences. Cela semble dû au grand nombre de pixels de l'arrière-plan comparé aux régions contenant des dents, qui donne une trop grande influence au fond dans le calcul de la corrélation. Si l'on se concentre sur une zone contenant des dents (Corr B), MLEM-TV et KL-TV atteignent le meilleur score de corrélation avec le volume de référence, ce qui correspond davantage à nos observations visuelles. Et en effet, si l'on regarde une zone du fond (Corr C), SIRT-TV obtient la meilleure corrélation, avec un écart avec les deux autres bien plus important que sur l'intégralité du volume. Cette remarque

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

est en accord avec les observations faites sur le fantôme de mâchoire dans la figure 4.10, où l'on voyait que les zones constantes étaient mieux reconstruites avec SIRT-TV. Pour résumer, MLEM-TV et KL-TV donnent les meilleurs résultats sur les structures qui nous intéressent, avec un léger avantage pour MLEM-TV.

	Métrique	FDK	MLEM	SIRT-TV	MLEM-TV	KL-TV
Faible dose	**CNR**	16.664	24.879	26.733	**28.163**	27.290
	Corr A	0.933	0.963	**0.974**	0.971	0.960
	Corr B	0.897	0.937	0.948	**0.954**	0.949
	Corr C	0.950	0.952	**0.969**	0.962	0.951
Très faible dose	**CNR**	13.999	25.798	27.275	**30.182**	29.308
	Corr A	0.921	0.960	**0.968**	**0.968**	0.953
	Corr B	0.817	0.932	0.929	**0.952**	0.950
	Corr C	0.921	0.946	**0.968**	0.957	0.943

Tableau 4.6 – CNR et corrélation des reconstructions des données expérimentales, en faible dose et très faible dose. Corr A est la valeur de la corrélation sur le volume entier, alors que Corr B est calculée sur une région 2D contenant des dents et Corr C sur une zone de l'arrière-plan.

Deux coupes axiales des reconstructions faibles doses sont montrées dans les figures 4.11 et 4.12. La variation totale réduit considérablement le bruit. L'arrière-plan n'est pas aussi uniforme que celui de la reconstruction avec une dose normale, mais comme dit précédemment, l'augmentation du paramètre TV aurait causé la perte de détails importants pour le diagnostic. Les images MLEM-TV et KL-TV sont très similaires, les bords des dents sont nets, alors que l'image SIRT-TV apparaît plus floue.

La coupe montrée dans la figure 4.12 contient du métal, connu pour créer des artefacts métalliques dans les reconstructions analytiques. Aucune méthode de réduction de ces artefacts n'a été utilisée dans les reconstructions. Comparées à celle de la figure précédente, la présence du métal dégrade la qualité et le contraste des reconstructions, surtout pour SIRT-TV. Nous pouvons remarquer que MLEM, même sans régularisation, est moins affecté par la présence de métal, contrairement à FDK. En particulier, les molaires gauches, pointées par les flèches rouges dans la figure 4.12 sont plus visibles. Cependant, les incisives sont moins nettes sur la reconstruction MLEM.

La figure 4.13 est un agrandissement de la première molaire mandibulaire gauche dans une coupe coronale. L'émail et la dentine peuvent être correctement distingués dans chacune des reconstructions. Le canal de la dent et l'os trabéculaire, respectivement identifiés par les flèches bleue et verte dans la figure 4.13 sont mieux reconstruits dans les images MLEM-TV et KL-TV. Le canal mandibulaire (flèche bleu clair) est clairement visible sur chaque reconstruction, y compris MLEM qui n'effectue aucun débruitage.

Les figures 4.14 à 4.16 représentent les reconstructions en très faible dose. Globalement, la qualité des images est moins bonne que dans le cas faible dose. Dans la figure 4.14, le bruit présent dans les reconstructions FDK et MLEM, et le flou de SIRT-TV masquent les canaux de la première molaire droite (flèches jaunes). Ces canaux sont visibles avec les méthodes de régularisation statistiques.

Les artefacts métalliques sont bien plus importants que dans les reconstructions faibles doses. L'inspection visuelle des reconstructions de la figure 4.15 confirme que les méthodes itératives atténuent naturellement ces artefacts sans correction spécifique. Les molaires gauches

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

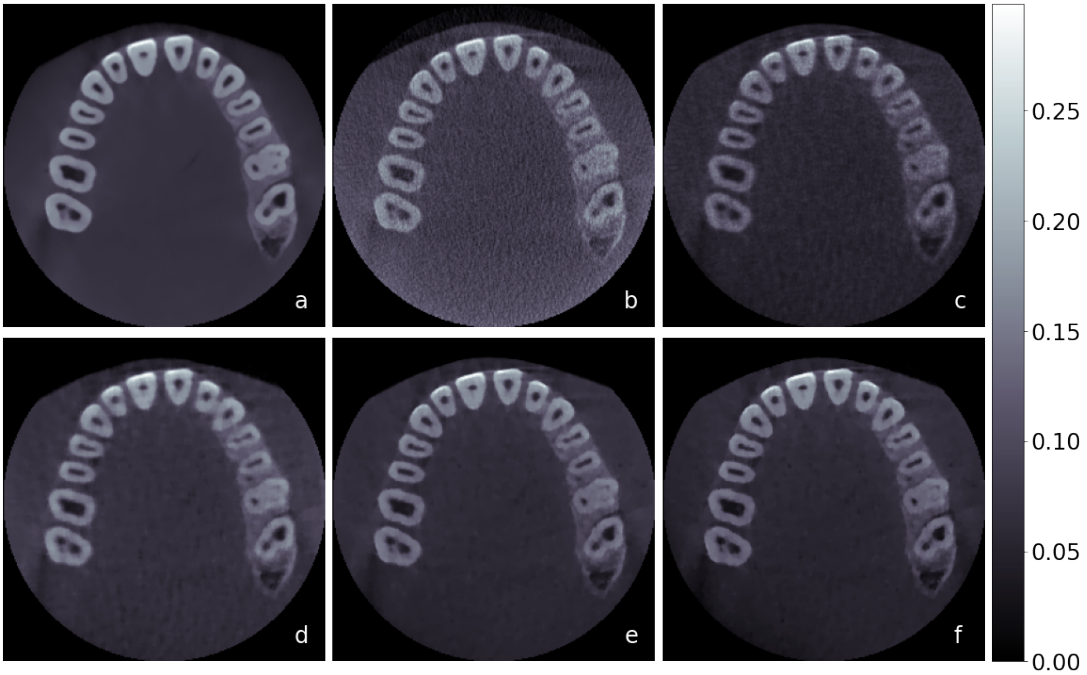


FIGURE 4.11 – Coupes axiales des reconstructions dentaires, en faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV, (f) KL-TV.

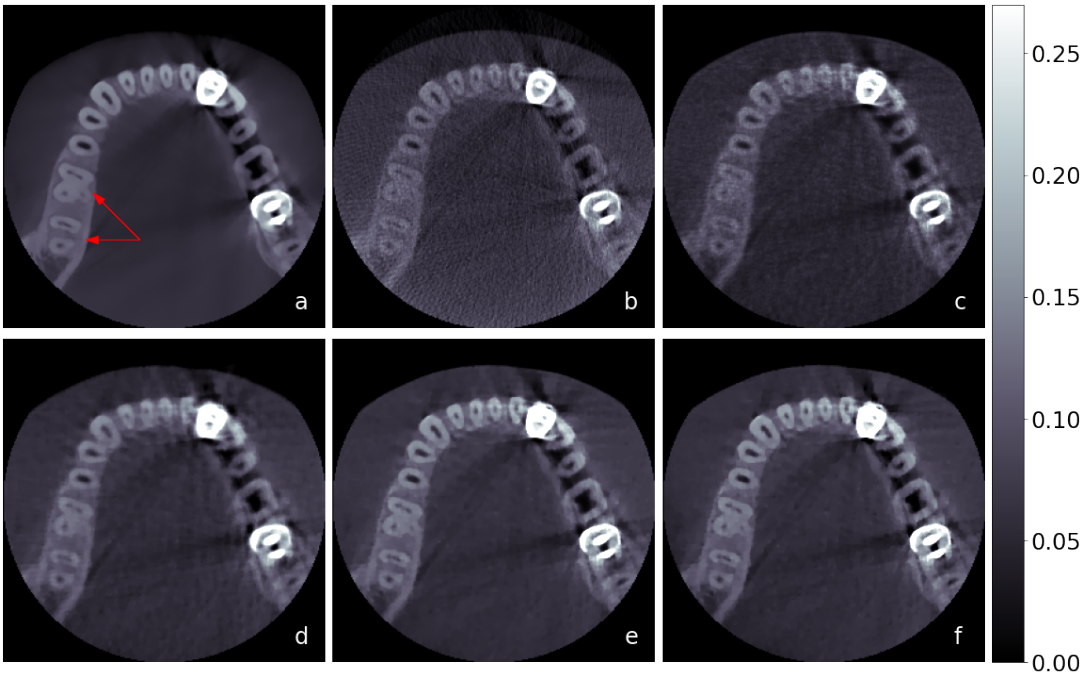


FIGURE 4.12 – Coupes axiales contenant du métal, extraites de reconstructions faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV, (f) KL-TV. Les flèches rouges indiquent les molaires gauches.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

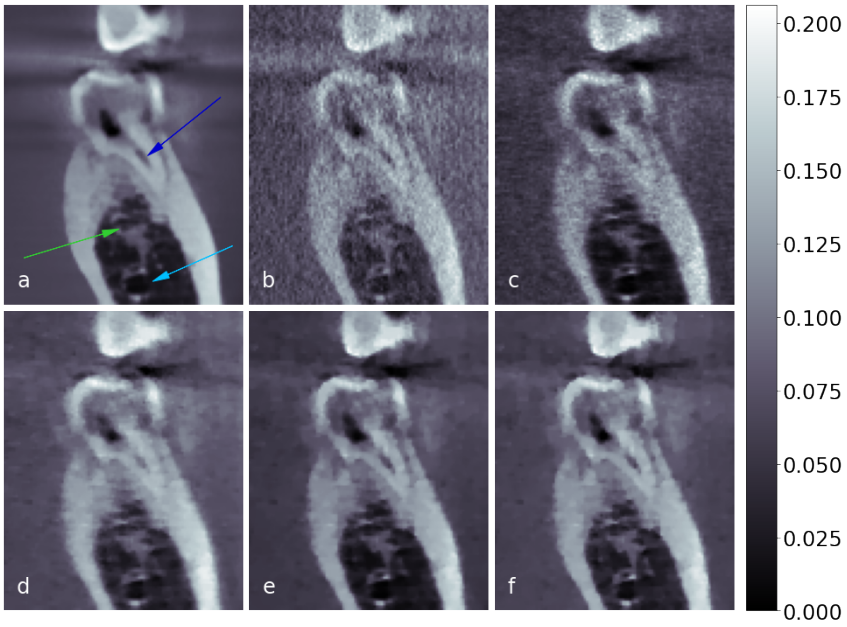


FIGURE 4.13 – Agrandissement d'une dent sur une coupe coronale, en faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV, (f) KL-TV. Les flèches bleues, vertes et bleu clair pointent respectivement vers le canal de la dent, l'os trabéculaire et le canal mandibulaire.

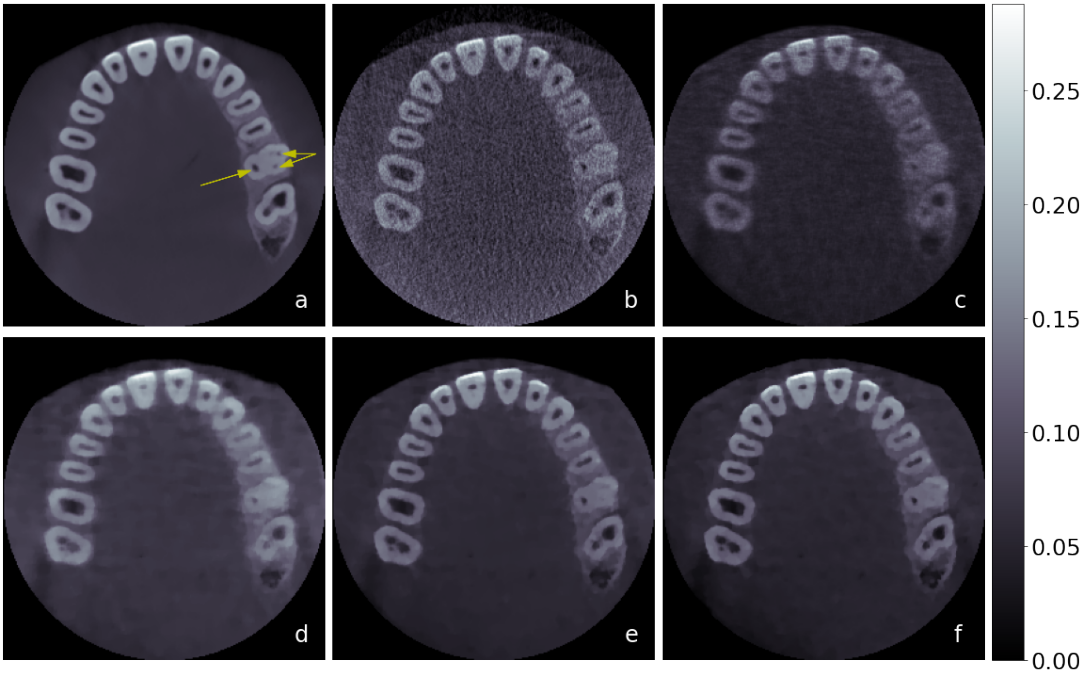


FIGURE 4.14 – Coupes axiales des reconstructions dentaires, en très faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV, (f) KL-TV. Les canaux radiculaires de la première molaire droite sont identifiés par les flèches jaunes.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

ne sont presque pas visibles dans la reconstruction FDK et sont à peine distinguables dans les images MLEM et SIRT-TV. Les canaux pointés par les flèches oranges ne sont visibles que dans les reconstructions MLEM-TV et KL-TV. Certains détails sont également perdus dans ces reconstructions par rapport au faible dose, comme le canal droit de la première molaire gauche (flèche blanche).

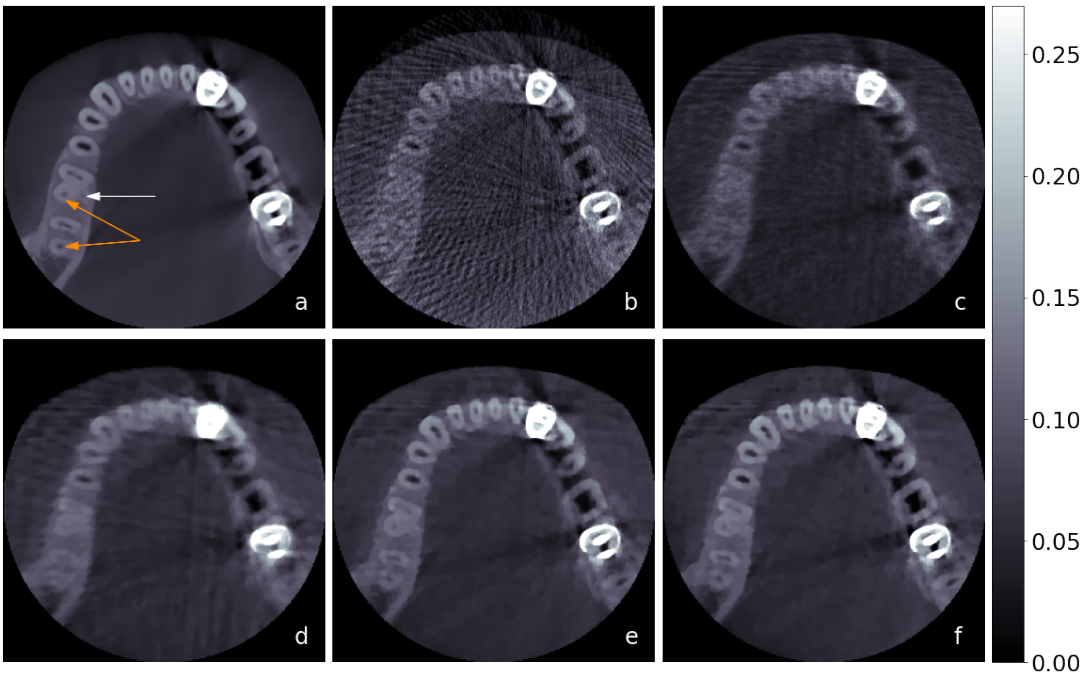


FIGURE 4.15 – Coupes axiales contenant du métal, en très faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV, (f) KL-TV. La flèche orange pointe les canaux radiculaires supprimés par MLEM et SIRT-TV, et la flèche blanche un canal qui n'est reconstruit par aucune méthode.

Un agrandissement de la première molaire mandibulaire gauche montrée en faible dose dans la figure 4.13 est visible dans la figure 4.16. Le canal mandibulaire (flèche bleu clair) est visible dans les trois reconstructions régularisées. FDK et MLEM reconstruisent des images dans lesquelles de nombreux détails sont perdus. Le canal de la dent (flèche bleue) de la reconstruction SIRT-TV n'est pas assez visible et la taille de la pulpe (flèche rose) est diminuée. Les algorithmes MLEM-TV et KL-TV reconstruisent plus fidèlement le volume, bien que quelques détails soient également perdus, en particulier au niveau de l'os trabéculaire.

VI Analyse et conclusion de l'étude

Nos résultats, tant pour les données simulées qu'expérimentales, démontrent clairement qu'une amélioration significative de la qualité de l'image peut être obtenue lorsque les modèles régularisés sont utilisés conjointement avec les méthodes itératives. Nous avons comparé trois algorithmes : SIRT-TV, MLEM-TV et KL-TV. L'algorithme SIRT-TV est basé sur un terme quadratique de fidélité aux données, correspondant à un bruit gaussien, alors que les deux autres utilisent la divergence de Kullback-Leibler, adaptée un bruit de Poisson. Au vu des résultats, la distribution de Poisson semble mieux adaptée à la modélisation des données faibles doses que la distribution gaussienne, bien qu'un modèle plus complet puisse également tenir compte des incertitudes gaussiennes de la détection. Les tests effectués en 2D avec projections parallèles,

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

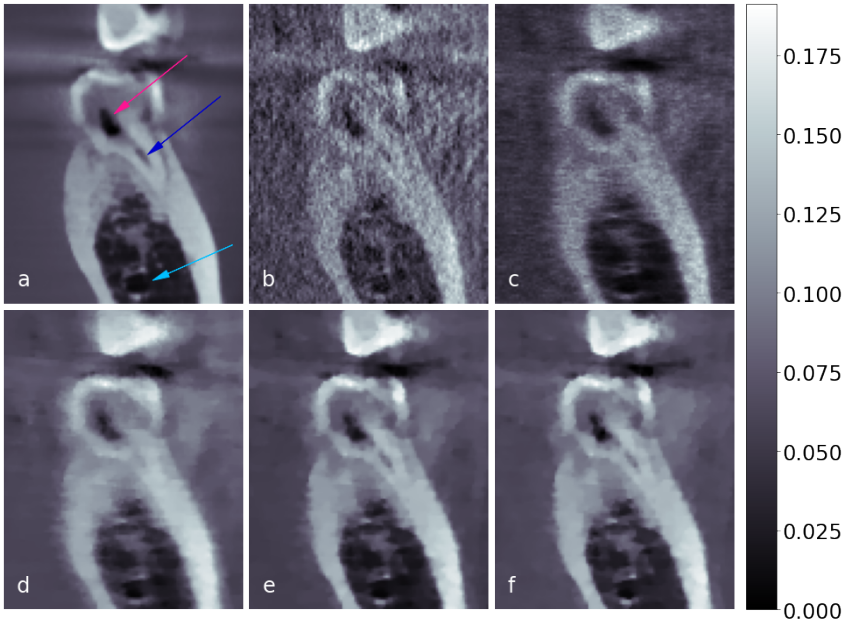


FIGURE 4.16 – Agrandissement d'une dent dans une coupe coronale, en très faible dose. (a) Vérité terrain, (b) FDK, (c) MLEM, (d) SIRT-TV, (e) MLEM-TV (f) KL-TV. Les flèches rose, bleue et bleu clair indiquent respectivement la pulpe, le canal de la dent et le canal mandibulaire.

en faisant varier le ratio bruit gaussien / bruit de Poisson, suggèrent que SIRT-TV devrait être utilisé lorsque le bruit gaussien prédomine. Nos expériences ont montré de meilleures performances pour les algorithmes KL-TV et MLEM-TV pour nos données expérimentales, ce qui confirme la nécessité d'inclure le bruit de Poisson lors de la modélisation de CBCT à faible dose.

Bien qu'un biais positif soit présent dans les régions à valeurs nulle de la reconstruction MLEM, aucun biais n'a été observé dans la reconstruction des données expérimentales. Une réduction efficace des artefacts en étoiles a été observée avec les méthodes itératives régularisées. Renforcés par la présence de métal, ces artefacts sont particulièrement visibles dans la reconstruction FDK standard. Les méthodes itératives les réduisent bien qu'ils restent visibles dans les basses fréquences. Ils ne peuvent être éliminés par une augmentation du paramètre TV sans une perte significative des détails des images. Ces artefacts pourraient, cependant, être davantage atténués avec un algorithme de réduction des artefacts métalliques.

Le tableau 4.7 résume les algorithmes utilisés dans cette étude, avec leur principaux avantages et inconvénients.

	Méthode	Fonctionnelle	Vitesse	Réduction artefacts	Réduction du bruit	Détails
FDK	Analytique	Aucune	++	Aucune	-	++
MLEM	Vraisemblance	$KL(Af, p)$	+	+	-	-
SIRT-TV	MAP-EM	$\frac{1}{2}\|Af - p\|_2^2 + \alpha TV(f)$	--	++	+	+
MLEM-TV	MAP-EM	$KL(Af, p) + \alpha TV(f)$	--	+++	+	++
KL-TV	PDHG	$KL(Af, p) + \alpha TV(f)$	-	+++	+	++

TABLEAU 4.7 – Résumé des cinq algorithmes. Nous rappelons sur quelles méthodes ils sont basés, la fonctionnelle qu'ils minimisent, leur temps de reconstruction, leur capacité à réduire le bruit, les artefacts et à restaurer correctement les détails.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Dans le CBCT dentaire, la distribution de la dose n'est pas uniforme dans tout le volume, car la dose est plus importante dans la région centrale où se trouve la région d'intérêt. La transcription mathématique de la dose reçue est la sensibilité, définie dans l'équation (4.2). Des valeurs plus élevées de la sensibilité sont observées dans la région d'intérêt et des valeurs plus faibles près des bords. De plus, en CBCT dentaire, une extension du volume est nécessaire en raison des projections tronquées et conduit à des divergences encore plus importantes. Cela conduit à une inhomogénéité dans le modèle mathématique que l'on ne rencontre pas dans les modèles à faisceau parallèle et qui a deux conséquences importantes sur les deux méthodes statistiques régularisées. La première est liée à la convergence des algorithmes KL-TV et MLEM-TV. Pour KL-TV, nous avons implémenté la version pré-conditionnée de l'algorithme PDHG de Pock et Chambolle (2011). Pour MLEM-TV nous avons introduit dans ce travail une version pré-conditionnée de l'algorithme qui empêche la divergence numérique lorsque le minimum de la sensibilité est petit. L'autre conséquence concerne la vitesse de convergence, le pré-conditionnement permet, dans ces deux cas, de l'accélérer. Pour le MLEM-TV original, la convergence de l'étape de débruitage TV dépend du pas de gradient τ dans (3.45), qui diminue avec le minimum de la sensibilité $s_{\min}$. Avec le pré-conditionnement, nous permettons à ce coefficient de varier dans le volume avec la sensibilité, ce qui permet une convergence plus rapide lors du débruitage de la partie centrale du volume, celle qui correspond au FoV.

En résumé, les inhomogénéités du modèle et le mauvais conditionnement résultant de la troncature des projections posent certains problèmes, que nous avons pu surmonter en mettant en œuvre des schémas d'optimisation numérique adaptés. Peu d'études abordent ces questions. Nous pouvons citer par exemple (Stsepankou *et al.*; 2012), où une distribution de données de Poisson et une régularisation TV sont considérées pour des données CBCT faibles doses. La reconstruction est effectuée avec l'algorithme MAP-TV de Green (1990). Cependant, cet algorithme n'est pas stable même si une convergence numérique peut être obtenue pour de petits paramètres de régularisation. La limite supérieure du paramètre de régularisation dépend du minimum de la sensibilité comme pour MLEM-TV. Cependant, dans leur travail, certains des problèmes auxquels nous sommes confrontés sont contournés car l'ensemble de données expérimentales est composé de projections complètes issues de la radiothérapie guidée par l'image.

En termes de temps de calcul, selon la configuration, MLEM-TV est soit comparable soit plus lent que KL-TV. La différence entre les deux est que KL-TV nécessite plus d'opérations de projection et de rétro-projection alors que MAP-EM effectue un débruitage TV complet à chaque itération. Ces conclusions sont différentes de ce qui a été montré dans le chapitre 2 et dans (Leuliet *et al.*; 2021) sur les reconstructions 2D à partir de projections parallèles et sur des données 3D de microscopie électronique expérimentale. Ceci est lié une fois de plus à la non-uniformité de la sensibilité dans le volume. La majeure partie du temps de calcul de MLEM-TV a été consacrée aux étapes de débruitage. Cette partie a été implémentée sur GPU en utilisant la bibliothèque CuPy de Python. Il est probable qu'une certaine accélération puisse être obtenue avec une implémentation plus adéquate, rendant les étapes de TV et de reconstruction comparables. L'utilisation d'OSEM au lieu de MLEM pourrait également accélérer la convergence.

La valeur du paramètre TV influence la qualité du débruitage et la précision de la reconstruction. Si le paramètre est trop grand, des détails de l'image seront supprimés, tandis que s'il est trop petit, on ne débruitera pas assez. Le choix de ce paramètre est donc déterminant pour la qualité de la reconstruction. Il existe une riche littérature sur les méthodes capables de calculer automatiquement le paramètre optimal. Elles sont basées sur le principe de divergence ou sur des estimateurs de risque. Nous avons essayé la règle présentée dans (Ito *et al.*; 2011), où les auteurs définissent le paramètre de régularisation en minimisant une fonctionnelle basée sur

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

la fonction de coût utilisée dans l'algorithme de reconstruction. Si cette méthode avait donné de bons résultats en géométrie parallèle 2D, dans notre cas, la valeur obtenue était instable. Il a été montré dans (Lucka *et al.*; 2018) que les méthodes automatiques ne sont pas nécessairement adaptées aux problèmes sévèrement mal posés. Cependant, nous avons observé dans nos expériences qu'avec une configuration assez similaire (même dose, nombre de projections, présence de métal ou non), le même paramètre TV pouvait être conservé pour tous les volumes. Ainsi, il serait possible de déterminer un nombre fixe de configurations avec leurs paramètres associés.

Les méthodes itératives ont certaines limites. Même si l'utilisation de processeurs graphiques permet d'accélérer les reconstructions 3D itératives, elles restent assez lentes pour une application en milieu clinique et nécessitent des ressources mémoire importantes. La régularisation TV que nous avons appliquée dans ce travail est largement utilisée dans l'état de l'art. Cependant, le TV est connue pour favoriser les zones constantes par morceaux et produire un effet "cartoon" dans l'image. Certaines petites structures peuvent être supprimées et la texture est modifiée.

L'objectif principal de ce travail était l'évaluation des méthodes de reconstruction 3D CBCT à l'aide de schémas itératifs et de régularisation TV en imagerie dentaire. Les schémas numériques ont été adaptés à la géométrie du faisceau conique et à la nature tronquée des projections. Nous avons appliqué avec succès quatre algorithmes itératifs à des fantômes 3D et à des données dentaires expérimentales. Le test effectué sur un ensemble de projections très faibles doses a permis d'évaluer les algorithmes dans une situation assez extrême. La régularisation TV améliore clairement la qualité de la reconstruction. Les algorithmes KL-TV et MLEM-TV ont donné les meilleures reconstructions en termes de temps de calcul et de précision de reconstruction.

Cependant, il faut souligner la difficulté de comparer les méthodes régularisées entre elles et de déterminer une méthode de comparaison qui soit juste. Nous pouvons premièrement mettre en avant la complexité de fixer un critère d'arrêt pour les itérations. L'étude des valeurs de la MSE ou du PSNR ne nous a pas permis d'arrêter les itérations après une valeur, ou un écart entre deux itérations consécutives, spécifiques qui soit stable pour toutes les méthodes. Un second point problématique est la fonction coût différente entre SIRT-TV et les deux autres méthodes régularisées. La valeur du paramètre TV est très différente, il est donc assez délicat de choisir des paramètres qui rendent les comparaisons justes. Les paramètres (nombre d'itérations et TV) ont été déterminés empiriquement, de sorte que la qualité d'image soit satisfaisante, tout en vérifiant que les valeurs de la MSE soient correctes. Il est possible que l'augmentation du nombre d'itérations de la bouche interne et/ou externe de SIRT-TV améliore les reconstructions avec cette méthode. Mais celle-ci étant déjà la plus lente à converger, l'augmentation de ces paramètres lui donnerait un désavantage considérable pour une utilisation clinique.

Par rapport à la reconstruction FDK standard, les méthodes itératives ont démontré une amélioration de la qualité de l'image dans la reconstruction de données CT faibles doses acquises avec un bruit photonique accru en raison du courant plus faible du tube à rayons X. Dans cette étude, nous avons choisi d'étudier les situations où la dose est réduite en diminuant le nombre de projections acquises. Avec une seule acquisition, la dose normale et la faible dose peuvent être obtenues simultanément en sous-échantillonnant la première afin d'obtenir la seconde. De cette façon, l'image faible dose peut être comparée à la référence obtenue en dose normale. Des travaux antérieurs sur d'autres applications faibles doses ont démontré que la réduction du courant est plus avantageuse par rapport à la réduction du nombre de projections (Zhao *et al.*; 2014). Notre objectif était d'étudier les algorithmes sur des données expérimentales,

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

tout en ayant une image de référence de l'objet.

Un inconvénient commun à toutes les méthodes itératives est le coût élevé de calcul que ce soit en termes de mémoire ou de temps. L'optimisation de l'implémentation ou du matériel et l'utilisation de calcul parallèle pourraient réduire les temps de reconstruction au point de les rendre acceptables pour une utilisation clinique.

Les avancées récentes en apprentissage profond, et plus particulièrement en traitement des images, ont permis leur application à la reconstruction tomographique. Les prochains chapitres traiteront de ces méthodes, et nous essaierons d'étudier dans quelles mesures elles peuvent remédier à certains défauts des limites des méthodes itératives.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Bibliographie

Aootaphao, S., Thongvigitmanee, S. S., Puttawibul, P. et Thajchayapong, P. (2022). Truncation effect reduction for fast iterative reconstruction in cone-beam CT. *BMC Medical Imaging*, 22(1):1–12. 69

Baba, R., Ueda, K. et Okabe, M. (2004). Using a flat-panel detector in high resolution cone beam CT for dental imaging. *Dentomaxillofacial radiology*, 33(5):285–290. 69

Bayaraa, T., Hyun, C. M., Jang, T. J., Lee, S. M. et Seo, J. K. (2020). A two-stage approach for beam hardening artifact reduction in low-dose dental CBCT. *IEEE Access*, 8:225981–225994. 69

Buades, A., Coll, B. et Morel, J.-M. (2005). A non-local algorithm for image denoising. *In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 2, pages 60–65. Ieee. 68

Chambolle, A. (2004). An algorithm for total variation minimization and applications. *Journal of Mathematical imaging and vision*, 20(1-2):89–97. 68

Chen, G., Zhu, F. et Ann Heng, P. (2015). An efficient statistical method for image noise level estimation. *In Proceedings of the IEEE International Conference on Computer Vision*, pages 477–485. 74

Dabov, K., Foi, A., Katkovnik, V. et Egiazarian, K. (2007). Image denoising by sparse 3D transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095. 68

Dang, H., Stayman, J. W., Sisniega, A., Zbijewski, W., Xu, J., Wang, X., Foos, D. H., Aygun, N., Koliatsos, V. E. et Siewerdsen, J. H. (2016). Multi-resolution statistical image reconstruction for mitigation of truncation effects : application to cone-beam CT of the head. *Physics in Medicine & Biology*, 62(2):539–559. 69, 70

Diwakar, M. et Kumar, M. (2018). A review on CT image noise and its denoising. *Biomedical Signal Processing and Control*, 42:73–88. 68, 69

Fan, L., Zhang, F., Fan, H. et Zhang, C. (2019). Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art*, 2(1):1–12. 68

Feldkamp, L. A., Davis, L. C. et Kress, J. W. (1984). Practical cone-beam algorithm. *Josa a*, 1(6):612–619. 68

Friot--Giroux, L., Peyrin, F. et Maxim, V. (2022). Iterative tomographic reconstruction with TV prior for low-dose CBCT dental imaging. *Physics in Medicine & Biology*. 68

Green, P. J. (1990). Bayesian reconstructions from emission tomography data using a modified EM algorithm. *IEEE Transactions on Medical Imaging*, 9(1):84–93. 88

Hsieh, J., Chao, E., Thibault, J., Grekowicz, B., Horst, A., McOlash, S. et Myers, T. (2004). A novel reconstruction algorithm to extend the CT scan field-of-view. *Medical physics*, 31(9): 2385–2391. 69

Humphries, T., Si, D., Coulter, S., Simms, M. et Xing, R. (2019). Comparison of deep learning approaches to low dose CT using low intensity and sparse view data. *In Medical Imaging 2019 : Physics of Medical Imaging*, volume 10948, pages 1048–1054. SPIE. 73

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Ibraheem, I. *et al.* (2012). Reduction of artifacts in dental cone beam CT images to improve the three dimensional image reconstruction. *J Biomed Sci Eng*, 5(8):409–15. 69

Ito, K., Jin, B. et Takeuchi, T. (2011). A regularization parameter for nonsmooth Tikhonov regularization. *SIAM Journal on Scientific Computing*, 33(3):1415–1438. 88

Leuliet, T., Friot--Giroux, L., Baaziz, W., Bretin, E., Ersen, O., Peyrin, F., Sixou, B. et Maxim, V. (2021). Efficiency of TV-regularized algorithms in computed tomography with Poisson-Gaussian noise. *In 2020 28th European Signal Processing Conference (EUSIPCO)*, pages 1294–1298. 88

Leuschner, J., Schmidt, M., Baguer, D. O. et Maass, P. (2021). LoDoPaB-CT, a benchmark dataset for low-dose computed tomography reconstruction. *Scientific Data*, 8(1):1–12. 69

Liang, X., Lambrichts, I., Sun, Y., Denis, K., Hassan, B., Li, L., Pauwels, R. et Jacobs, R. (2010). A comparative evaluation of cone beam computed tomography (CBCT) and multi-slice CT (MSCT). Part II : On 3D model accuracy. *European journal of radiology*, 75(2):270–274. 68

Liu, Y., Shangguan, H., Zhang, Q., Zhu, H., Shu, H. et Gui, Z. (2015). Median prior constrained TV algorithm for sparse view low-dose CT reconstruction. *Computers in biology and medicine*, 60:117–131. 73

Lucka, F., Proksch, T., Brune, C., Bissantz, N., Burger, M., Dette, H. et Wubbeeling, F. (2018). Risk estimators for choosing regularization parameters in ill-posed problems-properties and limitations. *Inverse Problems and Imaging*, 12(5):1121–1155. 89

Maltz, J. S., Bose, S., Shukla, H. P. et Bani-Hashemi, A. R. (2007). CT truncation artifact removal using water-equivalent thicknesses derived from truncated projection data. *In 2007 29th annual international conference of the IEEE engineering in medicine and biology society*, pages 2907–2911. IEEE. 70

Ohnesorge, B., Flohr, T., Schwarz, K., Heiken, J. et Bae, K. (2000). Efficient correction for CT image artifacts caused by objects extending outside the scan field of view. *Medical physics*, 27(1):39–46. 69

Pock, T. et Chambolle, A. (2011). Diagonal preconditioning for first order primal-dual algorithms in convex optimization. *In 2011 International Conference on Computer Vision*, pages 1762–1769. 71, 88

Rudin, L. I., Osher, S. et Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D : nonlinear phenomena*, 60(1-4):259–268. 68

Stsepankou, D., Arns, A., Ng, S., Zygmanski, P. et Hesser, J. (2012). Evaluation of robustness of maximum likelihood cone-beam CT reconstruction with total variation regularization. *Physics in Medicine & Biology*, 57(19):5955–5970. 88

Tsiklakis, K., Donta, C., Gavala, S., Karayianni, K., Kamenopoulou, V. et Hourdakis, C. J. (2005). Dose reduction in maxillofacial imaging using low dose Cone Beam CT. *European journal of radiology*, 56(3):413–417. 68

Van Aarle, W., Palenstijn, W. J., Cant, J., Janssens, E., Bleichrodt, F., Dabravolski, A., De Beenhouwer, J., Batenburg, K. J. et Sijbers, J. (2016). Fast and flexible X-ray tomography using the ASTRA toolbox. *Optics express*, 24(22):25129–25147. 74

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Wang, Z., Bovik, A. C., Sheikh, H. R. et Simoncelli, E. P. (2004). Image quality assessment : from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612. 72

Widmann, G., Bischel, A., Stratis, A., Kakar, A., Bosmans, H., Jacobs, R., Gassner, E.-M., Puelacher, W. et Pauwels, R. (2016). Ultralow dose dentomaxillofacial CT imaging and iterative reconstruction techniques : variability of Hounsfield units and contrast-to-noise ratio. *The British Journal of Radiology*, 89(1060):20151055. 72

Yu, L., Shiung, M., Jondal, D. et McCollough, C. H. (2012). Development and validation of a practical lower-dose-simulation tool for optimizing computed tomography scan protocols. *Journal of computer assisted tomography*, 36(4):477–487. 69

Zhao, Z., Gang, G. et Siewerdsen, J. (2014). Noise, sampling, and the number of projections in cone-beam CT with a flat-panel detector. *Medical Physics*, 41(6Part1):061909. 89

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

5 Reconstruction tomographique et apprentissage profond



I	**Introduction à l'apprentissage profond**		**96**
	1	Apprentissage automatique	96
	2	Apprentissage profond	98
	3	Apprentissage et paramètres d'un réseau	101
II	**Réseaux de neurones convolutifs**		**103**
	1	Couches de convolution	104
	2	Couches de pooling	105
	3	Architecture des CNN	106
III	**Apprentissage profond en CT**		**107**
	1	En post-traitement	108
	2	En pré-traitement	109
	3	Méthodes end-to-end	110
	4	Méthodes semi ou non-supervisées	112

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

L'apprentissage profond (*deep learning*) est un sous-ensemble de techniques d'apprentissage automatique (*machine learning*) qui s'intéresse aux algorithmes inspirés de la structure et du fonctionnement du cerveau, appelés réseaux neuronaux artificiels. Ils apprennent par l'exemple, et plus ils sont exposés à des données diversifiées, plus ils identifient des caractéristiques variées et sont capables de généraliser. Leurs applications en traitement de l'image (Voulodimos *et al.*; 2018), du langage (Young *et al.*; 2018) ou en robotique (Pierson et Gashler; 2017) ont considérablement fait avancer la recherche dans ces domaines.

Parmi les différentes tâches pour lesquelles l'apprentissage profond s'est montré particulièrement utile en traitement de l'image, nous pouvons mentionner le débruitage (Tian *et al.*; 2020), la super-résolution (Yang *et al.*; 2019), la segmentation (Minaee *et al.*; 2021) ou encore la classification (Rawat et Wang; 2017).

Du côté de l'imagerie médicale, il existe de nombreuses applications de l'apprentissage profond. Il a montré de bons résultats en diagnostic, analyse et correction d'images, et en chirurgie guidée par l'image. L'apprentissage profond est par exemple utilisé pour identifier automatiquement des tumeurs ou d'autres anomalies dans les images médicales, ou pour améliorer la qualité des images. Pour ce dernier point, il existe déjà de nombreuses applications dans tous les types d'imagerie : tomographie rayons X (Jin *et al.*; 2017), TEP (Reader *et al.*; 2020), IRM (Lundervold et Lundervold; 2019) et même en ultrasons (Van Sloun *et al.*; 2019).

Nous commençons par introduire les concepts fondamentaux de l'apprentissage profond, puis nous explorerons l'état de l'art de l'utilisation des méthodes d'apprentissage profond en tomographie. Les encadrés contiennent des exemples illustrant les notions introduites ou de nouvelles définitions, et les anglicismes seront utilisés après avoir été définis.

I Introduction à l'apprentissage profond

L'apprentissage profond est une sous-partie de l'apprentissage automatique, dont le but est de résoudre des problèmes en faisant de l'apprentissage sur des données. Cette partie a pour objectif de présenter les notions et le vocabulaire indispensables à la compréhension de l'apprentissage profond. Les principales sources de ce chapitre sont les livres (Goodfellow *et al.*; 2016), pour les aspects plus théoriques, et (Géron; 2022) qui présente le machine et l'apprentissage profond avec un point de vue pratique grâce à la librairie *Tensorflow*.

1 Apprentissage automatique

Les algorithmes d'apprentissage automatique apprennent à effectuer des tâches (classification, détection, régression, etc.) en analysant et en apprenant sur un grand ensemble de données. Le processus général est le suivant : on choisit un *modèle de prédiction* représenté par une fonction Γ_θ, on estime les paramètres optimaux θ^* de ce modèle par rapport à nos données.

Le modèle est choisi a priori en fonction de la tâche à effectuer : classifieurs pour la classification, polynômes pour une régression polynomiale par exemple. Ce modèle dépend de *paramètres* initialisés aléatoirement au début, puis la *phase d'apprentissage* ou *d'entraînement* optimise ces paramètres pour que le modèle puisse correctement effectuer la tâche demandée. Les modèles ont également des *hyperparamètres* qui sont choisis avant l'entraînement, en même temps que le modèle. Le choix de ces hyperparamètres est décisif dans les résultats de l'apprentissage du modèle.

Image de couverture : générée par l'IA Midjourney, prompt : "`Don't you cry tonight, there's a heaven above you --ar 4:3 --stylize 9000 --chaos 65 --v 3`", Guns N' Roses, *Don't cry*.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Dans une régression polynomiale, le modèle est un polynôme, les **paramètres** sont les coefficients du polynôme. Dans un réseau de neurones, le modèle est le type de réseau choisi et les paramètres sont les poids des neurones.
Le degré des polynômes ou le nombre de couches dans un réseau de neurones sont des **hyperparamètres**.

Type d'apprentissage

Les algorithmes d'apprentissage automatique peuvent être entraînés de différentes manières selon le type d'information disponible pour l'entraînement. Les quatre catégories principales sont : apprentissage supervisé, apprentissage non supervisé, apprentissage semi-supervisé et apprentissage par renforcement.

Apprentissage supervisé En apprentissage supervisé, les données sont étiquetées (ou annotées) : pour chaque donnée d'entrée x_i, on connait la donnée de sortie y_i correspondante. Les paires de données sont notées $\{(x_i, y_i)\}_i$, il s'agit de couples de *variables prédictives* (x_i) et de *variables cibles* (y_i). Le modèle doit donc apprendre à prédire y_i à partir de x_i, dans le but de déterminer la sortie la plus adéquate pour une nouvelle entrée x n'appartenant pas à l'ensemble d'entraînement. On cherche une fonction cible Γ_θ, dépendant des paramètres θ, qui minimise une fonction coût entre les prédictions et les cibles.

Un apprentissage supervisé suit le schéma suivant :

1. Choix du modèle, i.e. f, et initialisation aléatoire des paramètres θ.
2. Entraînement : recherche de paramètres optimaux θ^*, en s'appuyant sur les paires de données (x_i, y_i), qui minimisent une fonction coût représentant l'erreur entre les prédictions $\Gamma_\theta(x_i)$ et les cibles y_i.
3. Validation : la phase de validation consiste à tester Γ_θ sur des données n'ayant pas servi à l'entraînement, afin d'ajuster les hyperparamètres. Si besoin, on réitère l'étape 2 pour ré-entraîner avec les nouveaux hyperparamètres.
4. Test : une fois les hyperparamètres fixés et les paramètres optimaux θ^* déterminés, on teste la généralisation de Γ_{θ^*} sur de nouvelles données n'ayant été utilisées ni pour l'entraînement, ni la validation.

Afin d'effectuer les trois dernières étapes, il est donc nécessaire de séparer dès le début la base de données en trois parties disjointes qui serviront pour l'entraînement, pour la validation et pour le test.

Apprentissage non supervisé En apprentissage non supervisé, les variables cibles ne sont pas disponibles, le modèle doit apprendre à partir des variables prédictives uniquement. Il s'intéresse à la structure sous-jacente des données, et essaie d'extraire des classes ou groupes ayant des caractéristiques similaires. Les tâches complexes, pour lesquelles les données sont difficiles à étiqueter, voire impossibles, peuvent être étudiées par les algorithmes non-supervisés.

L'apprentissage semi-supervisé est à mi-chemin entre supervisé et non-supervisé : certaines données sont étiquetées, et d'autres non. Cette méthode permet de réduire le temps de préparation des données, tout en ayant un grand nombre de données à traiter et une efficacité supérieure au non-supervisé. L'apprentissage par renforcement concerne les algorithmes qui apprennent par essais et erreurs, afin de maximiser une récompense.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Dans ce manuscrit, nous ne détaillerons pas davantage ces deux types d'apprentissage car ils ne font pas partie du cadre d'étude de cette thèse. Cependant, nous tenions à présenter l'apprentissage semi-supervisé car son utilisation en imagerie médicale, où les données appariées peuvent être compliquées à obtenir, est importante.

Principaux défis en apprentissage supervisé

Durant le processus d'apprentissage, il faut prêter attention à certains détails qui peuvent compromettre la qualité des résultats. En dehors du choix décisif du modèle, d'autres variables peuvent impacter l'apprentissage, comme les données ou les hyperparamètres choisis.

Les données L'apprentissage automatique nécessite un grand nombre de données d'apprentissage pour être capable de généraliser correctement ce qu'il a appris à de nouvelles situations. Lorsque ce n'est pas le cas, on parle de sous-représentation, les erreurs d'entraînement et de tests sont alors importantes. De plus, les données doivent être représentatives de toutes les situations possibles et de bonne qualité.

Sur/sous-apprentissage On parle de *sur-apprentissage* lorsque le modèle entraîné est très performant sur la base d'apprentissage, mais ne généralise pas bien sur de nouvelles données. L'erreur d'entraînement est alors très faible, mais celle sur la base de validation est plutôt élevée. Le *sous-apprentissage* se produit lorsque le modèle ne s'adapte même pas correctement aux données d'entraînement.

La figure 5.1 illustre les effets du sous-apprentissage et sur-apprentissage. Dans le cas du sous-apprentissage, on voit que le modèle n'est pas adapté aux données. Pour le sur-apprentissage, le modèle est trop adapté aux données d'entraînement, et de nouvelles données ne seront pas correctement traitées. Un compromis entre biais [1] et variance [2] doit être trouvé pour ne pas se retrouver dans ces cas-là. Le sous-apprentissage peut être évité en complexifiant davantage le modèle ou en augmentant le temps d'entraînement. Pour limiter le sur-apprentissage, on peut ajouter une régularisation sur le modèle ou le simplifier, et augmenter la taille de la base de données.

2 Apprentissage profond

Un algorithme d'apprentissage profond est un algorithme d'apprentissage automatique dont la fonction cible Γ_θ est un réseau de neurones artificiel (ANN [3]). Les ANN sont des modèles d'apprentissage automatique inspirés de l'architecture neuronale du cerveau. Ils ont été introduits pour la première fois dans (McCulloch et Pitts; 1943), où ils sont utilisés comme démonstration de la résolution d'un problème dans le cerveau d'un animal à l'aide de connexions logiques entre ses neurones biologiques. Les neurones biologiques fonctionnent sous des impulsions électriques : des neurotransmetteurs sont envoyés d'un neurone à l'autre, et lorsqu'un neurone reçoit suffisamment de signal, il s'active.

Le modèle de McCulloch et Pitts (1943) utilise le même système et malgré sa simplicité, il permet tout de même de réaliser certaines opérations logiques comme on peut le voir dans la figure 5.2. Dans cette figure, le neurone s'active lorsque deux entrées sont positives. Chacune de ces opérations peut être combinée avec les autres pour donner lieu à une suite d'opérations pouvant rapidement devenir complexe, comme dans le cerveau. Les multiples connexions entre

1. Erreur entre prédictions et cibles
2. Dépendance aux données d'entraînement
3. *Artificial Neural Network*

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

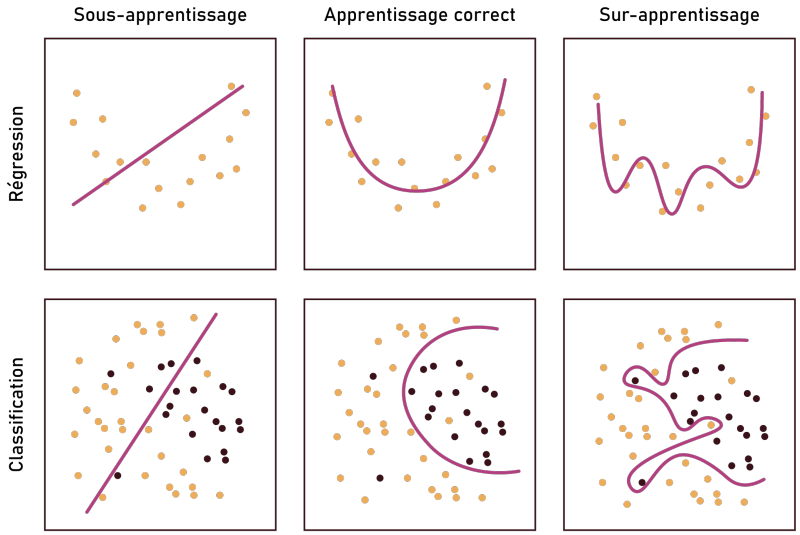


FIGURE 5.1 – Illustration de sous-apprentissage et de sur-apprentissage dans le cas d'une régression polynomiale et d'une classification sur les données d'entraînement.

un très grand nombre de neurones biologiques permettent de construire des architectures complexes.

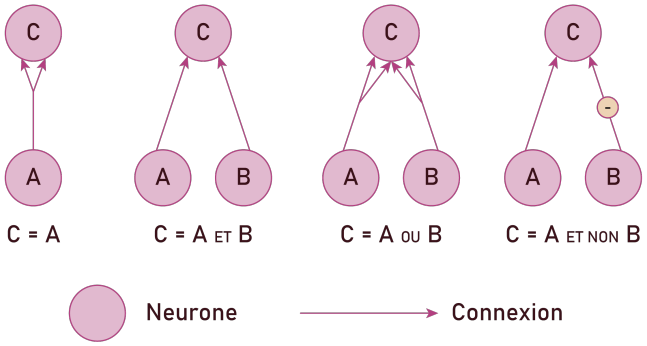


FIGURE 5.2 – Illustration des opérations logiques effectuées à l'aide de neurones artificiels. Figure inspirée de (Géron; 2022). Chaque flèche représente une connexion, et il faut au moins deux entrées positives pour activer le neurone.

L'architecture la plus simple de ANN est le *perceptron* introduit par Rosenblatt (1958). Un neurone seul ne peut effectuer que des tâches linéaires. Pour effectuer des tâches plus complexes, à l'instar du cerveau, on peut utiliser plusieurs neurones, combinés à des fonctions non linéaires. Ainsi, le perceptron utilise plusieurs neurones de McCulloch-Pitts. Ces neurones effectuent des sommes pondérées des entrées, ajoutent un biais, puis une fonction d'activation (non linéaire). Cette dernière, notée ρ, est utilisée pour autoriser au non l'activation des neurones. Les premières fonctions utilisées sont la fonction signe et la fonction de Heaviside définie par :

$$\text{Heaviside}(z) = \begin{cases} 0 & \text{si} \quad z < 0 \\ 1 & \text{si} \quad z \geq 0 \end{cases}.$$

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

La figure 5.3 représente un neurone de McCulloch-Pitts et son fonctionnement. Les poids w associés aux neurones correspondent à la probabilité d'utiliser une liaison plutôt qu'une autre. Le biais ajouté est noté b.

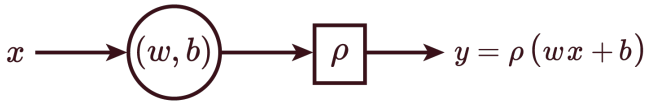


FIGURE 5.3 – Représentation des étapes suivies par une entrée x dans un neurone de McCulloch-Pitts. Le neurone multiplie x par le poids w et lui ajoute le biais b. La fonction d'activation ρ est ensuite appliquée à ce résultat pour donner y, la sortie.

La figure 5.4 schématise l'architecture d'un perceptron à deux entrées et trois sorties. Chacune des entrées passe par les trois neurones. Un perceptron peut déjà effectuer quelques tâches simples, mais si l'on veut complexifier la tâche, il est nécessaire d'augmenter la *profondeur* du réseau.

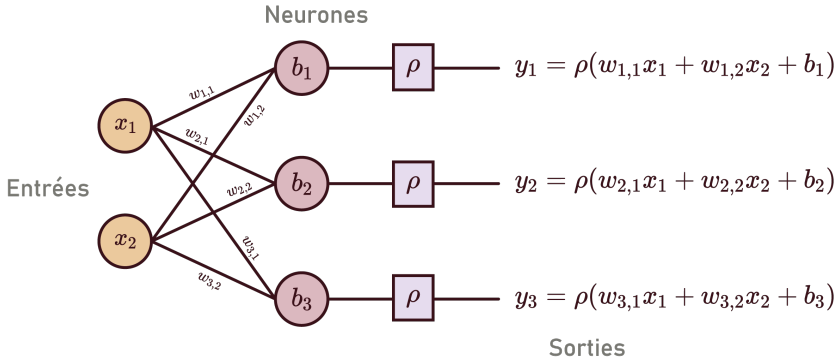


FIGURE 5.4 – Schéma d'un perceptron à deux entrées et trois sorties.

Le *perceptron multi-couche* [1] consiste ainsi en plusieurs couches de perceptrons. Dans cette architecture, la sortie des neurones d'une couche se retrouve en entrée des neurones de la couche suivante, jusqu'à la couche de sortie. Les couches intermédiaires sont appelées *couches cachées* et les réseaux contenant de telles couches sont les réseaux de neurones profonds (DNN [2]). L'empilement de couches et de fonctions d'activation non linéaires permet d'introduire de la non-linéarité et d'apprendre des fonctions plus complexes.

La figure 5.5 représente l'architecture d'un perceptron multi-couches, avec deux entrées, deux couches cachées avec respectivement 8 et 3 neurones, et une sortie. Mathématiquement, un MLP est une fonction définie par :

$$f : \mathbb{R}^d \to \mathbb{R}^n,\ x \mapsto y$$

avec d la dimension d'entrée du réseau et n la dimension de sortie. Dans ce type de réseau, un neurone d'une couche est connecté avec chaque neurone de la couche suivante, ont dit que les couches sont totalement connectées (*fully connected*) entre elles.

Les réseaux plus profonds peuvent apprendre des représentations plus complexes que les réseaux moins profonds. Cela peut contribuer à améliorer les performances du modèle

1. MLP : Multilayer Perceptron
2. *Deep Neural Network*

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

sur de nouvelles données. Comparé à l'apprentissage automatique, un des intérêts majeurs de l'apprentissage profond est que le traitement des données en entrée est moindre car la structure des réseaux de neurones est beaucoup plus flexible. Cela permet de déduire des informations sur les données avec un pré-traitement assez léger en amont. L'inconvénient est qu'il est plus délicat d'avoir une interprétabilité claire sur les résultats.

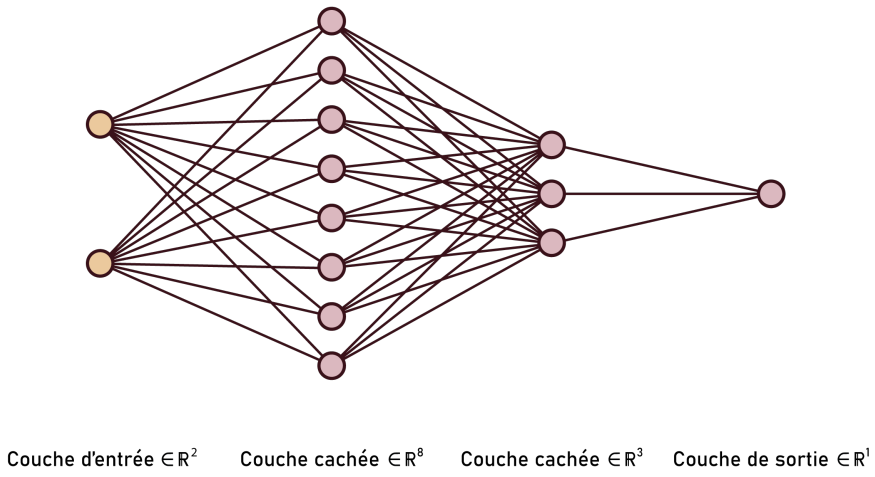


FIGURE 5.5 – Architecture d'un perceptron multi-couche.

Maintenant que les architectures simples de réseaux ont été introduites, nous pouvons revenir au principe de l'apprentissage profond. Comme précédemment pour l'apprentissage automatique, nous cherchons une fonction - ici Γ_θ, le réseau de neurones - qui soit capable de prédire une sortie y à partir d'une entrée x. Les réseaux ont des paramètres θ et des hyperparamètres comme les précédents exemples d'algorithmes d'apprentissage automatique.

3 Apprentissage et paramètres d'un réseau

Cette partie résume les phases d'apprentissage d'un réseau de neurones dans le cas d'apprentissage supervisé. On dispose alors de paires entrée/sortie (x_i, y_i).

Les paramètres sont les poids du réseau, $\theta = (w_k, b_k)$. Ils sont initialisés aléatoirement, puis ils sont modifiés au fur et à mesure de l'apprentissage pour minimiser la valeur de la fonction coût.

Fonctions coût

La fonction coût d'un réseau calcule une erreur entre les valeurs prédites en sortie du réseau et les valeurs de référence. Pendant l'entraînement, les poids du réseau sont optimisés pour minimiser cette fonction coût. Les fonctions coût les plus communes sont l'erreur quadratique moyenne (MSE), l'erreur moyenne absolue (MAE [1]), ou encore l'entropie croisée en classification. Le choix de la fonction coût est important et peut être décisif dans les résultats du réseau (Leuliet *et al.*; 2022).

1. *Mean Absolute Error*

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Optimisation

L'optimisation de ces paramètres s'effectue en deux temps : la propagation avant des données (*feedforward propagation*) et la rétro-propagation (*backward step*).

Propagation avant La propagation avant consiste à donner une entrée x_i au réseau, et à la faire passer à travers les différentes couches pour obtenir une estimation $\hat{y}_i$ en sortie du réseau. La fonction coût $\mathcal{L}$ calcule ensuite la différence entre le résultat obtenu $\hat{y}_i$ et la vérité terrain y_i pour mesure l'erreur de prédiction.

Rétro-propagation À partir de l'erreur de prédiction donnée par la fonction coût, l'étape de rétro-propagation met à jour les paramètres du réseau de sorte à minimiser cette erreur. L'optimisation de ces poids se fait avec un algorithme de type "descente de gradient" :

$$w \leftarrow w - \eta \frac{\partial \mathcal{L}(\hat{y}_i, y_i)}{\partial w} \tag{5.1}$$

avec η le pas de la descente de gradient, appelé *taux d'apprentissage* en apprentissage profond.

Le théorème de dérivation des fonctions composées permet de séparer la contribution de chaque couche dans le calcul de l'erreur finale :

$$\frac{\partial \mathcal{L}(\hat{y}_i, y_i)}{\partial w} = \frac{\partial \mathcal{L}(\hat{y}_i, y_i)}{\partial x} \times \frac{\partial x}{\partial y} \times \frac{\partial y}{\partial w}. \tag{5.2}$$

Un algorithme d'optimisation numérique est utilisé pour ensuite mettre à jour les poids du réseau, par exemple une descente de gradient stochastique (SGD), avec pour objectif la diminution de la valeur de la fonction coût.

En pratique, plusieurs échantillons sont utilisés en même temps : on donne en entrée du réseau un *batch* [1], et le calcul de l'erreur se fait sur l'ensemble du batch considéré. La rétro-propagation est effectuée, puis le processus est répété avant un nouveau batch jusqu'à ce que l'intégralité de la base de donnée soit passée à travers le réseau : ceci constitue ce qu'on appelle une époque (*epoch*). On effectue autant d'époques que nécessaire jusqu'à la convergence du réseau.

La descente de gradient classique est assez lente et de nombreux optimiseurs ont été proposé pour accélérer les calculs. Nous pouvons citer par exemple la méthode d'optimisation par momentum proposée par (Polyak; 1964) ou la méthode de Nesterov (Nesterov; 1983). La méthode la plus courante aujourd'hui est Adam (Kingma et Ba; 2014) et utilise un pas de descente adaptatif.

Fonctions d'activation

Les fonctions d'activation sont choisies au moment de la création de l'architecture du réseau, ce sont donc des hyperparamètres. Elles sont nécessaires pour ajouter de la non-linéarité au réseau : sans elles, le réseau serait réduit à une unique matrice et ne pourrait pas résoudre de problèmes non linéaires.

Les fonctions de Heaviside et signe permettent facilement d'activer le neurone ou non, et sont donc assez performantes pour la propagation avant. Cependant, leurs dérivées sont nulles presque partout, donc au moment du calcul du gradient, aucun paramètre ne peut être ajusté. D'autres fonctions d'activation ont été introduites, différentiables presque partout.

1. Groupe d'échantillons simultanément propagés dans le réseau

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Parmi les fonctions d'activation courantes, nous pouvons citer la sigmoïde, la tangente hyperbolique et la fonction LeakyReLU, représentées graphiquement dans la figure 5.6 et définies par :

$$\mathrm{sig}(x) = \frac{1}{1+e^{-x}} \qquad \tanh(x) = 2\,\mathrm{sig}(2x) - 1 \qquad \mathrm{LeakyReLU}_\lambda(x) = \max(\lambda x, x)$$

Les fonctions sigmoïde et tangente hyperbolique sont utilisées dans les problèmes de régression logistique, où les sorties doivent être dans $[0,1]$, ou en classification binaire où la sortie peut être interprétée comme la probabilité d'appartenir à la classe. La fonction LeakyReLU est une des fonctions les plus utilisées aujourd'hui malgré sa non-différentiabilité en 0.

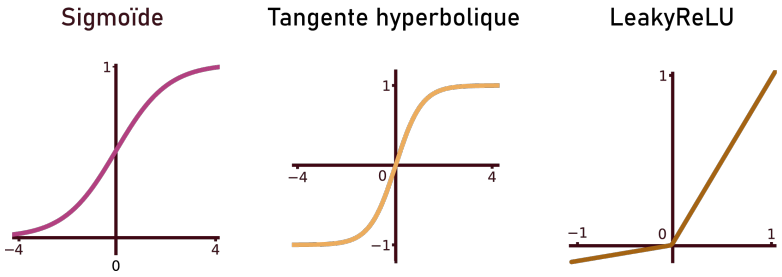


FIGURE 5.6 – Représentation graphique des fonctions d'activation sigmoïde, tangente hyperbolique et LeakyReLU pour un paramètre λ égal à 0.1.

Sur et sous-apprentissage

En apprentissage profond, les réseaux peuvent également souffrir de sur/sous-apprentissage. Le sous-apprentissage peut être dû à plusieurs facteurs : mauvais choix d'hyperparamètres, réseau trop simple, nombre de données insuffisant, etc.

Lorsque la base de données n'est pas assez grande, on peut faire de l'*augmentation de données*. La base de données est augmentée artificiellement en modifiant les données pour les multiplier, sans pour autant modifier leurs propriétés (Shorten et Khoshgoftaar; 2019). Par exemple, en traitement d'images, certaines transformations peuvent être appliquées aux images, telles que des symétries ou des rotations.

Pour empêcher le sur-apprentissage, on peut ajouter de la régularisation : *drop out* [1], *weight clipping* [2], etc. (voir d'autres exemples dans (Géron; 2022) par exemple).

De nos jours, les réseaux sont utilisés pour leur capacité à traiter tout type de données : images, sons, vidéos, texte, données statistiques, etc. Ils sont utilisés et performants dans de nombreux domaines. Nous nous intéresserons dans la suite à l'utilisation de l'apprentissage profond en traitement du signal et de l'image. Les réseaux de neurones convolutifs (CNN) sont particulièrement adaptés à cette modalité.

II Réseaux de neurones convolutifs

Comme les MLP, les réseaux de neurones convolutifs sont composés d'une série de couches de neurones. Cependant les MLP ne sont pas forcément adaptés au traitement d'images. En effet, ce dernier cherche à extraire des caractéristiques dans des données avec une structure spatiale,

1. Technique consistant à désactiver temporairement et aléatoirement certains neurones du réseau à chaque époque.
2. Seuillage de la norme et des valeurs du gradient lors de rétro-propagation.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

mais les MLP nécessiteraient que l'image soit aplatie en un vecteur, ce qui fait perdre toute cohérence spatiale des pixels. De plus, si les entrées du réseau sont des images, la dimension de l'entrée est très élevée. Par exemple, pour une image carrée de taille 256 pixels, l'entrée est de dimension $256^2 = 65\,536$. Si l'on veut utiliser une couche avec 1 000 neurones, ce qui reste peu par rapport à la dimension de l'entrée, on se retrouve avec plus de 65 millions de paramètres, seulement pour la première couche. L'utilisation de couches fully-connected devient alors très complexe, l'entraînement est lent et le réseau peu efficace.

1 Couches de convolution

L'architecture des réseaux de neurones convolutifs s'inspire du cortex visuel et permet de prendre en compte la topologie des données. Les neurones biologiques ont un champ réceptif assez petit, ils ne réagissent qu'aux stimuli visuels d'une région limitée du champ de vue. Les neurones des couches de convolution ne sont connectés qu'à certains neurones de la couche précédente, suivant le schéma des neurones biologiques. Dans un CNN, les images sont données en 2 ou 3 dimensions (avec plusieurs canaux 2D dans ce cas) et chaque couche de convolution détecte des caractéristiques particulières.

Plus précisément, à chaque couche de convolution est associé un *filtre de convolution*. Une couche est alors définie par la taille de son filtre, le *stride* et le *padding* utilisés pour appliquer le filtre.

Filtres Un filtre de convolution, aussi appelé noyau de convolution et noté K, se présente sous la forme d'une matrice, la hauteur et la largeur du filtre correspondent à la taille du champ réceptif et sont définies par la taille de la matrice. Les éléments de la matrice sont les poids à optimiser sur une couche. Ainsi, en prenant un filtre de taille 3×3, il y aura $9 + 1$ paramètres à optimiser (avec le biais) pour une sortie simple. Le nombre de paramètres est donc bien plus faible que pour les MLP. Pour une image I de taille $h \times w$ et un filtre de taille $p \times q$, la taille de l'image de sortie est $(h - p + 1) \times (w - q + 1)$. En pratique, les couches de convolution effectuent une corrélation entre le filtre et l'image. L'opération de corrélation discrète d'une image I par un noyau K est définie par :

$$\begin{aligned}(K * I)(i,j) &= \sum_m \sum_n I(m,n) K(i+m, j+n) \\ &= \sum_m \sum_n I(i+m, j+n) K(m,n).\end{aligned} \tag{5.3}$$

Stride Le *stride* est le pas spatial entre deux applications de filtres. Il joue sur la taille de l'image obtenue après convolution. Le *stride* peut être utilisé pour réduire la taille de l'image et augmenter la quantité d'informations extraites de l'entrée. Il peut également contribuer à rendre la couche convolutive plus efficace en évitant les calculs redondants.

Padding Le *padding* consiste à augmenter la taille de l'image. Il peut être utilisé pour conserver la taille de l'image après la convolution, où simplement donner des dimensions fixées à une image. Il permet également de réduire les effets d'*aliasing*. Lorsque la valeur des pixels ajoutés est fixée à zéro, on parle de *zero padding*. On peut également ajouter des pixels avec des valeurs miroirs si l'on ne veut pas que les zéros affectent le bord de l'image.

Exemples d'application Dans la figure 5.7, cinq filtres ont été appliqués à la photographie d'un graffiti représentant une souris. Les filtres utilisés sont les suivants :

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

(b) $K = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ (c) $K = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ (d) $K = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

(e) $K = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$ (f) $K = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$

Le *stride* a été choisi égal à un, et du *zero padding* a été utilisé pour conserver la même taille d'image. Le filtre (b) applique des moyennes verticales, les lignes verticales sont ainsi conservées. Le filtre horizontal (c) est la version horizontale du filtre précédent. Ces deux filtres peuvent être vus comme des moyennes 1D. Le filtre (d) fait une moyenne des pixels voisins pour chaque pixel, créant ainsi un effet de flou uniforme. Les filtres (e) et (f) sont des filtres utilisés en détection de contours, correspondant respectivement à un dérivateur et un Laplacien. Ces filtres sont connus et déjà utilisés en traitement d'images, les algorithmes d'apprentissage profond vont déterminer d'autres filtres permettant de faire ressortir les caractéristiques des images de la base de données. Nous pouvons remarquer que sur les figures (b), (c) et (d) les bords de l'image sont parfois plus sombres, ceci est dû au zero padding qui réduit la valeur de la convolution lorsque la moyenne pondérée est effectuée avec les zéros. Cet effet aurait été évité avec un padding "miroir".

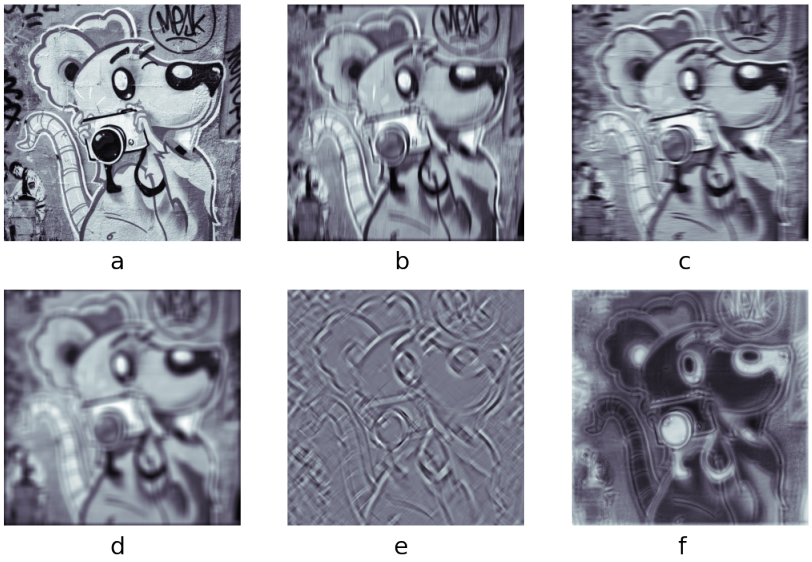


FIGURE 5.7 – Différents filtres appliqués à la photographie[1] d'un graffiti d'une souris (a). (b) Filtre vertical, (c) filtre horizontal, (d) flou uniforme, (e) gradient et (f) contours.

[1] Photographie prise le 12 Septembre 2022 à Porto. Graffeur : @mesk85

2 Couches de pooling

Lorsque les images sont de grandes dimensions ou lorsque l'on veut réduire la dimension pour extraire de nombreuses caractéristiques, des *couches de pooling* sont ajoutées après certaines couches de convolution. Appliquées dans la dimension spatiale, elles correspondent à un sous-échantillonnage de l'image. Il y a plusieurs façons d'effectuer le pooling, mais il faut d'abord

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

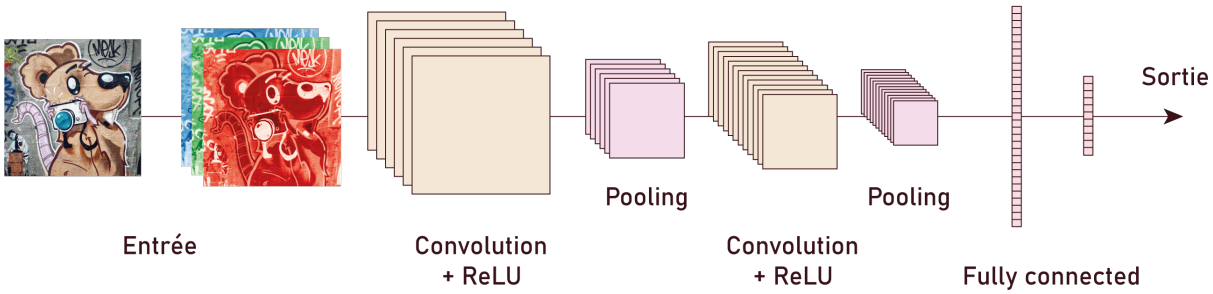


FIGURE 5.8 – Architecture typique d'un CNN utilisé en classification.

définir une taille de fenêtre : une fenêtre de taille 2 divisera la hauteur et la largeur de l'image par 2 et donc le nombre de pixels par 4. Une fois celle-ci choisie, on peut prendre la moyenne, le max ou la somme des valeurs présentes dans chaque fenêtre. Le plus utilisé actuellement est le max pooling car il permet de conserver les informations les plus fortes.

3 Architecture des CNN

Les CNN sont capables d'apprendre les relations spatiales des images en utilisant des couches convolutives. Ces couches sont placées les unes à la suite des autres, comme pour les MLP. Chaque couche apprend des caractéristiques et à la sortie des couches de convolution, une carte des caractéristiques (*feature map*) a été construite à partir de l'image d'entrée. Pendant les convolutions, plusieurs noyaux de convolution sont utilisés, permettant ainsi d'apprendre plusieurs caractéristiques différentes en même temps. Chaque noyau de convolution définit une caractéristique que l'on essaie de détecter dans l'image, en appliquant ces différents noyaux, on obtient différents canaux dans l'image résultante. Cela donne de la profondeur à la feature map (voir figure 5.8). Les couches superficielles d'un réseau neuronal apprennent les caractéristiques de base, telles que les lignes et les courbes dans une image. Quant aux couches les plus profondes, elles apprennent des caractéristiques plus complexes, telles que des formes et des motifs. Le schéma 5.8 représente un réseau typique de classification. Lorsque la sortie du réseau est une autre image, par exemple une version débruitée de l'image en entrée, on utilise des couches appelées "de déconvolution" qui augmentent la taille des images dans l'espace latent. Présenté dans la suite, le U-Net (Ronneberger *et al.*; 2015) est l'architecture typique pour ce genre de réseau.

Le premier CNN, LeNet, a été introduit par LeCun *et al.* (1998) pour faire de la reconnaissance de chiffres manuscrits sur la base de données MNIST [1]. Quelques années plus tard, Krizhevsky *et al.* (2017) ont utilisé leur réseau AlexNet, beaucoup plus large et profond que LeNet et ont remporté le challenge ImageNet de 2012. La performance de leur réseau provenait de l'utilisation du drop out (désactivation aléatoire de certains neurones pendant l'apprentissage) et de l'augmentation de données. En 2015, He *et al.* ont introduit le ResNet qui a surpassé en terme de performance tous les autres précédents réseaux. Ils ont utilisés des *skip connections* (cf flèche grises de la figure 5.9) initialement introduites pour lutter contre les valeurs trop faibles du gradient lors de la rétro-propagation, permettant ainsi d'accélérer le processus d'optimisation et d'améliorer le réseau. Les *skip connections* permettent de sauter certaines des couches du réseau neuronal et d'utiliser la sortie d'une couche comme entrée des couches suivantes.

1. http://yann.lecun.com/exdb/mnist/

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

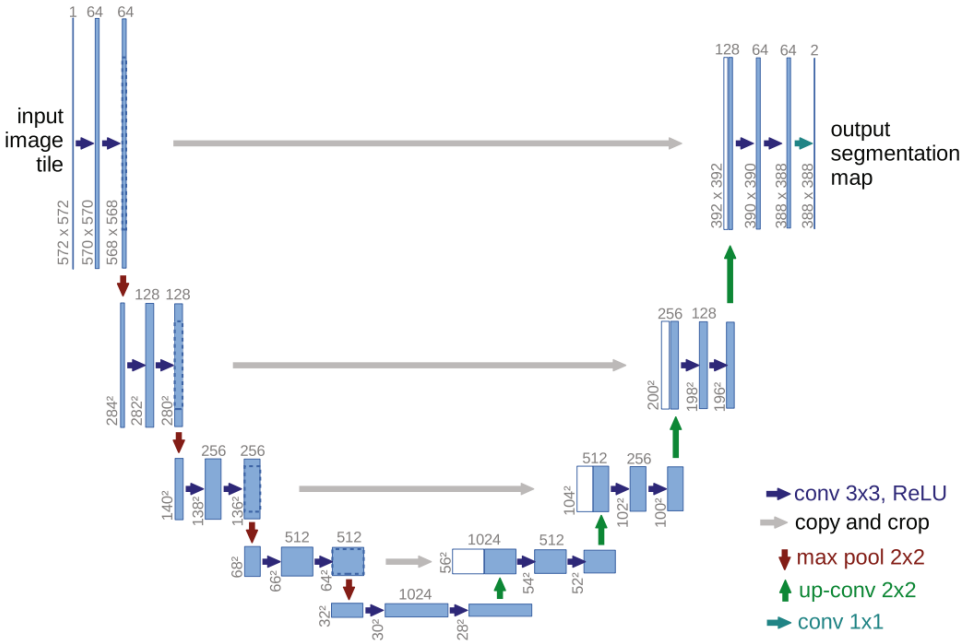


FIGURE 5.9 – Architecture du U-Net original (Ronneberger *et al.*; 2015). La taille des images/-feature maps est notée à gauche des blocs bleus représentant les feature map multi-canaux. Le nombre de canaux est écrit au-dessus de chaque bloc. Les blocs blancs sont les copies des blocs bleus avec pointillés qui sont concaténées au bloc bleu en sortie de la déconvolution. conv $d \times s$ correspond à une convolution avec un filtre de taille $d \times d$.

En 2015, Ronneberger *et al.* ont créé le U-Net pour la segmentation d'images biomédicales. L'architecture du réseau est présenté dans la figure 5.9. On reconnaît les couches de convolution et les couches de pooling sur la première moitié de ce schéma. L'image est convoluée et réduite jusqu'à atteindre 1024 canaux de taille 32×32 chacun dans l'espace latent. Ensuite, des couches de déconvolution sont utilisées pour petit à petit revenir à la taille initiale. Les auteurs ont ajouté des *skip connections*, représentées par les flèches grises, elles servent ici à conserver un maximum d'informations spatiales au moment de la déconvolution.

III Apprentissage profond en CT

Avec les avancées considérables que l'utilisation de l'apprentissage profond a permis en traitement de l'image, mais aussi en résolution de problèmes inverses (Arridge *et al.*; 2019), il est naturel de retrouver son application en imagerie médicale, et en particulier en reconstruction tomographique. Des réseaux sont déjà utilisés pour de nombreuses tâches en CT (Wang *et al.*; 2020) : pour améliorer la qualité des images en faible-dose (Chen *et al.*; 2017b, 2018), réduire les artefacts des reconstructions avec peu de projections (Jin *et al.*; 2017; Gupta *et al.*; 2018), les artefacts d'angle limité (Bubba *et al.*; 2019) ou encore les artefacts métalliques (Zhang et Yu; 2018; Liao *et al.*; 2019).

L'apprentissage profond peut être utilisé directement pour effectuer la reconstruction ou associé aux méthodes classiques. Cette partie cherche à présenter les différentes manières d'utiliser l'apprentissage profond en reconstruction tomographique et plus particulièrement dans le cas de faible dose. Nous parlerons plus en détail des méthodes supervisées. Nous pouvons séparer ces méthodes en deux groupes : celles qui utilisent une méthode de reconstruction

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

annexe, telle que FBP ou FDK, et qui appliquent les réseaux en pré ou post-traitement ; et celles qui incluent le processus de reconstruction au sein du réseau de neurones. Nous présenterons quelques méthodes non-supervisées dans la dernière sous-partie.

1 En post-traitement

Les méthodes de post-traitement cherchent à améliorer les reconstructions analytiques grâce à l'apprentissage profond. Les réseaux travaillent dans le domaine des images, n'utilisant ainsi pas les sinogrammes, et visent à améliorer leur qualité en réduisant le bruit et les artefacts présents dans les reconstructions analytiques. L'utilisation de l'apprentissage profond en débruitage d'images CT est justifiée par la difficulté de modéliser de manière exacte le bruit présent dans ces images. Les méthodes classiques de débruitage, qui reposent sur des a priori sur le type de bruit présent ne sont donc pas forcément capables de le supprimer correctement. De plus, le fait que les calculs se fassent sur les images (plutôt que sur les sinogrammes) présente deux avantages : l'utilisation de CNN est notamment adaptée pour cette tâche et la géométrie d'acquisition n'a pas besoin d'être prise en compte dans le réseau.

Ces réseaux sont plutôt appliqués en conjonction avec les méthodes analytiques car celles-ci ont l'avantage d'être particulièrement rapides comparées aux méthodes itératives. L'amélioration de la qualité de l'image devant être effectuée par le réseau, il n'est pas nécessaire de recourir à la régularisation qu'offrent les méthodes itératives.

Pour entraîner ce type de réseau, on dispose d'une base de données contenant des images LDCT (*Low-Dose CT)* et les images NDCT (*Normal Dose CT)* associées. On note $g \in \mathbb{R}^J$ une image LDCT, et $f \in \mathbb{R}^J$ l'image NDCT qui lui correspond, avec J le nombre de pixels. La relation entre ces deux images peut être formulée de la manière suivante :

$$g = \gamma(f) \tag{5.4}$$

avec $\gamma : \mathbb{R}^J \to \mathbb{R}^J$ l'opérateur représentant la dégradation des images LDCT par rapport aux images NDCT. Le problème consiste alors à identifier une fonction Γ_{θ^*} qui permette d'améliorer la qualité de l'image LDCT pour qu'elle soit semblable à une image NDCT. En prenant la norme ℓ_2 comme fonction coût, cela s'écrit :

$$\arg\min_{\theta} \|\Gamma_\theta(g) - f\|_2^2 . \tag{5.5}$$

La fonction Γ_θ est alors un réseau de neurones que l'on optimise sur les paramètres θ.

Les premières méthodes ayant utilisé des CNN en post-traitement en reconstruction tomographique sont FBPConvNet (Jin *et al.*; 2017) et RED-CNN (Chen *et al.*; 2017a). Les deux approches consistent à effectuer la rétro-projection filtrée sur des projections faibles-doses, puis entraînent le réseau à améliorer la qualité des reconstructions. Les deux réseaux utilisent du *Residual Learning*, la différence entre les deux méthodes réside dans le choix du réseau : FBPConvNet utilise un U-Net, et RED-CNN un auto-encodeur.

Residual Learning : Notre objectif ici est d'obtenir une fonction Γ_θ qui puisse estimer f à partir de g. L'idée du *residual learning* est d'apprendre à estimer la différence entre g et f, c'est-à-dire le résidu. Cela se fait tout simplement en ajoutant une *skip connection* entre l'entrée et la sortie du réseau. Le *residual learning* permet de diminuer les valeurs du gradient et de mieux conserver les détails de l'image d'origine.
Auto-Encodeur (AE) : Un AE est un réseau conçu pour compresser les informations dans la partie encodeur du réseau, puis les décompresser dans la partie décodeur du réseau. Un U-Net est un certain type d'AE avec des *skip connections* en plus.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Ces méthodes ont par la suite inspiré de nouveaux travaux améliorant les performances des U-Net. Nous pouvons citer l'article (Abascal *et al.*; 2021) dans lequel un U-Net résiduel utilise un a priori sur l'image en entrée, en plus de l'image dégradée. Gunduzalp *et al.* (2021) et Song *et al.* (2022) ont proposé des versions 3D du U-Net utilisé par Jin *et al.* (2017), permettant ainsi au réseau de prendre en compte les informations spatiales présentes en 3D. Au lieu d'appliquer des convolutions 2D, les U-Net 3D utilisent des filtres 3D pour effectuer les convolutions, les images d'entrée et de sortie sont des volumes 3D. Ces algorithmes nécessitent cependant beaucoup plus de ressources numériques que les réseaux 2D, puisque la taille des cartes de caractéristiques est plus grande. Il faut alors trouver des solutions pour que l'apprentissage puisse encore être possible. On peut, pour cela, travailler sur des *patchs* [1] plutôt que sur l'intégralité du volume et réduire la *taille de batch* [2] en entrée.

Ce type de méthodes sera utilisé dans le chapitre 6 pour améliorer la qualité d'image de volumes 3D dentaires reconstruits à partir d'un faible nombre de projections.

Les méthodes de post-traitement reposent sur une reconstruction analytique du volume. Des projections de faible qualité peuvent faire perdre des informations pendant la reconstruction, qui ne pourront pas être restaurées par le réseau. De plus, c'est dans les projections brutes que l'on trouve le plus d'informations, et celles-ci pourraient être utilisées avant la reconstruction.

2 En pré-traitement

Nous avons vu précédemment (cf chapitre 4) que la réduction de la dose pouvait se faire de deux manières différentes : en diminuant l'intensité du faisceau de rayons de X, ce qui induit un bruit photonique plus important sur les projections, ou en diminuant le nombre d'angles de vue, ce qui diminue le nombre de projections et crée des artefacts étoilés dans les reconstructions.

Dans le premier cas, le débruitage des projections est assez délicat à effectuer puisque ce sont des images dans le domaine de la transformée de Radon, les méthodes de débruitage classiques ne sont pas du tout adaptées. Ghani et Karl (2018a) se sont inspiré du DnCNN (Zhang *et al.*; 2017) et utilisent un CNN résiduel pour débruiter des sinogrammes corrompus par une combinaison de bruit gaussien et Poisson. Cependant, leur travail ne s'est appliqué que dans un cadre de simulations. La connaissance incomplète du type de bruit présent dans le sinogramme ne permet pas vraiment de générer des bases de données suffisamment réalistes pour effectuer l'entraînement et ensuite généraliser à des données réelles. De plus, l'acquisition de paires de sinogrammes bruités et non bruités est impossible à obtenir en pratique. Ainsi, les réseaux supervisés ne sont pas vraiment adaptés à cette tâche, nous verrons dans la suite que d'autres réseaux non-supervisés sont plus performants.

Dans le second cas, il faudrait augmenter le nombre de projections en effectuant de l'*inpaiting* sur le sinogramme. Cette tâche est également assez complexe, en particulier lorsque le nombre de projections est très faible, puisque l'introduction de mauvaises informations pourraient dégrader davantage la reconstruction. L'apprentissage profond pourrait être alors une solution : capable d'apprendre des distributions plus complexes que nos modélisations habituelles, il est une solution potentielle aux problèmes de traitement des sinogrammes. Dans (Ghani et Karl; 2018b), un CNN apprend à faire de l'*inpainting* sur le sinogramme, la reconstruction est ensuite effectuée avec une rétro-projection filtrée. La base de données pour l'apprentissage est simulée en sous-échantillonnant le nombre d'angles dans le sinogramme. Dong *et al.* (2019) ont proposé une méthode en deux étapes : le sinogramme est premièrement interpolé linéairement, puis un réseau vient améliorer la qualité de l'interpolation. Cette méthode a montré de bons résultats

1. Un patch de taille $d \times d \times d$ est un cube de côté d extrait d'un volume.
2. Nombre d'images ou de patchs utilisés dans un batch.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

en micro-CT sur des images d'os de rats.

Les deux méthodes présentées pour le moment ne sont basées que sur les données (images ou sinogrammes) et n'utilisent aucune connaissance de la physique d'acquisition. Ces deux approches peuvent d'ailleurs être utilisées conjointement : traitement du sinogramme avec un réseau, reconstruction, puis amélioration de la qualité de l'image avec un autre réseau. Une autre manière de traiter la reconstruction tomographique consiste à entraîner un réseau à faire l'intégralité de la reconstruction. Le même réseau peut alors effectuer lui-même les corrections sur le sinogramme, la reconstruction et la réduction des artefacts de l'image, ce sont les méthodes *end-to-end*.

3 Méthodes end-to-end

Méthodes apprenant l'opérateur inverse

Initialement introduit pour l'imagerie par IRM, AUTOMAP (Zhu *et al.*; 2018) est un réseau de neurones qui peut réaliser une cartographie entre les domaines des projections et des images comme une tâche d'apprentissage supervisé. Son processus de reconstruction étant assez général, il a été appliqué à la reconstruction CT par Liu *et al.* (2020). Son architecture commence avec deux couches complètement connectées, rendant ainsi son utilisation possible seulement sur des images et sinogrammes de taille limitée. Également avec des couches complètement connectées, nous pouvons citer le réseau iRadonMap (He *et al.*; 2020) qui a montré sa capacité de généralisation en ayant été entraîné uniquement sur des données d'images naturelles, les données médicales n'ont été utilisées que pour la phase de tests.

Dans (Li *et al.*; 2019), un réseau convolutif, iCT-Net, a été développé et entraîné pour reconstruire avec précision des images à partir de projections obtenues dans diverses conditions d'acquisition. Des reconstructions précises ont été réalisées dans le cas de faible nombre de projections et en tomographie intérieure (définition p. 38).

Kandarpa *et al.* (2022) ont implémenté deux réseaux basés respectivement sur des blocs denses complètement convolutifs (*Fully convolutional dense network)* et sur l'architecture d'un U-Net, pour reconstruire directement des images à partir d'un faible nombre de projections. Les très bons résultats obtenus sont entre autre dus à la concaténation, à différents niveaux de résolutions, d'images reconstruites avec FBP. L'utilisation d'estimations FBP de plus faible résolution permet aux réseaux, qui sont principalement utilisés pour le débruitage, d'apprendre la transformation plus complexe entre le sinogramme et le domaine de l'image.

Ces méthodes présentent deux inconvénients majeurs à leur utilisation en imagerie médicale. Le premier provient de l'utilisation de couches complètement connectées qui augmentent le nombre de paramètres à optimiser, augmentant la taille de la mémoire nécessaire et rendant l'application de telles méthodes uniquement en 2D pour le moment. Pour le second, puisque la reconstruction est effectuée par les couches complètement connectées, la connaissance de la géométrie du système n'est pas utilisée. Cela pose des questions sur la robustesse de ces méthodes. Pour ces deux raisons, ces réseaux nécessitent donc de grandes bases de données pour pouvoir généraliser correctement après l'entraînement. Or, les images médicales ne sont généralement pas disponibles en très grand nombre, ainsi pour l'instant, ces méthodes ont des performances moindres que celles présentées précédemment (Baguer *et al.*; 2020).

Unrolling

Les méthodes de *unrolling* proposent des algorithmes capables de prendre en compte la connaissance des propriétés du système d'acquisition tout en améliorant la qualité de l'image,

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

évitant ainsi les inconvénients majeurs des méthodes de post-traitement et des méthodes end-to-end précédentes.

Le principe est de reprendre le fonctionnement des méthodes itératives classiques en remplaçant les opérations obtenues par calcul de gradient ou d'opérateur proximal par des réseaux de neurones. Adler et Öktem (2018) ont publié une méthode générale, basée sur l'algorithme 1 de Chambolle-Pock (Chambolle et Pock; 2011) utilisant le unrolling pour résoudre des problèmes inverses. Nous rappelons la forme des itérations de l'algorithme PDHG (introduit et plus détaillé p. 51) :

$$\begin{cases} y^{(k+1)} \leftarrow \text{prox}_{\sigma}[F^*]\left(y^{(k)} + \sigma A \bar{f}^{(k)}\right) \\ f^{(k+1)} \leftarrow \text{prox}_{\tau}[G]\left(f^{(k)} - \tau\, {}^tA y^{(k+1)}\right) \\ \bar{f}^{(k+1)} \leftarrow f^{(k+1)} + \theta\left(f^{(k+1)} - f^{(k)}\right) \end{cases} \tag{5.6}$$

L'algorithme Learned Primal-Dual (Adler et Öktem; 2018) propose de remplacer les opérateurs proximaux et les différentes combinaisons linéaires par des réseaux convolutifs. Les itérations deviennent alors :

$$\begin{cases} y^{(k+1)} \leftarrow \Gamma_{\theta_d^{(k)}}\left(y^{(k)}, A\bar{f}^{(k)}, p\right) \\ f^{(k+1)} \leftarrow \Lambda_{\theta_p^{(k)}}\left(f^{(k)}, {}^tA y^{(k+1)}\right) \end{cases} \tag{5.7}$$

avec θ_p les paramètres du réseau primal et θ_d ceux du réseau dual. Ainsi, une itération de l'algorithme PDHG correspond à l'application successive des réseaux $\Gamma_{\theta_d^{(k)}}$ et $\Lambda_{\theta_p^{(k)}}$ De manière similaire à l'algorithme itératif, cette étape est répétée plusieurs fois : on donne en entrée du réseau une initialisation $f^{(0)}$ de f et les projections puis les réseaux convolutifs sont appliqués sur plusieurs itérations.

Contrairement aux réseaux end-to-end entièrement appris, ici les opérateurs A et ${}^tA$ sont connus et utilisés en tant que couches neuronales à poids fixes. Ces opérateurs ont été implémentés comme couches neuronales dans les librairies ODL (Adler *et al.*; 2017) et PYRO-NN (Syben *et al.*; 2019), par exemple. L'algorithme Learned Primal-Dual passe du domaine du sinogramme au domaine spatial plusieurs fois pendant l'entraînement et la reconstruction, à la manière des méthodes itératives classiques. Les entraînements, dans ce cas là, se font de manière supervisée.

Le réseau a été entraîné sur des fantômes d'ellipses et des fantômes humains, les résultats montrent que cette méthode surpasse les méthodes de post-traitement avec réseau, notamment lorsque peu de données sont disponibles. Cette méthode requiert cependant des ressources numériques conséquentes puisque les opérations de projection et de rétro-projection, qui sont généralement coûteuses, sont effectuées dans le réseau. Les applications en 3D et plus particulièrement en CBCT, où les projections doivent être prises en compte simultanément, sont donc freinées par cet inconvénient.

Pour éviter les contraintes liées au temps de calcul et à la mémoire, Tang *et al.* (2021) ont proposé un réseau stochastique de unrolling qui reconstruit le volume à partir de sous-ensembles des projections, à la manière de OSEM pour MLEM.

La recherche sur les méthodes de type unrolling n'en est encore qu'à ses débuts et les résultats préliminaires obtenus sur données faibles-doses sont assez prometteurs. L'entraînement peut se faire avec une base de données un peu moins importante que les autres méthodes end-to-end, mais les besoins en mémoire restent encore un point qui freine l'utilisation de unrolling pour tout type de reconstruction tomographique, à cause du nombre limité d'itérations.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

4 Méthodes semi ou non-supervisées

Les méthodes supervisées nécessitent des paires d'images, par exemple image bruitée et image non bruitée. Pour récupérer de telles paires, deux manières existent. La première consiste à acquérir les projections deux fois pour chaque patient, dans deux configurations de scan différentes. La création de la base de données appariées dans ce cas là nécessite tout de même des opérations de recalage. De plus, le plus grand inconvénient de cette méthode est la double exposition aux radiations des patients, difficile à justifier d'un point de vue médical. La seconde, plus rapide et plus facile à mettre en œuvre, simule les données de moins bonne qualité à partir de données réelles acquises dans de bonnes conditions. Il s'agit alors d'ajouter du bruit aux projections ou au volume reconstruit, de réduire le nombre de projections ou encore d'ajouter des artefacts. Si cette méthode paraît plus simple, n'ayant fait l'entraînement que sur données simulées, sa capacité de généralisation à des données réelles pose question. Le bruit ou les artefacts ajoutés sont-ils assez réalistes ? Ainsi, de nouvelles méthodes d'apprentissage profond non-supervisé sont apparues et nous en présentons quelques unes dans ce qui suit.

La plupart des méthodes présentées dans le cadre de l'apprentissage supervisé ont leur équivalent en non-supervisé. Il est important de noter que le but des méthodes non-supervisées n'est pas de surpasser les méthodes supervisées. En effet, celles-ci disposant de la vérité terrain comme objectif lors de l'apprentissage sont favorisées. Le but des méthodes non-supervisées est de proposer des algorithmes ayant des performances comparables à celles des autres réseaux, sans nécessiter de bases de données appariées.

En ce qui concerne le traitement du sinogramme, Zainulina *et al.* (2021) ont proposé un réseau qui n'utilise que les projections bruitées pour les débruiter : c'est une méthode d'apprentissage auto-supervisé (*self-supervised learning*). Pour débruiter une coupe du sinogramme 3D, le CNN exploite les informations contenues dans les coupes à proximité.

Pour le post-traitement, toutes les techniques de débruitage d'images (développées pour débruiter des images naturelles initialement) peuvent être utilisées. Récemment, Ulyanov *et al.* (2018) ont publié une méthode, nommée *Deep Image Prior* (DIP), pouvant efficacement débruiter une image sans entraînement préalable et sans connaissance de l'image non bruitée. Cette méthode reste assez lente, mais sa capacité à débruiter correctement une image avec un réseau de neurones non-supervisé l'a rendue assez populaire et de nombreuses études l'ont ensuite utilisée en tomographie (Gong *et al.*; 2018; Baguer *et al.*; 2020; Barutcu *et al.*; 2021).

Dans le domaine de l'image, le développement des GAN [1] (Goodfellow *et al.*; 2014) a permis d'augmenter le champ des possibles des méthodes d'apprentissage profond, en particulier en génération d'images. Ainsi, de même que les CNN, ces méthodes ont rapidement été appliquées à la reconstruction tomographique (Yi *et al.*; 2019).

Même si les GAN ont déjà été utilisés en apprentissage supervisé (Yang *et al.*; 2018), leur capacité à apprendre une distribution de probabilité en fait des méthodes particulièrement intéressantes en apprentissage non-supervisé. Park *et al.* (2019) ont entraîné un GAN à générer des images normales-doses à partir d'images faibles-doses non appairées. Introduits par Zhu *et al.* (2017), les Cycle-GAN entraînent simultanément deux GAN, l'un apprenant à générer des images dans l'espace B à partir de l'espace A, et l'autre l'inverse. Cette architecture de réseau améliore la qualité du GAN visant à réduire les artefacts et le bruit dans une image. Ces algorithmes ont déjà été utilisés pour réduire les artefacts de cone beam (Park et Ye; 2020) et en débruitage d'images faibles-doses (Kwon et Ye; 2021).

1. *Generative Adversarial Networks*

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Les méthodes présentées dans ce chapitre ne donnent qu'un léger aperçu de l'utilisation de l'apprentissage profond en reconstruction tomographique. De nombreuses autres méthodes, architectures de réseau et techniques dont nous n'avons pas parlé ont été développées. Pour résumer celles que nous avons listées, nous pouvons dire qu'elles se distinguent selon leur méthode d'entraînement (supervisée, non-supervisée) ou leur utilisation de l'apprentissage profond dans le processus de la reconstruction (post-processing, end-to-end, etc.). Chaque méthode présente ses avantages et ses inconvénients, ce sont l'objectif de la tâche à accomplir, la base de données utilisée et les ressources numériques disponibles qui peuvent amener à privilégier une méthode plutôt qu'une autre.

En post ou pré-traitement, l'apprentissage profond a montré des résultats très performants, surpassant les méthodes itératives, tout en se basant sur la reconstruction rapide FBP ou FDK. Ces méthodes ne travaillent que sur un seul domaine, les sinogrammes ou les images. Ainsi, même des réseaux relativement simples, tels que les U-Net, ont d'excellentes performances, particulièrement dans le cas d'apprentissage supervisé. L'idée des méthodes *end-to-end* est d'inclure le processus de reconstruction à l'intérieur du réseau pour créer un algorithme qui n'utilise que les projections brutes qui contiennent le maximum d'information. Les méthodes end-to-end qui ne comprennent pas l'opérateur de reconstruction exact dans leur processus, peuvent manquer de garantie sur la qualité de l'inversion et de sa capacité à généraliser lorsque la taille de la base de données d'entraînement n'est pas suffisante. De plus, les couches complètement connectées nécessaires à la reconstruction, augmentent le nombre de paramètres du réseau et rendent ces méthodes peu adaptées à la reconstruction 3D CBCT. Les algorithmes de *unrolling* prennent en compte la géométrie d'acquisition et permettent d'effectuer exactement les opérations de projection et rétro-projection. Cependant, en cone beam, leur application reste limité à des volumes de très petite taille puisqu'il faut prendre en compte toutes les projections à chaque opération.

En imagerie médicale, il est souvent assez compliqué d'obtenir des bases de données appairées. Leur simulation n'est par ailleurs pas toujours possible, voire peu représentative de la réalité. Ainsi, les méthodes peu ou non-supervisées ont vu leur popularité augmenter en imagerie médicale dernièrement. Cependant, excepté pour les méthodes auto-supervisées, il leur faut de grandes bases de données afin d'apprendre un maximum de caractérisation par elles-mêmes, et ceci peut être délicat à obtenir dans un tel domaine.

Le chapitre suivant s'intéresse à la mise en œuvre d'une méthode d'apprentissage profond de post-traitement supervisée en CBCT dentaire, et compare ses résultats avec les reconstructions analytiques de FDK (Feldkamp *et al.*; 1984) et les reconstructions itératives de KL-TV (Sidky *et al.*; 2012) décrites dans le chapitre 3.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Bibliographie

Abascal, J. F., Bussod, S., Ducros, N., Si-Mohamed, S., Douek, P., Chappard, C. et Peyrin, F. (2021). A residual U-Net network with image prior for 3D image denoising. *In 2020 28th European Signal Processing Conference (EUSIPCO)*, pages 1264–1268. IEEE. 109

Adler, J., Kohr, H. et Öktem, O. (2017). Operator discretization library (ODL). *Zenodo*. 111

Adler, J. et Öktem, O. (2018). Learned primal-dual reconstruction. *IEEE transactions on medical imaging*, 37(6):1322–1332. 111

Arridge, S., Maass, P., Öktem, O. et Schönlieb, C.-B. (2019). Solving inverse problems using data-driven models. *Acta Numerica*, 28:1–174. 107

Baguer, D. O., Leuschner, J. et Schmidt, M. (2020). Computed tomography reconstruction using deep image prior and learned reconstruction methods. *Inverse Problems*, 36(9):094004. 110, 112

Barutcu, S., Aslan, S., Katsaggelos, A. K. et Gürsoy, D. (2021). Limited-angle computed tomography with deep image and physics priors. *Scientific reports*, 11(1):1–12. 112

Bubba, T. A., Kutyniok, G., Lassas, M., März, M., Samek, W., Siltanen, S. et Srinivasan, V. (2019). Learning the invisible : a hybrid deep learning-shearlet framework for limited angle computed tomography. *Inverse Problems*, 35(6):064002. 107

Chambolle, A. et Pock, T. (2011). A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, 40(1):120–145. 111

Chen, H., Zhang, Y., Chen, Y., Zhang, J., Zhang, W., Sun, H., Lv, Y., Liao, P., Zhou, J. et Wang, G. (2018). LEARN : Learned experts' assessment-based reconstruction network for sparse-data CT. *IEEE transactions on medical imaging*, 37(6):1333–1347. 107

Chen, H., Zhang, Y., Kalra, M. K., Lin, F., Chen, Y., Liao, P., Zhou, J. et Wang, G. (2017a). Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE transactions on medical imaging*, 36(12):2524–2535. 108

Chen, H., Zhang, Y., Zhang, W., Liao, P., Li, K., Zhou, J. et Wang, G. (2017b). Low-dose CT via convolutional neural network. *Biomedical optics express*, 8(2):679–694. 107

Dong, X., Vekhande, S. et Cao, G. (2019). Sinogram interpolation for sparse-view micro-CT with deep learning neural network. *In Medical Imaging 2019 : Physics of Medical Imaging*, volume 10948, pages 692–698. SPIE. 109

Feldkamp, L. A., Davis, L. C. et Kress, J. W. (1984). Practical cone-beam algorithm. *Josa a*, 1(6):612–619. 113

Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc. xiii, 96, 99, 103

Ghani, M. U. et Karl, W. C. (2018a). CNN based sinogram denoising for low-dose CT. *In Mathematics in Imaging*, pages MM2D–5. Optica Publishing Group. 109

Ghani, M. U. et Karl, W. C. (2018b). Deep learning-based sinogram completion for low-dose CT. *In 2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, pages 1–5. IEEE. 109

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Gong, K., Catana, C., Qi, J. et Li, Q. (2018). Pet image reconstruction using deep image prior. *IEEE transactions on medical imaging*, 38(7):1655–1665. 112

Goodfellow, I., Bengio, Y. et Courville, A. (2016). *Deep learning*. MIT press. 96

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. et Bengio, Y. (2014). Generative adversarial networks. *arXiv preprint arXiv :1406.2661*. 112

Gunduzalp, D., Cengiz, B., Unal, M. O. et Yildirim, I. (2021). 3D U-NetR : Low Dose Computed Tomography Reconstruction via Deep Learning and 3 Dimensional Convolutions. *arXiv preprint arXiv :2105.14130*. 109

Gupta, H., Jin, K. H., Nguyen, H. Q., McCann, M. T. et Unser, M. (2018). CNN-based projected gradient descent for consistent CT image reconstruction. *IEEE transactions on medical imaging*, 37(6):1440–1453. 107

He, J., Wang, Y. et Ma, J. (2020). Radon inversion via deep learning. *IEEE transactions on medical imaging*, 39(6):2076–2087. 110

He, K., Zhang, X., Ren, S. et Sun, J. (2015). Deep residual learning for image recognition. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778. 106

Jin, K. H., McCann, M. T., Froustey, E. et Unser, M. (2017). Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26(9):4509–4522. 96, 107, 108, 109

Kandarpa, V., Perelli, A., Bousse, A. et Visvikis, D. (2022). LRR-CED : low-resolution reconstruction-aware convolutional encoder–decoder network for direct sparse-view CT image reconstruction. *Physics in Medicine & Biology*, 67(15):155007. 110

Kingma, D. P. et Ba, J. (2014). Adam : A method for stochastic optimization. *arXiv preprint arXiv :1412.6980*. 102

Krizhevsky, A., Sutskever, I. et Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90. 106

Kwon, T. et Ye, J. C. (2021). Cycle-free cyclegan using invertible generator for unsupervised low-dose ct denoising. *IEEE Transactions on Computational Imaging*, 7:1354–1368. 112

LeCun, Y., Bottou, L., Bengio, Y. et Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324. 106

Leuliet, T., Maxim, V., Peyrin, F. et Sixou, B. (2022). Impact of the training loss in deep learning–based CT reconstruction of bone microarchitecture. *Medical Physics*. 101

Li, Y., Li, K., Zhang, C., Montoya, J. et Chen, G.-H. (2019). Learning to Reconstruct Computed Tomography Images Directly From Sinogram Data Under A Variety of Data Acquisition Conditions. *IEEE transactions on medical imaging*, 38(10):2469–2481. 110

Liao, H., Lin, W.-A., Zhou, S. K. et Luo, J. (2019). ADN : artifact disentanglement network for unsupervised metal artifact reduction. *IEEE Transactions on Medical Imaging*, 39(3):634–643. 107

Liu, C., Huang, Y., Maier, J., Klein, L., Kachelrieß, M. et Maier, A. (2020). Robustness Investigation on Deep Learning CT Reconstruction for Real-Time Dose Optimization. *arXiv preprint arXiv :2012.03579*. 110

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Lundervold, A. S. et Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik*, 29(2):102–127. 96

McCulloch, W. S. et Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133. 98

Minaee, S., Boykov, Y. Y., Porikli, F., Plaza, A. J., Kehtarnavaz, N. et Terzopoulos, D. (2021). Image segmentation using deep learning : A survey. *IEEE transactions on pattern analysis and machine intelligence*. 96

Nesterov, Y. E. (1983). A method for solving the convex programming problem with convergence rate $O(1/k^2)$. *In Dokl. Akad. Nauk SSSR,*, volume 269, pages 543–547. 102

Park, H. S., Baek, J., You, S. K., Choi, J. K. et Seo, J. K. (2019). Unpaired image denoising using a generative adversarial network in x-ray ct. *IEEE Access*, 7:110414–110425. 112

Park, S. et Ye, J. C. (2020). Unsupervised Cone-Beam artifact removal using CycleGAN and spectral blending for adaptive radiotherapy. ISBI. 112

Pierson, H. A. et Gashler, M. S. (2017). Deep learning in robotics : a review of recent research. *Advanced Robotics*, 31(16):821–835. 96

Polyak, B. T. (1964). Some methods of speeding up the convergence of iteration methods. *Ussr computational mathematics and mathematical physics*, 4(5):1–17. 102

Rawat, W. et Wang, Z. (2017). Deep convolutional neural networks for image classification : A comprehensive review. *Neural computation*, 29(9):2352–2449. 96

Reader, A. J., Corda, G., Mehranian, A., da Costa-Luis, C., Ellis, S. et Schnabel, J. A. (2020). Deep learning for PET image reconstruction. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 5(1):1–25. 96

Ronneberger, O., Fischer, P. et Brox, T. (2015). U-net : Convolutional networks for biomedical image segmentation. *In International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer. xiii, 106, 107

Rosenblatt, F. (1958). The perceptron : a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386. 99

Shorten, C. et Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48. 103

Sidky, E. Y., Jørgensen, J. H. et Pan, X. (2012). Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm. *Physics in Medicine & Biology*, 57(10):3065–3091. 113

Song, X., Han, Y., Xi, X., Li, L., Zhu, L., Yang, S., Liu, M., Tan, S. et Yan, B. (2022). Preliminary denoising by 3D U-Net in image domain for low dose CT images. *In 2022 2nd International Conference on Bioinformatics and Intelligent Computing*, pages 367–370. 109

Syben, C., Michen, M., Stimpel, B., Seitz, S., Ploner, S. et Maier, A. K. (2019). Technical Note : PYRO-NN : Python reconstruction operators in neural networks. *Medical Physics*. 111

Tang, J., Mukherjee, S. et Schönlieb, C.-B. (2021). Stochastic primal-dual deep unrolling. *arXiv preprint arXiv :2110.10093*. 111

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Tian, C., Fei, L., Zheng, W., Xu, Y., Zuo, W. et Lin, C.-W. (2020). Deep learning on image denoising : An overview. *Neural Networks*, 131:251–275. 96

Ulyanov, D., Vedaldi, A. et Lempitsky, V. (2018). Deep image prior. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9446–9454. 112

Van Sloun, R. J., Cohen, R. et Eldar, Y. C. (2019). Deep learning in ultrasound imaging. *Proceedings of the IEEE*, 108(1):11–29. 96

Voulodimos, A., Doulamis, N., Doulamis, A. et Protopapadakis, E. (2018). Deep learning for computer vision : A brief review. *Computational intelligence and neuroscience*, 2018. 96

Wang, G., Ye, J. C. et De Man, B. (2020). Deep learning for tomographic image reconstruction. *Nature Machine Intelligence*, 2(12):737–748. 107

Yang, Q., Yan, P., Zhang, Y., Yu, H., Shi, Y., Mou, X., Kalra, M. K., Zhang, Y., Sun, L. et Wang, G. (2018). Low-dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss. *IEEE transactions on medical imaging*, 37(6):1348–1357. 112

Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J.-H. et Liao, Q. (2019). Deep learning for single image super-resolution : A brief review. *IEEE Transactions on Multimedia*, 21(12):3106–3121. 96

Yi, X., Walia, E. et Babyn, P. (2019). Generative adversarial network in medical imaging : A review. *Medical image analysis*, 58:101552. 112

Young, T., Hazarika, D., Poria, S. et Cambria, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational intelligence magazine*, 13(3):55–75. 96

Zainulina, E., Chernyavskiy, A. et Dylov, D. V. (2021). No-reference denoising of low-dose ct projections. *In 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 77–81. IEEE. 112

Zhang, K., Zuo, W., Chen, Y., Meng, D. et Zhang, L. (2017). Beyond a gaussian denoiser : Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155. 109

Zhang, Y. et Yu, H. (2018). Convolutional neural network based metal artifact reduction in X-ray computed tomography. *IEEE transactions on medical imaging*, 37(6):1370–1381. 107

Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R. et Rosen, M. S. (2018). Image reconstruction by domain-transform manifold learning. *Nature*, 555(7697):487–492. 110

Zhu, J.-Y., Park, T., Isola, P. et Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. *In Proceedings of the IEEE international conference on computer vision*, pages 2223–2232. 112

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

6 Application d'un réseau de neurones en post-traitement



I Post-traitement avec un U-Net 120
1 Description des réseaux utilisés 120
2 Entraînement des réseaux 123
II Évaluation des U-Net 125
1 Évolution des métriques pendant l'entraînement 125
2 Évaluation sur les métriques 126
3 Évaluation visuelle 127
4 Amélioration du U-Net-3D 129
III Comparaison avec une régularisation TV 130
1 Détail des algorithmes 131
2 Comparaison 131

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

L'objectif de ce chapitre est d'étudier sur un exemple de réseaux de neurones la capacité de ceux-ci à débruiter et à réduire les artefacts de reconstructions analytiques faibles doses en imagerie dentaire. Nous nous restreignons, dans cette étude, au cas de l'apprentissage supervisé et aux CNN de type U-Net. Nous testons différentes façons d'appliquer un U-Net en post-traitement, en considérant premièrement trois U-Net 2D entraînés selon les trois directions, puis un U-Net 3D.

Nous commençons par décrire les réseaux utilisés et leur entraînement. Une évaluation des performances des différents réseaux est ensuite faite avant de comparer leurs résultats avec des volumes reconstruits analytiquement et avec une méthode itérative.

I Post-traitement avec un U-Net

1 Description des réseaux utilisés

U-Net 2D

Initialement utilisé pour faire de la segmentation (Ronneberger *et al.*; 2015), le U-Net peut également être utilisé pour améliorer la qualité d'une reconstruction dégradée par la présence de bruit et d'artefacts (Jin *et al.*; 2017). Sa structure permet de récupérer les informations contextuelles de cette image dégradée afin de retrouver la qualité de l'image cible correspondante. Jin *et al.* (2017) ont utilisé un U-Net, appelé FBPConvNet, pour débruiter et réduire les artefacts d'images reconstruites par rétro-projection filtrée obtenues à partir d'un faible nombre de projections 2D parallèles. Nous utilisons une version légèrement modifiée du réseau qu'ils proposent sur nos reconstructions FDK.

Le réseau utilisé par Jin *et al.* est un U-Net résiduel de profondeur 5. Un zéro-padding est utilisé afin que l'image garde la même taille après convolution, seules les couches de pooling diminuent sa taille. Des couches de normalisation de batch (*Batch Normalization*) [1] sont utilisées après chaque convolution pour améliorer la stabilité et la rapidité de l'apprentissage (Ioffe et Szegedy; 2015).

Nos données étant différentes de celles utilisés par Jin *et al.*, nous avons effectué quelques modifications dans la structure du réseau par rapport à FBPConvNet, qui ont amélioré ses performances sur nos données :

- Les fonctions d'activation ReLU (*Rectified Linear Unit*, correspondant à LeakyReLU avec $\lambda = 0$) ont été remplacées par des fonctions Leaky ReLU qui ne s'annulent pas lorsque la sortie d'un neurone est négative. Ces fonctions d'activation permettent d'améliorer la stabilité de l'entraînement. Nous avons alors utilisé l'initialisation de He (He *et al.*; 2015) adaptée à ces activations (Géron; 2022).
- Les couches de déconvolution qui permettent de revenir à la taille de l'image originale, connues pour créer des artefacts en damier dans les zones de haute intensité (Odena *et al.*; 2016), ont été remplacées par des couches de sur-échantillonnage (*upsampling*) avec interpolation au plus proche voisin. Un exemple d'un tel artefact est visible dans la figure 6.1, à droite.
- Nous n'avons pas utilisé de seuil pour les valeurs du gradient lors de la rétro-propagation : cela créait des zones de flou (voir figure 6.1, à gauche) autour des objets en métal et diminuait trop la vitesse de convergence.

1. Les couches de normalisation de batch normalisent les données en entrée afin qu'elles aient une moyenne nulle et une variance égale à 1. Cela peut accélérer l'entraînement et améliorer la précision du réseau.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

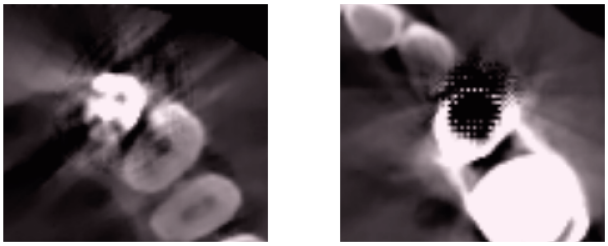


FIGURE 6.1 – Exemple d'artefacts présents dans les images de sortie des réseaux. A gauche : on observe un halo flou autour de la couronne en métal lorsque la valeur du gradient est seuillée à 10^{-2}. A droite : l'artefact en damier, caractéristique des couches de déconvolution dans les zones de haute intensité.

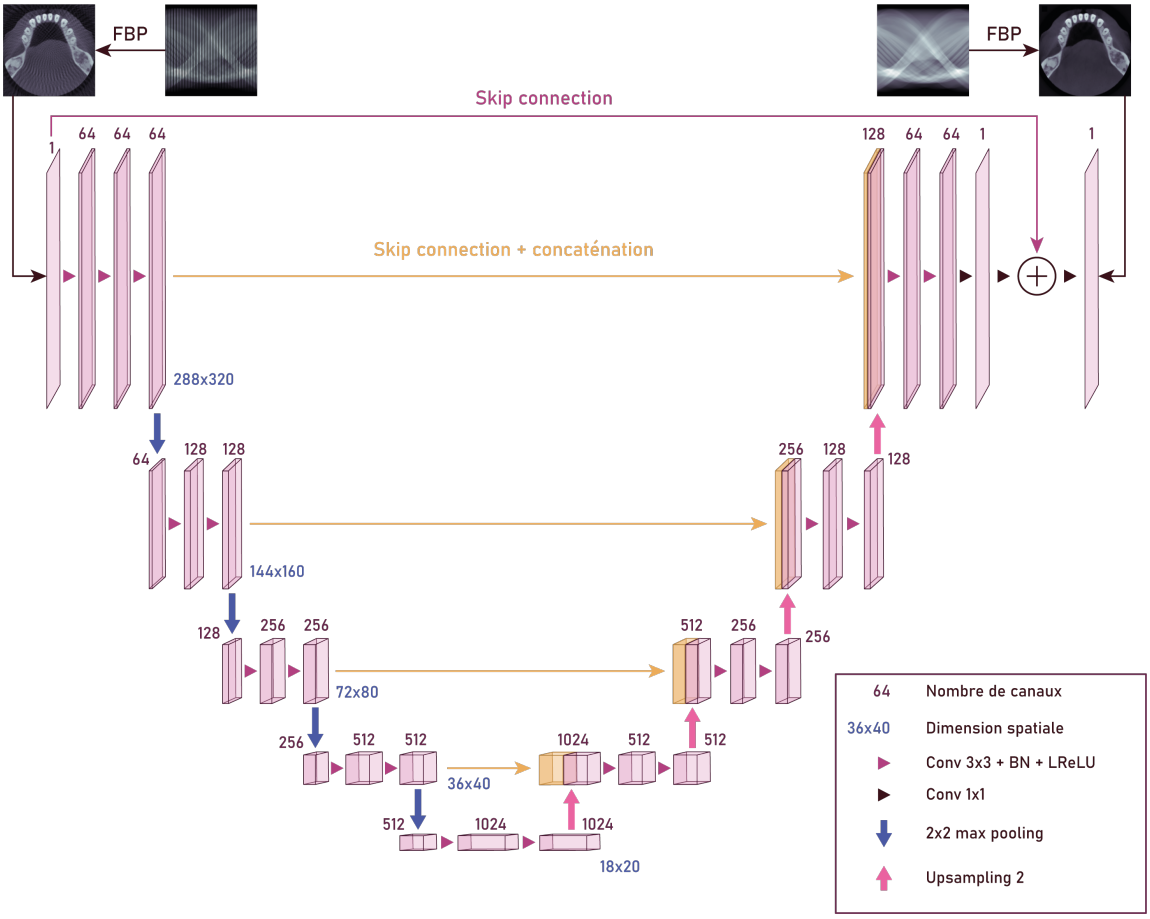


FIGURE 6.2 – Architecture du U-Net utilisé pour le post-traitement. Inspiré de (Jin *et al.*; 2017) avec modification de la fonction d'activation et des couches de upsampling.

L'architecture du réseau utilisé est montrée dans la figure 6.2. Le nombre total de paramètres entraînables est 31 389 569. La taille des images en entrée est d'environ 288×320 pixels.

U-Net MP

Un premier U-Net 2D a été entraîné sur les coupes axiales. Cependant, le traitement coupe par coupe du volume 3D donnait un traitement non corrélé entre les coupes et ne prenait pas en compte les informations que pouvaient apporter les autres directions. Nous avons

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

alors testé une configuration dans laquelle 3 U-Net sont entraînés, indépendamment les uns des autres, selon les différentes directions : axiale, sagittale et coronale, respectivement notés Γ_A, Γ_S et Γ_C. Nous considérons ensuite une moyenne pondérée de ces trois sorties comme reconstruction finale, à laquelle nous nous référons par U-Net-MP (Multi-Plans), parfois noté U-Net2.5D dans la littérature. Nous avons fait le choix de considérer une moyenne pondérée pour prendre en compte les performances différentes des réseaux : le réseau Γ_A donne les meilleurs résultats et Γ_S les moins bons, ainsi nous considérons une première approche où $f_{\text{U-Net-MP}} = 0.5 \times f_{\Gamma_A} + 0.3 \times f_{\Gamma_C} + 0.2 \times f_{\Gamma_S}$, avec f_Γ la sortie du réseau Γ. De plus amples détails sur les différences de performances entre les trois réseaux sont donnés dans le prochain paragraphe. Ce réseau permet d'intégrer la cohérence 3D sans pour autant trop augmenter le temps d'entraînement et la capacité de calcul nécessaire. Un schéma du principe de ce U-Net Multi-Plans est visible dans la figure 6.3.

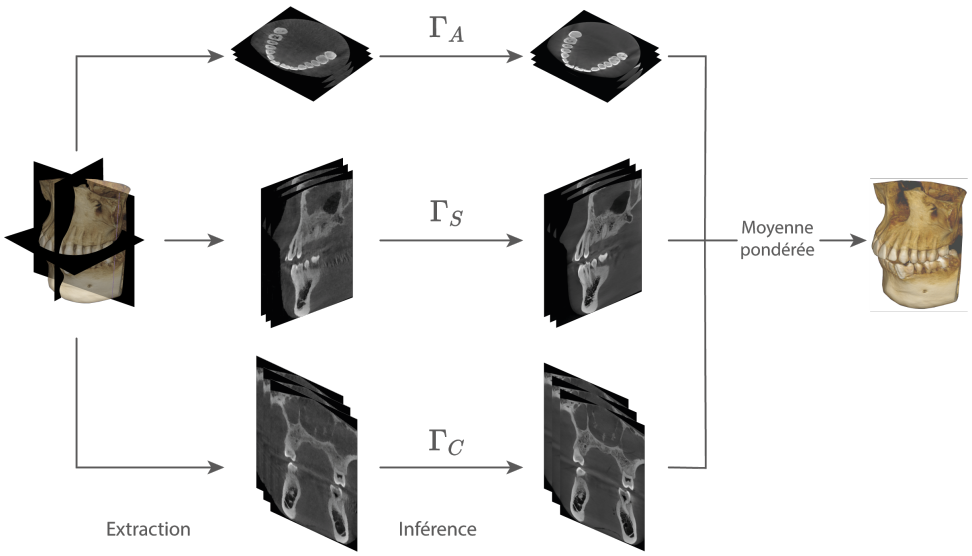


FIGURE 6.3 – Schéma du fonctionnement du U-Net-MP. Trois réseaux Γ_A, Γ_S et Γ_C sont respectivement entraînés sur l'ensemble des coupes axiales, sagittales et coronales. Pour l'inférence, le volume passe, coupe par coupe, par chacun des trois réseaux, puis les trois volumes obtenus sont moyennés pour obtenir le volume final.

Différents paramètres peuvent être à l'origine des différences entre les performances des réseaux 2D. Une première explication peut provenir de la direction du faisceau de rayons pendant les acquisitions et de la reconstruction en elle-même. Sans considérer l'angle de cône, les rayons X qui traversent le patient sont mieux adaptés à la reconstruction des coupes axiales (comme l'indique la reconstruction exacte que l'on obtient dans le plan de la source). La reconstruction FDK est donc effectuée coupe par coupe, avec un filtre rampe 1D appliqué en axial. Ainsi, les coupes axiales des reconstructions, normales et faibles doses, sont de meilleure qualité que les coupes coronales et sagittales. Deuxièmement, nous expliquons la différence de performance entre sagittal et coronal par la différence de répartition du bruit qui impacte ces deux directions : un SNR plus faible est présent en coronal car de par la forme du crâne, les rayons dans la direction sagittale traversent plus de matière que ceux dans la direction coronale. Les reconstructions FDK [1] sont ainsi corrompues par un bruit directionnel, et de légères rayures verticales sont visibles sur les coupes coronales. Celles-ci ne sont pas visibles sur les coupes

1. Encore plus particulièrement en faible dose.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés



(a) (b) (c)

FIGURE 6.4 – (a) Illustration des rayures verticales présentes dans une coupe coronale d'une reconstruction FDK faible dose. (b) et (c) Exemple d'une coupe axiale (et son agrandissement) d'un volume auquel on a appliqué un U-Net entraîné sur les coupes sagittales. Image légèrement saturée pour permettre une meilleure visualisation en impression. On peut remarquer les rayures verticales correspondant à la direction d'entraînement.

sagittales, mais elles créent des différences entre ces coupes. Ces rayures sont correctement corrigées par le réseau Γ_C, mais les différences d'intensité entre les coupes sagittales, présentes dans les reconstructions FDK, ne sont pas corrigées par Γ_S puisque l'entraînement s'effectue par coupes 2D. Un exemple de telles rayures est visible dans la figure 6.4a.

Les figures 6.4b et 6.4c montrent une coupe axiale d'un volume auquel on a appliqué un U-Net entraîné sur les coupes sagittales. Les différences d'intensité entre des coupes sagittales successives sont bien visibles. Ceci illustre l'intérêt de considérer des réseaux 2.5D qui vont lisser ces imperfections.

U-Net-3D

Une architecture de U-Net-3D a été implémentée en reprenant l'architecture du U-Net 2D utilisé précédemment. Nous avons remplacé les couches de convolution et upsamplig 2D par leur équivalent en 3D. Afin de conserver des ressources de calcul nécessaires raisonnables et équivalentes au 2D, nous avons divisé le nombre de filtres de convolution par 2 partout. Le nombre de paramètres entraînables est ainsi de 21 185 153. Ce nombre est inférieur à celui du U-Net 2D (qui était d'environ 31 millions), mais sans réduire le nombre de filtres de convolution, le réseau atteignait plus de 84 millions de paramètres. Dans cette configuration, l'écart entre le nombre de paramètres des U-Net 2D et 3D aurait rendu la comparaison entre les deux moins juste, même si comme nous l'expliquons dans les discussions, rendre la comparaison entre réseau 2D et 3D juste reste une tâche difficile.

Toujours avec l'objectif d'effectuer les entraînements dans un moindre coût numérique, les réseaux ont été entraînés sur des patchs de taille $160 \times 160 \times 160$ [1] et non pas sur l'intégralité des volumes. Deux cents patchs par volume ont été extraits de manière stochastique pour l'entraînement.

2 Entraînement des réseaux

Les U-Net 2D ont été entraînés de manière supervisée et les vérités terrains ont été obtenues en reconstruisant les volumes analytiquement à partir d'acquisitions à dose normale. En entrée

1. Taille de patch maximale qui rentrait en mémoire sur la machine de calcul utilisée.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

de réseau, sont données les reconstructions analytiques obtenues en considérant une projection sur cinq, ce qui équivaut à 20% de la dose.

Base de données

Nous disposons d'une base de données anonymisée de 32 patients des deux sexes, âgés de 18 à 70 ans. Les projections coniques ont été acquises en dose normale, sous des conditions de shortscan par un scanner CS 8100SC 3D. Chaque volume contient un nombre moyen de $320 \times 280 \times 280$ voxels, de 300 µm de côté, selon la taille du champ de vue choisie.

Nous avons séparé la base de données en trois parties. Vingt-cinq patients ont été utilisés pour l'entraînement et six autres pour la validation et l'ajustement des hyper-paramètres. Le dernier volume a été utilisé pour les tests, il n'a jamais été vu par le réseau pendant la phase d'entraînement. La figure 6.5 montre des coupes de quatre patients différents tirés de la base de données. Le dernier patient a été choisi pour illustrer les résultats et faire les comparaisons avec les reconstructions analytiques et itératives dans la section III.

Les volumes ont été normalisés pour que leur valeurs soient comprises entre -1 et 1. La même normalisation a été appliquée à l'ensemble des volumes, c'est-à-dire que la valeur maximale des volumes de la base de données, a été utilisée pour diviser les valeurs des volumes, indépendamment de la valeur maximale de chacun.

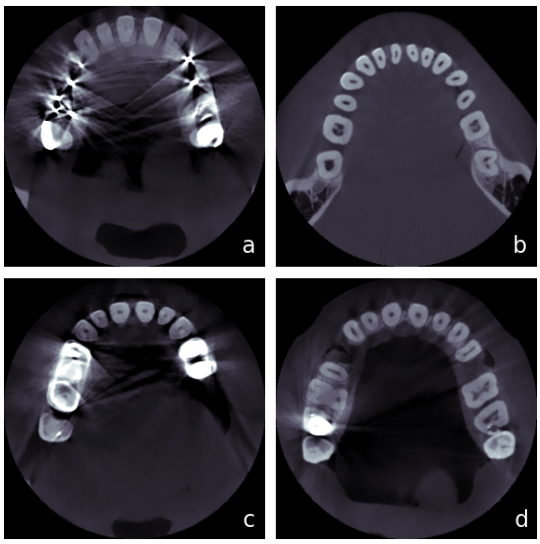


FIGURE 6.5 – (a-c) Coupes de patients de la base de données d'entraînement. (d) Coupe du patient utilisé pour les tests de la section III.

Hyperparamètres choisis

Les réseaux 2D ont été entraînés sur 50 époques, avec une taille de batch égale [1] à 10. Le U-Net-3D a été entraîné sur 100 époques, la taille de batch a été fixée [2] à 1. L'optimiseur choisi est Adam et après une étude de la valeur optimale du taux d'apprentissage sur le jeu de données de validation [2], nous l'avons fixé à 10^{-3} pour les coupes axiales et coronales, et à 10^{-2} pour les coupes sagittales. Nous n'avons pas eu le temps d'optimiser le taux d'apprentissage pour le U-Net-3D, celui-ci a été fixé à 10^{-3}. La moyenne de l'erreur absolue (MAE) a été utilisée comme

1. Nombre maximal pour la mémoire disponible.
2. A partir des valeurs de la MAE (Erreur absolue moyenne) et de la MSE (Erreur quadratique moyenne).

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

fonction coût, des tests avec l'erreur quadratique moyenne ayant abouti a des images un peu trop lisses. Nous avons gardé le paramètre par défaut $\lambda = 0.3$ de la plupart des librairies, pour les fonctions d'activation Leaky ReLU.

Le taux d'apprentissage ne reste pas fixe tout au long de l'entraînement : une valeur initiale lui est donnée, puis il suit une décroissance exponentielle selon les époques. Cela permet d'affiner légèrement l'apprentissage en réduisant le pas à l'approche de l'optimum : à la dernière époque, le taux vaut $0.9\times$ le taux initial.

Implémentation numérique

Les entraînements ont été réalisés avec un processeur Intel Xeon Gold 6226R et un GPU Nvidia Tesla V100 [1].

La reconstruction FDK ainsi que les opérations de projection et de rétro-projection ont été effectuées à l'aide de la bibliothèque ASTRA (Van Aarle *et al.*; 2016). Les modèles de U-Net utilisés ont été implémentés avec Tensorflow et la librairie Keras (Chollet *et al.*; 2015).

II Évaluation des U-Net

Pour évaluer la robustesse de nos réseaux, une validation croisée *Leave One Out* a été appliquée sur l'ensemble de la base donnée. Cela signifie que 32 réseaux ont été entraînés (par type de réseau), avec à chaque fois un volume différent pour le test.

Par soucis d'accélérer largement le temps total de calcul, nous avons choisi de conserver les taux d'apprentissage déterminés précédemment. Nous avons en effet constaté une très faible différence dans les résultats lors de l'optimisation initiale de ce seul hyperparamètre, pour la plage de valeur testée [2]. Il est ainsi probable que l'optimiser pour chacun des jeux de données d'entraînement/validation n'aurait eu que très peu d'influence sur les métriques d'évaluation sur le jeu de données tests. Cela revient ainsi à ne pas optimiser les hyperparamètres pour chacun des entraînements, ce qui semble acceptable pour l'objectif de cette étude.

Le fait d'avoir des données de validation pour chaque configuration permet de choisir les poids de l'époque présentant la meilleure valeur de MAE afin d'évaluer le réseau sur le volume test.

1 Évolution des métriques pendant l'entraînement

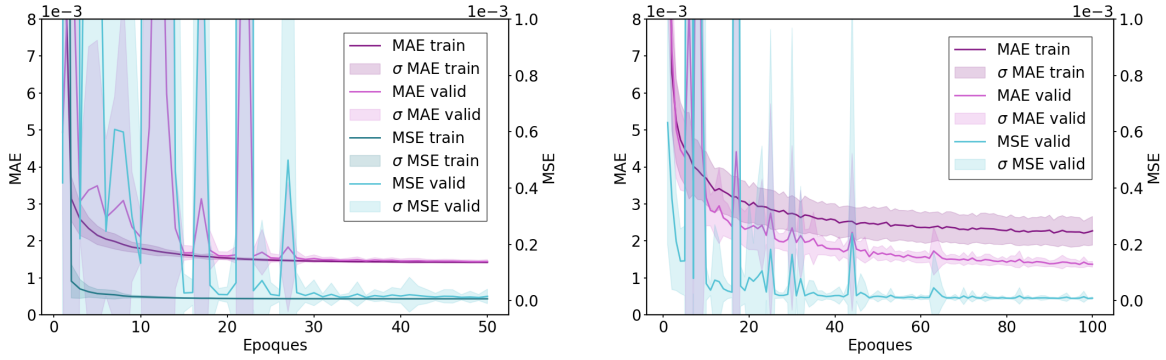
La moyenne et l'écart-type de l'évolution des valeurs de MSE et de MAE sur les données d'entraînement et de validation du U-Net Γ_A par époques, sont représentées dans la figure 6.6a. Les courbes pour les autres réseaux (Γ_S et Γ_C) sont très similaires. Les valeurs sur les données d'entraînement sont assez fluctuantes sur les 25 premières époques, mais elles se stabilisent ensuite et convergent à l'époque 50.

Les courbes de la moyenne et de l'écart-type de la MAE, du U-Net-3D, sur les données d'entraînement et de validation ainsi que celles pour la MSE évaluée sur les données de validation sont données dans la figure 6.6b. En observant la MAE en validation, nous voyons qu'elle continue de décroître après l'époque 50, justifiant ainsi la nécessité d'augmenter le nombre d'époques. Elle semble cependant être bien stabilisée à l'époque 100, voire 80. Nous pouvons également remarquer que les valeurs de la MAE en validation sont plus faibles que celles de l'entraînement. Ceci est dû au fait que nous avons appliqué un zéro padding sur

1. Ces travaux ont bénéficié d'un accès aux moyens de calcul de l'IDRIS au travers de l'allocation de ressources 2021-AD011012734 attribuée par GENCI.
2. Entre 10^{-4} et 10^{-2}.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

les images pour calculer les métriques pour la validation, par simplicité, pour que la taille des images soit divisible par [1] $2^4 = 16$. La prise en compte de ces pixels ajoute une constante additive quantifiable au calcul de la MAE, sans réelle importance si l'on ne s'intéresse qu'à la convergence.



(a) Réseau Γ_A

(b) Réseau U-Net-3D

FIGURE 6.6 – Moyennes et écart-type (σ) des courbes de MSE et MAE évaluées sur les données d'entraînement et de validation, en fonction des époques, des 32 entraînements du réseau Γ_A et du réseau U-Net-3D.

2 Évaluation sur les métriques

Une fois les réseaux entraînés, nous avons gardé le jeu de paramètres donnant une valeur de la MAE la plus faible sur les données de validation, parmi les 50 époques (100 en 3D), et ceci de façon indépendante entre les entraînements. Ces réseaux ont été appliqués aux volumes servant de tests, permettant ainsi d'obtenir les inférences sur les 32 volumes différents, avec les réseaux Γ_A, Γ_C, Γ_S, U-Net-MP et U-Net-3D. Les métriques entre les volumes prédits par les réseaux et les volumes de référence ont été calculées et les moyennes et écart-type des valeurs obtenues sont visibles dans le tableau 6.1. Le U-Net-MP présente les meilleures métriques suivi d'assez près par le U-Net axial. Les U-Net entraînés sur les coupes sagittales ont de moins bons résultats. Le U-Net-3D obtient de bons résultats, un peu moins bons que le Γ_A et le U-Net-MP, mais avec des valeurs très proches.

	Γ_A	Γ_C	Γ_S	U-Net MP	U-Net 3D
NRMSE	0.0441 ± 0.0064	0.0480 ± 0.0039	0.0668 ± 0.0483	**0.0432 ± 0.0069**	0.0473 ± 0.0070
PSNR	56.642 ± 1.3301	55.848 ± 1.0149	53.768 ± 3.0785	**56.827 ± 1.4045**	56.051 ± 1.1786
SSIM	0.9982 ± 0.0003	0.9977 ± 0.0005	0.9968 ± 0.0007	**0.9983 ± 0.0003**	0.9979 ± 0.0003

TABLEAU 6.1 – Moyennes et écart-types des valeurs des métriques sur les volumes d'évaluation pour les 32 entraînements de la validation croisée.

1. 4 représentant la profondeur du réseau.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

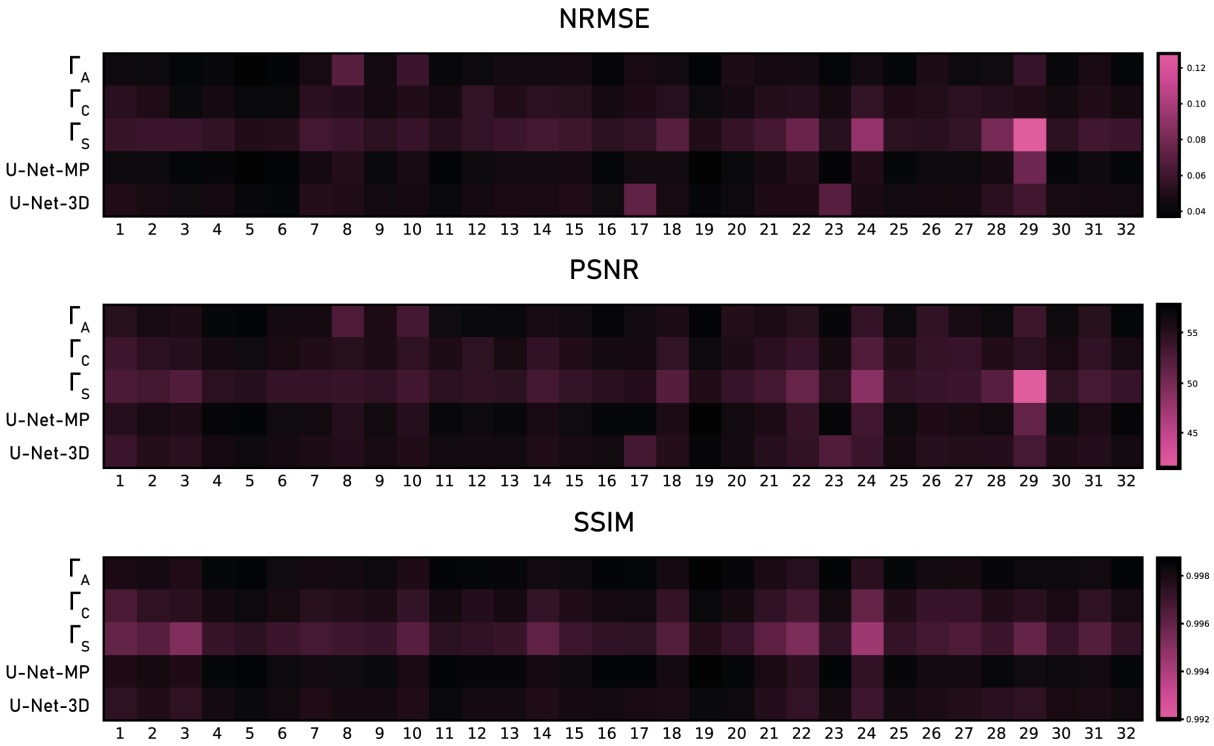


FIGURE 6.7 – Visualisation des valeurs des métriques sur chacun des volumes. Plus la couleur d'un volume est rose, moins la métrique associée est bonne.

3 Évaluation visuelle

Les valeurs des métriques pour chacun des volumes sont représentées dans la figure 6.7. Les trois images correspondent aux trois métriques et chaque ligne correspond à un réseau. Les 32 volumes sont alignés horizontalement. Cette représentation confirme les moins bonnes valeurs des métriques pour le U-Net Γ_S, avec en particulier le 29e volume qui a de mauvais résultats en termes de NRMSE et PSNR, et globalement de mauvaises performances par rapport au SSIM.

La figure 6.8 présente des coupes des volumes ayant eu les meilleures et les moins bonnes valeurs de métriques. Sur toutes les métriques, les volumes ayant eu les meilleurs résultats sont ceux des patients 19 et 5. La mâchoire du patient 19 ne contient aucun objet en métal, et celle du patient 5 seulement un petit morceau de métal visible dans la coupe axiale. Nous pouvons remarquer que pour les volumes avec les moins bonnes métriques, tous contiennent de nombreux objets métalliques, dans la mâchoire maxillaire et la mâchoire mandibulaire. Des fils en métal peuvent être observés dans le bas de la mâchoire mandibulaire du patient 24, visibles dans les coupes coronales et sagittales. Ainsi, en dehors des couronnes ou implants en métal, nous supposons que ces fils en métal peuvent être sources d'erreur puisqu'aucun autre patient n'a ce type de caractéristiques, donc la base de données manque de représentation de ce métal et des artefacts engendrés. De même, les patients 24 et 29 ont des implants, visibles en particulier dans la coupe sagittale du patient 29. Ce type de vis n'est pas beaucoup présent dans notre base de données, et peut également souffrir de manque de représentation.

La présence de métal dans certains patients rend également compliquée la normalisation des volumes en entrée des réseaux. En dehors des couches de normalisation de batch utilisées après les couches de convolution, il est nécessaire de normaliser les données avant l'entraînement (LeCun *et al.*; 1998). Les volumes ont initialement été normalisés entre -1 et 1 avant de passer à travers les réseaux. Cependant, les hautes intensités du métal font que finalement, la plupart

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

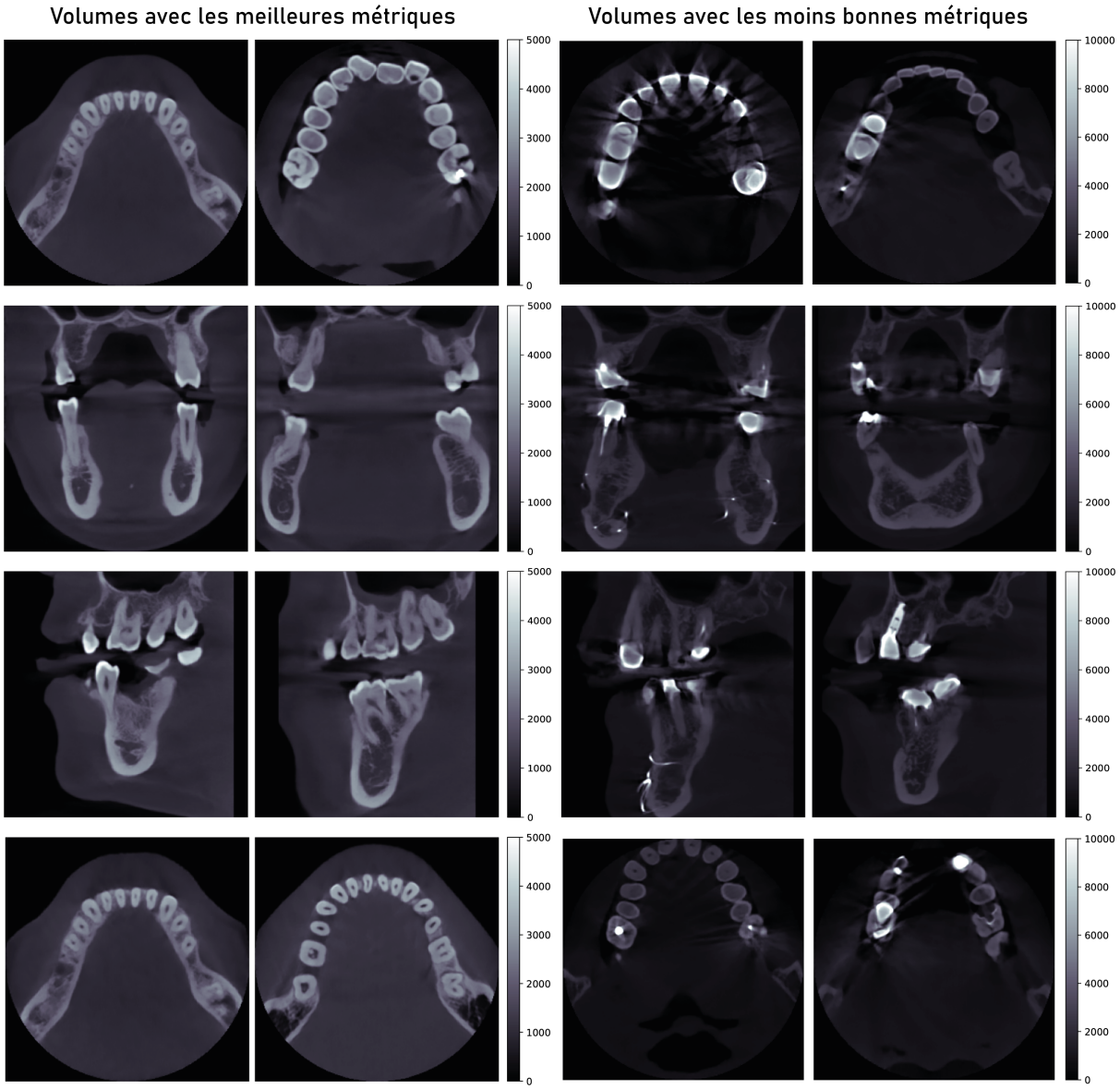


FIGURE 6.8 – Exemples de coupes de volumes en sortie des réseaux. Les deux premières colonnes correspondent aux meilleures valeurs des métriques, et les deux dernières aux moins bonnes. Chaque ligne correspond à un réseau. Première ligne de gauche à droite : coupes axiales des patients 19, 5, 24 et 8 en sortie du réseau Γ_A. Seconde ligne : patients 19, 5, 24 et 26 en sortie de Γ_C. Troisième ligne : coupes sagittales en sortie de Γ_S des patients 19, 5, 24 et 29. Dernière ligne : coupes axiales des patients 19, 6, 17 et 23 en sortie du réseau U-Net-3D.

des valeurs des volumes sont comprises entre -0.1 et 0.1. Ces grandes différences de valeur entre les voxels représentant du métal, et les autres voxels qui sont majoritaires, peuvent poser des problèmes lors de l'entraînement et diminuer les performances des réseaux. Nous avons, par ailleurs, essayé d'entraîner les réseaux après avoir seuillé les données pour limiter les hautes valeurs du métal, mais les résultats n'étaient pas satisfaisants.

Seulement trois volumes ont des valeurs supérieures à 30 000 HU : les volumes 1, 26 et 29, avec le volume 29 atteignant les plus hautes valeurs. Ceci pourrait expliquer pourquoi les métriques (NRMSE et PSNR) sont moins bonnes pour ce volume que pour les autres.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

4 Amélioration du U-Net-3D

Les performances légèrement moindres du U-Net-3D par rapport à Γ_A et U-Net-MP peuvent avoir différentes sources :

- L'apprentissage par patch ne permet pas de profiter du maximum d'information contextuelle disponible lorsque l'on considère le volume dans son intégralité.
- La diminution du nombre de filtres par rapport aux réseaux 2D diminue le nombre de paramètres entraînables comme nous l'avons vu dans la section 1. On a donc un réseau avec moins de paramètres, pour une tâche plus complexe - traiter des volumes 3D -, ce qui rend la comparaison avec les U-Net 2D discutable.

Nous n'avons pas pu tester l'impact de l'apprentissage par patch sur les performances du U-Net-3D par manque de mémoire lors de l'entraînement. Nous avons, en revanche, essayé d'entraîner un réseau 3D avec la même architecture que celui utilisé en 2D, c'est-à-dire avec 64 filtres dès la première couche de convolution, au lieu de 32. Par manque de temps en fin de thèse, ce test n'a été effectué que sur un seul entraînement, celui qui utilise les 31 premiers volumes pour l'entraînement/validation et le volume 32 en test.

Les réseaux 3D précédents reprenaient la structure des réseaux U-Net-2D avec un nombre de filtres inférieur afin de limiter l'écart entre le nombre de paramètres des réseaux 2D et 3D. Nous obtenions alors environ 21 millions de paramètres, contre 84 millions en gardant la même structure de réseau [1]. Ce réseau est noté U-Net-3D+ dans la suite.

Un réseau avec la même structure (ie. même nombre de filtres) que les U-Net 2D a été entraîné. L'augmentation du nombre de filtres a nécessité de réduire la taille des patch utilisés pendant l'entraînement. Nous avons donc utilisé des patchs de taille $80 \times 80 \times 80$ pour cette expérience.

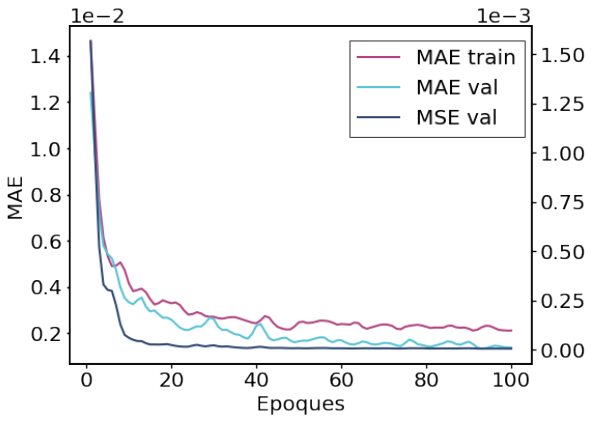


FIGURE 6.9 – Courbes de la fonction coût (MAE train) sur les 100 époques, et des MAE et MSE évaluées sur les données de validation (val) pour le réseau U-Net-3D+.

La figure 6.9 montre l'évolution de la fonction coût et des métriques pour cet entraînement. La MSE en validation semble avoir convergé relativement tôt (60e époque). Une légère décroissance peut cependant être encore observée sur la courbe de MAE en validation, suggérant que l'augmentation du nombre d'époques pourrait être envisagée.

Le tableau 6.2 résume les valeurs de la NRMSE, du PSNR et de la SSIM évaluées sur le volume 32, pour les différentes sorties des réseaux. Les valeurs pour le U-Net-MP sont également données à titre de comparaison. Parmi les U-Net-3D, c'est le U-Net-3D+ qui présente les

1. C'est-à-dire une baisse d'environ 32% du nombre de paramètres par rapport aux U-Net 2D, et une augmentation de 270% dans le second.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

	U-Net-MP	U-Net-3D	U-Net-3D+
NRMSE	**0.0374**	0.0387	0.0385
PSNR	53.69	53.36	**53.94**
SSIM	**0.9968**	0.9965	0.9967

Tableau 6.2 – Valeurs des métriques évaluées sur le patient 32, pour les volumes en sortie du U-Net-3D, du U-Net-3D+, ainsi que le U-Net-MP pour comparer.

meilleurs valeurs, surpassant même le U-Net-MP en termes de PSNR. Le U-Net-3D garde des valeurs très proches de celles obtenues avec les deux autres méthodes, mais légèrement moins bonnes.

La figure 6.10 illustre une coupe axiale contenant du métal et un agrandissement d'une coupe coronale des deux volumes obtenus à partir des réseaux 3D. Les images sont très ressemblantes et il est difficile de leur trouver des différences. Nous pouvons tout de même remarquer que les artefacts métalliques semblent plus atténués par le U-Net-3D dans les zones en dehors de la mâchoire (cf flèche rose), alors que ceux présents au niveau des dents sont moins visibles dans l'image en sortie du U-Net-3D+ (flèche bleue).

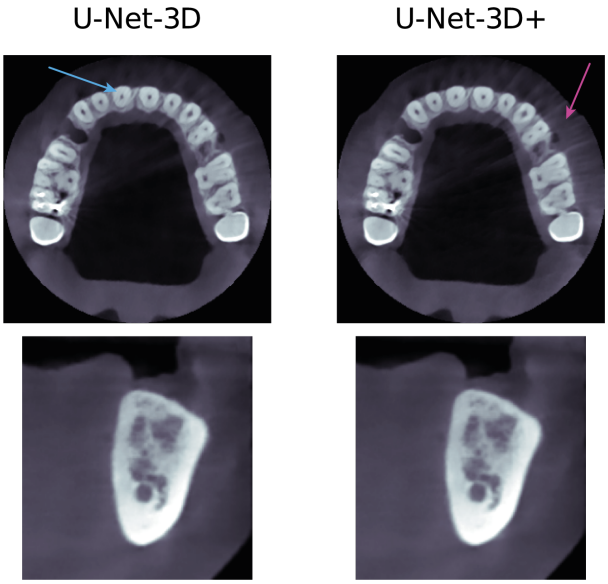


Figure 6.10 – Coupe axiale et agrandissement d'une coupe coronale des volumes en sorties des réseaux U-Net-3D et U-Net-3D+.

III Comparaison avec une régularisation TV

L'objectif de cette partie est de comparer différentes méthodes de reconstruction vues précédemment dans le cas de la tomographie à faible dose : la reconstruction analytique FDK, la reconstruction itérative KL-TV avec régularisation TV, et le post-traitement de la reconstruction FDK à l'aide d'un U-Net. Au vu des récents résultats très encourageants en deep learning, nous espérons que cette dernière méthode obtienne des performances comparables, si ce n'est meilleures, que l'algorithme itératif, en des temps beaucoup plus rapides.

Nous appliquons ces trois méthodes à des données faibles doses où une réduction de la

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

dose a été simulée en diminuant le nombre de projections (1 sur 5) par rapport aux acquisitions CBCT en dose normale dont nous disposions. Nous étudions leur capacité à réduire les artefacts étoilés et le bruit sur les reconstructions dans le cas du CBCT dentaire. Une partie des résultats montrés dans cette section a été présentée au colloque francophone GRETSI (Friot--Giroux *et al.*; 2022).

1 Détail des algorithmes

Les deux méthodes de reconstruction utilisées dans cette comparaison ont été présentées précédemment. FDK (Feldkamp *et al.*; 1984) effectue une reconstruction analytique très rapide mais également très sensible au bruit et au faible nombre de projections.

L'algorithme pré-conditionné KL-TV (Sidky *et al.*; 2012) est un algorithme primal-dual itératif qui minimise une fonctionnelle constituée de la divergence de Kullback-Leibler pour l'attache aux données, et de la variation totale pour la régularisation. Ainsi que nous l'avons vu dans les chapitres 3 et 4, cette régularisation TV réduit efficacement les artefacts en étoile et le bruit des reconstructions en faible dose, mais peut supprimer certains détails si le paramètre TV utilisé est trop élevé. Pour conserver un équilibre entre préservation des détails et atténuation des artefacts, nous avons choisi d'effectuer deux reconstructions avec les paramètres suivants : $\alpha = 0.05$ et $\alpha = 0.1$. Nous noterons les deux versions de l'algorithme correspondant par KL-TV$_\alpha$ avec la valeur de α correspondante. Les valeurs en dessous de 0.05 donnent des reconstructions encore trop bruitées, et au dessus de 0.1 le lissage induit une trop grande perte de détails.

Afin de limiter les artefacts dus aux projections tronquées, les projections ont été étendues avec un cylindre pour la reconstruction FDK, et la taille du volume reconstruit a été doublée pour KL-TV. Sur chaque côté, 150 voxels ont été ajoutés et après la reconstruction, le volume est tronqué pour ne conserver que la région d'intérêt (ROI).

2 Comparaison

De même façon que précédemment, les réseaux ont été entraînés sur les données de 25 patients, 6 patients ont été utilisés pour choisir les paramètres donnant les valeurs de MAE de validation les plus faibles sur l'ensemble des époques. Pour le test, nous avons utilisé les données du patient 32, et les résultats des différentes méthodes sur ce volume sont présentés dans les figures suivantes. Dans chaque figure on retrouve la reconstruction FDK à dose normale qui représente la vérité terrain, les différentes reconstructions faibles-doses : FDK, KL-TV$_{0.05}$, KL-TV$_{0.1}$, et les volumes obtenus après reconstruction FDK faible-dose et post-traitement avec U-Net axial (Γ_A), U-Net-MP, U-Net-3D et U-Net-3D+. Effectivement, même si nous n'avons pas effectué l'évaluation de ce dernier réseau sur l'intégralité des volumes, les bonnes performances qu'il présente sur le volume 32 sont intéressantes à comparer.

La figure 6.11 illustre une coupe axiale dans laquelle une couronne sur la seconde molaire maxillaire gauche en métal génère des artefacts métalliques. Ces artefacts, bien plus présents lorsque peu de projections sont utilisées, sont particulièrement visibles dans les reconstructions FDK et KL-TV$_{0.05}$, alors que les trois autres méthodes les réduisent considérablement. Les canaux de la première molaire gauche (indiqués par une flèche bleue) sont mieux résolus avec les U-Net qu'avec les autres méthodes.

La figure 6.12 montre un agrandissement de la coupe axiale de la figure 6.11, permettant de comparer un peu mieux des volumes en sortie des différents réseaux. Les réseaux Γ_A et U-Net-3D+ offrent les meilleures images, avec un débruitage conservant correctement les structures. Les contours des objets dans le volume en sortie du U-Net-3D sont un peu moins nets que ceux des autres. Dans le volume traité par le U-Net-MP, on remarque un quadrillage, assez visible

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

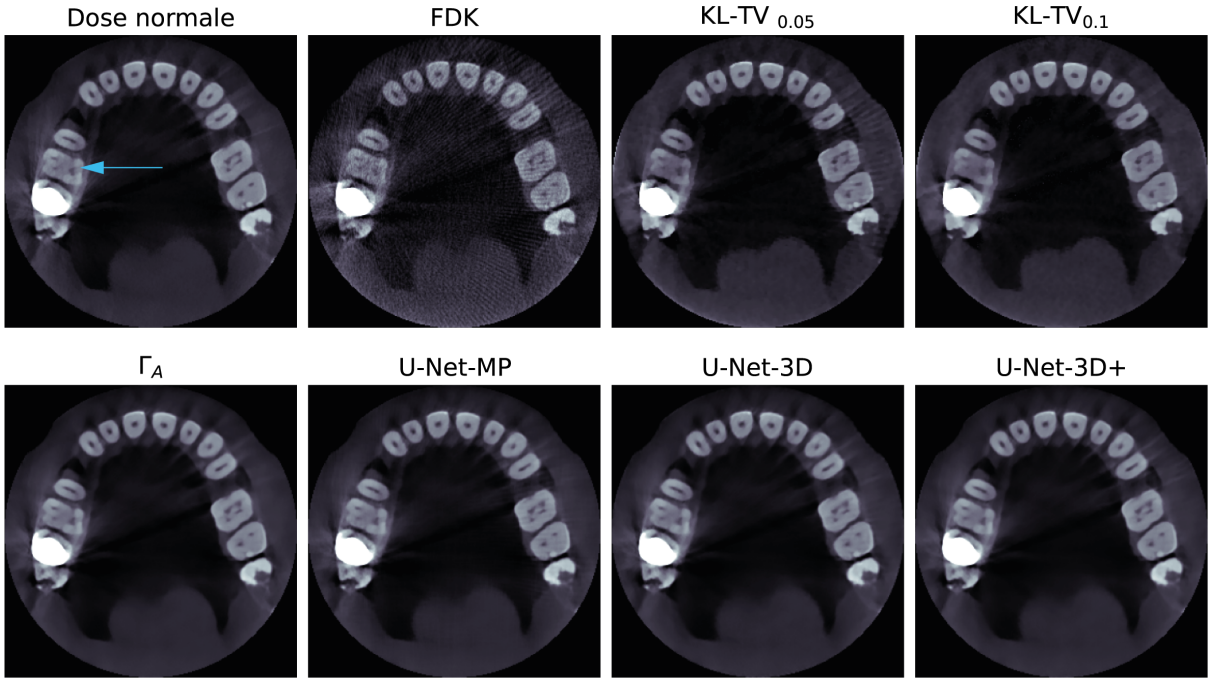


FIGURE 6.11 – Reconstructions d'une coupe axiale contenant une couronne en métal.

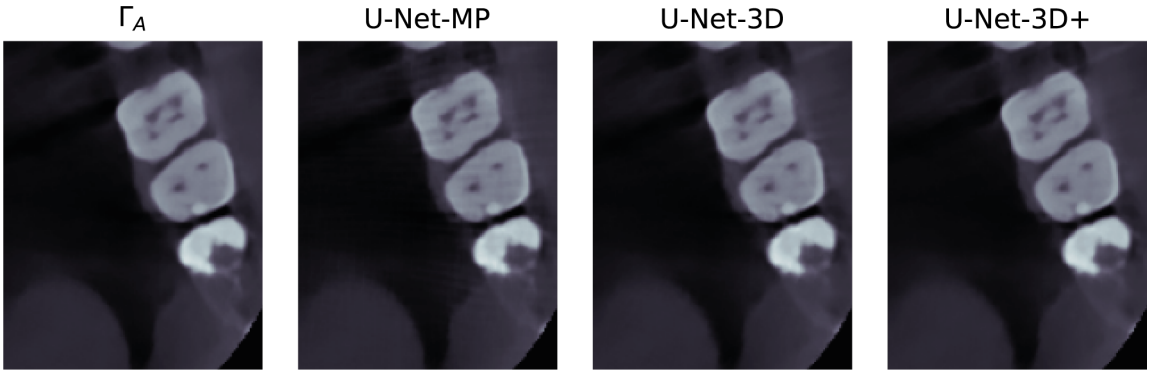


FIGURE 6.12 – Agrandissement de la coupe axiale présentée dans la figure 6.11, provenant des volumes en sortie des réseaux.

sur les zones plus claires des dents, qui provient des inhomogénéités du volume en sortie de Γ_S.

La figure 6.13 est un agrandissement sur une coupe axiale ne contenant cette fois-ci pas de métal. Le ligament alvéolaire dentaire (flèche orange) n'est pas visible sur les reconstructions itératives, et il apparaît un peu plus net sur la reconstruction obtenue avec le U-Net-MP que sur celles provenant des autres U-Net.

Dans les figures 6.14 et 6.15, on s'intéresse respectivement à la préservation de la taille du canal de la dent (flèche jaune) et de la restauration du canal mandibulaire (flèche rose). La base du canal n'est pas visible sur les reconstructions itératives et la structure de l'os alvéolaire est supprimée. Le U-Net-MP semble mieux conserver la base du canal de la dent que les autres réseaux. Les contours du canal mandibulaire sont également assez flous avec les reconstructions itératives. Le traitement coupe par coupe du U-Net axial laisse apparaître de légères rayures horizontales qui sont lissées par le U-Net-MP et les U-Net-3D.

Globalement, toutes les méthodes testées améliorent la qualité de l'image par rapport à la

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

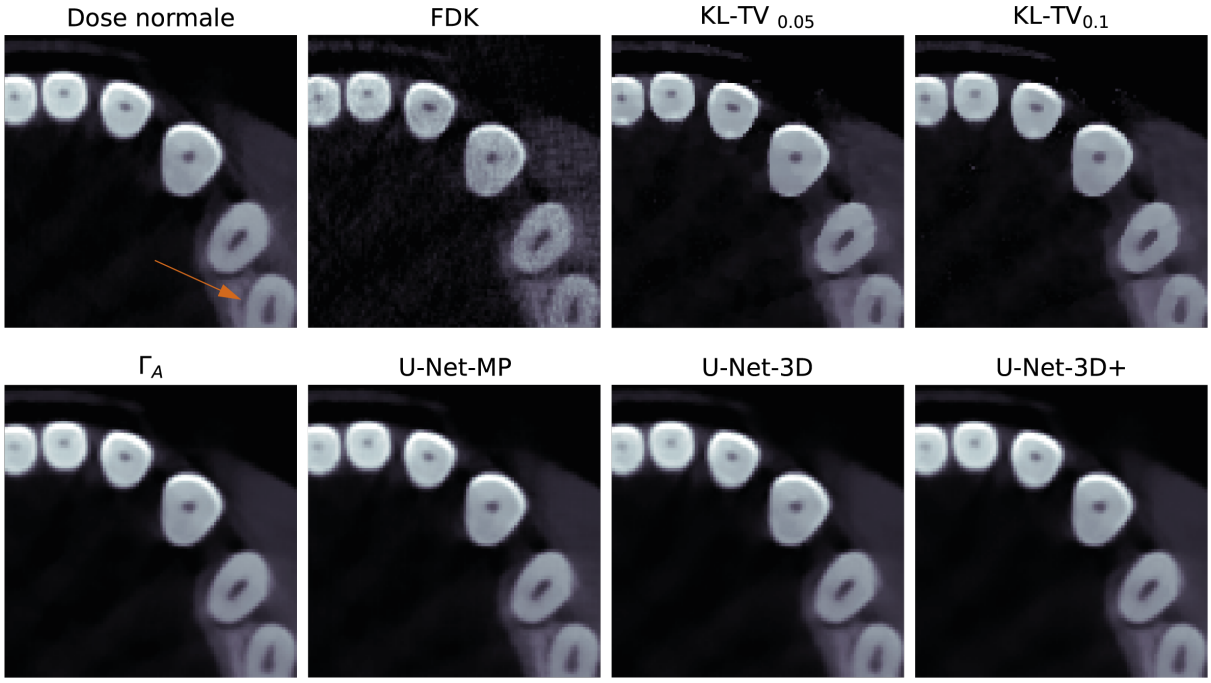


FIGURE 6.13 – Agrandissements de reconstructions d'une coupe axiale ne contenant pas de métal. La flèche orange pointe vers le ligament alvéolaire dentaire de la deuxième pré-molaire droite.

reconstruction FDK, les images sont moins bruitées et les artefacts métalliques atténués, surtout avec les réseaux de neurones. Les deux reconstructions avec l'algorithme KL-TV illustrent bien le compromis entre préservation des détails et réduction du niveau de bruit : les images reconstruites avec KL-TV$_{0.1}$ sont plus lisses mais quelques détails que l'on pourrait deviner avec KL-TV$_{0.05}$ (tels que la base du canal de la dent de la figure 6.14) sont complètement supprimés. Les méthodes de post-traitement avec U-Net permettent d'obtenir des images bien plus précises et proches de la réalité, même si une légère perte de texture peut être observée à certains endroits. Les U-Net-MP et 3D permettent de corriger les incohérences d'intensité entre deux coupes axiales successives que l'on peut observer sur les coupes sagittales et coronales des reconstructions avec le U-Net axial.

Afin de comparer quantitativement les reconstructions, nous avons utilisé le PSNR, le NRMSE et le SSIM auparavant présentés dans la section II.2 du chapitre 4. Malgré la difficulté de comparer voxel par voxel les volumes issus des reconstructions itératives avec les autres volumes, comme expliqué dans le chapitre 4, nous avons choisi d'utiliser ces métriques après avoir effectué une transformation linéaire des niveaux des gris de ces volumes pour que leurs valeurs correspondent à celle des autres volumes. Cependant, les valeurs des métriques sur le volume total étant à considérer avec précaution ; nous avons donc également calculé les métriques sur les trois ROIs présentés dans la suite et ayant pu être transformés de manière plus précise.

Le tableau 6.3 synthétise les valeurs des métriques utilisées pour comparer les méthodes sur différents ROIs. Nous les avons appliquées sur le volume total et sur les trois ROIs correspondant respectivement aux trois agrandissements présentés dans les figures 6.13, 6.14 et 6.15. Les métriques des deux reconstructions obtenues avec les algorithmes itératifs sont meilleures que celles de la reconstruction avec FDK et ont des résultats très similaires entre elles. La reconstruction avec $\alpha = 0.1$ a de meilleurs PSNR et NRMSE, mais le SSIM avec $\alpha = 0.05$ qui lisse

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

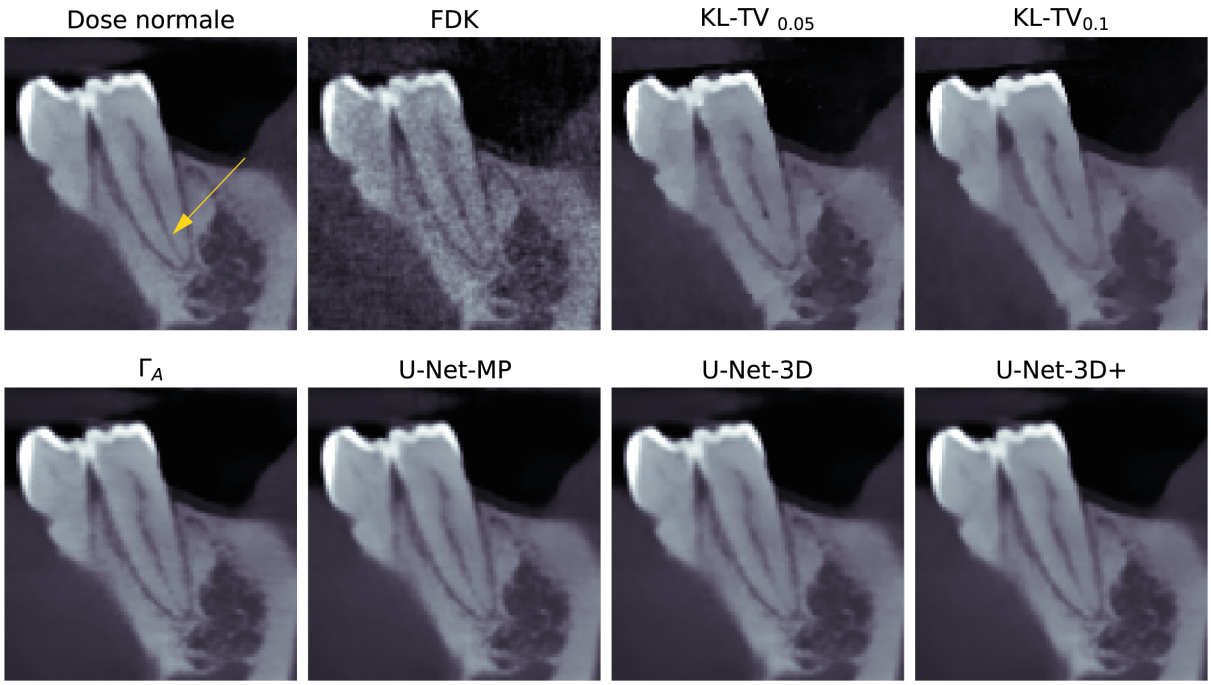


FIGURE 6.14 – Agrandissements de reconstructions d'une coupe sagittale, sur la deuxième prémolaire mandibulaire gauche.

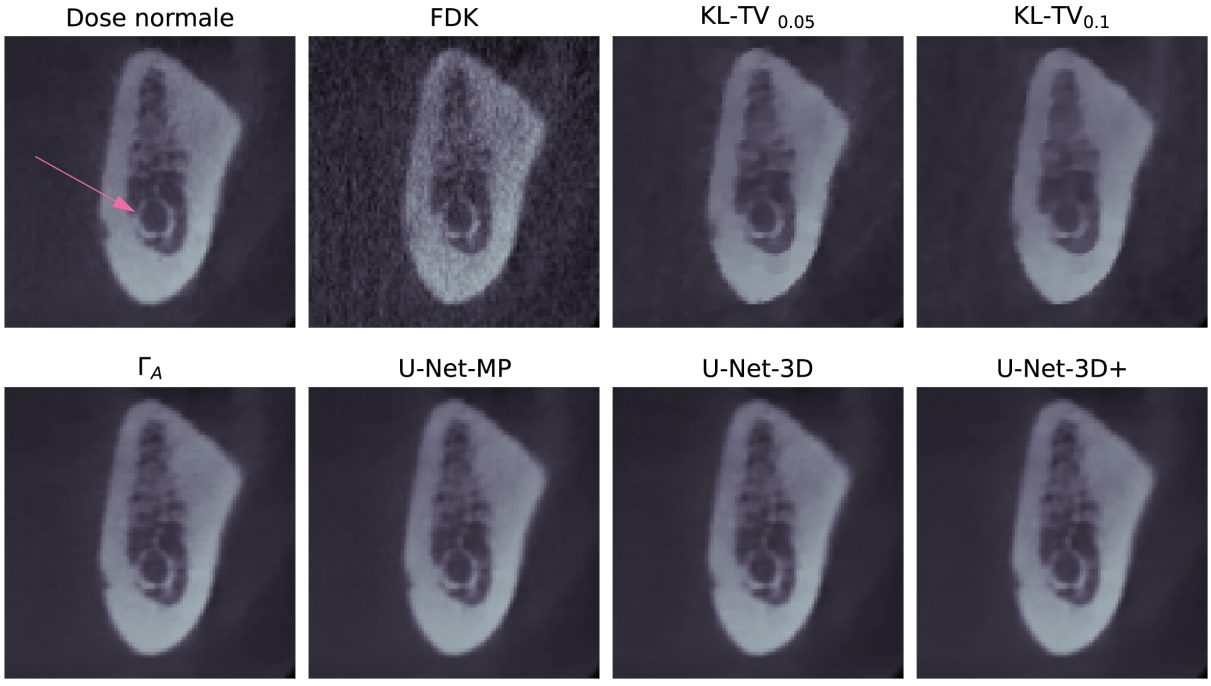


FIGURE 6.15 – Agrandissement de reconstructions d'une coupe coronale, sur une zone de la mâchoire contenant le canal mandibulaire.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

	NRMSE				PSNR				SSIM			
Méthodes	Total	ROI 1	ROI 2	ROI 3	Total	ROI 1	ROI 2	ROI 3	Total	ROI 1	ROI 2	ROI 3
FDK	0.169	0.136	0.116	0.154	40.58	41.77	41.11	40.37	0.932	0.923	0.931	0.922
KL-TV$_{0.05}$	0.136	0.129	0.095	0.140	42.46	42.25	42.83	41.24	0.961	0.956	0.951	0.915
KL-TV$_{0.1}$	0.134	0.128	0.095	0.137	42.62	42.34	42.85	41.41	0.962	0.956	0.950	0.913
Γ_A	0.037	0.035	0.034	**0.031**	53.80	53.67	51.80	54.02	**0.997**	0.995	0.994	**0.997**
U-Net-MP	0.037	**0.034**	**0.033**	**0.031**	53.69	**53.91**	51.95	54.04	**0.997**	**0.996**	**0.995**	**0.997**
U-Net-3D	0.039	0.038	0.035	0.033	53.36	52.86	51.44	53.63	0.996	0.993	0.994	0.996
U-Net-3D+	**0.036**	0.035	**0.033**	**0.031**	**53.94**	53.64	**52.09**	**54.18**	**0.997**	0.995	**0.995**	**0.997**

Tableau 6.3 – Valeurs des métriques sur volume entier et les trois ROIs.

moins les images est supérieur. Les méthodes d'apprentissage profond surpassent largement les résultats des autres méthodes, notamment pour le U-Net-MP et le U-Net-3D+. Les artefacts produits par les projections tronquées pourraient être à l'origine de la différence observée entre les valeurs NRMSE des méthodes itératives et des U-Net. Malgré l'augmentation de la taille du volume, les projections tronquées engendrent une augmentation de l'intensité de certains voxels.

Concernant les temps de calcul, il a fallu 35 minutes pour reconstruire itérativement les volumes avec KL-TV, en effectuant 500 itérations. Les réseaux 2D ont été entraînés en 1 heure et 35 minutes, et le réseau 3D en 6 heures et 18 minutes. Le volume avec reconstruction FDK et post-traitement avec U-Net est obtenu en moins d'une minute en 2D, et moins de trois minutes en 3D (sur GPU).

La reconstruction itérative permet d'améliorer la qualité visuelle par rapport à FDK, mais la perte de détails cruciaux, tels que la longueur des canaux dentaires, engendrée par la régularisation TV sont un frein pour leur utilisation clinique. De plus, les résultats dépendent du choix du paramètre de régularisation qui doit être ajusté en fonction du niveau de bruit présent dans les projections. Les temps de calcul assez longs et les ressources numériques nécessaires présentent également des inconvénients dans l'utilisation de ces méthodes.

Nous observons une réduction des artefacts métalliques, dans toutes les images, bien qu'aucun algorithme de réduction de ces artefacts n'ait été utilisé. La qualité pourrait être améliorée avec de tels algorithmes, en particulier sur les post-traitements avec U-Net qui souffrent des différences de plages de valeurs entre les différentes coupes ou patch. Le post-traitement avec les U-Net donne de très bons résultats en des temps suffisamment courts pour une utilisation médicale. Les comparaisons et l'évaluation effectuées montrent que la reconstruction FDK + U-Net-MP conserve bien les détails nécessaires à un diagnostic correct, tout en ayant un aspect visuel agréable. Ce dernier point est important car le confort apporté par une image non bruitée rend le diagnostic plus simple.

Le U-Net-3D utilisé dans cette étude, ne présente pas de meilleurs résultats que le U-Net-MP, ni le U-Net axial Γ_A. De possibles raisons de cette moins bonne performance ont été citées dans la section 4 : entraînement sur patch, manque de profondeur ou nombre d'époques trop faible. Le système d'apprentissage par patch ne permet pas d'avoir autant de contexte spatial que le permettrait un apprentissage sur l'intégralité du volume. Cependant, les ressources numériques nécessaires à de tels entraînements n'étaient pas disponibles, malgré l'utilisation d'un super-calculateur, et nous n'avons pas pu faire ces tests durant cette thèse. En revanche, les

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

expériences effectuées ont montré que l'augmentation de la profondeur améliorait les résultats. Ces observations ont été appuyées par les comparaisons plus approfondies effectuées dans la dernière section avec le U-Net-3D+. Les résultats de ce dernier réseau sont comparables à ceux du U-Net-MP et les volumes obtenus de bonne qualité. Il nous faut également rappeler que par manque de temps, les hyperparamètres des réseaux 3D n'ont pas été optimisés, et que l'optimisation de ceux-ci ne pourrait qu'améliorer les bons résultats que nous obtenons déjà. Enfin, il est possible qu'utiliser des images de références obtenues avec des méthodes itératives améliore aussi les résultats, les volumes reconstruits avec FDK souffrant de plus d'artefacts.

Nous pouvons souligner le fait que la comparaison entre réseau 2D et 3D est une tâche particulièrement difficile étant donné que pour avoir le même nombre de paramètres pour chaque, la structure du réseau doit être modifiée. Or, cette structure est importante dans la capacité du réseau à représenter la fonction liant image faible dose et image à dose normale.

Ainsi, ce chapitre nous a permis de faire une première exploration de l'utilisation des méthodes d'apprentissage profond en reconstruction CBCT dentaire. Une étude plus approfondie permettrait de réellement pouvoir évaluer la performance (impact du nombre de données, du nombre de paramètres, etc.) pour obtenir une comparaison 2D/3D plus juste. Les expériences menées montrent que ces méthodes peuvent considérablement améliorer la qualité de reconstructions obtenues analytiquement à partir de projections acquises à faible dose, surpassant les méthodes itératives régularisées avec variation totale.

Les résultats obtenus sont très encourageants, et nous espérons qu'il serait possible d'avoir encore de meilleures performances en utilisant des méthodes incluant le processus de reconstruction au sein du réseau. Au vu de la complexité de la géométrie conique, et du faible nombre de données à notre disposition, notre choix se porterait plutôt sur des méthodes de type unrolling, que sur les méthodes apprenant l'opération de reconstruction. Néanmoins, la prise en main difficile des librairies permettant d'inclure les opérateurs de tomographie dans les réseaux, ainsi que les ressources numériques requises pour ces méthodes, et plus particulièrement en CBCT, rendent très lent le développement de telles méthodes. De récents articles suggèrent de ne pas considérer l'intégralité des projections pendant les reconstructions, par itérations, mais plutôt un nombre limité de projections, en utilisant les projections par cycle à la manière d'OSEM (Tang *et al.*; 2021). Les même auteurs proposent également une méthode où la reconstruction est faite en sous-résolution pour diminuer la mémoire requise (Tang; 2022).

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Bibliographie

Chollet, F. *et al.* (2015). Keras. 125

Feldkamp, L. A., Davis, L. C. et Kress, J. W. (1984). Practical cone-beam algorithm. *Josa a*, 1(6):612–619. 131

Friot--Giroux, L., Leuliet, T., Peyrin, F. et Maxim, V. (2022). Reconstruction tomographique faible dose en imagerie dentaire : régularisation TV versus UNet. *In XXVIIIème Colloque Francophone de Traitement du Signal et des Images (GRETSI 2022)*. 131

Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc. 120

He, K., Zhang, X., Ren, S. et Sun, J. (2015). Delving deep into rectifiers : Surpassing human-level performance on imagenet classification. *In Proceedings of the IEEE international conference on computer vision*, pages 1026–1034. 120

Ioffe, S. et Szegedy, C. (2015). Batch normalization : Accelerating deep network training by reducing internal covariate shift. *In International conference on machine learning*, pages 448–456. PMLR. 120

Jin, K. H., McCann, M. T., Froustey, E. et Unser, M. (2017). Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26(9):4509–4522. xiv, 120, 121

LeCun, Y., Bottou, L., Bengio, Y. et Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324. 127

Odena, A., Dumoulin, V. et Olah, C. (2016). Deconvolution and checkerboard artifacts. *Distill*. 120

Ronneberger, O., Fischer, P. et Brox, T. (2015). U-net : Convolutional networks for biomedical image segmentation. *In International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer. 120

Sidky, E. Y., Jørgensen, J. H. et Pan, X. (2012). Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm. *Physics in Medicine & Biology*, 57(10):3065–3091. 131

Tang, J. (2022). Operator sketching for deep unrolling networks. *arXiv preprint arXiv :2203.11156*. 136

Tang, J., Mukherjee, S. et Schönlieb, C.-B. (2021). Stochastic primal-dual deep unrolling. *arXiv preprint arXiv :2110.10093*. 136

Van Aarle, W., Palenstijn, W. J., Cant, J., Janssens, E., Bleichrodt, F., Dabravolski, A., De Beenhouwer, J., Batenburg, K. J. et Sijbers, J. (2016). Fast and flexible X-ray tomography using the ASTRA toolbox. *Optics express*, 24(22):25129–25147. 125

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Conclusion

Cette thèse avait pour objectif d'étudier différentes méthodes de reconstruction tomographique afin de les appliquer dans le cas de l'imagerie dentaire à faible dose. Pour cela, nous avons adapté des méthodes itératives utilisant une régularisation avec variation totale aux spécificités de la géométrie conique majoritairement utilisée en dentaire, et nous avons évalué les performances de réseaux de neurones en post-traitement de reconstructions analytiques.

Nous avons commencé par introduire les principes physiques et mathématiques de la reconstruction tomographique, tout en insistant sur les caractéristiques de l'imagerie dentaire CBCT. De premières méthodes de reconstruction ont été présentées : les reconstructions analytiques, rapides et efficaces en conditions idéales, mais sensibles au bruit et aux projections en faible nombre ; et certaines reconstructions itératives permettant de prendre en compte la distribution du bruit dans les projections.

Dans un second temps, nous avons étudié les algorithmes de reconstruction tomographique itératifs incluant une régularisation TV, adaptés à une distribution gaussienne ou Poisson pour le bruit. Nous les avons appliqués au fantôme de Shepp-Logan 2D en étudiant leur performance sur un bruit mixte gaussien-Poisson. Ces résultats préliminaires ont permis de mettre en avant l'importance du choix du paramètre de régularisation dans le résultat de la reconstruction.

Ces algorithmes ont ensuite été adaptés et appliqués progressivement à la géométrie conique avec les fantômes de Shepp-Logan et de mâchoire, puis à des données expérimentales acquises sur un fantôme anthropomorphique. Ces expériences ont mis en valeur les avantages et les inconvénients de ces méthodes. Elles permettent d'un côté de réduire efficacement les artefacts et le bruit présent dans les reconstructions analytiques faibles-doses, avec un compromis important entre conservation des détails et réduction du bruit, compromis géré par le paramètre de régularisation. D'un autre côté, ces méthodes nécessitent d'importantes ressources de calcul numérique, en particulier en CBCT et dans le cas de projections tronquées, et les reconstructions sont relativement lentes pour une utilisation clinique. De plus, la régularisation TV utilisée dans ces méthodes n'est pas forcément la plus adaptée à l'imagerie dentaire : elle favorise les zones constantes par morceaux et est assez efficace quand il s'agit de reconstruire de grosses structures telles que les dents, mais a tendance à effacer les détails particulièrement fins (canaux de dents, structure d'os). Cette étude a également mis en avant la difficulté de comparer les différentes méthodes itératives régularisées entre elles. La fonction coût différente entre SIRT-TV et les deux autres, ou encore la différence d'architecture entre l'algorithme KL-TV et les autres rendent délicat l'obtention d'une comparaison juste.

Nous pouvons lister plusieurs manières d'améliorer le travail effectué sur les méthodes itératives. Tout d'abord, il serait possible d'accélérer les calculs en utilisant OSEM (Hudson et Larkin; 1994), ou l'algorithme de Chambolle-Pock stochastique (Chambolle *et al.*; 2018) qui n'utilisent qu'une partie des projections à chaque itération. L'ensemble des projections choisies change à chaque itération de sorte à ce que la totalité des projections soit prise en compte régulièrement. Ces méthodes pourraient réduire les temps de calcul et la mémoire nécessaire. Du point de vue de la régularisation, la variation totale étudiée ici, n'est pas adaptée pour

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

reconstruire les structures très fines. Les reconstructions itératives régularisées par TV souffrent ainsi de la perte de petits détails tels que la structure des os, ou certains canaux. Il serait donc intéressant d'utiliser d'autres régularisations. Ye *et al.* (2018) ont proposé un algorithme itératif dans lequel la régularisation est effectuée par un réseau de neurones.

Un deuxième axe d'étude d'amélioration de la reconstruction tomographique a été étudié dans cette thèse : l'application des méthodes d'apprentissage profond, particulièrement efficaces en traitement d'images depuis quelques années. Nous avons synthétisé quelques travaux présents dans l'état de l'art, en expliquant qu'elles devaient être adaptées en fonction du type de données à traiter. Ainsi, disposant d'une base de données de projections acquises en dose normale, nous avons décidé d'utiliser des méthodes d'apprentissage supervisé, en simulant des acquisitions faibles doses à partir des données à notre disposition. Nous avons pris une architecture relativement simple mais qui a déjà prouvé son efficacité dans de nombreux domaines d'application : les U-Net. Plusieurs U-Net ont été entraînés à réduire les artefacts présents dans les reconstructions analytiques effectuées à partir des données faibles-doses. Nos expériences ont montré de très bons résultats pour le U-Net multi-plans, simple et relativement rapide à entraîner, aboutissant à des reconstructions précises et de qualité, permettant d'effectuer plus facilement de meilleurs diagnostics médicaux.

Les bonnes performances des réseaux de neurones doivent tout de même être mises en perspective. La phase d'entraînement est cruciale dans la capacité du réseau à généraliser ce qu'il a appris sur de nouvelles données. Notre base de données, constituée de 32 patients, a déjà montré quelques exemples de cas où la sous-représentation des données a pu poser des problèmes. Pour un usage clinique de ces algorithmes, il est donc nécessaire de les ré-entraîner avec une plus grande base de données.

Les expériences menées dans cette thèse donnent un aperçu des bonnes capacités des méthodes d'apprentissage profond à améliorer l'aspect de reconstructions analytiques. Ces méthodes sont néanmoins justement basées sur une reconstruction analytique. Dans nos simulations de faible dose, nous avons choisi un nombre de projections suffisamment élevé pour que les reconstructions analytiques restent correctes, malgré la corruption par le bruit et les artefacts. Nous pouvons facilement imaginer que dans des cas plus extrêmes, la reconstruction analytique ne permette pas de récupérer assez d'informations pour qu'un réseau puisse restaurer correctement les images par la suite. Une façon d'éviter ce problème serait d'inclure le processus de reconstruction au réseau de neurones, en utilisant les réseaux de unrolling. La nécessité de très grandes capacités numériques, plus particulièrement en cone beam, fait qu'en pratique, ces algorithmes sont encore difficiles à mettre en œuvre. De récents articles laissent cependant espérer une possible utilisation des méthodes, en considérant par exemple une approche de type OSEM dans lequel les projections ne seraient pas toutes utilisées simultanément (Tang *et al.*; 2021) ; ou encore en utilisant une version sous-échantillonnée des volumes et des opérateurs de projections/rétro-projection (Tang; 2022).

Nous rappelons également qu'une large variété de réseau, est aujourd'hui disponible et que cette thèse n'a étudié que des U-Net. L'entraînement d'autres réseaux, tels que les GAN ou les *Transformers*, pourrait étayer cette étude sur les capacités des réseaux de neurones à améliorer la qualité d'images CT faibles-doses. Enfin, nous pourrions effectuer d'autres tâches simultanément au débruitage, et par exemple prendre en compte le flou dans les images et le corriger avec de la déconvolution.

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Bibliographie

Chambolle, A., Ehrhardt, M. J., Richtárik, P. et Schonlieb, C.-B. (2018). Stochastic primal-dual hybrid gradient algorithm with arbitrary sampling and imaging applications. *SIAM Journal on Optimization*, 28(4):2783–2808. 139

Hudson, H. et Larkin, R. (1994). Accelerated image reconstruction using ordered subsets of projection data. *IEEE Transactions on Medical Imaging*, 13(4):601–609. 139

Tang, J. (2022). Operator sketching for deep unrolling networks. *arXiv preprint arXiv :2203.11156*. 140

Tang, J., Mukherjee, S. et Schönlieb, C.-B. (2021). Stochastic primal-dual deep unrolling. *arXiv preprint arXiv :2110.10093*. 140

Ye, D. H., Srivastava, S., Thibault, J.-B., Sauer, K. et Bouman, C. (2018). Deep residual learning for model-based iterative ct reconstruction using plug-and-play framework. *In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6668–6672. IEEE. 140

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Liste des communications

Publication dans un journal

Friot--Giroux Louise, Peyrin Françoise et Maxim Voichiţa. (2022). Iterative tomographic reconstruction with TV prior for low-dose CBCT dental imaging. *Physics in Medicine & Biology*, *67* : 205010. URL : https://iopscience.iop.org/article/10.1088/1361-6560/ac950c

Participations à des congrès

Leuliet Théo, Friot--Giroux Louise, Baaziz Walid, Bretin Elie, Ersen Ovidiu, Peyrin Françoise, Sixou Bruno et Maxim Voichiţa. (2021). Efficiency of TV-regularized algorithms in computed tomography with Poisson-Gaussian noise. *2020 28th European Signal Processing Conference (EUSIPCO)*, p. 1294-1298. URL : https://ieeexplore.ieee.org/abstract/document/9287762

Friot--Giroux Louise, Leuliet Théo, Peyrin Françoise et Maxim Voichiţa. (2022). Reconstruction tomographique faible dose en imagerie dentaire : régularisation TV versus UNet. *XXVIIIème Colloque Francophone de Traitement du Signal et des Images (GRETSI 2022)*. URL : http://gretsi.fr/data/colloque/pdf/2022_friotgiroux900.pdf

Friot--Giroux Louise, Peyrin Françoise et Maxim Voichiţa. (2023). Reconstruction tomographique faible dose en imagerie dentaire : régularisation TV versus UNet. *1er Colloque Français d'Intelligence Artificielle en Imagerie Biomédicale (IABM 2023)*. (Résumé accepté)

Friot--Giroux Louise, Peyrin Françoise et Maxim Voichiţa. (2023). Enhancement of low-dose tomographic volumes in dental imaging by CNNs. *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI 2023)*. (Résumé soumis)

Communication orale

Friot--Giroux Louise, Peyrin Françoise, Maxim Voichiţa, Ripoche Xavier et Maury Colombe. Reconstruction tomographique avancée en imagerie dentaire. *3e journée industrielle du LabEx PRIMES, 20 Juin 2022*.

Autres

Création d'un jeu vidéo : TomoVision (2020). Friot--Giroux Louise, Tohier Céline (game designer), Six Jonathan (game programmer), Illien Alizée (graphiste) et Rozlonkowski Nicolas (sound designer). *Scientific Game Jam, Lyon, 2020.*

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

INSTITUT NATIONAL DES SCIENCES APPLIQUÉES LYON

Département de la Formation par la Recherche et des Études Doctorales (**FEDORA**)

Bâtiment INSA direction, 1er étage
37, av. J. Capelle
69621 Villeurbanne Cédex
fedora@insa-lyon.fr

Référence : TH0940_FRIOT--GIROUX

L'INSA Lyon a mis en place une procédure de contrôle systématique via un outil de détection de similitudes (logiciel Compilatio). Après le dépôt du manuscrit de thèse, celui-ci est analysé par l'outil. Pour tout taux de similarité supérieur à 10%, le manuscrit est vérifié par l'équipe de FEDORA. Il s'agit notamment d'exclure les auto-citations, à condition qu'elles soient correctement référencées avec citation expresse dans le manuscrit.

Par ce document, il est attesté que ce manuscrit, dans la forme communiquée par la personne doctorante à l'INSA Lyon, satisfait aux exigences de l'Etablissement concernant le taux maximal de similitude admissible.

INSA LYON
Campus LyonTech La Doua
20, avenue Albert Einstein - 69621 Villeurbanne cedex - France
Tél. +33 (0)4 72 43 83 83 - Fax +33 (0)4 72 43 85 00
www.insa-lyon.fr

membre de UNIVERSITÉ DE LYON

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés

FOLIO ADMINISTRATIF

THÈSE DE L'INSA LYON, MEMBRE DE L'UNIVERSITÉ DE LYON

NOM : FRIOT--GIROUX

DATE DE SOUTENANCE : 07/02/2023

PRÉNOM : Louise

TITRE : Méthodes de reconstruction avancées en tomographie dentaire par faisceau conique

NATURE : Doctorat

NUMÉRO D'ORDRE : 2023ISAL0008

ÉCOLE DOCTORALE: Électronique, Électrotechnique, Automatique (EEA)

SPÉCIALITÉ : Traitement du Signal et de l'Image

RÉSUMÉ : La tomographie par rayons X est une technique très utilisée en imagerie médicale pour observer les structures anatomiques des patients. A partir de mesures de l'absorption des rayons X par les tissus, un processus de reconstruction permet d'obtenir des images du volume observé. En imagerie dentaire, le CBCT utilise un faisceau de rayons X conique et permet d'imager rapidement le volume 3D de la mâchoire. L'exposition aux rayons X est cependant dangereuse pour la santé, et l'enjeu principal de la tomographie X est d'obtenir des images de bonne qualité avec une dose minimale.
Les méthodes analytiques offrent des reconstructions rapides mais peu robustes lorsque les projections sont bruitées ou en faible nombre. Les méthodes itératives donnent la possibilité de prendre en compte la modélisation physique du problème et des connaissances a priori sur le volume. Récemment, l'essor des méthodes d'apprentissage profond et leur application au traitement des images ont suggéré de nouvelles pistes de recherche en tomographie.
L'objectif de cette thèse est d'implémenter des méthodes de reconstruction fiables et efficaces pour la reconstruction tomographique faible dose en imagerie dentaire.
Trois algorithmes itératifs, appliquant une régularisation par variation totale, ont été implémentés et appliqués sur fantômes et données réelles faibles doses. Les résultats ont montré la supériorité des méthodes itératives sur les méthodes analytiques.
Nous avons également testé les méthodes d'apprentissage profond en reconstruction tomographique. Des réseaux ont été entraînés de manière supervisée pour améliorer la qualité des images reconstruites avec faible dose. Les résultats ont été comparés avec ceux d'une méthode itérative.
Différentes méthodes de reconstruction tomographique ont été appliquées sur fantômes et données réelles. Les méthodes itératives ont montré de meilleurs résultats que la méthode analytique. Les méthodes d'apprentissage profond, s'appliquent sur reconstruction analytiques et surpassent les méthodes itératives en termes de temps de reconstruction, avec une meilleure qualité d'images.

MOTS-CLÉS : Reconstruction tomographique, Imagerie médicale, Imagerie dentaire, Méthodes itératives, Deep Learning

LABORATOIRE DE RECHERCHE : Centre de Recherche en Acquisition et Traitement de l'Image pour la Santé (CREATIS)

DIRECTRICE DE THÈSE : Voichiţa MAXIM
CO-DIRECTRICE DE THÈSE : Françoise PEYRIN

PRÉSIDENT DU JURY : Laurent DESBAT

COMPOSITION DU JURY :
Alexandre BOUSSE
Emil SIDKY
Émilie CHOUZENOUX
Laurent DESBAT

Cette thèse est accessible à l'adresse : https://theses.insa-lyon.fr/publication/2023ISAL0008/these.pdf
© [L. Friot--Giroux], [2023], INSA Lyon, tous droits réservés