

# Nouvel algorithme d’inversion robuste pour le RADAR GPR

Matthieu GALLET<sup>1</sup>, Ammar MIAN<sup>1</sup>, Guillaume GINOLHAC<sup>1</sup>, Nickolas STELZENMULLER<sup>2</sup>

<sup>1</sup>Université Savoie Mont-Blanc, LISTIC, Annecy, France

<sup>2</sup>Géolithe, Grenoble, France

matthieu.gallet, ammar.mian, guillaume.ginolhac@univ-smb.fr, nickolas.stelzenmuller@geolithe.com

**Résumé** – Dans ce papier, nous proposons une nouvelle méthode d’inversion dans le but d’améliorer la détection d’objets enfouis. Pour cette détection, on utilise un RADAR GPR (Ground Penetrating Radar) qui émet une onde qui va traverser le sous sol et se réfléchir sur d’éventuels objets enterrés. A cause du mouvement du RADAR la réponse de ces objets a une forme d’hyperbole. L’approche proposée dans ce papier est basé sur un modèle convolutif avec dictionnaire et une matrice rang faible. Le problème d’optimisation utilise une norme d’Huber à la place de la norme classique pour avoir une meilleure robustesse au bruit qui est très présent dans une image GPR. Nous testons notre approche sur un jeu de données réelles fourni par la société Géolithe et montrons l’apport de la norme de Huber par rapport à la norme classique.

**Abstract** – In this paper, we propose a new inversion method to improve the detection of buried objects. For this detection, we use a GPR RADAR which transmits a wave that will cross the ground and reflect on possible buried objects. Because of the motion of the RADAR the response of these objects has a hyperbolic shape. The approach proposed in this paper is based on a convolutional model with dictionary and a low rank matrix. The optimization problem uses a Huber norm instead of the classical norm to have a better robustness to noise which is very present in a GPR image. We test our approach on a real dataset provided by Geolithe and show the interest of the Huber norm compared to the classical one.

## 1 Introduction

Le RADAR GPR (Ground Penetrating Radar) a pour but de détecter des objets enfouis dans le sol en utilisant une onde électromagnétique qui se propage dans le sous sol [1]. A cause du déplacement du RADAR, la réponse des objets a une forme d’hyperbole. Il est alors intéressant pour les utilisateurs de classifier ces objets en utilisant cette forme [2–4]. Pour utiliser pleinement les algorithmes de classification, il est nécessaire d’avoir accès à des données labellisées. Néanmoins cette étape est souvent délicate pour les experts car le Rapport Signal à Bruit est très faible dans une image GPR en particulier à cause de la présence des contributions des différentes couches du sous sol créant un signal ayant une grande amplitude, mais aussi de par la présence d’artefacts d’acquisition. Il est donc primordiale d’améliorer la qualité de ces images en les débruitant et en supprimant les réponses des différentes couches horizontales.

Pour faire cette étape de débruitage, il existe plusieurs approches comme par exemple en utilisant des techniques de traitement de signal [5–9] ou alors des approches d’inversion provenant par exemple du traitement d’images [10–12]. Plus particulièrement, on s’intéresse ici au travail proposé dans [13] qui considère le problème bien connu de R-PCA [14] avec dictionnaire [15] pour améliorer la détection de tuyaux enfouis. Dans ce travail, la réponse des objets provient d’un modèle convolutif utilisant un dic-

tionnaire construit à partir d’hyperboles génériques (par manque de place, nous n’aborderons pas la construction de ce dictionnaire dans ce papier). La réponse des différentes couches du sous-sol, aussi appelée clutter, est volontairement séparée du signal d’intérêt reconstruit, en étant supposée contenue dans une matrice rang faible. Les résultats sont plutôt prometteurs mais l’image obtenue est encore assez bruitée et sensible aux données aberrantes (comme un dés-alignement causé par une erreur de trajectoire du RADAR). Nous proposons alors de remplacer la norme  $L_2$  dans le terme de l’attache aux données par la norme d’Huber [16]. Nous utilisons ensuite l’algorithme Alternating Direction Method of Multipliers (ADMM) pour résoudre le problème d’optimisation. Finalement, nous testons notre algorithme et le comparons avec l’algorithme de [13] sur un jeu de données réelles fourni par la société Géolithe.

## 2 Modèle des données

Comme le montre la figure 1, un radargramme B-Scan peut être décomposé comme suit :

$$y(x, t) = s(x, t) + cl(x, t) + n(x, t) \quad (1)$$

avec  $n(x, t)$  le bruit,  $cl(x, t)$  les réponses du sol et des différentes couches et  $s(x, t)$  les réponses des différents objets d’intérêt. Cette dernière contribution se modélise

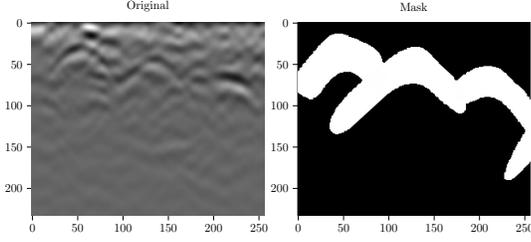


FIG. 1: Radargramme B-Scan original (à gauche) et image correspondante des masques des hyperboles créés par les objets recherchés.

comme une somme de différentes hyperboles ce qui donne [13]:

$$s(x, t) = \sum_{k=1}^K c_k(x, t) * h_k(x, t) \quad (2)$$

où  $h_k(x, t)$  est l'image d'une hyperbole et  $c_k(x, t)$  le coefficient correspondant. Dans cet article, nous avons considéré un modèle convolutif qui permet de simplifier grandement la construction du dictionnaire. Comme le signal  $y(x, t)$  est échantillonné dans les deux dimensions, (1) se réécrit en forme matricielle :

$$\mathbf{Y} = \sum_{k=1}^K \mathbf{C}_k \otimes \mathbf{H}_k + \mathbf{L} + \mathbf{N}, \quad (3)$$

où les dimensions des matrices sont égales à celles de l'image B-scan, c'est-à-dire  $N_x \times N_t$  et  $\otimes$  est le produit convolutif matriciel. La matrice  $\mathbf{L}$  contient les réponses du sol et des couches horizontales et on suppose donc qu'elle est de faible rang. La matrice  $\mathbf{N}$  est principalement composée des contributions des différents bruits. Les matrices  $\{\mathbf{H}_k\}_{1,K}$  sont construites à partir de toutes les hyperboles possibles ce qui constitue notre dictionnaire. Par manque de place, nous ne décrivons pas comment ce dictionnaire est construit mais certains détails sont disponibles dans [13]. Les matrices  $\{\mathbf{C}_k\}_{1,K}$  contiennent les cartes des coefficients permettant de localiser les objets d'intérêt.

### 3 Inversion Robuste basée sur la norme de Huber

Pour améliorer la qualité de l'image originale, nous devons donc estimer les matrices  $\mathbf{L}$  et  $\{\mathbf{C}_k\}_{1,K}$  à partir de  $\mathbf{Y}$ . Comme le nombre d'objets d'intérêt dans une image B-scan est limité, nous ajoutons une contrainte de parcimonie sur les matrices  $\{\mathbf{C}_k\}_{1,K}$  et comme  $\mathbf{L}$  est supposé être de faible rang, nous ajoutons également une contrainte de rang dans le problème d'optimisation en utilisant une norme nucléaire. Ce problème d'optimisation est celui de [14] sans dictionnaire et est équivalent à celui présenté dans [15]. En se basant sur le modèle (3), un premier problème d'optimisation a été proposé dans [13]

sous la forme suivante:

$$\operatorname{argmin}_{\{\mathbf{C}_k\}, \mathbf{L}} \|\mathbf{Y} - \sum_{k=1}^K \mathbf{C}_k \otimes \mathbf{H}_k - \mathbf{L}\|_2^2 + \lambda \sum_{k=1}^K \|\mathbf{C}_k\|_1 + \kappa \|\mathbf{L}\|_*, \quad (4)$$

où  $\lambda$  agit sur la parcimonie du résultat et  $\kappa$  promeut le côté rang faible de la matrice  $\mathbf{L}$ . Ce problème a été résolu en introduisant des variables auxiliaires pour rentrer dans le cadre de l'ADMM.

En inspectant la fonction à minimiser (4), on note que le terme d'attachement aux données  $\|\mathbf{Y} - \sum_{k=1}^K \mathbf{C}_k \otimes \mathbf{H}_k - \mathbf{L}\|_2^2$  utilise la norme L2. Il est bien connu que la norme L2 est sensible aux données aberrantes et que si c'est le cas, la performance de la régression peut être fortement diminuée. Pour contrer cette problématique, il est alors intéressant de robustifier ce terme de la fonction d'optimisation. Une méthode courante dans les méthodes robustes [16] consiste alors à utiliser la fonction de Huber :

$$\mathcal{H}_\delta(x) = \begin{cases} |x|^2 & , |x| \leq \delta \\ 2\delta|x| - \delta^2 & , |x| > \delta \end{cases} \quad (5)$$

qui permet de définir la norme de Huber d'un vecteur  $\mathbf{x} \in \mathbb{C}^p$ :

$$\|\mathbf{x}\|_{\mathcal{H}}^2 = \sum_{i=1}^p \mathcal{H}_\delta(x_i). \quad (6)$$

On se retrouve ainsi avec le nouveau problème de minimisation suivant:

$$\operatorname{argmin}_{\{\mathbf{C}_k\}, \mathbf{L}} \|\mathbf{Y} - \sum_{k=1}^K \mathbf{C}_k \otimes \mathbf{H}_k - \mathbf{L}\|_{\mathcal{H}}^2 + \lambda \sum_{k=1}^K \|\mathbf{C}_k\|_1 + \kappa \|\mathbf{L}\|_*. \quad (7)$$

Pour résoudre ce problème, nous nous plaçons également dans le cadre de l'ADMM et réécrivons en introduisant des variables auxiliaires  $\mathbf{M}$  et  $\{\mathbf{S}_k\}$  le problème sous la forme suivante:

$$\begin{cases} \text{minimisation} & \|\sum_{k=1}^K \mathbf{H}_k \otimes \mathbf{C}_k + \mathbf{L} - \mathbf{Y}\|_{\mathcal{H}}^2 \\ & + \kappa \|\mathbf{M}\|_* + \lambda \sum_{k=1}^K \|\mathbf{S}_k\|_1 \\ \text{sous la contrainte} & \mathbf{M} = \mathbf{L} \\ & \forall k, \mathbf{S}_k = \mathbf{C}_k \end{cases} \quad (8)$$

Dans ce problèmes les variables  $\{\mathbf{S}_k\}$  permettent de découpler le terme d'attache aux données de la contrainte de parcimonie. De la même façon, la variable auxiliaire  $\mathbf{M}$  découple l'attache aux données à la minimisation de la norme nucléaire. Cette approche permet alors une résolution en sous-problèmes plus simples à traiter individuellement. Le Lagrangien augmenté du problème est ainsi le suivant:

$$\begin{aligned} \mathcal{L} = & \|\mathbf{L} + \sum_{k=1}^K \mathbf{H}_k \otimes \mathbf{C}_k - \mathbf{Y} + \mathbf{U}_L\|_{\mathcal{H}}^2 \\ & + \lambda \sum_{k=1}^K \|\mathbf{S}_k\|_1 + \kappa \|\mathbf{M}\|_* \\ & + \frac{\rho_S}{2} \|\sum_{k=1}^K \mathbf{C}_k - \mathbf{S}_k + \mathbf{U}_{k,S}\|_2^2 \\ & + \frac{\rho_L}{2} \|\mathbf{M} - \mathbf{L} + \mathbf{U}_L\|_2^2 \\ & - \frac{\rho_L}{2} \|\mathbf{U}_L\|_2^2 - \frac{\rho_S}{2} \|\sum_{k=1}^K \mathbf{U}_{k,S}\|_2^2, \end{aligned} \quad (9)$$

où  $\mathbf{U}_L$ ,  $\mathbf{U}_{k,S}$  sont les variables duales de  $\mathbf{L}$  et  $\mathbf{S}_k$  et  $\rho_L$ ,  $\rho_S$  sont les pénalités associées aux contraintes. Dans la suite du développement, nous avons fixé  $\kappa = 1$ . Nous présentons dans la suite les différentes étapes de l'ADMM pour chaque étape d'estimation des paramètres.

### Mise à jour $\{\mathbf{C}_k\}$

Pour éviter le calcul de la convolution, nous transposons ce problème d'optimisation dans le domaine fréquentiel avec une simple transformée de Fourier [15]. De plus, nous vectorisons toutes les matrices, ce qui conduit aux nouveaux paramètres suivants, comme par exemple  $\mathbf{y}^f = \text{vec}(FFT(\mathbf{Y}))$ ,  $\mathbf{h}_k^f = \text{vec}(FFT(\mathbf{H}_k))$ , ...<sup>1</sup>. On définit aussi la matrice  $\mathbf{H}^f = (\text{diag}(\mathbf{h}_1^f) \dots \text{diag}(\mathbf{h}_K^f))$ , les vecteurs  $\mathbf{c}^f = (\mathbf{c}_1^f \dots \mathbf{c}_K^f)^T$ ,  $\mathbf{z}_k^f = \mathbf{s}_k^f + \mathbf{u}_{k,S}^f$ ,  $\mathbf{z}^f = (\mathbf{z}_1^f \dots \mathbf{z}_K^f)^T$  et  $\mathbf{x}^f = \mathbf{y}^f - \mathbf{l}^f - \mathbf{u}_L$ . Avec ces nouvelles variables, le problème d'optimisation se réduit à :

$$\mathbf{c}^{f(t+1)} = \arg \min_{\mathbf{c}^f} \frac{\rho_L}{2} \|\mathbf{H}^f \mathbf{c}^f - \mathbf{x}^f\|_{\mathcal{H}}^2 + \frac{\rho_S}{2} \|\mathbf{c}^f - \mathbf{z}^f\|_2^2 \quad (10)$$

Ne pouvant pas obtenir de solutions analytiques, nous devons calculer l'estimer par une méthode numérique comme une descente de gradient. Pour le calcul du gradient du second terme, nous obtenons le même résultat que dans l'inversion classique [13], tandis que pour le gradient du premier terme, la dérivation conduit à  $\mathbf{H}^{fH} \mathbf{a}$  où chaque élément  $\mathbf{a}_i = \Psi_\delta [(\mathbf{H}^f \mathbf{c}^f - \mathbf{x}^f)_i]^2$ . Finalement, le gradient de (10) par rapport à  $\mathbf{c}^f$  est :

$$\Delta_{\mathbf{c}^f} \mathcal{L} = \rho_L \mathbf{h}^{fH} \odot (\mathbf{1}_K \otimes \mathbf{a}) + \rho_S (\mathbf{c}^f - \mathbf{z}^f), \quad (11)$$

où  $\mathbf{h}^f = \text{vec}(\mathbf{H}^f)$ . L'estimation de  $\mathbf{c}^{f(t+1)}$  s'effectue donc par une descente de gradient :

$$\mathbf{c}_{j+1}^{f(t+1)} = \mathbf{c}_j^{f(t+1)} - \frac{\alpha_c}{t+1} \Delta_{\mathbf{c}^f} \mathcal{L}. \quad (12)$$

### Mise à jour $\mathbf{L}$

Dans ce cas, le problème d'optimisation devient :

$$\mathbf{L}^{(t+1)} = \arg \min_{\mathbf{L}} \|\mathbf{L} + \sum_{k=1}^K \mathbf{H}_k \otimes \mathbf{C}_k - \mathbf{Y} + \mathbf{L}\|_{\mathcal{H}}^2 + \frac{\rho_L}{2} \|\mathbf{M} - \mathbf{L} + \mathbf{U}_L\|_2^2 \quad (13)$$

Le problème est simplifié en utilisant les nouvelles variables  $\mathbf{A} = \sum_{k=1}^K \mathbf{H}_k \otimes \mathbf{C}_k - \mathbf{Y}$  et  $\mathbf{V} = \mathbf{M} + \mathbf{U}_L$  :

$$\mathbf{L}^{(t+1)} = \arg \min_{\mathbf{L}} \|\mathbf{A} + \mathbf{L}\|_{\mathcal{H}}^2 + \frac{\rho_L}{2} \|\mathbf{V} - \mathbf{L}\|_2^2 \quad (14)$$

A partir de ce problème d'optimisation, nous pouvons utiliser l'opérateur proximale de Huber [17] :

$$\text{prox}_{\gamma, \mathcal{H}_\delta}(x) = \begin{cases} \frac{x}{\gamma+1} & \text{if } |x| < \delta(\gamma+1) \\ x - \delta \gamma \text{sign}(x) & \text{sinon} \end{cases} \quad (15)$$

<sup>1</sup>L'opérateur  $\text{vec}(\mathbf{M})$  transforme une matrice  $N_x \times N_y$  en un vecteur  $N = N_x N_y$

<sup>2</sup> $\Psi_\delta(x)$  est la dérivé de (5) et on a  $\Psi_\delta(x) = x$ , si  $|x| \leq \delta$  et  $\delta \text{sign}(x)$  sinon ( $\text{sign}(x)$  est égal à 1 lorsque  $x = 0$  et égal à 0 sinon).

ce qui nous donne comme mise à jour pour  $\mathbf{L}$  :

$$\mathbf{L}^{(t+1)} = \text{prox}_{1/\rho_L, \mathcal{H}_\delta}(\mathbf{V} + \mathbf{A}) - \mathbf{A} \quad (16)$$

### Mise à jour de $\{\mathbf{S}_k\}$ , $\mathbf{M}$ , $\{\mathbf{U}_{k,S}\}$ et $\mathbf{U}_L$

Ces étapes restent les mêmes que dans le problème inverse classique [13] et donc  $\{\mathbf{S}_k\}^{(t+1)} = \mathcal{S}_{\lambda/\rho_S}(\mathbf{C}_k - \mathbf{U}_{k,S})$  et  $\mathbf{M} = \mathcal{T}_{\kappa/\rho_L}(\mathbf{L} - \mathbf{U}_L)$  où  $\mathcal{S}_{\lambda/\rho_S}$  est l'opérateur proximal de seuillage doux et  $\mathcal{T}_{1/\rho_L}$  est l'opérateur proximal de seuillage sur les valeurs singulières [14]. A la fin de la boucle, les variables duales sont mises à jour avec  $\mathbf{U}_L^{(t+1)} = \mathbf{U}_L^{(t)} + \sum_{k=1}^K \mathbf{H}_k \otimes \mathbf{C}_k^{(t+1)} - \mathbf{Y} + \mathbf{L}^{(L+1)}$  et  $\mathbf{U}_{k,S}^{(t+1)} = \mathbf{U}_{k,S}^{(t)} + \mathbf{S}_k^{(t+1)} - \mathbf{C}_k^{(t+1)}$ .

## 4 Simulations

Le jeu de données complet fourni par la société geolithe est composé de 1000 radargrammes d'une taille moyenne de  $(N_x, N_y) = (4000, 800)$  pixels associés à un masque d'étiquettes donnant les hyperboles correspondant aux objets d'intérêt. Les cibles sont étiquetées par des experts en géophysique. Dans cette section, nous étudions l'intérêt de prendre une norme de Huber au lieu de la norme L2 dans le problème d'optimisation. L'image originale et l'image du masque sont présentées dans la Fig. 1. Pour l'algorithme ADMM, nous avons fixé  $\rho_S$  à 500 et  $\rho_L$  à 250, le seuil de Huber à 750 et le nombre d'itérations du gradient à 5 et un pas de gradient  $\alpha_c = 5e-4$ . Le nombre total d'itérations de l'ADMM est fixé à 500. Nous comparons les résultats de l'inversion robuste avec ceux de l'inversion classique de [13] pour  $\lambda = 0.4$ . Les résultats sont présentés dans la figure 2. Nous montrons également le contenu de la matrice  $\mathbf{L}$  pour chaque méthode. Nous remarquons que la reconstruction est meilleure avec l'approche robuste. En particulier, la matrice  $\mathbf{L}$  contient moins d'hyperboles lorsqu'elle est traitée avec la norme robuste. Pour avoir une conclusion plus quantitative, nous étudions la robustesse de notre algorithme au bruit. Nous considérons un bruit additif Gaussien de moyenne nulle et de variance comprise entre 0,01 et 10. Nous calculons plusieurs métriques : l'EQM, le PSNR (Peak Signal to Noise Ratio) et le SSIM (Structural Similarity Index) par rapport à la variance du bruit. Les résultats sont donnés dans la figure 3. Pour toutes les métriques, la méthode robuste donne de meilleures performances ce qui montre sa meilleure robustesse au bruit additif. Les mêmes conclusions ont été tirées lors de l'étude en présence d'un bruit multiplicatif.

## 5 Conclusion

Nous avons développé un algorithme d'inversion adapté aux images GPR en utilisant un modèle convolutif et la norme de Huber afin d'être plus robuste au bruit ou aux différents problèmes (artefacts d'acquisition et fortes

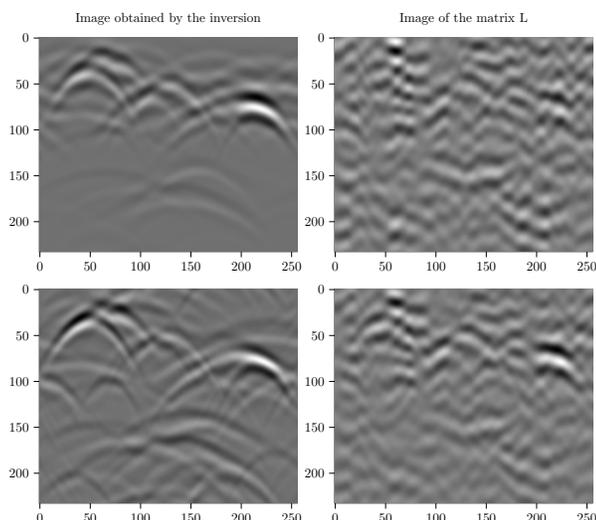


FIG. 2: Résultats de l'inversion robuste avec l'image reconstruite (en haut à gauche) et la matrice  $L$  (en haut à droite). Même chose pour l'inversion classique sur le bas de l'image. Le paramètre de sparsité est fixé à  $\lambda = 0.4$  et le nombre d'itérations de l'ADMM est de 500.

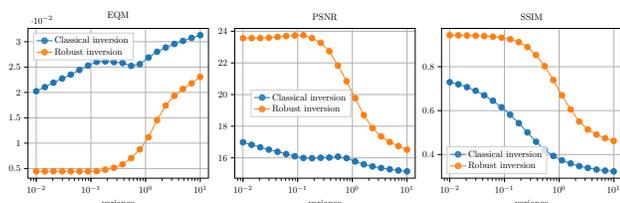


FIG. 3: EQM (gauche), PSNR (milieu) et SSIM (droite) en fonction de la variance du bruit avec l'image originale et le bruit additif.  $\lambda = 0.4$  et le nombre d'itérations est de 20.

réponses de cibles perturbatrices) souvent rencontrés dans les images GPR. Les résultats de cette approche robuste sur un jeu de données réelles ont montré le grand intérêt de ce nouvel algorithme.

## References

- [1] David J Daniels, *Ground Penetrating Radar*, IEE, 2004.
- [2] Xisto L. Travassos, Sérgio L. Avila, and Nathan Ida, "Artificial Neural Networks and Machine Learning techniques applied to Ground Penetrating Radar: A review," *Applied Computing and Informatics*, vol. 17, no. 2, pp. 296–308, Jan. 2020, Publisher: Emerald Publishing Limited.
- [3] Mostafa Elsaadouny, Jan Barowski, and Ilona Rolfes, "ConvNet Transfer Learning for GPR Images Classification," in *2020 German Microwave Conference (GeMiC)*, Mar. 2020, pp. 21–24, ISSN: 2167-8022.
- [4] Zhi-Kang Ni, Jun Pan, Cheng Shi, Shengbo Ye, Di Zhao, and Guangyou Fang, "DL-Based Clutter Removal in Migrated GPR Data for Detection of Buried Target," vol. 19, pp. 1–5, 2022, Conference Name: IEEE Geoscience and Remote Sensing Letters.
- [5] W. Ng, T. Chan, H.C. So, and K.C. Ho, "Particle filtering based approach for landmine detection using Ground Penetrating Radar," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, no. 11, pp. 3739–3755, Nov 2008.

- [6] A.M. Zoubir, I.J. Chant, C.L. Brown, B. Barkat, and C. Abeynayake, "Signal processing techniques for landmine detection using impulse ground penetrating radar," *Sensors Journal, IEEE*, vol. 2, no. 1, pp. 41–51, Feb 2002.
- [7] A. Manandhar, P.A. Torrione, L.M. Collins, and K.D. Morton, "Multiple-instance hidden markov model for gpr-based landmine detection," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 53, no. 4, pp. 1737–1745, April 2015.
- [8] Q. Hoarau, G. Ginolhac, A. M. Atto, and J. M. Nicolas, "Robust adaptive detection of buried pipes using GPR," *Signal Processing*, vol. 132, pp. 293–305, Mar. 2017.
- [9] V. Kovalenko, A.G. Yarovoy, and L.P. Lighthart, "A novel clutter suppression algorithm for landmine detection with GPR," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 11, pp. 3740–3751, Oct 2007.
- [10] Fok Hing Chi Tivive, Abdesselam Bouzerdoum, and Canicious Abeynayake, "Gpr signal classification with low-rank and convolutional sparse coding representation," in *2017 IEEE Radar Conference (RadarConf)*, 2017, pp. 1352–1356.
- [11] Wenbin Shao, Abdesselam Bouzerdoum, and Son Lam Phung, "Sparse representation of gpr traces with application to signal classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 7, pp. 3922–3930, 2013.
- [12] Fabio Giovanneschi, Kumar Vijay Mishra, Maria Antonia Gonzalez-Huici, Yonina C. Eldar, and Joachim H. G. Ender, "Dictionary learning for adaptive gpr landmine classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 10036–10055, 2019.
- [13] Guillaume Terrasse, Jean-Marie Nicolas, Emmanuel Trouvé, and Émeline Drouet, "Sparse decomposition of the GPR useful signal from hyperbola dictionary," in *2016 24th European Signal Processing Conference (EUSIPCO)*, 2016, pp. 2400–2404.
- [14] Emmanuel J. Candès, Xiaodong Li, Yi Ma, and John Wright, "Robust principal component analysis?," *J. ACM*, vol. 58, no. 3, jun 2011.
- [15] Brendt Wohlberg, "Convolutional sparse representations with gradient penalties," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 6528–6532.
- [16] P.J. Huber and E.M. Ronchetti, *Robust Statistics*, Wiley, 2009.
- [17] P. L. Combettes and J.-C. Pesquet, "Proximal thresholding algorithm for minimization over orthonormal bases," *SIAM Journal on Optimization*, vol. 18, no. 4, pp. 1351–1376, 2007.