Title: A Rank-Deficient and Sparse Penalized Optimization Model for Compressive Indoor Radar Target Localization

Journal: REV Journal on Electronics and Communications

We received comments from three reviewers. Following are the corrections and revisions in response to the comments and suggestions made by the reviewers.

Responses to Reviewer B Comments

Comment/Suggestion of Reviewer B	Response/Action
In this paper, authors consider the problem of wall clutter removal and target image reconstruction in through-wall radar imaging. To this end, the problem is formulated in terms of a joint low-rank and sparse regularized optimization problem where low-rank structure and sparseness are to take into account the effect of the wall antenna signals and target pixels. The results on synthetic and real data showed that the proposed approach outperforms several existing ones. The paper is well written and the idea is clear. The following problems are concerned:	

Comment/Suggestion of Reviewer B	Response/Action
• Intuitively why does joint solving the problem of wall clutter removal and target image reconstruction yield better result than the sequential tasks?	- Background: in indoor radar imaging, the received target signals are corrupted by strong wall returns. Thus, for target localization, the problem of wall clutter mitigation and target image formation needs to be solved. By formulating such task as a joint low-rank and sparse regularized optimization problem, we can simultaneously capture the wall interference and reconstruct the target image, instead of performing multistage processing of signal estimation, wall clutter suppression, and target image reconstruction. - Intuitively, jointly modeling the wall and target image components through a full optimization model could lead to an optimal solution, thereby yielding improved performances. In other words, by incorporating further prior knowledge into the model, we hope to improve the model performance. In comparison, independent solving such tasks sequentially through multi-stage could be affected by suboptimality and uncertainly; the performance of wall clutter mitigation and target image formation are sensitive to the estimation error arising in the signal recovery stage. - In the revised paper, the discussion on the key motivation of the joint solving wall clutter removal and target image reconstruction, instead of considering such tasks independently is now further highlighted in Paragraph 4, Introduction Section. The revised text now reads: *"Instead of performing multistage independently, the key idea of the proposed approach in this paper is to perform wall clutter mitigation and target image reconstruction in CS TWRI simultaneously through an optimization model. This optimization model is formulated by incorporating two intrinsic signal structures: (1) low-dimensional structure of wall clutter and (2) the sparsity profile of the target scene. The former structure is due to the fact that the electromagnetic reflections from the front wall received along the antenna array are highly correlated. As a result, if the wall antenna signals are arranged as columns of a matrix, this matrix is low-rank. The later attribute of the model is because target pixels occupy only a small region in the form image. In other words, the target image is sparse. Intuitively, we could perform these two important tasks even better if we represent the model more precisely and completely. By incorporating further prior knowledge into the model, we hope to improve the model performance."*

Comment/Suggestion of Reviewer B	Response/Action
• In my opinion, comparison of different methods on synthetic data is also necessary to assess performance of the proposed approach.	- Considering the suggestion of the reviewer, in the revised paper, more experiments have been conducted on synthetic data to compare the performances of the proposed approach and the different existing methods. The simulation results are presented in Subsection IV-A3. The discussion now reads as follows: - "*This experiment aims to evaluate the performance of the proposed low-rank and sparse approach in comparison with other existing CS-based imaging methods. For comparison, we implement the existing direct CS and multistage CS-based approaches for wall clutter mitigation and target image reconstruction using 50% reduced dataset acquired in the same manner as in Subsection IV-A2. While the direct CS method reconstructs a target image from the raw reduced data vector $\boldsymbol{y}$, the existing multistage approaches first perform data recovery, then employ a spatial filtering [15] or a subspace projection technique [16] for wall-clutter mitigation, i.e., removing the wall component $\boldsymbol{Z}^w$, and finally reconstruct a target image by solving the l_1-norm regularized problem in (10) using the wall-clutter subtracted data.* - *Fig. 4 shows the target images formed by the different imaging methods. Without clutter removal, the direct CS method forms an image shown in Fig. 4(a). Clearly, strong wall clutter dominates the targets, making target localization very difficult. On the other hand, by using the spatial filtering for wall clutter migration, the multistage CS approach produces the target image, shown in Fig. 4(b), in which strong wall clutter has been significantly mitigated, but this image still contains a high level of false alarms. Fig. 4(c) presents the image reconstructed by the multistage CS method, but the subspace projection is used, instead of the spatial filtering, for wall clutter mitigation. The quality of the target image is enhanced slightly due to better level of clutter mitigation, i.e., the subspace projection outperforms the spatial filtering in terms of clutter suppression. The proposed rank-deficient and sparse approach yields the image depicted in Fig. 4(d), where the target pixels are further enhanced and clutter pixels are alleviated considerably.* - *The TCR values of the target images, shown in Fig. 4, formed by the different CS-based methods are computed and listed in Table II. As we expect from the visual interpretation, the proposed joint nuclear-norm and l_1-norm approach significantly enhances the image quality in terms of TCR; it yields the target image with a TCR value of 42.33 dB, the highest TCR value among those of the evaluated CS-based imaging approaches.*"

Comment/Suggestion of Reviewer B	Response/Action
• In terms of image reconstruction quality, it is clear from the figures that the proposed approach provides better result than the others. However, how can we localize the target after target image reconstruction since, for example, from Fig.5 (c) and (d), it is difficult to know what is target without ground-truth data?	- It is worth noting that the focus of this paper is to solve the problem of image reconstruction in the presence of strong wall clutter. Once the target image is obtained, target detection and localization techniques can be applied. Hence, high quality of target image leads to improved target detection. - Considering the suggestion from the reviewer, to compare the detection performances among the different techniques, we now apply a thresholding detector to the formed images shown in Fig. 5[*]. Here the threshold is determined by partitioning the input image into two classes of object and background; the threshold used for classification is found by maximizing the between-class variance. The result shown in Fig. 7(d) demonstrates that the target can be easily localized by the proposed approach, without prior knowledge of the ground-truth. - The relevant discussion has been added to the revised paper. The text now reads "*To evaluate the capability of target detection by different wall clutter mitigation and image reconstruction methods, we apply a thresholding technique to the target images formed by the different methods shown in Fig. 6. Here, the input image is partitioned into two classes: object and background. The threshold used for classification is found by maximizing the between-class variance. Fig. 7 presents the detection results after applying the same threshold value to the form images. As expected, without wall clutter mitigation, the direct CS fails to localize the target as demonstrated in Fig. 7(a). By contrast, it can be observed from Figs. 7(b) and (c) that the multistage CS approach can detect the target with a high level of false alarms. The detection result in Fig. 7(d) shows that the proposed model is able to localize the target well, with no appearance of false alarms*."

[*] Fig. 6 in the revised paper.

Responses to Reviewer C Comments

Comment/Suggestion of Reviewer C	Response/Action
In this paper, the authors proposed a rank-deficient and sparsity regularized optimization model to address two important problems of wall clutter mitigation and target image formation in compressive indoor radar imaging. The task of wall clutter suppression and target image reconstruction is formulated as a composite nuclear and l_1-penalized minimization problem and an iterative algorithm based on the proximal forward-backward splitting technique is developed, which captures wall clutter and yields an indoor target image simultaneously.	
+ How many kinds of noise may appear in Indoor Radar Target Localization? Only Gaussian white noise was concerned in this work.	- In indoor radar target imaging, there are two major types of interferences. The first interference is due to the electromagnetic reflections from the wall, known as *wall clutter*. The second type is the sensing noise, which is considered to follow the complex Gaussian in the model. While the wall clutter is the major factor that affects the imaging performances and thereby being treated in this paper, we assume the second type of noise to follow the Gaussian distribution as in several other research works [8, 9, 10, 11].

Comment/Suggestion of Reviewer C	Response/Action
+ How is the regularization parameter chosen in this work?	- Background: The proposed algorithm requires the selection of two regularization parameters γ and λ. Parameter γ controls the importance of the low-rank term and uses in the singular value thresholding to estimate the low-rank wall clutter matrix. This regularization parameter is used as that in the low-rank matrix completion approaches [23, 25], in which γ is typically set to $\gamma = 10^{-2}\Vert\mathcal{A}^*(\boldsymbol{y})\Vert_2$. Here, because Subproblem (23) is a low-rank matrix completion problem, we keep the same setting for γ and found that it is suitable for through-wall radar imaging. Similarly, λ guarantees the sparsity level of the target image $\boldsymbol{s}$, which is similar to that in the CS l_1 minimization algorithms [5, 6, 18, 19], where λ is set to $\lambda = \kappa\Vert\boldsymbol{D}^H\boldsymbol{y}\Vert_\infty$; the factor κ can be chosen as $\kappa = 10^{-2}$ for high probability of precise reconstruction. Here, this typical setting for λ is kept for the l_1 minimization in the image formation problem (29). - In the paper, the relevant text in Subsection IV-A2 has been revised for clarity. The text now reads: "*Using $\boldsymbol{y}$ an $\boldsymbol{D}$, clutter mitigation and target image reconstruction can be performed with Algorithm 1. This algorithm requires three parameters γ, λ, and tol, which need to be selected appropriately. Parameter γ controls the importance of the low-rank term and uses in SVT to estimate the low-rank wall clutter matrix, see Step 3 in Algorithm 1. Setting γ to a very large value, e.g., $\gamma = \Vert\mathcal{A}^*(\boldsymbol{y})\Vert_2$, leads to the solution of $\boldsymbol{Z}^w$ being rank 0, whereas choosing a small value, e.g., $\gamma = 10^{-4}\Vert\mathcal{A}^*(\boldsymbol{y})\Vert_2$ makes the algorithm converge very slowly. The values $10^{-4}\Vert\mathcal{A}^*(\boldsymbol{y})\Vert_2$ and $\Vert\mathcal{A}^*(\boldsymbol{y})\Vert_2$ can be regarded as the lower and upper bounds for γ, respectively. Here, in the experiments, γ was set to $\gamma = 10^{-2}\Vert\mathcal{A}^*(\boldsymbol{y})\Vert_2$. While γ controls the rank of matrix $\boldsymbol{Z}^w$, the parameter λ guarantees the sparsity level of the target image $\boldsymbol{s}$. For $\lambda \geq 10^{-2}\Vert\boldsymbol{D}^H\boldsymbol{y}\Vert_\infty$, the unique solution to Problem (15) for $\boldsymbol{s}$ is the zero vector. In the following experiments, λ was set to $\lambda = 10^{-2}\Vert\boldsymbol{D}^H\boldsymbol{y}\Vert_\infty$.*"
+ Do the authors need to estimate the level of noises?	- The proposed model does not estimate the level of noise explicitly. However, the noise level is bounded by the first term in the objective function (15), i.e., $\frac{1}{2}\Vert\boldsymbol{y} - [\mathcal{A}(\boldsymbol{Z}^w) + \boldsymbol{D}\boldsymbol{s}]\Vert_2^2$.
+ Why did the authors set up the simulation and experiment with different settings? We expect to see the same scenario in both the simulation and experiment.	- The performance of the proposed model was evaluated using both simulations and experiments. In fact, both simulations and experiments use the same setting of stepped-frequency synthetic aperture radar (SAR) with monostatic mode for data collection. The main differences are the number of antennas, frequency range, and number of behind-wall targets. The motivation is we aim to test the robustness of the proposed model under different scenarios, that may happen in various potential applications.

Responses to Reviewer D Comments

Comment/Suggestion of Reviewer D	Response/Action
The paper proposed a rank-deficient and sparse penalized optimization method for through-wall radar imaging in the presence of structured wall clutter. Experiments are conducted on both simulated and real radar data to evaluate the performance of the proposed method. The paper is well-written and the experimental results look promising.	
Major comments: • The conference [1] should be cited in the journal version itself. [1] "Van Tang, Ha, and Van-Giang Nguyen. A Rank-Deficient and Sparse Penalized Optimization Model for Compressive Indoor Radar Imaging. 2019 3rd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom). IEEE, 2019."	- In the revised paper, the reference has been cited as [17]. The relevant text in Paragraph 5, Introduction Section reads: "*The idea of joint wall clutter mitigation and image formation, and preliminary results have been presented in [17].*"
• Similarities and differences with the conference paper [1] should be elaborated in details.	- In the revised paper, the similarities and differences between the paper and the conference [1] are discussed in Paragraph 5, Introduction Section. The relevant text now reads: - "*The idea of joint wall clutter mitigation and image formation, and preliminary results have been presented in [17]. This paper extends this work in three respects: model formulation, iterative algorithm, and experimental evaluation. The problem formulation is described completely in this paper, for both full and compressive sensing operations. Furthermore, the problem formulation is discussed and compared with the two existing techniques of DS beamforming and multistage CS-based models, which highlights the advantages of the proposed model. In terms of algorithm design, this paper presents rigorous steps for solving the joint nuclear-norm and l_1-norm regularized least squares (LS) minimization problem, based on the proximal forward-backward splitting framework [18]–[20]. This generic technique, its application to TWRI, and how the proximal evaluations of the two key operators, namely singular value thresholding and soft-thresholding to overcome the challenging nonsmooth nature of the penalty terms are detailed in this paper. Algorithm analysis, its convergence, and computational complexity are discussed. Extensive simulations and experiments are conducted to evaluate the performance of the proposed model. In addition, performance comparisons with several state-of-the-art methods are described and analyzed in different CS settings.*"
• Although the paper is well organized, many sections are copied from [1]. The author should rewrite the paper to avoid self-plagiarism.	- The paper has been revised carefully to avoid the repetitions from the conference [1].

Comment/Suggestion of Reviewer D	Response/Action
• The objective function $f(\boldsymbol{Z}, \boldsymbol{s})$ (15) is not convex in both $(\boldsymbol{Z}, \boldsymbol{s})$ in general. The function $f(\boldsymbol{Z}, \boldsymbol{s})$ is only convex in term of one variable ($\boldsymbol{Z}$ or $\boldsymbol{s}$) when the other is fixed (i.e., constant). Therefore, the claim "Algorithm 1 terminates after it reaches a local optimum. This local point is also the global one because the problem is convex" is not true. Please clarify and correct it.	- In the paper, the relevant issue has been clarified and corrected. The text in Paragraph 5, Subsection III-B has been revised for correctness. The text now reads "*It is worth noting here that in the minimization, the wall clutter matrix is estimated via SVT applied to the data matrix in which the recent estimated target component has been fixed and subtracted. Likewise, the scene image is reconstructed by applying the shrinkage operator to the measurement vector in which the recent estimated wall component is fixed and segregated. Algorithm 1 terminates after it reaches a local optimum. We implement this termination condition as the change of the cost function is very small (Step 5).*"

Comment/Suggestion of Reviewer D	Response/Action
Minor comments: • Introduction: it would be nicer to give some insight in the advantages and disadvantages of the state-of-the-arts.	- The introduction section has been revised in which the key features of the state-of-the-art approaches have been discussed, and their benefits and shortcomings have been analyzed. The relevant discussion in Paragraphs 2 and 3 in Introduction Section now reads: - "*Conventionally, TWRI techniques require a complete dataset to generate an image of indoor targets using backprojection, such as delay-and-sum beamforming [1, 2, 4]. In other words, such techniques are effective for image formation only for the case in which all the antennas and frequencies are available for data acquisition. However, this data collection mode makes data acquisition prolonged and system storage ineffective. To accelerate data collection and provide high-resolution imaging, several TWRI approaches have been considered using the compressive sensing (CS) framework [5]–[7]. As CS is a powerful signal processing technique that allows compressive sampling and precise reconstruction of sparse signals, it has been applied to TWRI for image formation from far reduced measurements [8]–[10]. Using CS, the task of image formation is formulated as an l_1 penalized minimization problem, in which the l_1 regularizer is used to promote the sparseness of the target scene. It has been shown that this minimization model is suitable for the situations where strong wall clutter has been completely removed prior to image reconstruction through background subtraction. Having the access to the background scene, however, is impossible in many practice operations. In fact, the presence of wall clutter causes the l_1 -penalized approaches ineffective; they reconstruct only the pixels belonging to wall clutter that tend to dominate the target pixels, making target detection very difficult.* - *To alleviate wall interferences, the problem of target image formation in conjunction with wall clutter mitigation has been considered in several CS-based studies that consist of two major stages [11]–[14]. The first stage performs wall clutter mitigation, followed by image formation in the second stage. In the wall clutter suppression stage, a full data volume needs to be estimated from the reduced dataset before spatial filtering [15] or subspace projection [16] techniques are applied to the estimated data for wall clutter removal. The wall-clutter free data are then used in the second stage for image formation through an l_1 minimization. Due to the multistage signal processing, these CS-based approaches may be affected by suboptimality and uncertainly; the performances of wall clutter mitigation and target image formation are sensitive to the estimation error arising in the signal recovery stage.*"

Comment/Suggestion of Reviewer D	Response/Action
• The experimental results show convergence rate only, therefore the computational complexity of the proposed method should be presented.	- Considering the suggestion of the reviewer, the computational complexity of the proposed algorithm has been evaluated and presented in the revised paper. The relevant text in Paragraph 4, Subsection III-B now reads: "*It can be observed that Algorithm 1 performs gradient computation (Step 2), followed by the proximal evaluations of SVT operator for wall clutter estimation (Step 3) and shrinkage operator for target image reconstruction (Step 4). Evaluations of these two operators are the most time-consuming steps and thereby forming the computational complexity of Algorithm 1. The computational complexity of the SVT operation in Step 3 is $\mathcal{O}(MN^2)$, and the time complexity of the shrinkage operator in Step 4 is $\mathcal{O}(Q)$. Thus, the overall computation complexity of each iteration is $\mathcal{O}(MN^2+Q)$.*"
• Fig 3.b shows the value of the cost function versus the number of iterations t. Theoretically, the value of f should be zero when the iteration t go to the infinity, because of the minimization. The value, however, approximates ~ 0.6 . 10^7 when the iteration t, it is too big which means that the algorithm fails to minimize the cost function f. What wrong with that value?	- During the minimization, we evaluate the cost function using Equation (15). It comprises the least squares term, the nuclear-norm term and the l1-norm term. These terms play their trade-off roles with the regularization parameters. Although the objective is to find the lowest rank matrix $\boldsymbol{Z}^w$, and the sparest vector $\boldsymbol{s}$, these solutions must also follow the data fidelity induced the least squares term. - With the setting of the regularization parameters γ and λ, we observe that the value of the cost function is monotone-decreasing. In other words, the algorithm yields the target image and captures the wall clutter by minimizing the objective function of the optimization problem, though the value of the cost function is large. The large value of the cost function is mainly due to the contribution of the nuclear-norm term (sum of the singular values of the wall clutter matrix $\boldsymbol{Z}^w$), which tend to be big values.

Finally, we gratefully thank the Editor-in-Chief, Associate Editor, and three reviewers for the constructive and detailed comments that helped us improve the paper. We appreciate your time and feedback.

*********** End ***********