Utilisation de dictionnaires parcimonieux incohérents pour la suppression d'artefact en ligne

Matthieu CHANCEL[1, 2] Hacheme AYASSO[1] Michel DESVIGNES[1] Jean-Michel VIGNOLLE[2] Eric LESPESSAILLES[3]

[1]Univ. Grenoble Alpes, CNRS, Grenoble INP*, GIPSA-lab,38000 Grenoble, France
*Institute of Engineering Univ. Grenoble Alpes

[2]Trixell, 460 rue du Pommarin, 38430 Moirans, France

[3]Centre Hospitalier Universitaire d'Orléans, 14 Av. de l'Hôpital, 45100 Orléans, France

Résumé – Parmi les perturbations qui peuvent apparaitre lors de l'acquisition de radiographies X, l'artefact "en ligne" se présente sous la forme d'une ligne sombre suivie d'une plus claire. Il apparait aléatoirement avec une amplitude variable selon les colonnes de l'image. Sa correction est complexe car chaque pixel contient à la fois de l'information clinique pertinente et cet artefact. Les approches traditionnelles ne sont pas adaptées à sa correction. Nous présentons une nouvelle méthode de suppression d'artefact "en ligne" basée sur un modèle joint de contenu clinique et d'artefact sous forme de dictionnaires parcimonieux : SParse Line ElimiNation with Dual Incoherent Dictionaries (SPLENDID). Le modèle clinique est appris afin d'etre faiblement corrélé avec le modèle d'artefact. Nous comparons notre méthode avec différentes solutions proposées dans la littérature. Des mesures réalisées sur un large jeu de données de radiographies montrent la supériorité de notre méthode face aux approches traditionnelles.

Abstract – Among the anomalies that may arise during X-ray image acquisition, the line artifact manifests as a dark line followed by a lighter one. It appears randomly with a variable amplitude depending on the image columns. Its correction is challenging, as each pixel simultaneously contains both relevant clinical information and the artifact. Traditional approaches are not well-suited for its removal. We propose a novel approach for line artifact suppression based on a joint model of clinical content and artifact representation using sparse dictionaries: SParse Line ElimiNation with Dual Incoherent Dictionaries (SPLENDID). The clinical model is learned to be weakly correlated with the artifact model. We compare our method with various solutions from the literature. Experimental evaluations on a large dataset of X-ray images demonstrate the superiority of our method over the traditional approaches.

1 Introduction

L'imagerie par rayons X est la modalité la plus utilisée en routine clinique dans le domaine médical. Elle est aussi de plus en plus utilisée dans le domaine de la détection et du contrôle qualité en milieu industriel. L'objet d'étude doit être imagé sans déformations ou perturbations. Dans des conditions extrêmes, des artefacts, comme l'artefact "en ligne" [7] (figure 1.a), peuvent apparaître sporadiquement à des positions aléatoires, avec des variations de forme et de niveau. Ils pénalisent la bonne lecture du contenu clinique par les praticiens. L'artefact "en ligne" se présente sous la forme d'une ligne sombre suivie d'une plus claire.

A notre connaissance, cet artefact n'est pas traité dans la littérature, mais plusieurs méthodes ont été proposées pour traiter des artefacts plus ou moins proches.

Le premier type de méthodes modélise l'artefact, l'estime sur l'image avant de le soustraire à celle-ci. La plupart des méthodes exploite la signature spectrale de l'artefact et l'isole par un filtrage basé sur un *a priori*. Sasada et al. [6] propose ainsi une méthode rapide pour corriger la grille anti-diffusion, artefact de forme connue sur des contenus cliniques presque constants. Cette approche laisse des parties de l'artefact visibles, car l'hypothèse sous-jacente de ces méthodes est celle d'un contenu clinique basse-fréquence. Cette hypothèse est parfois fausse, en particulier lors de la présence de vis (figure 1.a), agrafe, cathéter ou d'instrument d'exploration. Un second type de méthodes propose une modélisation conjointe du contenu clinique et de l'artefact. Chen et al. [1] proposent la concaténation d'une représentation par dictionnaires des modèles d'artefact et de contenu clinique. Sans *a priori* sur l'artefact, avec des dictionnaires appris indépendamment l'un de l'autre, l'estimation conjointe supprime parfois des informations cliniques en les attribuant à l'artefact. Cotte et al. [2] imposent un dictionnaire d'artefact qui diminue ces erreurs. Mais des caractéristiques similaires aux 2 dictionnaires mènent à des ambiguïtés d'estimation et à des erreurs de corrections.

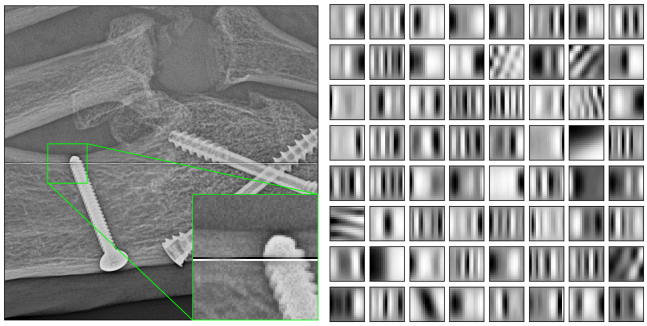
Pour résoudre ce problème, nous proposons une modélisation conjointe en ajoutant une incohérence structurée entre le contenu clinique et l'artefact. L'estimation du contenu clinique n'interfère plus avec celle de l'artefact qui est mieux estimé. Appliquée à la correction de l'artefact en ligne sur des radiographies, nous montrons sa supériorité sur les méthodes antérieures [1, 2, 6].

2 Méthodologie proposée

2.1 Modèle de représentation de l'image

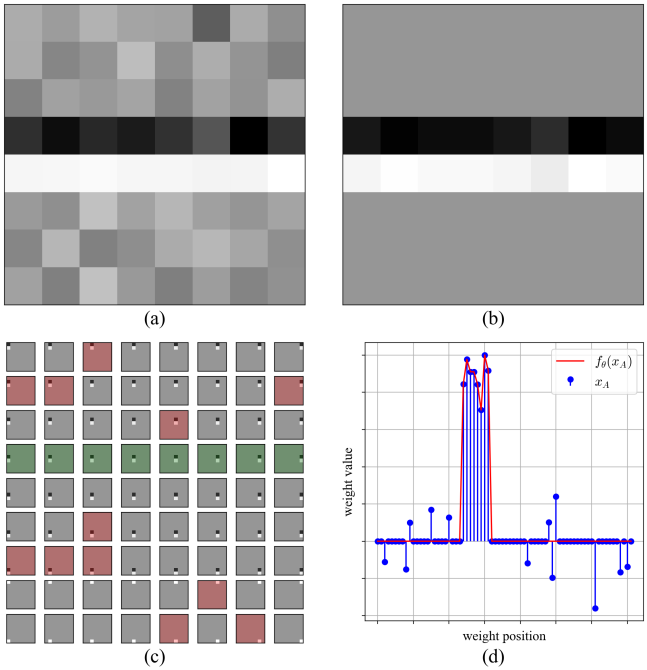
Nous proposons un modèle d'observation linéaire additif utilisé par [1, 2] car il est bien adapté pour représenter les images radiographiques médicales en présence d'artefacts :

$$\mathcal{Y} = \mathcal{C} + \mathcal{A} \tag{1}$$



(a) (b)

FIGURE 1 : (a) : Image avec artefact en ligne (zoom dans le coin inférieur droit), (b) : Atomes du dictionnaire clinique.



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

FIGURE 2 : (a) : Patche (8×8) avec du bruit et un artefact "en ligne", (b) : Estimation de l'artefact $\tilde{a}_i$ (c) Dictionnaire d'artefact associé, en vert les bons atomes utilisés en résolvant l'équation (3) et en rouge ceux utilisées dû au bruit (d) : Encodage avant et après l'application de f_θ, les erreurs de l'encodage ont disparu.

Avec $\mathcal{Y} \in \mathbb{R}^n$ l'image observée, $\mathcal{C} \in \mathbb{R}^n$ le contenu clinique, $\mathcal{A} \in \mathbb{R}^n$ l'artefact.

L'estimation de l'artefact étant plus facile que celle du contenu clinique, nous choisissons de restaurer ce dernier $\tilde{\mathcal{C}}$ par soustraction de l'artefact estimé $\hat{\mathcal{A}}$ de l'image d'observée $\mathcal{Y}$:

$$\tilde{\mathcal{C}} = \mathcal{Y} - \hat{\mathcal{A}} \tag{2}$$

Introduire un modèle uniquement pour l'artefact $\mathcal{A}$ conduit à des erreurs d'estimation en raison des caractéristiques partagées avec le contenu clinique. Par conséquent, nous proposons un cadre d'estimation conjointe pour $\mathcal{C}$ et $\mathcal{A}$ où nous choisissons 2 dictionnaires parcimonieux, un pour le contenu clinique et un pour les artefacts, comme [2]. Le modèle devient alors :

$$\mathcal{Y} = D_\mathcal{C} x_\mathcal{C} + D_\mathcal{A} x_\mathcal{A} + \epsilon = Dx + \epsilon \tag{3}$$

Avec $D_\mathcal{C} \in \mathbb{R}^{n \times k_C}$ la matrice de caractéristiques du contenu clinique (dictionnaire, dont les éléments stockés comme vecteurs sont appelés atomes), $D_\mathcal{A} \in \mathbb{R}^{n \times k_A}$ le dictionnaire de l'artefact. $x_\mathcal{C} \in \mathbb{R}^{k_C}$ et $x_\mathcal{A} \in \mathbb{R}^{k_A}$ sont respectivement les vecteurs de poids inconnus qui encodent la représentation de $\mathcal{C}$ et $\mathcal{A}$ sur leur dictionnaire associé $D_\mathcal{C}$ et $D_\mathcal{A}$. $\epsilon \in \mathbb{R}^n$ représente les erreurs des modèles.

Sous sa forme réduite, $D = \left[D_\mathcal{C} \mid D_\mathcal{A}\right] \in \mathbb{R}^{n \times (k_C + k_A)}$ est la concaténation de $D_\mathcal{C}$ et $D_\mathcal{A}$ et $x = \left[x_\mathcal{C} \mid x_\mathcal{A}\right]^T \in \mathbb{R}^{k_C + k_A}$ le vecteur de poids encodant y sur les deux dictionnaires.

L'estimation jointe de $x_\mathcal{C}$ et $x_\mathcal{A}$ (équation (3)) est cruciale pour préserver la reconstruction de l'artefact ($\hat{\mathcal{A}} = D_\mathcal{A} x_\mathcal{A}$) et éviter que celle-ci ne capture du contenu clinique. L'estimation est résolue à l'aide d'un algorithme classique de représentation parcimonieuse en pseudo-norme ℓ_0 [3].

Apprendre des dictionnaires à partir d'images radiologiques de haute résolution et résoudre l'équation (3) demandent des ressources de calcul importantes. Pour résoudre ce problème, une approche par patch est utilisée. L'image $\mathcal{Y}$ est décomposée en P patches $y_i \in \mathbb{R}^{n_p}$ en utilisant une transformation linéaire $R_i \in \mathbb{R}^{n_p \times n}$ avec ou sans recouvrement, stockés comme vecteurs de $Y \in \mathbb{R}^{n_p \times P}$.

Le modèle de l'équation (3) est résolu pour chaque patch. La correction de l'artefact est ensuite effectuée pour chaque patch (algo.1)

2.2 Dictionnaire d'artefact et l'*a priori* associé

Pour s'adapter aux variations de niveau de l'artefact "en ligne", nous proposons le dictionnaire D_A Fig. 2.b (illustré pour des patches de 8×8) dont chaque élément est la fonction mère des ondelettes d'Haar verticale de la largeur d'une colonne. Elles représentent toutes les positions 2D possibles de cette fonction et elles sont groupées par ligne.

Avec cette structure, en l'absence de bruit, le vecteur x_{A_i}, la représentation d'un artefact sur un patch i est une fonction porte idéale. Sur l'exemple figure 2 avec des patches 8×8, si le patch était sans bruit, les éléments du vecteur x_{A_i} numéro 25 à 32 seraient à 1, toutes les autres seraient à 0, indiquant que les 8 éléments en vert figure 2.c du dictionnaire sont les seuls utilisés et estimant parfaitement l'artefact.

En présence de bruit, les 8 éléments en rouge (figure 2.c) sont eux aussi utilisés, et les éléments en bleu (figure 2.d) du vecteur x_{A_i} numéro 25 à 32 ont des valeurs proches de 1, toutes les autres ont des valeurs nulles ou faibles.

Une fonction non linéaire f_θ sur l'encodage de l'artefact x_{A_i} filtre le bruit, détecte la représentation de l'artefact sous forme de porte (en rouge Fig. 2.d.) et corrige l'estimation de l'artefact $\tilde{a}_i = D_A f_\theta(x_{A_i})$

Algorithme 1 : SPLENDID

Entrées : $\mathcal{Y}, D = \left[D_C \mid D_A\right], L_0$
Sorties : $\mathcal{C}$
$\tilde{\mathcal{A}} \leftarrow \vec{0}$
pour $i = 1 \rightarrow P$ **faire**
  $y_i \leftarrow R_i \mathcal{Y}$
  $\hat{x} = \arg\min_x \|Dx - y_i\|_2^2$ s.t $\|x\|_0 \leq L_0$
  $\tilde{a}_i = D_A f_\theta(\hat{x}_{A_i})$
  $\tilde{\mathcal{A}} \leftarrow \tilde{\mathcal{A}} + R_i^{-1} \tilde{a}_i$
fin
$\tilde{\mathcal{C}} = \mathcal{Y} - \tilde{\mathcal{A}}$

2.3 Apprentissage du dictionnaire de contenu clinique

Contrairement au modèle de l'artefact, la variété du contenu clinique incite à apprendre le dictionnaire clinique D_C avec des méthodes d'apprentissage comme Sahoo et al. [5].

Cependant, lors de l'apprentissage, des caractéristiques communes avec l'artefact peuvent se retrouver aussi dans le dictionnaire de contenu clinique. Dans ce cas, la résolution de l'équation (3) provoque des erreurs d'encodage de l'artefact en le partageant avec le contenu clinique et la correction de l'artefact n'est pas complète.

Ramirez et al. [4] proposent une pénalité d'incohérence pendant l'apprentissage des dictionnaires dans le cadre de classification de contenus cliniques. Nous adaptons ce type de pénalité à notre problème : les atomes cliniques appris seront faiblement corrélés avec ceux du dictionnaire d'artefact.

Soit $D_C \in \mathbb{R}^{n_p \times k_c}$ le dictionnaire clinique à apprendre sur une base de données stockée dans $Y_l \in \mathbb{R}^{n_p \times S}$. Chaque y_i est un échantillon de la base de données stocké sous forme de vecteur dans Y_l, et S est le nombre d'échantillons. Soit $D_A \in \mathbb{R}^{n_p \times k_a}$ le dictionnaire d'artefact conçu précédemment.

En fixant $X_C^{(t)}$ la solution parcimonieuse minimisant l'erreur entre Y_l et $D_C X_C$ à l'itération t, avec un schéma classique de minimisation alternée pour l'apprentissage de dictionnaire, nous écrivons :

$$D_C^{(t+1)} \triangleq \underset{D_C \in \mathcal{S}}{\operatorname{argmin}} \|Y_l - D_C X_C^{(t)}\|_F^2 + \lambda \|D_A^T D_C\|_F^2 \quad (4)$$

Avec $D_C^{(t+1)}$ le dictionnaire mis à jour, $\| \bullet \|_F$ désignant la norme de Frobenius, $\mathcal{S}$ un ensemble de contraintes et $\lambda \in \mathbb{R}^+$ un scalaire, équilibrant le poids de la pénalité dans l'équation (4). Nous utilisons une stratégie de descente de coordonnées pour l'estimation de $D_C^{(t+1)}$ comme dans [5] où un atome $d_{C_j}^{(t+1)}$ est calculé à l'aide d'autres atomes $d_{C_i}^{(t)}, \forall i \neq j$. Nous trouvons

$$d_{C_j}^{(t+1)} = \left(\|X_{C_j}^{(t)}\|^2 I_{n_p} + \lambda D_A D_A^T \right)^{-1} E_j^{(t)} \left(X_j^{(t)} \right)^T \quad (5)$$

où $E_j^{(t)} = Y_l - \sum_{i \neq j} d_{C_i}^{(t)} X_i^{(t)}$ et $I_{n_p} \in \mathbb{R}^{n_p \times n_p}$ est une matrice d'identité.

3 Application

3.1 Aperçu

Nous évaluons notre méthode sur un artefact "en ligne" simulé avec une amplitude variable, correspondant à des artefacts réels. Il est ajouté à une image de pied clinique extraite d'une base de 952 images de pieds et de mains, fournie par le Centre Hospitalier Universitaire d'Orléans[1]. L'artefact présente des variations d'intensité qui sont difficiles à prendre en compte.

Nous comparons, d'un point de vue qualitatif et quantitatif, notre proposition aux trois méthodologies [1, 2, 6] de suppression d'artefacts décrites dans l'introduction après une adaptation à l'artefact "en ligne".

3.2 Experimentations

Le dictionnaire d'artefact conçu manuellement (Fig. 2.b pour une taille 16×16) comporte 272 atomes pour des patches de taille 16×16. Le nombre L_0 d'atomes utilisés pour représenter parfaitement un artefact est la largeur du patch.

Le dictionnaire clinique de 256 atomes (Fig. 1.b) est appris à partir de 0.1×10^6 patches de taille (16×16), extraits aléatoirement de la base de pieds et mains, avec une pénalité d'incohérence (relative à la norme de la base d'apprentissage) sur le dictionnaire d'artefact λ de valeur 0.71.

Nous avons comparé les différentes méthodes avec des artefacts "en ligne" simulés à des niveaux, formes et positions aléatoires sur les 952 images cliniques de pieds et de mains.

L'algorithme OMP [3] est utilisé pour résoudre l'équation (3) de codage parcimonieux pour chaque patch extrait. La parcimonie choisie est $L_0 = 32$ au total, 16 atomes au moins pour encoder l'artefact sur D_A, 16 atomes supplémentaires pour l'encodage du contenu clinique sur D_C .

Pour une étude quantitative, nous utilisons le PSNR entre l'image corrigée (contenu clinique $\tilde{\mathcal{C}}$ et la vérité terrain $\mathcal{C}$), ainsi qu'une mesure de qualité (SafeGuarding ou SFM) définie par (avec $\| \bullet \|_0$ désignant la pseudo-norme ℓ_0) :

$$\text{SFM} = \frac{\|\tilde{\mathcal{C}} - \mathcal{C}\|_0}{n} \quad (6)$$

Cette mesure indique le pourcentage de pixels erronés. Il est en effet important de ne pas modifier à tort le contenu clinique, en particulier en l'absence d'artefact.

3.3 Résultats et discussions

La figure 3 présente la correction par les 4 méthodes sur une des images de la base. La première ligne illustre la correction finale. La deuxième ligne montre l'erreur du contenu clinique obtenu par rapport à la vérité terrain.

Notre méthode offre une meilleure extraction de l'artefact (Fig. 3.b & Fig. 3.g) ainsi qu'une bonne correction, sans modifier le contenu clinique existant.

La méthode [1] ne s'adapte pas aux variations de l'intensité (Fig. 3.c & Fig. 3.h) et laisse une partie de l'artefact. Elle modifie aussi des travées osseuses horizontales du contenu clinique réel.

La méthode [2] sous-estime l'artefact en raison de la forte corrélation entre les dictionnaires clinique et d'artefact (Fig. 3.d & Fig. 3.i). La majeure partie de l'artefact a été capturée par le dictionnaire clinique qui partage certaines caractéristiques comme les lignes horizontales avec l'artefact.

La méthode [6] sur-estime l'artefact (Fig. 3.e) car elle manque d'*a priori* sur le contenu clinique, entraînant la suppression d'informations cliniques (Fig. 3.j). Elle ne s'adapte pas aux variations de l'artefact.

D'un point de vue quantitatif, l'histogramme décumulatif de l'erreur (Fig. 3.e) montre que notre méthode produit moins

[1] La cohorte de patients a été constituée lors d'un essai clinique dans le cadre du « Programme Hospitalier de Recherches Cliniques » (PHRC, Programme de Recherche IDRCB 2008-A01422-53) financé par le Ministère français de la Santé. L'objectif initial de cette recherche était de comparer les performances diagnostiques et métrologiques de la radiographie conventionnelle par rapport aux radiographies numériques haute résolution pour détecter le rétrécissement de l'espace articulaire et les érosions dans la polyarthrite rhumatoïde (PR). Tous les patients remplissaient les critères révisés de 1987 du Collège américain de rhumatologie pour la PR. L'autorisation de mener l'étude a été obtenue auprès du comité d'éthique (CPP Tours-2008-A01422-53). Tous les patients ont donné leur consentement écrit.

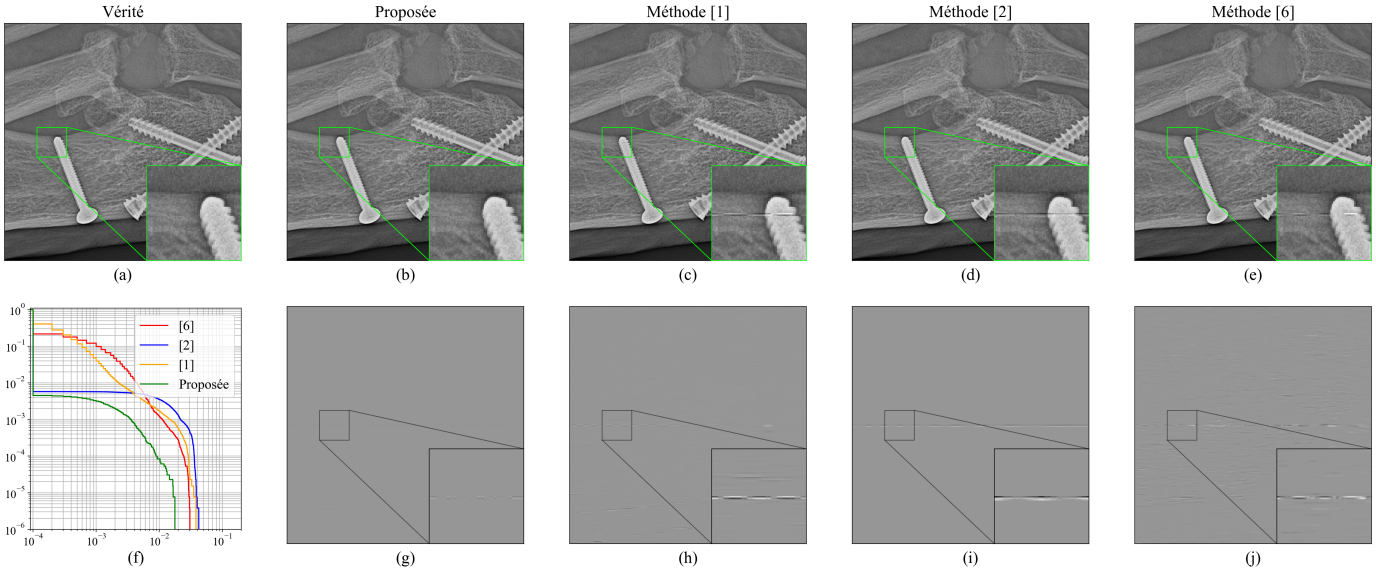


FIGURE 3 : Comparaison de chaque méthode (organisée par colonnes) appliquée à la suppression d'un artefact "en ligne" sur une image radiographique d'un pied. La première ligne compare la correction obtenue (zoom en bas à droite) à la vérité terrain (a), la deuxième ligne montre l'erreur du contenu clinique (zoom en bas à droite) et (e) l'histogramme décumulatif normalisé de l'erreur.

d'erreurs, avec un pourcentage plus faible par rapport aux autres méthodes.

TABLE 1 : Mesures de PSNR et SFM sur les 952 images

	SPLENDID	[1]	[2]	[6]
PSNR (dB) ↗	**67.3**	60.1	61.4	57.5
SFM (%) ↘	**0.5**	26.1	**0.5**	23.93

Les métriques moyennes du PSNR et de SFM sur l'ensemble du jeu de données (Tableau 1) montrent que notre méthode surpasse les autres méthodes sur ces 2 mesures. La méthode [2] présente une SFM extrêmement basse, ce qui est essentiel en imagerie clinique. Au contraire, la méthode [6], très rapide à calculer (environ 0,1 seconde) par rapport à notre méthode (environ 1 seconde), obtient de mauvais résultats en termes de SFM, tout comme la méthode [1].

4 Conclusions

Nous avons présenté une méthode de suppression d'artefacts de ligne sur des images radiographiques médicales basée sur la modélisation conjointe du contenu clinique et de l'artefact avec une incohérence structurée entre les 2 modélisations. Elle donne de meilleurs résultats par rapports aux approches traditionnelles.

Les travaux futurs se concentreront sur l'accélération de la résolution du modèle conjoint, l'apprentissage sur des ensembles de données plus grands et l'extension à d'autres artefacts, ainsi que sur l'étude optimale du paramètre λ équilibrant la pénalité d'incohérence.

Références

[1] Yang CHEN, Luyao SHI, Qianjing FENG, Jian YANG, Huazhong SHU, Limin LUO, Jean-Louis COATRIEUX et Wufan CHEN : Artifact Suppressed Dictionary Learning for Low-Dose CT Image Processing. *IEEE Transactions on Medical Imaging*, 33(12):2271–2292, décembre 2014.

[2] Florian COTTE, Michel DESVIGNES, Hacheme AYASSO et Jean-Michel VIGNOLLE : A sparse dictionary representation approach for anti-scattering grid artifact removal in X-ray images. *Biomedical Signal Processing and Control*, 86:105247, septembre 2023.

[3] Y.C. PATI, R. REZAIIFAR et P.S. KRISHNAPRASAD : Orthogonal matching pursuit : recursive function approximation with applications to wavelet decomposition. *In Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, pages 40–44 vol.1, novembre 1993.

[4] Ignacio RAMIREZ, Pablo SPRECHMANN et Guillermo SAPIRO : Classification and clustering via dictionary learning with structured incoherence and shared features. *In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 3501–3508, juin 2010.

[5] Sujit Kumar SAHOO et Anamitra MAKUR : Dictionary Training for Sparse Representation as Generalization of K-Means Clustering. *IEEE Signal Processing Letters*, 20(6):587–590, juin 2013.

[6] Ryoji SASADA, Masahiko YAMADA, Shoji HARA, Hideya TAKEO et Kazuo SHIMURA : Stationary grid pattern removal using 2D technique for moire-free radiographic image display. page 688, San Diego, CA, mai 2003.

[7] Alisa I. WALZ-FLANNIGAN, Kimberly J. BROSSOIT, Dayne J. MAGNUSON et Beth A. SCHUELER : Pictorial Review of Digital Radiography Artifacts. *RadioGraphics*, 38(3):833–846, mai 2018.