

# Méthodes numériques pour une reconstruction et un recalage combinés en Mammographie Numérique 3D par Tomosynthèse (MNT/DBT)

Alex Delalande, Baptiste Doyen, Lorenzo Croissant

29 novembre 2018

## Résumé

En imagerie médicale, particulièrement dans le cadre de la tomographie, des problèmes inverses mal posés nécessitent l'élaboration de méthodes numériques efficaces, rapides et incorporant des *a priori* pour mener à bien leur résolution. L'article de Yang et al.<sup>[2]</sup> présente deux nouvelles approches pour combiner deux de ces problèmes, la reconstruction et le recalage, dans le cadre de la Mammographie Numérique 3D par Tomosynthèse. Ce rapport réalise une synthèse critique de l'article et offre une évaluation des différentes méthodes proposées.

## 1 Introduction

Le cancer du sein est le cancer le plus commun en France avec 60,000 cas par an. Depuis 2004 la France procède à un dépistage organisé pour toutes les femmes de plus de 50 ans. Ce dépistage est réalisé par une mammographie digitale, une modalité d'acquisition 2D par rayons-X, et son examen par un radiologue. Malgré l'efficacité démontrée du dépistage, les tissus fibrokystiques (denses) du sein ont des taux d'absorption élevés et rendent la détection de lésions difficile. Le taux de rappel est donc fort, ce qui exerce un stress chez les patients et fait perdre du temps et de l'argent au système de soin. La Mammographie Numérique 3D par Tomosynthèse (MNT/DBT) est une nouvelle méthode d'acquisition qui promet d'augmenter la spécificité et la sensibilité de l'examen de dépistage du cancer du sein.

La tomosynthèse est par définition la reconstruction d'un volume à partir de projections (dans le cas de la MNT, des radios à faibles doses) prises avec des angles différents. Cette reconstruction numérique permet de synthétiser des images dans des plans d'intérêt sans être gêné par les tissus denses situés hors du plan. Ceci permet de plus facilement distinguer les lésions cachées sous des tissus denses et aussi de différencier les lésions bénignes et malignes sans faire de rappel. Améliorer les méthodes de reconstruction offre donc une opportunité d'encore améliorer la performance de la MNT.

Afin de présenter une critique de l'article de Yang et al.<sup>[2]</sup>, nous définirons d'abord mathématiquement les problèmes de reconstruction d'un volume et de recalage de plusieurs acquisitions, puis nous étudions les méthodes numériques proposées : traitement séparé, traitement itératif, et enfin traitement joint. Nous aborderons enfin leur implémentation et la discussion de leur résultats.

## 2 Principaux concepts

### 2.1 Acquisition Tomographique

La tomosynthèse est un problème de reconstruction d'image 3D à partir de projections 2D. Notons  $f_r$  le volume 3D étudié, défini à valeur dans  $\mathbb{R}^{D_r}$  (représentation discrète du volume continu de  $\mathbb{R}^3$  étudié) et  $p := \{p_i\}_i^n$  l'ensemble des  $n$  projections acquises, avec  $p \in \mathbb{R}^{n \times D_a}$ .

Les dimensions  $D_r$  et  $D_a$  sont respectivement les dimensions de l'image réelle et de chaque acquisition. Notons  $A \in \mathbb{R}^{n \times D_a \times D_r}$  le tenseur représentant le système d'acquisition, c'est-à-dire tel que  $p = Af_r$ .  $A$  induit une perte d'information inhérente à l'acquisition et porte la grille d'acquisition de  $f_r$ . Si l'on dispose d'un opérateur d'interpolation  $R$  dans  $\mathbb{R}^3$  on peut définir de manière équivalente  $p = AR(f_r)$ , où  $R(f_r)$  est l'interpolation continue de l'image définie comme discrète  $f_r$  et  $AR(f_r)$  sont les projections de  $R(f_r)$ .

## 2.2 Tomosynthèse & Reconstruction

La tomosynthèse (ou reconstruction) consiste en retrouver  $f_r$  à partir de  $p$ , ou en trouver une approximation  $f$  telle que  $Af \approx p$ . Il s'agit d'un problème mal posé sous-déterminé car  $A$  est une projection sur un sous-espace strictement plus petit que l'espace d'origine.

Différentes résolutions ont été proposées pour traiter ce problème. Du côté des approches analytiques, les méthodes dites de *Back Projection* sont les plus communes et reposent sur l'étude de  $f_r$  et de  $p$  dans un autre domaine, comme le domaine fréquentiel pour la *Filtered Back Projection*. Ces méthodes sont efficaces mais souffrent des artefacts dus à l'échantillonnage de l'acquisition qui réduisent la qualité de leur reconstructions. *A contrario*, ce problème peut également être résolu par des méthodes itératives. La modélisation la plus classique revient à minimiser une distance entre la reconstruction projetée et la projection acquise, et peut être posé comme suit :

$$f^* := \underset{f}{\operatorname{argmin}} \frac{1}{2} \|Af - p\|^2, \text{ sujet à } f \in \mathbb{R}^{D_r}.$$

Il s'agit de l'approche choisie par Yang et al.<sup>[2]</sup> pour réaliser leurs reconstructions. L'objectif d'optimisation étant en moindres carrés, il s'agit d'un problème convexe, mais *a priori* non strictement convexe car  $A^T A$  n'est pas nécessairement inversible. Ce problème est usuellement résolu par une méthode itérative du premier ordre ou du deuxième ordre, comme ce sera le cas ici.

## 2.3 Recalage

Le sein développe souvent des néoplasmes pré-cancéreux (*e.g.* carcinomes canalaire *in situ*) qui doivent être surveillés par le radiologue. Pour ce faire, il compare les images de différents rendez-vous de suivi. La variabilité inter-image associée rends l'analyse difficile, mais peut être facilitée numériquement au moyen de méthodes de recalage. On considère deux images  $p_1$  et  $p_2$ , acquises sur la même machine et le même patient à deux moments différents. On suppose que  $p_1$  est le résultat de l'acquisition sous  $A$  de  $f_1$ , soit  $p_1 = Af_1$ , et que  $p_2 = Af_2$ . On suppose aussi  $f_1$  et  $f_2$  sont liées par une transformation paramétrique  $T_\theta$ , telle que  $f_1 = T_\theta f_2$ .

La partie recalage est généralement traitée en choisissant une des deux approches suivantes afin de réduire le problème à une forme connue dans la littérature : paramétrique (*e.g.* rigide, affine, polynomiale, spline) ou déformable (*e.g.* flux optique, déformation élastique ou fluide). Yang et al.<sup>[2]</sup> ont choisi de travailler sur deux types de transformations paramétriques : affines et B-Splines.

### 3 Méthodes numériques

#### 3.1 Approche séquentielle

L'approche traditionnelle en tomographie traite les problèmes de reconstruction et de recalage séparément. Découpler les deux problèmes est l'approche intuitive, mais n'utilise pas l'information potentiellement contenue dans un problème pour aider la résolution de l'autre. Les étapes de reconstruction reviennent à traiter séparément les deux problèmes de reconstruction :

$$\hat{f}_1 = \underset{f_1}{\operatorname{argmin}} \left( \frac{1}{2} \|Af_1 - p_1\|_F^2 \right) \quad (3.1)$$

$$\hat{f}_2 = \underset{f_2}{\operatorname{argmin}} \left( \frac{1}{2} \|Af_2 - p_2\|_F^2 \right), \quad (3.2)$$

de celui du recalage qui revient à estimer le paramètre de la transformation  $\theta$  en résolvant :

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \left( \frac{1}{2} \|T_\theta(\hat{f}_2) - \hat{f}_1\|_F^2 \right) \quad (3.3)$$

Considérons une solution  $f_1^*$  au problème d'optimisation 3.1, elle doit annuler le gradient  $A^T A f_1^* = A^T p_1$ . Il n'y a pas de garantie sur l'existence d'un inverse de l'application  $A^T A : \mathbb{R}^{D_r} \rightarrow \mathbb{R}^{D_r}$  et son calcul peut-être numériquement coûteux. Cependant, on peut trouver un  $f_1^*$  satisfaisant la condition d'annulation du gradient en utilisant une méthode itérative, souvent une descente de gradient.

On peut ici noter que le problème de l'équation 3.1 aurait pu être régularisé afin d'imposer certains comportements au volume reconstruit  $f_1^*$ , portant par exemple sur la variation totale de son gradient (régularisation la plus courante) pour minimiser la longueur des contours de la forme reconstruite et observer une reconstruction "lisse".

#### 3.2 Une nouvelle approche itérative

La première approche originale proposée par Yang et al. suppose que l'estimation de la transformation  $T_\theta$  permet d'améliorer la reconstruction et inversement. Dans leur première méthode, Yang et al. combinent les deux tâches dans une approche itérative, où l'on alterne des étapes de minimisation des deux problèmes. A l'étape  $k$ , on réalise  $j$  étapes de minimisation indépendante des équations 3.1 et 3.2 dont on obtient  $\hat{f}_{1,k}$  et  $\hat{f}_{2,k}$  puis on estime  $\theta$  en résolvant

$$\hat{\theta}_k = \underset{\theta}{\operatorname{argmin}} \left\| T_\theta(\hat{f}_{2,k}) - \hat{f}_{1,k} \right\|_F^2.$$

On met alors à jour  $\hat{f}_{1,k+1}$  avec  $T_{\hat{\theta}_k}(\hat{f}_{2,k})$ , ce qui permet de lier les reconstructions et recalage. On répète ce processus jusqu'à satisfaction d'un critère de convergence. Yang et al. proposent pour résoudre les problèmes de reconstruction d'implémenter une méthode du deuxième ordre quasi-Newton comme L-BFGS<sup>[1]</sup>.

L'approche itérative consiste donc à résoudre partiellement le problème de reconstruction, d'utiliser les résultats partiels obtenus pour ensuite résoudre le problème de recalage puis de répéter ce processus un certain nombre d'étapes en ré-initialisant à chaque fois les variables du problème de reconstruction avec les estimées partielles.

La boucle extérieure de cette méthode indexée par  $k$  est équivalente à l'approche séquentielle vue plus haut. La différence tient ici au fait qu'on stoppe le processus de reconstruction à l'étape  $j$  avant sa convergence, que l'on calcule  $\hat{\theta}_k$  puis ensuite que l'on recommence le même processus de reconstruction avec comme paramètres initiaux de l'étape  $k + 1$  :  $f_{1,init} = T_{\hat{\theta}_k}(\hat{f}_{2,k})$  et  $f_{2,init} = \hat{f}_{2,k}$ . Afin de conserver l'information issue de  $\hat{f}_{1,k}$  et  $\hat{f}_{2,k}$ , Yang et al.<sup>[2]</sup> proposent aussi d'initialiser avec  $f_{1,init} = \frac{1}{2}(T_{\hat{\theta}_k}(\hat{f}_{2,k}) + \hat{f}_{1,k})$ .

Le premier type d'initialisation permet d'introduire dans le processus de reconstruction de l'information issue des paramètres  $\hat{\theta}_k$  et donc de permettre un couplage du type entrée-sortie des deux problèmes : la sortie du recalage est réutilisée en entrée de la reconstruction. Le second type permet aussi ce couplage et préserve en partie l'information transmise par  $\hat{f}_{1,k}$ .

### 3.3 Une nouvelle approche simultanée

Yang et al.<sup>[2]</sup> poussent l'idée de complémentarité entre les tâches de reconstruction et de recalage encore plus loin en décidant de résoudre ces deux tâches simultanément, c'est-à-dire en réduisant la reconstruction et l'estimation de transformation à un seul et même problème d'optimisation. Le problème formulé par Yang et al.<sup>[2]</sup> est le suivant :

$$\hat{f}_1, \hat{\theta} = \underset{f_1, \theta}{\operatorname{argmin}} \frac{1}{2} \left( \|Af_1 - p_1\|_F^2 + \|AT_\theta(f_1) - p_2\|_F^2 \right). \quad (3.4)$$

Notons  $J(f_1, \theta)$  la fonction objectif de 3.4. Une solution doit nécessairement annuler le gradient de  $J$ . Yang et al.<sup>[2]</sup> montrent qu'on a les dérivées partielles suivantes :

$$\frac{\partial J}{\partial f_1} = A^T(Af_1 - p_1) + T_\theta^* A^T(AT_\theta f_1 - p_2) \quad (3.5)$$

$$\frac{\partial J}{\partial \theta} = (A \frac{dT_\theta}{d\theta} f_1)^T (AT_\theta f_1 - p_2) \quad (3.6)$$

avec  $T_\theta^*$  l'opérateur adjoint de  $T_\theta$  et  $dT_\theta/d\theta$  la dérivée de la transformation par rapport au paramètre de transformation  $\theta$ , approchée en pratique par différences finies.

Il faut noter que l'introduction de la transformation  $T_\theta$  dans le problème 3.4 rend l'objectif non-convexe en  $\theta$  si  $T_\theta$  ne dépend pas linéairement de  $\theta$ . Ceci rend une optimisation jointe de  $f_1$  et  $\theta$  impossible. Yang et al.<sup>[2]</sup> procède donc à une minimisation alternée découplée, procédure la plus courante pour ce genre d'objectif. Il s'agit simplement d'une méthode itérative ou le passage de l'itération  $n$  à l'itération  $n + 1$  est fait de la manière suivante :

$$\hat{f}_{1,n+1} = \underset{f_1}{\operatorname{argmin}} J(f_1, \hat{\theta}_n) \quad (3.7)$$

$$\hat{\theta}_{n+1} = \underset{\theta}{\operatorname{argmin}} J(\hat{f}_{1,n+1}, \theta) \quad (3.8)$$

Ce schéma est suivi en pratique en utilisant à nouveau la méthode L-BFGS pour chaque minimisation découplée et en utilisant les gradients présentés en équations 3.5 et 3.6.

On peut se demander si la minimisation étant découplée et alternée, elle ne revient pas à effectuer une méthode itérative avec un nombre "infini" d'étapes de minimisation sur  $f_1$  et  $f_2$ . En réalité, ces deux approches diffèrent car dans l'approche itérative, le recalage intervient dans la reconstruction seulement au niveau de l'initialisation des volumes, tandis qu'ici le recalage est toujours présent puisque qu'on estime un seul volume  $f_1$  et on utilise sa version recalée pour

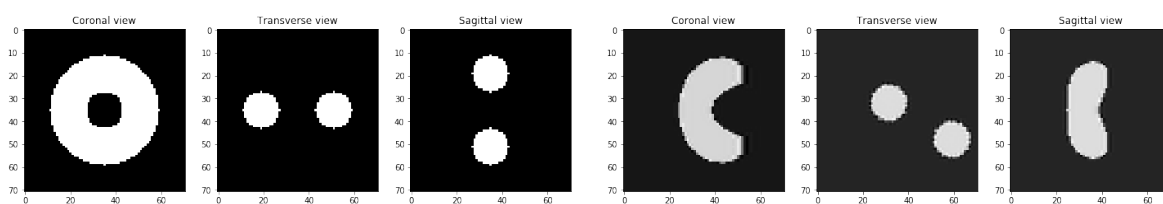
représenter le deuxième volume  $f_2$ . Cette idée se traduit avec la présence des transformations  $T_\theta$  dans l'équation 3.5 du gradient du coût par rapport à l'image  $f_1$  qui impose la prise en compte du recalage dans chaque étape de l'optimisation du problème de reconstruction 3.7. De la même manière, la reconstruction est omniprésente dans le problème de recalage 3.8 car un seul volume est estimé et le recalage doit permettre une bonne reconstruction des deux volumes étudiés avec ce seul volume estimé. Là encore, cette idée se traduit avec la présence de la matrice de projection  $A$  dans le gradient du coût par rapport au paramètre de recalage  $\theta$  dans l'équation 3.6.

## 4 Implémentation des méthodes

Nous avons implémenté<sup>1</sup> les trois méthodes avec les deux mêmes transformations que Yang et al.<sup>[2]</sup>, affine (rigide) et B-spline (non-rigide). Les transformations affines sont des combinaisons d'une rotation  $R$  et d'une translation  $t$  telle que  $T_\theta(x) := Rx + t$ . Cette transformation a donc 6 paramètres en  $3D$ , par simplicité nous nous sommes restreints à une rotation autour d'un axe du repère, et ainsi à 4 degrés de liberté (1 de rotation et 3 de translation). Ce sont ces degrés de liberté que l'on va chercher à estimer par les méthodes de recalage. Les splines sont des fonction polynomiales par morceaux, dont les B-splines sont une sous-famille. Leur caractéristiques sont régies par un ensemble de points de contrôle à interpoler, qu'il faudra donc estimer par le recalage. Les splines sont un modèle déformable et permettent de modéliser des déformations beaucoup plus complexe et hétérogènes que les affines.

### 4.1 Résultats avec des transformations affines

Suivant l'exemple de Yang et al.<sup>[2]</sup> nous avons commencé par travailler sur un fantôme toroïdal. Ne disposant pas de matériel médical, ce fantôme est une simulation numérique d'une acquisition tomographique, de taille  $70 \times 70 \times 70$  avec seulement deux valeurs d'intensité (0, 1). Nous avons généré un tore que nous avons ensuite projeté selon plusieurs angles avec le tenseur  $A$  et ainsi obtenir des projections réalistes. Ce volume réel  $f_{r,1}$  est présenté dans la figure 1a. Pour générer une seconde image  $f_{r,2}$  nous avons appliqué une transformation affine avec une rotation de 30 autour de l'axe  $Ox$  puis d'une translation par  $[0, -5, -10]$  à notre tore puis nous avons simulé une acquisition, qui est visible dans 1b.

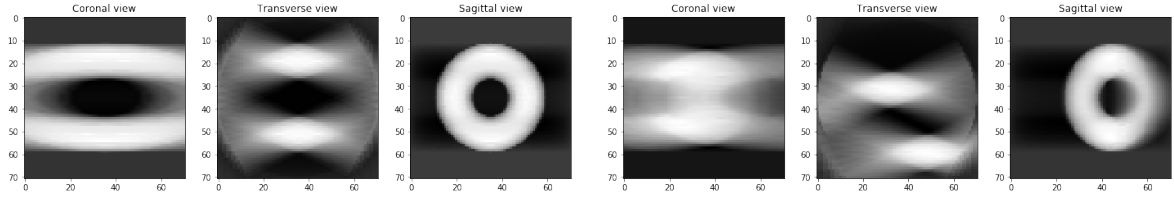


(a) coupe coronale, transverse et sagittale de  $f_{r,1}$       (b) coupe coronale, transverse et sagittale de  $f_{r,2}$

FIGURE 1: Les images réelles que nous allons tenter de reconstruire.

Nous avons utilisé le package python TomoPy pour générer le tenseur  $A$ . Ce package permet de simuler la projection de rayons  $X$  à travers un objet  $3D$  sur un détecteur  $2D$ . Pour obtenir  $A^T$  utilisée par exemple dans le gradient nous avons utilisé une méthode de *backward projection*. La méthode de reconstruction itérative produit les résultats visibles dans la figure 2.

1. Code disponible ici



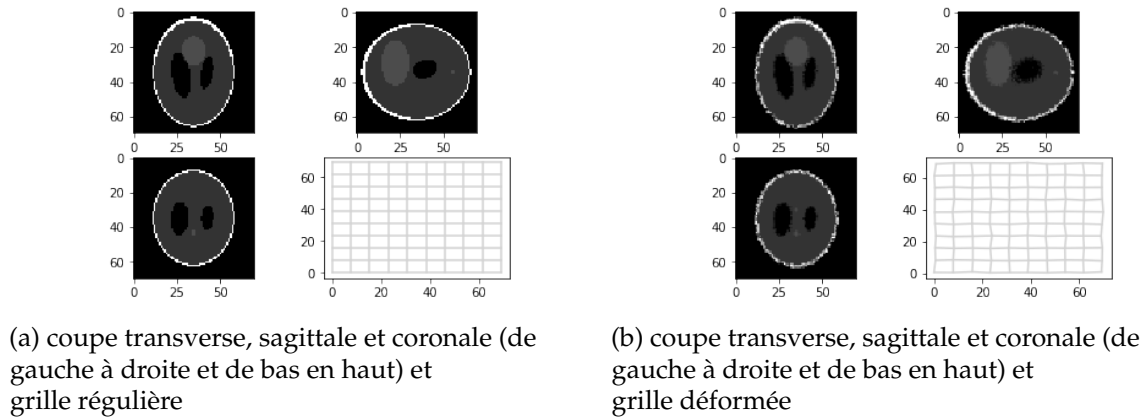
(a) coupe coronale, transverse et sagittale de  $\hat{f}_1$  (b) coupe coronale, transverse et sagittale de  $\hat{f}_2$

FIGURE 2: Les résultats de la méthode séquentielle (reconstruction).

On obtient donc ici des résultats très similaires à ceux de l'article. On remarque également que ces résultats peuvent être obtenus avec un nombre assez limité d'étapes de convergence (au plus une vingtaine), ce qui pourrait justifier l'utilisation d'un critère d'optimisation simple sans terme de régularisation. Enfin, le recalage simple nous permet également de retrouver assez fidèlement les paramètres de la transformation. Numériquement, on obtient :  $[0, -4.81, -10.29]$  pour le vecteur de translation et 29.85 degrés pour l'angle de rotation au bout de seulement 8 itérations de la descente de gradient.

## 4.2 Résultats avec des B-splines (transformations non-rigides)

L'image "ground-truth" choisie ici est celle du fantôme de *Shepp-Logan* dans sa version 3D (générée grâce à TomoPy). Bien qu'utilisé généralement comme modèle de la tête humaine, il s'agit d'un jeu de données classique en reconstruction. Il nous permet également de tester<sup>2</sup> les méthodes sur des images avec plusieurs zones d'intensités différentes (des ellipsoïdes dans le cas de ce fantôme). L'image générée est ici de dimension  $70 \times 70 \times 70$ . Pour effectuer la transformation, nous avons utilisé le package python SimpleITK. De plus, les points de contrôle ont été choisis au hasard. Une fois la transformation effectuée, on détermine ensuite les valeurs d'intensités aux noeuds de la grille régulière grâce à une interpolation trilinéaire.



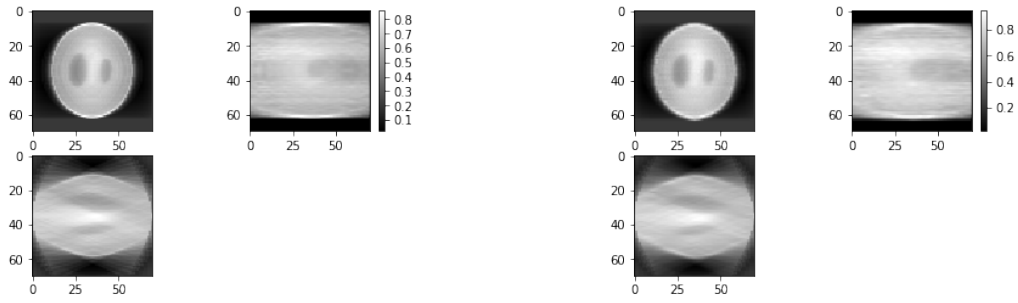
(a) coupe transverse, sagittale et coronale (de gauche à droite et de bas en haut) et grille régulière

(b) coupe transverse, sagittale et coronale (de gauche à droite et de bas en haut) et grille déformée

FIGURE 3: Coupes et grilles de  $f_r$  sans et avec transformation B-spline.

Avec une approche de reconstruction par minimisation des moindres carrés, nous obtenons les images dans la figure 5 :

2. Le code est disponible ici.



(a) coupes transverse, sagittale et coronale (de gauche à droite et de bas en haut) de  $\hat{f}_1$

(b) coupes transverse, sagittale et coronale (de gauche à droite et de bas en haut) de  $\hat{f}_2$

FIGURE 4: Coupes de  $\hat{f}$  avant et après transformation B-spline (transformation non-rigide)

L'approche itérative ne nous a pas permis d'obtenir des résultats de reconstruction significativement supérieurs par rapport à la méthode simple des moindres carrés dans le cas d'une projection non-transformée. Néanmoins, dans le cas d'une projection transformée, les résultats sont bien supérieurs comme le montre les images reconstruites suivantes :



(a) coupes coronale, transverse et sagittale (dans l'ordre) de  $\hat{f}_{2,seq}$

(b) coupes coronale, transverse et sagittale (dans l'ordre) de  $\hat{f}_{2,iter}$

FIGURE 5: Coupes de  $\hat{f}$  obtenues après reconstruction séquentielle et itérative

## 5 Discussion

Pour conclure, on peut formuler les remarques suivantes quant aux **résultats et évaluation de Yang et al.**<sup>[2]</sup>. D'abord, les algorithmes d'optimisation itératifs choisis dépendent toujours de paramètres (nombre d'itérations, critère de convergence) et d'une initialisation. On a pu remarquer lors de nos implémentations que les performances pouvaient "s'échanger" entre les méthodes selon ces paramètres, montrant ainsi une forte dépendance à ces grandeurs et réduisant la robustesse des résultats. On peut aussi remarquer que la méthode simultanée reconstruit plus fidèlement les volumes  $f_1$  et  $f_2$  en termes de formes mais elle génère des artefacts de textures qui sont nuisibles à la détection des lésions dans le cadre du dépistage du cancer du sein

Ensuite, les expériences proposées par Yang et al.<sup>[2]</sup> ne proposent pas d'évaluation des recalages. Ce type d'évaluation est impossible pour l'IRM mammaire réel (on ne connaît pas la transformation réelle) mais l'est pour les expériences sur les volumes simulés. Il est ainsi dommage de ne pas avoir accès à une évaluation quantitative de ces transformations. Au-delà de l'évaluation du recalage, un prochain travail pourrait étudier l'influence du type et du niveau de transformation géométrique sur l'efficacité des méthodes de reconstruction-recalage couplés. En effet, on sent bien que si la transformation induit une différence très forte entre les deux volumes, l'idée de reconstruction couplée perd en pertinence.

Deuxièmement, on peut aussi porter les remarques suivantes sur les **méthodes proposées par Yang et al.**<sup>[2]</sup>. Premièrement, les méthodes d'optimisation employées peuvent être coûteuses en calcul pour des données de haute résolution et de nombreuses approximations sont employées dans la méthode simultanée. Ces deux éléments pourraient amener en pratique à avoir des reconstructions plus longues et moins bonne sur des données haute résolution.

Ensuite, le couplage proposé est vu comme comme une augmentation du nombre de projections et donc permettant une amélioration des reconstructions. On a pu voir qu'en effet les reconstructions pouvaient s'avérer meilleures sous certains critères mais il faut noter que ce couplage peut diminuer les différences entre les volumes  $f_1$  et  $f_2$ , différences essentielles au suivi des lésions. Pour palier ce problème, les fonctions de coût pourraient intégrer cet objectif de détection des lésions (avec une régularisation adaptée par exemple).

## Références

- [1] Liu, D. C. and J. Nocedal  
1989. On the limited memory bfgs method for large scale optimization. *Mathematical Programming*, 45(1) :503–528. :1989
- [2] Yang, G., J. H. Hipwell, D. J. Hawkes, and S. R. Arridge  
2013. Numerical methods for coupled reconstruction and registration in digital breast tomosynthesis. *CoRR*, abs/1307.6008. :2013