

Alternating Minimization Algorithm for Speckle Reduction with Shifting Technique

Hyenkyun Woo and Sangwoon Yun*

Abstract—Speckle (multiplicative noise) in synthetic aperture radar (SAR) makes it difficult to interpret the observed image. Due to edge preserving feature of total variation (TV), variational models with TV regularization have attracted much interest in reducing the speckle. Algorithms based on the augmented Lagrangian function have been proposed to efficiently solve speckle reduction variational models with TV regularization. But these algorithms require inner iterations or inverses involving the Laplacian operator at each iteration. In this paper, we adapt Tseng’s alternating minimization algorithm with shifting technique to efficiently remove the speckle without any inner iterations or inverses involving the Laplacian operator. The proposed method is very simple and highly parallelizable, and so it is very efficient to despeckle huge size SAR images. Numerical results show that our proposed method outperforms the state-of-the-art algorithms for speckle reduction variational models with the TV regularizer in terms of the CPU time.

Index Terms—Convex optimization, Alternating minimization, Total variation, Speckle, Synthetic aperture radar, Multiplicative noise, Denoising

I. INTRODUCTION

Synthetic aperture radar on a satellite is very useful for monitoring the ground of the Earth, irrespective of the weather condition [10], [28]. The received backscattered signal of SAR is corrupted by the speckle. We assume that the speckle is fully developed, i.e., the number of point scatterers in a resolution cell is large [11], [22]. The intensity of the observed signal at each pixel is the product of a backscattering coefficient (or a radar cross section) of the target and the speckle which is (negative) exponentially distributed with unit mean [31], [42]. Since the standard deviation of the speckle is equal to the mean of it, the signal-to-noise ratio (SNR) of the observed signal is only one (i.e. 0 dB) [4]. Unlike the standard optical imaging system, such as digital camera, the noise level of SAR is very high and so the speckle in SAR makes it hard to interpret valuable information, such as edges and patterns, from the observed signal. Therefore, the reduction of the speckle from the observed signal is a meaningful preprocessing for applications, such as segmentation and feature extraction.

Copyright (c) 2010 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org. This work was supported by Priority Research Centers Program(2009-0093827) and Basic Science Research Program(2010-0510-1-3) through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology. *Asterisk indicates corresponding author*

H. Woo is with Department of Mathematical Sciences, Seoul National University, Seoul 151-742, Korea e-mail: hyenkyun@gmail.com. S. Yun is with Department of Mathematics Education, Sung Kyun Kwan University, Jongro-gu, Seoul 110-745, Korea e-mail:sangwoony@gmail.com.

Recently, variational models [2], [6], [15], [20], [25], [30], [37], [39], [45] with the total variation regularizer have attracted much interest in removing the speckle due to edge preserving feature of TV [34] (many sharp edges exist in urban areas). Among them, the model based on maximum a posteriori (MAP) criterion for a logarithmic transformed image [6], [37] and the I-divergence model [39] are a convex variational model for speckle reduction. Algorithms based on the augmented Lagrangian function with nonlinear block-Gauss-Seidel method have been proposed to efficiently solve aforementioned convex variational models for speckle reduction [6], [39]. However, inner iterations [6] or inverses involving the Laplacian operator [39] are required depending on the formulation of the augmented Lagrangian function. The computational cost of inner iterations and evaluation of inverses is rather expensive.

In this paper, we propose a general framework, which combines an alternating minimization algorithm (AMA) [41] with shifting technique, to solve convex variational models for speckle reduction. The shifting technique significantly improves the performance of the AMA for speckle reduction; see Section III-B. The proposed algorithm has a simple closed form solution at each iteration. This is a main advantage of our algorithm over other state-of-the-art algorithms [6], [39] since no inner iterations and no inverses involving the Laplacian operator are required. Numerical results show that our proposed algorithm outperforms the state-of-the-art algorithms for speckle reduction in terms of the CPU time. The peak signal-to-noise ratio (PSNR) of the proposed algorithm is comparable to that of the state-of-the-art algorithms.

A. Variational Models for Speckle Reduction

In this subsection, we briefly describe three variational models for speckle reduction.

Let $\Omega \subset \mathbb{R}^2$, b be the observed radar image intensity. Then we want to find the unknown true image intensity, i.e. radar cross section (RCS) $\tilde{u} : \Omega \rightarrow \mathbb{R}^+$ from the observed image intensity:

$$b = \tilde{u}\eta, \quad (1)$$

where η is the speckle (multiplicative noise). Since the noise level of the speckle is relatively high, the conventional SAR system reduces the noise variance of the observed signal by multi-looking process, i.e., averaging the observed signals from slightly different angles of the same resolution cell (L -look SAR image). We note that the same effect is achieved by averaging neighboring pixels in the image domain. It is known that the probability density function (PDF) of the speckle η

of the L -look SAR image is given by the following Gamma distribution:

$$p(\eta) = \frac{L^L \eta^{L-1}}{\Gamma(L)} e^{-L\eta} H(\eta), \quad (2)$$

where H is Heaviside function, the mean value of the speckle η is 1 ($\mathbf{E}(\eta) = 1$), and the variance of η is $1/L$ [31]. The corresponding conditional probability density function of the observed pixel intensity b given its mean value \tilde{u} , which is the RCS of the target, is

$$p(b|\tilde{u}) = \frac{1}{\Gamma(L)} \left(\frac{L}{\tilde{u}}\right)^L b^{L-1} e^{-\frac{Lb}{\tilde{u}}}.$$

For the observed image b , a Bayesian analysis using MAP criterion with the TV regularizer (prior) leads to the following energy minimization problem [2]:

$$\min_u \int_{\Omega} \left(\log u + \frac{b}{u} \right) dx + \frac{\tilde{\lambda}}{L} \int_{\Omega} |\nabla u| dx, \quad (3)$$

where $\tilde{\lambda} > 0$. Unfortunately, this model is nonconvex. In order to resolve the nonconvexity, logarithmic transformation is commonly used [6], [25], [37] (i.e. $u_l = \log(u)$). Then the resulting variational model is expressed as follows:

$$\begin{aligned} u_l^* &= \arg \min_{u_l} \left\{ \int_{\Omega} (u_l + b e^{-u_l}) dx + \frac{\tilde{\lambda}}{L} \int_{\Omega} |\nabla u_l| dx \right\}, (4) \\ u^* &= e^{u_l^*}, \end{aligned}$$

where $\tilde{\lambda} > 0$. Notice that, since the logarithmic function is monotonically increasing, the location of edges does not change, and thus the gradient operator is applied to logarithmic transformed images; see [25] for more details. We call the above model (4) as the exponential model. But the logarithmic transformation expands the values in the dark area, while compressing those in the bright area. Also, the mean value is biased [24]. Recently, Steidl and Teuber [39] show that the solution of the classical I-divergence model:

$$\min_u \int_{\Omega} (u - b \log u) dx + \frac{\tilde{\lambda}}{L} \int_{\Omega} |\nabla u| dx \quad (5)$$

is theoretically equal to the solution of the exponential model (4). The main advantage of the model (5) is that it does not need a nonlinear transformation. Even though the theoretical solution of the model (5) is equal to that of the model (4), the numerical performance of the state-of-the-art algorithm depends on the model (i.e. model-specific); see Section IV.

B. Overview

The paper is organized as follows. In Section II, we review recently proposed algorithms. In Section III, we introduce our proposed method and show its convergence. In Section IV, we report our numerical results for speckle reduction. We give our conclusions in Section V.

II. DISCRETE FORMULATION AND RELATED WORKS

In this section, we review recently developed algorithms for solving speckle reduction models.

For notational convenience, we use vector notation, i.e., the 2D $M \times N$ image is columnwise stacked into a vector, for the rest of the paper. Therefore, the image u is a vector in $U = (0, C]^n$, where $n = MN$ and C is a positive real number. The observed noisy image b is a vector in $(0, \tilde{C}]^n$ with a positive scalar $\tilde{C} \geq C$. The i -th component of u is denoted by u_i . $\mathbf{1}$ denotes a vector of ones in \mathbb{R}^n . In the sequel, the multiplication and division of vectors are performed in component-wise. Hence the exponential model (4) and the I-divergence model (5) are expressed as the following discrete formulations respectively:

$$\min_{u \in \log U} \langle u + b e^{-u}, \mathbf{1} \rangle + \lambda \|\nabla u\|, \quad (6)$$

where $\lambda = \frac{\tilde{\lambda}}{L}$ and $\|\nabla u\| = \sum_{i=1}^n \|(\nabla u)_i\|_2$ with $(\nabla u)_i \in \mathbb{R}^2$,

$$\min_{u \in U} \langle u - b \log u, \mathbf{1} \rangle + \lambda \|\nabla u\|. \quad (7)$$

The Neumann boundary condition is used for the discrete gradient operator ∇ . Note that the adjoint of the gradient operator ∇ is the negative divergence operator $-\text{div}$, i.e. $\langle p, \nabla u \rangle = \langle -\text{div} p, u \rangle$. For more details, see [9].

The following lemma shows that the exponential model (6) and the I-divergence model (7) have a unique optimal solution.

Lemma II.1. *The exponential model (6) and the I-divergence model (7) have a unique minimizer.*

Proof. If we add the indicator function of the bounded constraint ($\delta_{\tilde{U}}(u) = 0$ if $u \in \tilde{U}$, $\delta_{\tilde{U}}(u) = \infty$ if $u \notin \tilde{U}$ with $\tilde{U} = \log U$ for (6) and $\tilde{U} = U$ for (7)) to the objective function of each model ((6) and (7)) and drop the bound constraint, then the objective function of each reformulation of (6) and (7) is proper, lower semicontinuous, level bounded [33]. Therefore, by [33, Theorem 1.9], the set of minimizers of (6) and that of (7) are nonempty and compact. In addition, each reformulation of (6) and (7) is strictly convex. Hence the exponential model (6) and the I-divergence model (7) have a unique minimizer. ■

The next lemma shows that the minimizer u_{exp}^* of the exponential model (6) is bounded below by the minimum component of $\log b$ and above by the maximum component of $\log b$, and the minimizer u_{idiv}^* of the I-divergence model (7) is bounded below by the minimum component of b and above by the maximum component of b . Its proof is omitted since the boundedness can be obtained by proceeding as in the proof of [39, Proposition 3.1].

Lemma II.2. *For all $i = 1, \dots, n$,*

$$\log b_{\min} \leq (u_{\text{exp}}^*)_i \leq \log b_{\max}, \quad b_{\min} \leq (u_{\text{idiv}}^*)_i \leq b_{\max}, \quad (8)$$

where $b_{\min} = \min_{i \in \{1, \dots, n\}} b_i$ and $b_{\max} = \max_{i \in \{1, \dots, n\}} b_i$.

By Lemma II.2, without loss of generality, we assume that $U = [b_{\min}, C]^n$ in the sequel.

A. Algorithms for the Exponential Model (6)

Bioucas-Dias and Figueiredo [6] proposed to apply the alternating direction method of multipliers (ADMM) [16], [17], [35] to solve the following reformulation of the exponential model (6):

$$u_l^* = \arg \min_{u \in \log U} \{ \langle u + be^{-u}, \mathbf{1} \rangle + \lambda \|\nabla d\| \mid d = u \}, \quad u^* = e^{u_l^*}. \quad (9)$$

They called this algorithm as the multiplicative image denoising by augmented Lagrangian (MIDAL). The framework of the MIDAL can be expressed as follows:

$$\begin{cases} u^{k+1} = \arg \min_{u \in \log U} \langle u + be^{-u}, \mathbf{1} \rangle + \frac{\alpha}{2} \|u - d^k - p^k\|_2^2 \\ d^{k+1} = \arg \min_d \lambda \|\nabla d\| + \frac{\alpha}{2} \|d - u^{k+1} + p^k\|_2^2 \\ p^{k+1} = p^k - (u^{k+1} - d^{k+1}). \end{cases} \quad (10)$$

They proposed to use Newton's method to approximately obtain u^{k+1} and to use Chambolle's algorithm [8] to approximately get d^{k+1} at each iteration. Note that the first step of (10) has a closed form solution in terms of the Lambert W function [6], [12], but it is difficult to evaluate this closed form solution.

Huang et al. [25] applied a variable splitting to the penalized formulation of the form (9), i.e., added the penalty term $\rho \|d - u\|_2^2$ to the objective function and removed the linear constraint. The minimizer of the penalized problem converges to (9) only when ρ approaches infinity. Hence this causes numerical difficulties since the penalized problem becomes severely ill-conditioned when ρ is very large. Besides this drawback, their method requires inner iterations as the MIDAL does.

B. Algorithms for the I-divergence Model (7)

Steidl and Teuber [39] proposed to compute the minimizer of the model (7) by applying Douglas-Rachford splitting technique [16] (or equivalently ADMM [17], [35]). Actually, they considered to apply the ADMM to the following three reformulations:

$$\min_{u \in U} \{ \langle u - b \log u, \mathbf{1} \rangle + \lambda \|z\| \mid z = \nabla u \}. \quad (11)$$

$$\min_{u \in U} \{ \langle u - b \log u, \mathbf{1} \rangle + \lambda \|\nabla d\| \mid d = u \}. \quad (12)$$

$$\min_{u \in U} \left\{ \langle u - b \log u, \mathbf{1} \rangle + \lambda \|z\| \mid \begin{pmatrix} u \\ z \end{pmatrix} = \begin{pmatrix} I \\ \nabla \end{pmatrix} \mathbf{d} \right\}, \quad (13)$$

where I denotes the identity matrix with dimension defined within the context. But the authors of [39] commented that the performance of directly applying the framework of the MIDAL to (12) is poor. They also commented that, if the discrete cosine transform (DCT) is used, the ADMM for (13) performs best. Also, this one has the advantage that it requires no inner iterations which are required for (11) and (12) when the ADMM is applied. Hence, in the following, we describe the framework of the ADMM applied for (13). We call it

as the ADMM-III. The framework of the ADMM-III can be expressed as follows:

$$\begin{cases} d^{k+1} = \arg \min_d \left\| \begin{pmatrix} p_1^k \\ p_2^k \end{pmatrix} + \begin{pmatrix} I \\ \nabla \end{pmatrix} d - \begin{pmatrix} u^k \\ z^k \end{pmatrix} \right\|_2^2 \\ u^{k+1} = \arg \min_{u \in U} \{ \langle u - b \log u, \mathbf{1} \rangle + \frac{\alpha}{2} \|p_1^k + d^{k+1} - u\|_2^2 \} \\ z^{k+1} = \arg \min_z \{ \lambda \|z\| + \frac{\alpha}{2} \|p_2^k + \nabla d^{k+1} - z\|_2^2 \} \\ \begin{pmatrix} p_1^{k+1} \\ p_2^{k+1} \end{pmatrix} = \begin{pmatrix} p_1^k \\ p_2^k \end{pmatrix} + \begin{pmatrix} I \\ \nabla \end{pmatrix} d^{k+1} - \begin{pmatrix} u^{k+1} \\ z^{k+1} \end{pmatrix}. \end{cases}$$

We can express the above framework in a simplified form:

$$\begin{cases} d^{k+1} = (I - \Delta)^{-1} (u^k - p_1^k - \text{div}(z^k - p_2^k)) \\ u^{k+1} = \frac{1}{2} \left(p_1^k + d^{k+1} - \mu \mathbf{1} + \sqrt{(p_1^k + d^{k+1} - \mu \mathbf{1})^2 + 4\mu b} \right) \\ u^{k+1} = \mathcal{P}_U(u^{k+1}) \\ z^{k+1} = \text{shrink}(p_2^k + \nabla d^{k+1}, \mu \lambda) \\ \begin{pmatrix} p_1^{k+1} \\ p_2^{k+1} \end{pmatrix} = \begin{pmatrix} p_1^k \\ p_2^k \end{pmatrix} + \begin{pmatrix} I \\ \nabla \end{pmatrix} d^{k+1} - \begin{pmatrix} u^{k+1} \\ z^{k+1} \end{pmatrix}, \end{cases} \quad (14)$$

where $\mathcal{P}_U(u)$ denotes the projection of u onto the set U , $\mu = \frac{1}{\alpha}$, Δ is the Laplacian operator, and the shrinkage operator is defined as in [43]:

$$z_i = \text{shrink}(a_i, c) = \max(\|a_i\|_2 - c, 0) \frac{a_i}{\|a_i\|_2},$$

where $a_i \in \mathbb{R}^2$ and $c \in \mathbb{R}$. We note that the similar approach as the ADMM-III has been applied for Poisson noise reduction models [19], [36], [44].

III. ALTERNATING MINIMIZATION ALGORITHM WITH SHIFTING TECHNIQUE

In this section, we describe an alternating minimization algorithm (AMA) studied in [41] for solving speckle reduction models, i.e., the exponential model (6) and the I-divergence model (7), by using separable convex reformulations. And we show the convergence of the iterates generated by the AMA when it is applied to solve reformulations of (6) and (7). Also we introduce shifting technique to improve the stability and the performance of the AMA when it is applied for speckle reduction variational models.

A. Alternating Minimization Algorithm

In this subsection, we describe briefly the AMA for solving the following separable convex minimization problem:

$$\begin{aligned} \min_{u, z} \quad & f(u) + g(z) \\ \text{subject to} \quad & \nabla u = z, \quad u \in \tilde{U} \end{aligned} \quad (15)$$

where $f : \tilde{U} \rightarrow \mathbb{R}$ is a twice differentiable function, $g(z) = \lambda \|z\|$, and \tilde{U} is a closed convex set in \mathbb{R}^n . The formulation (15) with $f(u) = \langle u + be^{-u}, \mathbf{1} \rangle$ and $\tilde{U} = \log U$ is the separable convex reformulation of the exponential model (6). The formulation (15) with $f(u) = \langle u - b \log u, \mathbf{1} \rangle$ and $\tilde{U} = U$ is the separable convex reformulation of the I-divergence model (7).

We assume that f is strongly convex with modulus $\sigma > 0$ on \tilde{U} , i.e.,

$$\nabla_u^2 f(u) \succeq \sigma I, \quad \forall u \in \tilde{U},$$

which means that $\nabla_u^2 f(u) - \sigma I$ is positive semidefinite. This together with Taylor expansion implies that, for all $v, w \in \tilde{U}$,

$$f(w) \geq f(v) + \langle \nabla_v f(v), w - v \rangle + \frac{\sigma}{2} \|w - v\|_2^2. \quad (16)$$

The Lagrangian function and the augmented Lagrangian function for (15) are respectively

$$\mathcal{L}(u, z, p) := f(u) + g(z) + \langle p, z - \nabla u \rangle \quad (17)$$

and

$$\mathcal{L}_\alpha(u, z, p) := f(u) + g(z) + \langle p, z - \nabla u \rangle + \frac{\alpha}{2} \|z - \nabla u\|_2^2. \quad (18)$$

Then the framework of the AMA for solving (15) can be expressed as follows:

$$\begin{cases} u^{k+1} = \arg \min_{u \in \tilde{U}} \mathcal{L}(u, z^k, p^k) \\ z^{k+1} = \arg \min_z \mathcal{L}_\alpha(u^{k+1}, z, p^k) \\ p^{k+1} = p^k + \alpha(z^{k+1} - \nabla u^{k+1}), \end{cases} \quad (19)$$

where α is a constant satisfying the condition:

$$0 < \alpha < \frac{2\sigma}{\|\Delta\|_2}. \quad (20)$$

Note that $\|\Delta\|_2$ is the maximum eigenvalue of the Laplacian operator and the value of $\|\Delta\|_2$ is less than or equal to 8 when the Neumann boundary condition is used [17].

In the following lemma, we show that the functions $\langle u + be^{-u}, \mathbf{1} \rangle$ of (6) and $\langle u - b \log u, \mathbf{1} \rangle$ of (7) are strongly convex on $\log U$ and U respectively. This lemma will be used in Lemma III.3. In what follows,

$$f_1(u) := \langle u + be^{-u}, \mathbf{1} \rangle \quad \text{and} \quad f_2(u) := \langle u - b \log u, \mathbf{1} \rangle.$$

Lemma III.1. *The function $f_1(u)$ is strongly convex with modulus $\frac{\min_j b_j}{C}$ on $\log U$ and the function $f_2(u)$ is also strongly convex with modulus $\frac{\min_j b_j}{C^2}$ on U .*

Proof. $f_1(u)$ is twice differentiable on the set $\log U$ and $\nabla_u^2 f_1(u) = \text{Diag}(b \circ e^{-u})$ for all $u \in \log U$, where $\text{Diag}(v)$ denotes a diagonal matrix whose diagonal entries consist of the components of v . Hence $\nabla_u^2 f_1(u) \succeq \sigma_1 I$, where $\sigma_1 := \frac{\min_j b_j}{C}$, for all $u \in \log U$. Therefore $f_1(u)$ is strongly convex with modulus σ_1 on $\log U$.

$f_2(u)$ is twice differentiable on the set U and $\nabla_u^2 f_2(u) = \text{Diag}(b \circ u^{-2})$ for all $u \in U$. Hence $\nabla_u^2 f_2(u) \succeq \sigma_2 I$, where $\sigma_2 := \frac{\min_j b_j}{C^2}$, for all $u \in U$. Therefore $f_2(u)$ is strongly convex with modulus σ_2 on U . ■

The next lemma shows that the dual problem of the exponential model (6) and that of the I-divergence model (7) have an optimal solution. In other words, the exponential model (6) and the I-divergence model (7) have an optimal Lagrange multiplier vector corresponding to the linear constraint $\nabla u = z$.

Lemma III.2. *The dual problem of the exponential model (6) has an optimal solution and that of the I-divergence model (7) has an optimal solution.*

Proof. The Lagrangian dual function of the exponential model (6) and that of the I-divergence model (7) are expressed as follows:

$$\begin{aligned} h(p) &= \inf_{u \in \tilde{U}, z} f(u) + g(z) + \langle p, z - \nabla u \rangle \\ &= - \sup_{u \in \tilde{U}} \{ \langle -\text{div } p, u \rangle - f(u) \} - \sup_z \{ \langle -p, z \rangle - g(z) \} \\ &= -f^*(-\text{div } p) - g^*(-p), \end{aligned}$$

where $f = f_i$, $i = 1, 2$ and $f^*(y) := \sup_{u \in \tilde{U}} \{ \langle y, u \rangle - f(u) \}$. Hence the dual problems of the exponential model (6) and the I-divergence model (7) have the following formulation:

$$\begin{aligned} \min_p \quad & f^*(-\text{div } p) \\ \text{subject to} \quad & \|p\|_* \leq \lambda, \end{aligned} \quad (21)$$

where $\|\cdot\|_*$ is a dual norm defined by $\|x\|_* = \max_{\|y\| \leq 1} x^T y$. By Lemma III.1, f is strongly convex, and hence f^* is finite, convex, and differentiable on the set $\{p \mid \|p\|_* \leq \lambda\}$ (see [32, §12 and Theorem 26.3]). Therefore the dual problem (21) has an optimal solution. ■

We note that (15) is a linearly constrained convex minimization problem, and so the strong duality (i.e., there is no duality gap) is satisfied; see also [39].

The next lemma gives two inequalities which are crucial to prove the convergence of the AMA when it is applied to solve the exponential model (6) and the I-divergence model (7). The proof of this lemma is given in Appendix A.

Lemma III.3. *Let $\{u^k\}$, $\{z^k\}$, $\{p^k\}$ be sequences generated by the AMA when it is applied to solve the problem (15) with $f = f_i$ for $i = 1, 2$. Then we have the following two inequalities*

$$\begin{aligned} & \sigma \|u^{k+1} - u^*\|_2^2 - \alpha \|\nabla u^{k+1} - \nabla u^*\|_2^2 \\ & \leq \langle p^{k+1} - p^*, \nabla u^{k+1} - \nabla u^* \rangle \\ & \quad - \alpha \langle z^{k+1} - z^*, \nabla u^{k+1} - \nabla u^* \rangle, \end{aligned} \quad (22)$$

where (u^*, z^*) is a solution of the problem (15), p^* is a solution of its dual problem (21), $\sigma = \sigma_i$ for $i = 1, 2$, and

$$\langle p^{k+1} - p^*, z^{k+1} - z^* \rangle \leq 0. \quad (23)$$

The following theorem establishes the global convergence of the AMA when it is applied to solve the problem (15) with $f = f_i$ for $i = 1, 2$. It can be proved by proceeding as in the statement of [41, page 134] with Lemma III.3. But we include the proof for completion.

Theorem III.4. *Let $\{u^k\}$, $\{z^k\}$, $\{p^k\}$ be sequences generated by the AMA with $0 < \alpha < \sigma/4$ when it is applied to solve the problem (15) with $f = f_i$ for $i = 1, 2$. Then u^k converges to the unique optimal solution u^* .*

Proof. Let (u^*, z^*) be a solution of the problem (15) and p^* be a solution of its dual problem (21). By summing (22) and

(23), using $\nabla u^* = z^*$, and rearranging, we obtain

$$\begin{aligned} & \sigma \|\hat{u}^{k+1}\|_2^2 - \alpha \|\nabla \hat{u}^{k+1}\|_2^2 + \langle \hat{p}^{k+1}, z^{k+1} - \nabla u^{k+1} \rangle \\ & \quad + \alpha \langle \hat{z}^{k+1}, \nabla \hat{u}^{k+1} \rangle \\ & = \sigma \|\hat{u}^{k+1}\|_2^2 - \alpha \|\nabla \hat{u}^{k+1}\|_2^2 + \frac{1}{\alpha} \langle \hat{p}^{k+1}, p^{k+1} - p^k \rangle \\ & \quad + \alpha \langle \hat{z}^{k+1}, \nabla \hat{u}^{k+1} \rangle \\ & \leq 0, \end{aligned}$$

where $\hat{p}^{k+1} = p^{k+1} - p^*$, $\hat{u}^{k+1} = u^{k+1} - u^*$, $\hat{z}^{k+1} = z^{k+1} - z^*$, and the equality follows from $p^{k+1} = p^k + \alpha(z^{k+1} - \nabla u^{k+1})$. This together with the identities

$$\begin{aligned} 2\langle \hat{p}^{k+1}, p^{k+1} - p^k \rangle & = \|p^{k+1} - p^k\|_2^2 + \|\hat{p}^{k+1}\|_2^2 - \|\hat{p}^k\|_2^2 \\ \frac{1}{\alpha^2} \|p^{k+1} - p^k\|_2^2 & = \|\nabla \hat{u}^{k+1}\|_2^2 + \|\hat{z}^{k+1}\|_2^2 \\ & \quad - 2\langle \hat{z}^{k+1}, \nabla \hat{u}^{k+1} \rangle, \end{aligned}$$

where the second equality uses $p^{k+1} = p^k + \alpha(z^{k+1} - \nabla u^{k+1})$ and $\nabla u^* = z^*$, implies that

$$\begin{aligned} & \|\hat{p}^k\|_2^2 \\ & \geq \|\hat{p}^{k+1}\|_2^2 - \alpha^2 \|\nabla \hat{u}^{k+1}\|_2^2 + 2\sigma\alpha \|\hat{u}^{k+1}\|_2^2 \\ & \quad + \alpha^2 \|\hat{z}^{k+1}\|_2^2 \\ & \geq \|\hat{p}^{k+1}\|_2^2 + 2\alpha(\sigma - 4\alpha) \|\hat{u}^{k+1}\|_2^2 + \alpha^2 \|\hat{z}^{k+1}\|_2^2, \end{aligned}$$

where the second inequality uses the Cauchy-Schwarz inequality and $\|\nabla\|_2^2 = \|\Delta\|_2 \leq 8$. Since the choice of k and p^* are arbitrary, by the condition $0 < \alpha < \sigma/4$, we obtain that

$$\begin{aligned} \|p^k - p^*\|_2^2 & \geq \|p^{k+1} - p^*\|_2^2 \\ & \quad + 2\alpha(\sigma - 4\alpha) \|u^{k+1} - u^*\|_2^2 + \alpha^2 \|z^{k+1} - z^*\|_2^2 \end{aligned} \quad (24)$$

holds for all $k = 0, 1, \dots$ and any solution p^* of the dual problem (21). Therefore (24) implies that $\{p^k\}$ is bounded, $\sum_{k=0}^{\infty} \|u^k - u^*\|_2^2 < \infty$ and $\sum_{k=0}^{\infty} \|z^k - z^*\|_2^2 < \infty$, so that $u^{k+1} \rightarrow u^*$, and $z^{k+1} \rightarrow z^* = \nabla u^*$. Hence u^k converges to the unique optimal solution u^* . ■

B. Shifting Technique

In this subsection, we introduce the shifting technique to significantly improve the performance of the AMA.

First of all, we describe the framework of the AMA applied to solve the exponential model (6) and the I-divergence model (7). In what follows, $\Psi(x) = \log x$ for the exponential model (6) and $\Psi(x) = x$ for the I-divergence model (7).

The framework can be simply expressed as follows:

$$\begin{cases} u^{k+1} = \mathcal{P}_{\Psi(U)} \left(\Psi \left(\frac{b}{\mathbf{1} + \text{div } p^k} \right) \right), \\ z^{k+1} = \text{shrink}(\nabla u^{k+1} - \frac{p^k}{\alpha_i}, \frac{\lambda}{\alpha_i}), \\ p^{k+1} = p^k + \alpha_i (z^{k+1} - \nabla u^{k+1}), \end{cases} \quad (25)$$

where $0 < \alpha_1 < \frac{\sigma_1}{4}$ with $\sigma_1 = \frac{\min_j b_j}{C}$ for the exponential model (6) and $0 < \alpha_2 < \frac{\sigma_2}{4}$ with $\sigma_2 = \frac{\min_j b_j}{C^2}$ for the I-divergence model (7). Hence we do not require any inner iterations and avoid computing any inverses involving the Laplacian operator. Those are advantages over the state-of-the-art algorithms such as the MIDAL (10) and ADMM-III (14) introduced in Section II.

When the AMA is applied to solve the exponential model (6) and the I-divergence model (7), each strongly convex modulus σ_i , which depends on $\min_j b_j$ by Lemma III.1, of the fidelity terms is crucial for the performance of the proposed algorithm. But, it is observed that the minimum value of the measurement data is very close to 0 ($\min_j b_j \approx 0$), especially when L is small. Note that if $L = 1$ then the PDF of the speckle is (negative) exponential distribution with unit mean and unit variance (see (2) and Fig. 1 (c), (d)). Therefore, the minimum value of the speckled image is overwhelmed by the speckle, especially when L is small. For example, the minimum value of the speckled images in Fig. 6 (b) and (c) is 0. Hence, if we directly apply the AMA to solve the exponential model (6) or the I-divergence model (7) with images such as Fig. 6 (b) and (c), then the AMA converges very slowly.

In order to overcome this drawback, we can add Tikhonov regularization $\frac{1}{2\delta} \|u\|_2^2$ with $\delta > 0$ to each fidelity term of the exponential model (6) and the I-divergence model (7). But Tikhonov regularization essentially changes variational models. Also, inner iterations are required for the exponential model (6) with Tikhonov regularization. We can also overcome that drawback by adding the proximal term $\frac{1}{2\delta} \|u - u^k\|_2^2$ to the first step (minimization of the Lagrangian function) of the AMA (19) [46]. This modified version is also known as the relaxed AMA (rAMA), which is equivalent to the primal-dual hybrid gradient (PDHG) algorithm [9], [18], [48]. When the rAMA is applied to the exponential model (6), the only difference from the AMA is the first step of the AMA (19), i.e.,

$$u^{k+1} = \arg \min_{u \in \log U} \langle u + be^{-u}, \mathbf{1} \rangle + \langle p^k, z^k - \nabla u \rangle + \frac{1}{2\delta} \|u - u^k\|_2^2. \quad (26)$$

There is a closed form solution for (26) in terms of the Lambert W function [6], [12] but it is difficult to evaluate this closed form solution. Hence we need an algorithm, such as Newton's method, to obtain u^{k+1} as did for the MIDAL (10).

On the contrary, when the rAMA is applied to the I-divergence model (7), the first step of the rAMA has a simple closed form solution:

$$u^{k+1} = \mathcal{P}_U \left(\frac{\hat{p}^k + \sqrt{(\hat{p}^k)^2 + 4\delta b}}{2} \right), \quad (27)$$

where $\hat{p}^k = u^k - \delta(\mathbf{1} + \text{div } p^k)$. Hence the rAMA can also avoid any inner iterations and inversions for the I-divergence model.

From the Kuhn-Tucker Theorem [32, Corollary 28.3.1], the optimal solution u^* is

$$u^* = \Psi \left(\frac{b}{\mathbf{1} + \text{div } p^*} \right). \quad (28)$$

Since $\tilde{u} \approx \Psi^{-1}(u^*)$ where \tilde{u} is the original image, the above equality (28) together with (1) implies that the corresponding speckle η can be characterized by

$$\eta = \mathbf{1} + \text{div } p^* = \frac{b}{\Psi^{-1}(u^*)}.$$

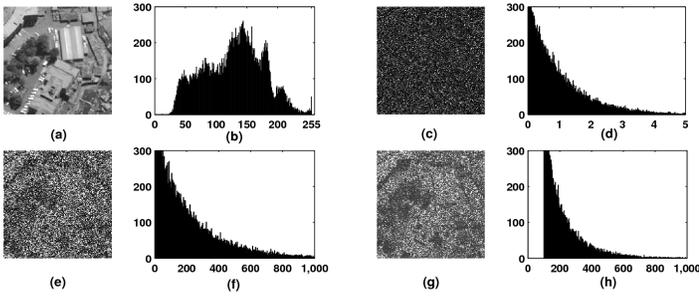


Fig. 1: (a) Original image \tilde{u} , (b) Histogram of \tilde{u} , (c) Speckle η , (d) Histogram of the speckle η , (e) Speckled shifted image $(\tilde{u} + T\mathbf{1})\eta$, (f) Histogram of the speckled shifted image $(\tilde{u} + T\mathbf{1})\eta$, (g) Shifted speckled image $\tilde{u}\eta + T\mathbf{1}$, and (h) Histogram of the shifted speckled image $\tilde{u}\eta + T\mathbf{1}$. Note that $L = 1$, $T = 100$, and $\langle (\tilde{u} + T\mathbf{1})\eta, \mathbf{1} \rangle = \langle \tilde{u}\eta + T\mathbf{1}, \mathbf{1} \rangle \approx 227n$.

Hence the speckle η satisfies the unit mean property, i.e. $\langle \eta, \mathbf{1} \rangle / n = 1$, since $\langle \text{div } p^*, \mathbf{1} \rangle = 0$; see also [39] for the case of the I-divergence model.

If we shift the speckled image by T (i.e. $\tilde{b} = b + T\mathbf{1}$) then, due to the unit mean property of the speckle, the following equality holds:

$$\langle \Psi^{-1}(u^*)\eta + T\mathbf{1}, \mathbf{1} \rangle = \langle (\Psi^{-1}(u^*) + T\mathbf{1})\eta, \mathbf{1} \rangle.$$

We note that, even though $\Psi^{-1}(u^*)\eta + T\mathbf{1}$ and $(\Psi^{-1}(u^*) + T\mathbf{1})\eta$ are equal in average, the minimum values of both terms are different, i.e. the minimum value of $\tilde{b} = \Psi^{-1}(u^*)\eta + T\mathbf{1}$ is $\min_j \Psi^{-1}(u_j^*)\eta_j + T$ and the minimum value of $(\Psi^{-1}(u^*) + T\mathbf{1})\eta$ is $\min_j (\Psi^{-1}(u_j^*) + T)\eta_j$; see Fig. 1.

In homogeneous region of the original image, we get the despeckled image $\Psi^{-1}(u^*)$ by the following averaging process:

$$\begin{aligned} \Psi^{-1}(\tilde{u}^*) &= \mathbf{E}[\Psi^{-1}(u^*)\eta + T\mathbf{1}] \\ &= \mathbf{E}[(\Psi^{-1}(u^*) + T\mathbf{1})\eta] \\ &= \Psi^{-1}(u^*) + T\mathbf{1}, \end{aligned} \quad (29)$$

where $\Psi^{-1}(\tilde{u}^*)$ is the solution for the shifted speckled image \tilde{b} . Note that we assume that $\Psi^{-1}(u^*) + T\mathbf{1}$ and η are statistically independent [31]. The above equality (29), i.e., the invariance with respect to shifting, is satisfied by various denoising filters [7], [27], [40]. In axiomatic scale-space theory of Alvarez et al. [1], [3], it is known as the *gray-level shift invariance*.

Similar to the results of Lemma II.2, it can be shown that $\Psi^{-1}(\tilde{u}^*)$ satisfies the following bound condition:

$$\Psi^{-1}(\tilde{u}^*) \in U + T\mathbf{1}.$$

When $\lambda = 0$, $\Psi^{-1}(u^*) = \Psi^{-1}(\tilde{u}^*) - T\mathbf{1}$ since $\Psi^{-1}(\tilde{u}^*) = \mathcal{P}_{U+T\mathbf{1}}(\tilde{b})$ (see Algorithm 1). Hence, we can roughly consider that the minimizer of the exponential model (6) and that of the I-divergence model (7) are approximately obtained from the shifted speckled image \tilde{b} as follows:

$$\Psi^{-1}(u^*) \approx \Psi^{-1}(\tilde{u}^*) - T\mathbf{1}.$$

The main purpose of the shifting technique is to increase the modulus σ_i which is crucial for the performance of the AMA; see Fig. 3. Since each modulus σ_i depends on $\min_j b_j$ and C (see Lemma III.1), we can obtain larger modulus by the shifting technique (i.e. adding T to the received backscattered signal of SAR). In other words, for the exponential model (6),

$$\sigma_1 = \frac{\min_j b_j + T}{C + T}.$$

It is easy to show that σ_1 is an increasing function of $T \geq 0$ and $\sigma_1 \rightarrow 1$ as $T \rightarrow \infty$. For the I-divergence model (7),

$$\sigma_2 = \frac{\min_j b_j + T}{(C + T)^2}.$$

It is easy to prove that σ_2 has a maximum value at $T = C - 2\min_j b_j$. Since the given number of looks L is very small in practical applications [26], [31], we consider that $L = 1, 3$ ($\min_j b_j \approx 0$) in our numerical experiments in Section IV. Therefore, if we choose $T \gg \min_j b_j$ then the parameters σ_1 and σ_2 mainly depend on the shifting parameter T .

In Algorithm 1, we describe formally our proposed speckle reduction method. We call this speckle reduction method as *Alternating Minimization Algorithm with Shifting Technique* (AMAST).

Algorithm 1 AMAST

Input: Given the noisy image b

Initialization: Let $\tilde{b} = b + T\mathbf{1}$ with $T > 0$, $u^{-1} = 0$, $p^{-1} = 0$, and $u^0 = \arg \min_{v \in \Psi(U+T\mathbf{1})} \|v - \Psi(\tilde{b})\|_2^2$.

while $\|\Psi^{-1}(u^k) - \Psi^{-1}(u^{k-1})\|_2 > \varepsilon \|\Psi^{-1}(u^{k-1})\|_2$ **do**

Update $u^{k+\frac{1}{2}} = \Psi(\tilde{b}/(1 + \text{div } p^k))$

Project $u^{k+1} = \arg \min_{v \in \Psi(U+T\mathbf{1})} \|v - u^{k+\frac{1}{2}}\|_2^2$

Update $z^{k+1} = \text{shrink}(\nabla u^{k+1} - \frac{p^k}{\alpha_i}, \frac{\lambda}{\alpha_i})$

Update $p^{k+1} = p^k + \alpha_i(z^{k+1} - \nabla u^{k+1})$

end while

SOLUTION : $\Psi^{-1}(u^*) = \Psi^{-1}(u^{k+1}) - T\mathbf{1}$

IV. EXPERIMENTAL RESULTS

In this section, we report numerical experiments for speckle reduction with twelve test images, which include eight remote sensing images, in Fig. 2 and two real SAR images in Fig. 8 (a) and (c) from <http://www.sandia.gov/radar/sar-data.html>. We compare our proposed algorithm with the MIDAL [6], ADMM-III [39], rAMA (PDHG [48]), and probabilistic patch based (PPB) filter which is studied by Deledalle et al. [13]. The PPB is a nonlocal means based algorithm for speckle reduction. Note that there are two versions for the PPB, i.e. the iterative PPB (PPBit) and non-iterative PPB (PPBnit).

All algorithms are implemented in C language with interface to MATLAB through the mex function. All runs are performed on a laptop with Intel i7-640LM CPU (2.13 - 2.93GHz) and 8GB Memory. The Operating System is 64bit Linux. We note that the MIDAL is implemented by us based on the guideline



Fig. 2: Dataset for the numerical experiments - twelve different test images; the size of images varies from 256×256 to 1500×1500 . The grayscale of the images is $[0, 255]$.

L	METHOD	Exponential Model				I-Divergence Model			
		λ	α_1	δ	T	λ	α_2	δ	T
1	MIDAL	1.0/L	1.0	-	1	-	-	-	-
	ADMM-III	-	-	-	-	1.0/L	0.01	-	1
	rAMA	1.0/L	2.0	0.2	1	1.0/L	0.05	10	1
	AMAST	1.0/L	0.043	-	30	1.0/L	$1.5e-4$	-	30
	AMAST-a	1.0/L	α_1^k	-	30	1.0/L	α_2^k	-	30
3	MIDAL	1.5/L	1.5	-	1	-	-	-	-
	ADMM-III	-	-	-	-	1.5/L	0.01	-	1
	rAMA	1.5/L	2.0	0.2	1	1.5/L	0.05	10	1
	AMAST	1.3/L	0.06	-	30	1.3/L	$2.4e-4$	-	30
	AMAST-a	1.3/L	α_1^k	-	30	1.3/L	α_2^k	-	30

TABLE I: The parameter values for numerical experiments.

in [6] and the ADMM-III is implemented by us based on the MATLAB version which is provided by the authors of [39]. The PPB is available from [14].

A. Results on Synthetic Images

In this subsection, we report numerical experiments for speckle reduction with twelve synthetic images. All the test images are corrupted by the speckle (L -look SAR image, (2)). We note that, in practical applications, the given number of looks L is very small [26], [31]. Therefore, we choose $L = 1, 3$. We terminate all the algorithms except the PPB by the following stopping condition:

$$\|\Psi^{-1}(u^{k+1}) - \Psi^{-1}(u^k)\|_2 \leq \varepsilon \|\Psi^{-1}(u^k)\|_2, \quad (30)$$

where $\varepsilon = 3 \times 10^{-4}$. Note that the PPB is terminated by the fixed number of iterations; see [13], [14] for details.

The structure of the proposed method (see Algorithm 1) is really simple. But we need to compute log for each image pixel at each iteration for the exponential model. Therefore, instead of using standard C library, we use *Fastlog* [38], which is faster but less accurate. In Table I, we summarize values for the parameters, i.e. λ, α, δ of all the algorithms except the PPB.

For the MIDAL, $\alpha = \lambda L$ and the number of Newton iterations is four for the first step in (10) as recommended in [5], [6]. The split Bregman method [21] is used for the second step in (10) instead of Chambolle's algorithm [8] (the

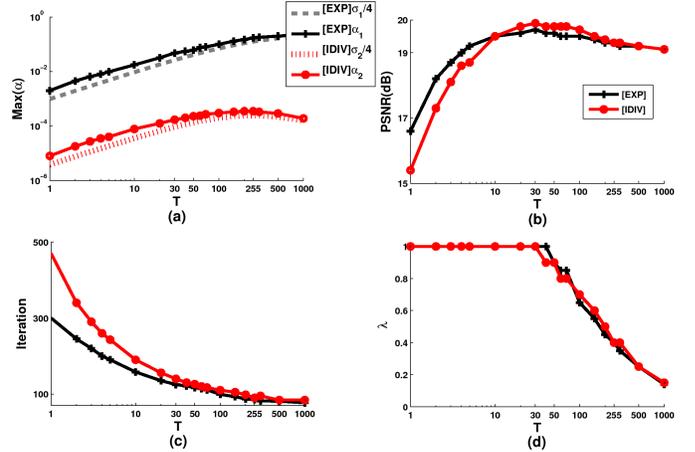


Fig. 3: (a) The comparison of the empirical maximum value of α_i and the theoretical bound $\sigma_i/4$ with respect to T . Note that the empirical α_i is getting close to the theoretical bound as T increases. (b) PSNR vs. T , (c) #iterations vs. T , and (d) λ vs. T . The speckled image ($\min_j b_j = 0$ and $C = 255$) in Fig. 6 (b) is used.

split Bregman method is faster than Chambolle's algorithm) and the number of iterations for the split Bregman method is ten.

For the ADMM-III, we set $\alpha = 0.01$ (this parameter is optimized for our experiments) and use the DCT to evaluate the inverses of the operator $I - \Delta$.

For the rAMA, we set $\alpha = 2.0, \delta = 0.2$, the number of Newton iterations is four when u^{k+1} is updated (26) for the exponential model, and set $\alpha = 0.05, \delta = 10$ for the I-divergence model.

For the MIDAL, ADMM-III, and rAMA, we do not need to shift the speckled image, but the minimum value of it is zero for some images, such as speckled images in Fig. 6, and so we shift speckled images by one as did in [20].

For the proposed method, the shifting parameter T is related to the performance of the algorithm. As shown in Fig. 3, the algorithm converges very slowly when T is too small. We note that it takes 5883 iterations to obtain 19.54dB for the exponential model and takes 7524 iterations to obtain 19.69dB for the I-divergence model when $T = 1$ and $\varepsilon = 10^{-5}$ is used in (30). It is observed that if the value of the parameter T increases then the number of iterations decreases. The quality of the recovered image is worse when T is too large. Based on our numerical experience, it would be better to choose $T \in [10, 100]$. We note that it would be better to choose small λ when T is large. From Fig. 3 (a), we can choose a slightly higher value of α_i than the theoretical bound $\sigma_i/4$ without fail to converge. We choose $T = 30$ for all experiments based on Fig. 3 (b). When $T = 30$, the theoretical upper bound of α_1 is 2.6×10^{-2} for the exponential model (6) and that of α_2 is 0.9×10^{-4} for the I-divergence model (7). But we can choose $\alpha_1 = 4.3 \times 10^{-2}$ and $\alpha_2 = 1.5 \times 10^{-4}$ for $L = 1$ and we can select $\alpha_1 = 6.0 \times 10^{-2}$ and $\alpha_2 = 2.4 \times 10^{-4}$ for $L = 3$; see Table I.

It is conservative to use the constant step size α_i , especially

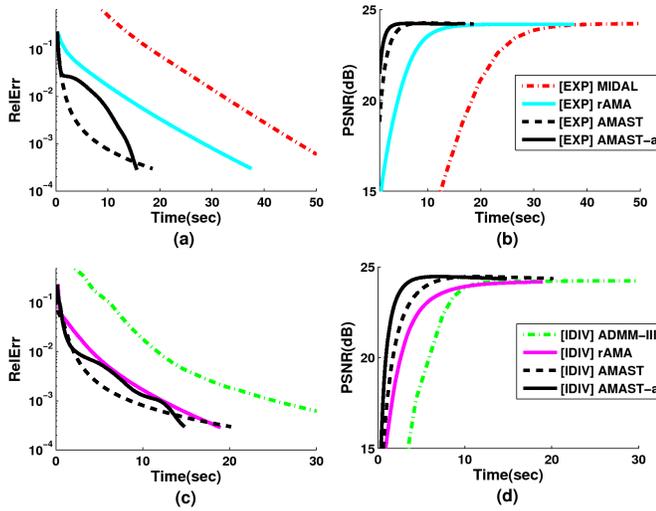


Fig. 4: (a) Relative error vs. CPU time for the exponential model, (b) PSNR vs. CPU time for the exponential model, (c) Relative error vs. CPU time for the I-divergence model, and (d) PSNR vs. CPU time for the I-divergence model. Even though the intermediate relative error of the AMAST-a is larger than that of the AMAST, the AMAST-a is stable and the total computational cost of it is reduced (less number of iterations) when it is compared to that of the AMAST. The speckled image in Fig. 7 (c) is used.

at the initial stage. Hence we update α_i adaptively at the initial stage to accelerate the convergence of our proposed algorithm; see Fig. 4. We update α_i adaptively as follows:

$$\alpha_i^k = \alpha_i \times 10^{0.3 \max((Q-k)/Q, 0)}, \quad (31)$$

where $Q = 150$ for $L = 1$ and $Q = 100$ for $L = 3$. We call this method as “AMAST-a”.

In Fig. 4, we show the relative error and the PSNR (on the speckled image in Fig. 7 (c)) for each model with respect to the CPU time. Even though the intermediate relative error of the AMAST-a is larger than that of the AMAST, the AMAST-a is stable and the total computational cost of it is reduced (less number of iterations) when it is compared to that of the AMAST.

Table II reports the PSNR, the CPU time, and the number of iterations of the MIDAL, ADMM-III, rAMA, AMAST(-a), PPBit for solving speckle reduction models with twelve different images in Fig. 2. We note that the PSNR is defined by

$$10 \log_{10} \left(\frac{255^2 n}{\|\tilde{u} - \Psi^{-1}(u^*)\|_2^2} \right)$$

where n is the size of the image, \tilde{u} is the original image, and $\Psi^{-1}(u^*)$ is the recovered image. Based on average CPU time, the AMAST-a is at least two times faster than MIDAL and rAMA for the exponential model and is faster than ADMM-III and rAMA for the I-divergence model. The PSNR of AMAST(-a) is comparable to that of the MIDAL, ADMM-III, and rAMA. The AMAST-a for the exponential model is the fastest algorithm. We note that the performance of the AMAST(-a) does not much depend on the form of the

variational model. The PSNR of the PPBit is better than all the other algorithms when $L = 3$. Especially, if the image, such as the “Barbara” image, has periodic patterns then the PSNR of the PPBit is at least 1.0dB higher than that of all the other algorithms. But the PPBit is too slow.

In Fig. 5, we compare the AMAST-a with the MIDAL for the exponential model (6), the ADMM-III for the I-divergence model (7), and the nonlocal means based PPB [13] on the “Barbara” image. The AMAST-a is faster than all the other algorithms. The PSNR of the AMAST-a is even better than that of the PPB when $L = 1$. When $L = 3$, the PPB recovers periodic patterns better than all the other algorithms.

In Fig. 6, we compare the AMAST-a with the MIDAL for the exponential model (6), the ADMM-III for the I-divergence model (7), and the nonlocal means based PPB on the “Remote4” image. The AMAST-a is faster than all the other algorithms. The PSNR of the AMAST-a is comparable to that of all the other algorithms. Since there is no periodic pattern, the PSNR of the PPB is less than that of the AMAST-a. As commented in [13], the PPB well preserves edges, but small and thin structures, such as cars on the road, are removed.

In Fig. 7, we compare the AMAST-a and rAMA with the MIDAL for the exponential model (6), the ADMM-III for the I-divergence model (7), and the nonlocal means based PPB on the “Remote8” image. For the exponential model, the AMAST-a is at least 2.4 times faster than the MIDAL and rAMA with comparable PSNR. For the I-divergence model, the proposed algorithm AMAST-a is the fastest algorithm with better PSNR. The PPB is very slower than all the other algorithms.

B. Results on real SAR Images

METHOD	Exponential Model				I-Divergence Model			
	λ	α_1	δ	T	λ	α_2	δ	T
MIDAL	1.5/L	1.5	-	1	-	-	-	-
ADMM-III	-	-	-	-	1.5/L	0.004	-	1
rAMA	1.5/L	2.0	0.2	1	1.5/L	0.05	10	1
AMAST	1.2/L	0.055	-	70	1.2/L	8e-5	-	70
AMAST-a	1.2/L	α_1^k	-	70	1.2/L	α_2^k	-	70

TABLE III: The parameter values for numerical experiments on real SAR images in Fig. 8. Note that we set $L=3.5$ based on the estimated number of looks of the images in Fig. 8 (a) and (c).

In this subsection, we report numerical experiments on speckle reduction with two real SAR images in Fig. 8 (a) and (c). The raw data of each real SAR image in Fig. 8 is in the grayscale $[0, 2^{16}]$. Even though the dynamic range of the real SAR image is very high, the meaningful information is usually in a low grayscale region except the bright scatterers. Therefore we set $U = [0, 16 \times 255]$ (see also [23], [47]). From homogeneous regions of real SAR images (white squares in Fig. 8 (a) and (c)), we estimate the equivalent number of looks (ENL) [31]; ENL=3.47 for (a) and ENL=3.51 for (b) in Fig. 8.

L	IMAGE	Exponential model				I-Divergence model				Nonlocal
		MIDAL[6]	rAMA	AMAST	AMAST-a	ADMM-III[39]	rAMA	AMAST	AMAST-a	PPBit[13]
		PSNR/Time/Iter	PSNR/Time/Iter	PSNR/Time/Iter	PSNR/Time/Iter	PSNR/Time/Iter	PSNR/Time/Iter	PSNR/Time/Iter	PSNR/Time/Iter	PSNR/Time
1	Barbara	20.8/1.2/17	20.7/1.6/96	20.9/0.8/220	20.9/0.6/154	20.7/0.7/55	20.7/0.7/128	21.0/0.8/246	21.0/0.6/179	20.8/33.0
	Boat	21.7/5.5/17	21.7/7.1/98	21.7/3.7/210	21.7/2.5/146	21.7/3.9/54	21.7/3.3/132	21.7/3.8/242	21.7/2.7/175	21.5/102.7
	House	22.0/1.2/17	22.1/1.7/97	22.0/0.8/208	22.0/0.5/146	22.1/0.6/51	22.2/0.8/138	22.0/0.8/239	22.0/0.6/174	21.6/33.2
	Lena	23.3/5.6/17	23.3/7.1/97	23.3/3.8/215	23.4/2.7/155	23.3/3.9/55	23.3/3.4/130	23.4/4.2/243	23.4/2.9/176	23.4/102.7
	Remote1	19.5/4.3/17	19.4/5.7/96	19.5/3.4/236	19.5/2.5/169	19.4/3.3/56	19.4/2.7/124	19.7/3.5/280	19.7/2.7/212	19.3/89.8
	Remote2	19.7/6.0/17	19.7/7.3/98	19.5/4.2/203	19.5/2.8/142	19.7/4.7/52	19.7/3.8/135	19.6/4.1/239	19.6/3.0/172	19.6/110.7
	Remote3	19.4/1.5/17	19.4/2.2/97	19.5/1.1/211	19.5/0.7/146	19.4/0.9/53	19.4/1.0/129	19.6/1.0/244	19.6/0.8/177	19.1/41.7
	Remote4	19.6/5.9/18	19.5/6.9/95	19.5/4.3/241	19.5/3.1/175	19.5/3.7/52	19.5/3.2/124	19.8/4.3/267	19.8/3.3/200	19.2/102.7
	Remote5	18.8/5.8/18	18.7/6.5/95	18.6/4.6/256	18.6/3.4/194	18.7/4.2/59	18.7/3.0/113	18.9/4.6/302	18.9/3.6/235	18.6/101.5
	Remote6	21.0/13.7/17	20.9/16.4/97	20.8/8.8/224	20.8/6.1/157	20.9/13.2/57	20.9/7.9/130	21.1/9.9/264	21.1/7.4/198	20.5/235.2
3	Remote7	21.9/25.3/17	21.7/24.9/96	21.8/15.5/230	21.8/10.8/161	21.8/24.2/56	21.7/12.2/126	22.2/16.7/262	22.2/12.7/198	21.2/359.0
	Remote8	22.0/55.6/17	21.9/60.1/97	22.3/31.6/219	22.3/22.5/154	21.9/65.4/53	21.9/29.4/131	22.6/34.5/250	22.6/25.3/183	21.4/842.3
	Average	20.8/11.0/17	20.8/12.3/97	20.8/6.9/223	20.8/4.9/158	20.8/10.7/54	20.8/6.0/128	21.0/7.4/257	21.0/5.5/190	20.5/179.6
	Barbara	22.2/1.3/17	22.1/1.1/62	22.1/0.5/120	22.1/0.3/86	22.1/0.4/33	22.1/0.5/83	22.2/0.5/139	22.2/0.4/98	23.4/34.3
	Boat	23.9/5.7/17	23.9/4.5/62	23.7/2.1/114	23.7/1.7/94	23.9/2.3/32	23.8/2.2/84	23.7/2.2/132	23.7/1.6/93	23.9/103.3
	House	24.5/1.2/17	24.6/1.1/64	24.3/0.5/111	24.4/0.4/96	24.6/0.4/31	24.6/0.5/89	24.3/0.4/127	24.3/0.4/93	24.7/34.3
	Lena	25.6/5.7/17	25.6/4.5/62	25.4/2.1/115	25.4/1.8/90	25.6/2.4/33	25.5/2.2/84	25.4/2.2/135	25.4/1.6/95	26.1/103.3
	Remote1	21.1/4.5/18	21.1/3.5/60	21.0/1.9/124	21.0/1.5/93	21.1/1.9/32	21.1/1.7/78	21.1/2.0/150	21.1/1.4/104	21.0/90.6
	Remote2	21.2/6.4/18	21.2/4.8/63	21.0/2.3/112	21.0/1.9/90	21.2/2.8/32	21.2/2.4/85	21.0/2.4/128	21.0/1.7/91	21.2/111.4
	Remote3	20.8/1.6/18	20.7/1.4/61	20.7/0.6/118	20.7/0.5/85	20.7/0.5/32	20.7/0.6/80	20.8/0.6/138	20.8/0.5/97	20.6/42.9
3	Remote4	21.6/5.9/18	21.6/4.3/60	21.5/2.3/122	21.5/1.8/91	21.6/2.2/31	21.6/2.1/80	21.6/2.4/146	21.6/1.7/102	21.3/103.4
	Remote5	20.7/5.8/18	20.7/4.1/59	20.5/2.4/128	20.5/1.8/92	20.7/2.5/34	20.7/1.9/70	20.6/2.6/161	20.6/1.9/116	21.0/102.3
	Remote6	22.7/14.8/18	22.7/10.0/61	22.5/5.1/122	22.5/3.8/93	22.7/7.7/33	22.7/5.1/83	22.6/5.6/147	22.6/4.1/103	22.4/235.1
	Remote7	23.8/26.5/18	23.8/16.1/61	23.7/8.5/122	23.7/6.6/93	23.8/14.2/33	23.8/8.3/80	23.8/9.7/147	23.8/7.3/109	23.7/360.5
	Remote8	24.2/55.6/17	24.2/37.5/61	24.2/18.6/121	24.2/15.6/101	24.2/42.4/32	24.2/19.0/84	24.4/20.2/140	24.3/14.8/102	23.9/838.0
	Average	22.7/11.2/18	22.7/7.7/61	22.6/3.9/119	22.6/3.1/92	22.7/6.6/32	22.7/3.9/82	22.6/4.2/141	22.6/3.1/100	22.8/180.0

TABLE II: Comparison of the performance of the AMAST(-a) and rAMA with that of the MIDAL for the exponential model, the ADMM-III for the I-divergence model, and the PPBit. The proposed method, AMAST-a outperforms all the other algorithms in terms of the CPU time, while achieving comparable PSNR. The AMAST-a for the exponential model is the fastest algorithm for speckle reduction.

Therefore, we set $L=3.5$.¹ In Table III, we report parameters used to despeckle real SAR images in Fig. 8 (a) and (c). For the AMAST-a, we use the adaptive update scheme in (31) with $Q = 100$.

The iterative probabilistic patch based filter (PPBit) seems to remove the speckle in homogeneous regions well and preserve edges and small bright scatterers well. But it attenuates thin strip lines in Fig. 8 (a). It seems that the visual quality of the AMAST-a is similar to that of the rAMA, MIDAL, and ADMM-III. The AMAST-a is at least two times faster than the rAMA, MIDAL, and ADMM-III and at least 37 times faster than the PPBit.

Note that, for better visualization, all images in Fig. 8 are displayed after square root transformation is applied.

V. CONCLUSION

In this paper, we have proposed the alternating minimization algorithm with shifting technique (AMAST(-a)) for convex speckle reduction variational models - the exponential model and the I-divergence model - and have shown that our algorithm outperforms the rAMA, MIDAL, ADMM-III, and PPBit for speckle reduction in terms of the CPU time. The AMAST-a for the exponential model is the fastest method. Also, the AMAST(-a) obtains comparable PSNR when it is compared to the MIDAL, rAMA for the exponential model (6) and the ADMM-III, rAMA for the I-divergence model (7). We

¹Note that real SAR images in Fig. 8 are the absolute values of the complex raw data and it is theoretically known that the PDF of the speckle is the Rayleigh distribution rather than the Gamma distribution. But we empirically find that the Gamma distribution with $L = 3.5$ does better approximate the true noise distribution. See also [47].

note that the AMAST(-a) has computational advantages over other state-of-the-art algorithms, since no inner loops and no computation of inverses involving the Laplacian operator are required.

Our proposed algorithm can be applied to solve speckle reduction variational models with different regularizers other than local total variation (for example, nonlocal regularization [13], [29], [39], [46]). When we update values u, z, p by using (25), each component of those values can be evaluated in parallel. Therefore, the AMAST(-a) can be effectively accelerated on parallel hardware, and hence can be efficiently applied for huge size SAR images. These can be future research topics.

APPENDIX A PROOF OF LEMMA III.3

Let (u^*, z^*) be a solution of the problem (15) and p^* be a solution of its dual problem (21). Then (u^*, z^*, p^*) is a saddle point of the Lagrangian function (17) and so we have

$$\mathcal{L}(u^*, z^*, p) \leq \mathcal{L}(u^*, z^*, p^*) \leq \mathcal{L}(u, z, p^*) \quad \forall u \in \tilde{U}, z, p. \quad (32)$$

By the convexity of \tilde{U} , the first step in (19), and the second inequality of (32), we have that

$$\begin{aligned} f(u^{k+1}) - \langle p^k, \nabla u^{k+1} \rangle &\leq f(\tilde{u}_\beta^*) - \langle p^k, \nabla \tilde{u}_\beta^* \rangle, \\ f(u^*) + g(z^*) + \langle p^*, z^* - \nabla u^* \rangle &\leq f(\tilde{u}_{1-\beta}^*) + g(z^*) \\ &\quad + \langle p^*, z^* - \nabla \tilde{u}_{1-\beta}^* \rangle, \end{aligned}$$

where $\tilde{u}_\beta^* = (1 - \beta)u^{k+1} + \beta u^*$ and $\tilde{u}_{1-\beta}^* = \beta u^{k+1} + (1 - \beta)u^*$ with $0 < \beta \leq 1$. Adding the above two inequalities and

rearranging yield that

$$\begin{aligned} & f(u^{k+1}) - f(\tilde{u}_\beta^*) + f(u^*) - f(\tilde{u}_{1-\beta}^*) \\ & \leq \beta \langle p^k - p^*, \nabla u^{k+1} - \nabla u^* \rangle. \end{aligned}$$

This together with (16) and Lemma III.1 implies that

$$\begin{aligned} & \beta \langle \nabla_u f(\tilde{u}_\beta^*) - \nabla_u f(\tilde{u}_{1-\beta}^*), u^{k+1} - u^* \rangle + \sigma \beta^2 \|u^{k+1} - u^*\|_2^2 \\ & \leq \beta \langle p^k - p^*, \nabla u^{k+1} - \nabla u^* \rangle. \end{aligned}$$

This together with the mean value theorem, the convexity of \tilde{U} , the strong convexity of f , and the fact that $\tilde{u}_\beta^* - \tilde{u}_{1-\beta}^* = (1 - 2\beta)(u^{k+1} - u^*)$ yields that

$$\begin{aligned} & (\beta - 2\beta^2)\sigma \|u^{k+1} - u^*\|_2^2 + \sigma \beta^2 \|u^{k+1} - u^*\|_2^2 \\ & = (\beta - \beta^2)\sigma \|u^{k+1} - u^*\|_2^2 \\ & \leq \beta \langle p^k - p^*, \nabla u^{k+1} - \nabla u^* \rangle. \end{aligned}$$

By dividing both sides of the above inequality by β and taking $\beta \rightarrow 0$, we have that

$$\sigma \|u^{k+1} - u^*\|_2^2 \leq \langle p^k - p^*, \nabla u^{k+1} - \nabla u^* \rangle. \quad (33)$$

By subtracting $\alpha \|\nabla u^{k+1} - \nabla u^*\|_2^2$ from both sides, we obtain that

$$\begin{aligned} & \sigma \|u^{k+1} - u^*\|_2^2 - \alpha \|\nabla u^{k+1} - \nabla u^*\|_2^2 \\ & \leq \langle p^k - p^*, \nabla u^{k+1} - \nabla u^* \rangle - \alpha \|\nabla u^{k+1} - \nabla u^*\|_2^2 \\ & = \langle p^{k+1} - p^*, \nabla u^{k+1} - \nabla u^* \rangle \\ & \quad - \alpha \langle z^{k+1} - z^*, \nabla u^{k+1} - \nabla u^* \rangle, \end{aligned}$$

where the equation uses $\nabla u^* = z^*$ and the third step in (19), which is (22).

Next, by the second step in (19), we have that

$$\begin{aligned} & g(z^{k+1}) + \langle p^k, z^{k+1} \rangle + \frac{\alpha}{2} \|z^{k+1} - \nabla u^{k+1}\|_2^2 \\ & \leq g(\tilde{z}_\beta^*) + \langle p^k, \tilde{z}_\beta^* \rangle + \frac{\alpha}{2} \|\tilde{z}_\beta^* - \nabla u^{k+1}\|_2^2 \end{aligned}$$

where $\tilde{z}_\beta^* = (1 - \beta)z^{k+1} + \beta z^*$ with $0 < \beta \leq 1$. Using the convexity of g with $\tilde{z}_\beta^* = (1 - \beta)z^{k+1} + \beta z^*$ and rearranging terms in the above inequality yield that

$$\begin{aligned} & g(z^{k+1}) - (1 - \beta)g(z^{k+1}) - \beta g(z^*) \\ & \leq \beta \langle p^k, z^* - z^{k+1} \rangle \\ & \quad + \beta (\langle \alpha \nabla u^{k+1}, z^{k+1} - z^* \rangle + \langle \alpha z^{k+1}, z^* - z^{k+1} \rangle) \\ & \quad + \frac{\alpha}{2} (\beta^2 \|z^*\|_2^2 - 2\beta^2 \langle z^{k+1}, z^* \rangle + \beta^2 \langle z^{k+1}, z^{k+1} \rangle) \\ & = \beta \langle p^{k+1}, z^* - z^{k+1} \rangle + \frac{\alpha \beta^2}{2} \|z^{k+1} - z^*\|_2^2, \quad (34) \end{aligned}$$

where the equality uses the third step in (19).

By the second inequality of (32), we have that

$$\begin{aligned} & f(u^*) + g(z^*) + \langle p^*, z^* - \nabla u^* \rangle \\ & \leq f(u^*) + g(\tilde{z}_{1-\beta}^*) + \langle p^*, \tilde{z}_{1-\beta}^* - \nabla u^* \rangle, \end{aligned}$$

where $\tilde{z}_{1-\beta}^* = \beta z^{k+1} + (1 - \beta)z^*$ with $0 < \beta \leq 1$. Using the convexity of g and rearranging terms in the above inequality yield that

$$g(z^*) - \beta g(z^{k+1}) - (1 - \beta)g(z^*) \leq \beta \langle p^*, z^{k+1} - z^* \rangle. \quad (35)$$

Adding (34) and (35) and rearranging imply that

$$\beta \langle p^{k+1} - p^*, z^{k+1} - z^* \rangle \leq \frac{\alpha \beta^2}{2} \|z^{k+1} - z^*\|_2^2.$$

Dividing both sides of the above inequality by β and taking $\beta \rightarrow 0$ yield (23).

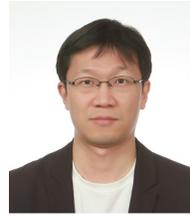
ACKNOWLEDGMENTS

We thank anonymous referees for their detailed comments and suggestions to improve the paper. Also, we thank G. Steidl and T. Teuber for providing us the code (ADMM-III) in [39].

REFERENCES

- [1] L. ALVAREZ, P.L. LIONS, AND J.M. MOREL, "Axioms and fundamental equations of image processing", *Arch. Rational Mech. Anal.*, 123 (1993) pp. 199-257.
- [2] G. AUBERT AND J.F. AUJOL, "A variational approach to remove multiplicative noise", *SIAM J. Appl. Math.*, 68 (2008), pp. 925-946.
- [3] G. AUBERT P. KORNPORST, "Mathematical problems in Image Processing", Second Edition, Springer-Verlag, 2006.
- [4] R. BAMLER, "Principles of synthetic aperture radar", *Surveys in Geophysics*, 21 (2000), pp. 147-157.
- [5] J.M. BIOUCAS-DIAS AND M.A.T. FIGUEIREDO, "Total variation restoration of speckled images using a split-Bregman algorithm", *Proc. the IEEE ICIP'09*, 2009.
- [6] J.M. BIOUCAS-DIAS AND M.A.T. FIGUEIREDO, "Multiplicative noise removal using variable splitting and constrained optimization", *IEEE Trans. Image Processing*, 19 (2010), pp. 1720-1730.
- [7] T.F. CHAN, S. OSHER, AND J. SHEN, "The digital TV filter and nonlinear denoising", *IEEE Trans. Image Processing*, 10 (2001), pp. 231-241.
- [8] A. CHAMBOLLE, "An algorithm for total variation minimization and applications", *J. Math. Imaging and Vis.*, 20 (2004), pp. 89-97.
- [9] A. CHAMBOLLE AND T. POCK, "A first-order primal-dual algorithm for convex problems with applications to Imaging", *J. Math. Imaging and Vis.*, 40 (2011), pp. 120-145.
- [10] M. CHENEY AND B. BORDEN, *Fundamentals of Radar Imaging*, SIAM, 2009.
- [11] M. J. COLLINS AND J.M. ALLAN, "Modeling and Simulation of SAR Image Texture", *IEEE Trans. Geoscience and Remote Sensing*, 47 (2009), pp. 3530-3546.
- [12] R. CORLESS, G. GONNET, D. D. HARE, D. JEFFREY, AND D. KNUTH, "On the Lambert W function", *Advances in Computational Mathematics*, 5 (1996), pp. 329-359.
- [13] C.A. DELEDALLE, L. DENIS, AND F. TUPIN, "Iterative weighted maximum likelihood denoising with probabilistic patch-based weights", *IEEE Trans. Image Processing*, 18 (2009), pp. 2661-2672.
- [14] "The probabilistic patch-based filter", Available: <http://perso.telecom-paristech.fr/~deledalle/ppb.php>.
- [15] L. DENIS, F. TUPIN, J. DARBON, AND M. SIGELLE, "SAR Image regularization with fast approximation discrete minimization", *IEEE Trans. Image Processing*, 18 (2009), pp. 1588-1600.
- [16] J. ECKSTEIN AND D. P. BERTSEKAS, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators", *Math. Prog.*, 55 (1992), pp. 293-318.
- [17] E. ESSER, "Applications of Lagrangian-based alternating direction methods and connections to split Bregman", Preprint, 2009.
- [18] E. ESSER, X. ZHANG, AND T. CHAN, "A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science", *SIAM J. Imaging Sci.*, 3 (2010), pp. 1015-1046.
- [19] M.A.T. FIGUEIREDO AND J.M. BIOUCAS-DIAS, "Restoration of Poissonian images using alternating direction optimization", *IEEE Trans. Image Processing*, 19 (2010), pp. 3133-3145.
- [20] S. DURAND, J. FADILI, AND M. NIKOLOVA, "Multiplicative noise removal using L1 fidelity on frame coefficients", *J. Math Imaging Vis.*, 36 (2010) pp. 201-226.
- [21] T. GOLDSTEIN AND S. OSHER, "The split Bregman method for L1 regularized problems", *SIAM J. Imaging Sci.*, 2 (2009), pp. 323-343.
- [22] J. GOODMAN, "Some fundamental properties of speckle", *J. Optical Society of America*, 66 (1976), pp. 1145-1150.
- [23] W. H. HENSLEY AND A. W. DOERRY, "Viewing GFF format SAR images with Matlab", *Sandia Report*, (2006).

- [24] D. HOEKMAN, "Speckle ensemble statistics of logarithmically scaled data", *IEEE Trans. Geoscience and Remote Sensing*, 29 (1991), pp. 180-182.
- [25] Y. M. HUANG, M. K. NG, AND Y. W. WEN, "A new total variation method for multiplicative noise removal", *SIAM J. Imaging Sci.*, 2 (2009), pp. 20-40.
- [26] C. R. JACKSON AND J. R. APEL, *Synthetic Aperture Radar Marine User's Manual*, 2005, Available at <http://www.sarusersmanual.com>.
- [27] D.T. KUAN, A.A. SAWCHUK, T.C. STRAND, AND P. CHAVEL, "Adaptive noise smoothing filter for images with signal-dependent noise", *IEEE Trans. Pattern Anal. Mach. Intell.*, 7 (1985), pp. 165-177.
- [28] S. LEE, "Overview of KOMPSAT-5 program, mission, and system", *IEEE Int. Geoscience and Remote Sensing Symposium*, (2010), pp. 797-800.
- [29] B. LU, L. ZHANG, AND F. XING, "SAR speckle reduction based on nonlocal means method", *Second Int. Conf. on Computer modeling and simulation*, (2010), pp. 156-159.
- [30] F. LI, M. K. NG, AND C. SHEN, "Multiplicative noise removal with spatially varying regularization parameters", *SIAM J. Imaging Sci.*, 3 (2010), pp. 1-20.
- [31] C. OLIVER AND S. QUEGAN, *Understanding Synthetic Aperture Radar Imaging*, SciTech Publishing, 2004.
- [32] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, 1970.
- [33] R.T. ROCKAFELLAR AND R.J.-B. WETS, *Variational Analysis*, Springer-Verlag, New York, 1998.
- [34] L. RUDIN, S. OSHER, AND E. FATEMI, "Nonlinear total variation based noise removal algorithms", *Physica D*, 60 (1992), pp. 259-268.
- [35] S. SETZER, "Operator splitting, Bregman methods and frame shrinkage in image processing", *Int. J. Comput. Vision*, 92 (2011), pp. 265-280.
- [36] S. SETZER, G. STEIDL, AND T. TEUBER, "Deblurring Poissonian images by split Bregman techniques", *J. Vis. Commun. Image R.*, 21 (2010), pp. 193-199.
- [37] J. SHI AND S. OSHER, "A nonlinear inverse scale space method for a convex multiplicative noise model", *SIAM J. Imaging Sci.*, 1 (2008), pp. 294-321.
- [38] L. DE SORAS, *Fast log function*, http://www.flipcode.com/archives/Fast_log_Function.shtml.
- [39] G. STEIDL AND T. TEUBER, "Removing multiplicative noise by Douglas-Rachford splitting", *J. Math Imaging Vis.*, 36 (2010), pp. 168-184.
- [40] C. TOMASI AND R. MANDUCHI, "Bilateral filtering for gray and color images", *Proc. of the IEEE Int. Conf. on Computer Vision*, Jan. 1998, pp. 839-846.
- [41] P. TSENG, "Applications of a splitting algorithm to decomposition in convex programming and variational inequalities", *SIAM J. Control and Optimization*, 29 (1991), pp. 119-138.
- [42] M. TUR, C. CHIN, AND J.W. GOODMAN, "When is speckle noise multiplicative", *Applied Optics*, 21 (1982), pp. 1157-1159.
- [43] Y. WANG, J. YANG, W. YIN, AND Y. ZHANG, "A new alternating minimization algorithm for total variation image reconstruction", *SIAM J. Imaging Sci.*, 1 (2008), pp. 248-272.
- [44] C. WU, J. ZHANG, AND X.-C. TAI, "Augmented Lagrangian method for total variation restoration with non-quadratic fidelity", *Inverse Problems and Imaging*, 5 (2011), pp. 237-261.
- [45] L. XIAO, L.L. HUANG, AND Z.H. WEI, "A Weberized total variation regularization-based image multiplicative noise removal algorithm", *EURASIP J. Adv. in Signal Processing*, 2010, Article ID 490384, 15 pages, doi:10.1155/2010/490384
- [46] S. YUN AND H. WOO, "Linearized proximal alternating minimization algorithm for motion deblurring by nonlocal regularization", *Pattern Recognition*, 44 (2011), pp. 1312-1326.
- [47] S. YUN AND H. WOO, "A new multiplicative denoising variational model based on m-th root transformation", *submitted to IEEE Trans. on Image Processing*, 2011.
- [48] M. ZHU AND T. CHAN, "An efficient primal-dual hybrid gradient algorithm for total variation image restoration", UCLA CAM-Report, 2008.



inverse problems.

Hyenkyun Woo He is a senior researcher at Department of Mathematical Sciences in Seoul National University, Seoul, Korea. He received the B.S. degree in electrical engineering from the Sogang University, Seoul, Korea, in 1996, and the M.S., and Ph.D degrees in mathematics from the Yonsei University, Seoul, Korea in 1998, and 2007, respectively. He was with IDIS and Samsung Electronics and developed several video surveillance systems, and image processing algorithms. His research interests include image and video signal processing, and



Sangwoon Yun He is in with Department of Mathematics Education in Sung Kyun Kwan University. He received Ph.D. in Mathematics from University of Washington in 2007. His research interest is convex and nonsmooth optimization, variational analysis, image processing.



Fig. 5: Performance comparison of the AMAST-a with the MIDAL [6] for the exponential model, with the ADMM-III [39] for the I-divergence model, and the nonlocal means based PPB [13]. The given numbers are PSNR/CPU time. The proposed algorithm AMAST-a is faster than all the other algorithms. The PSNR of the AMAST-a is even better than that of the PPB when $L = 1$. When $L = 3$, the PPB recovers periodic patterns better than all the other algorithms.

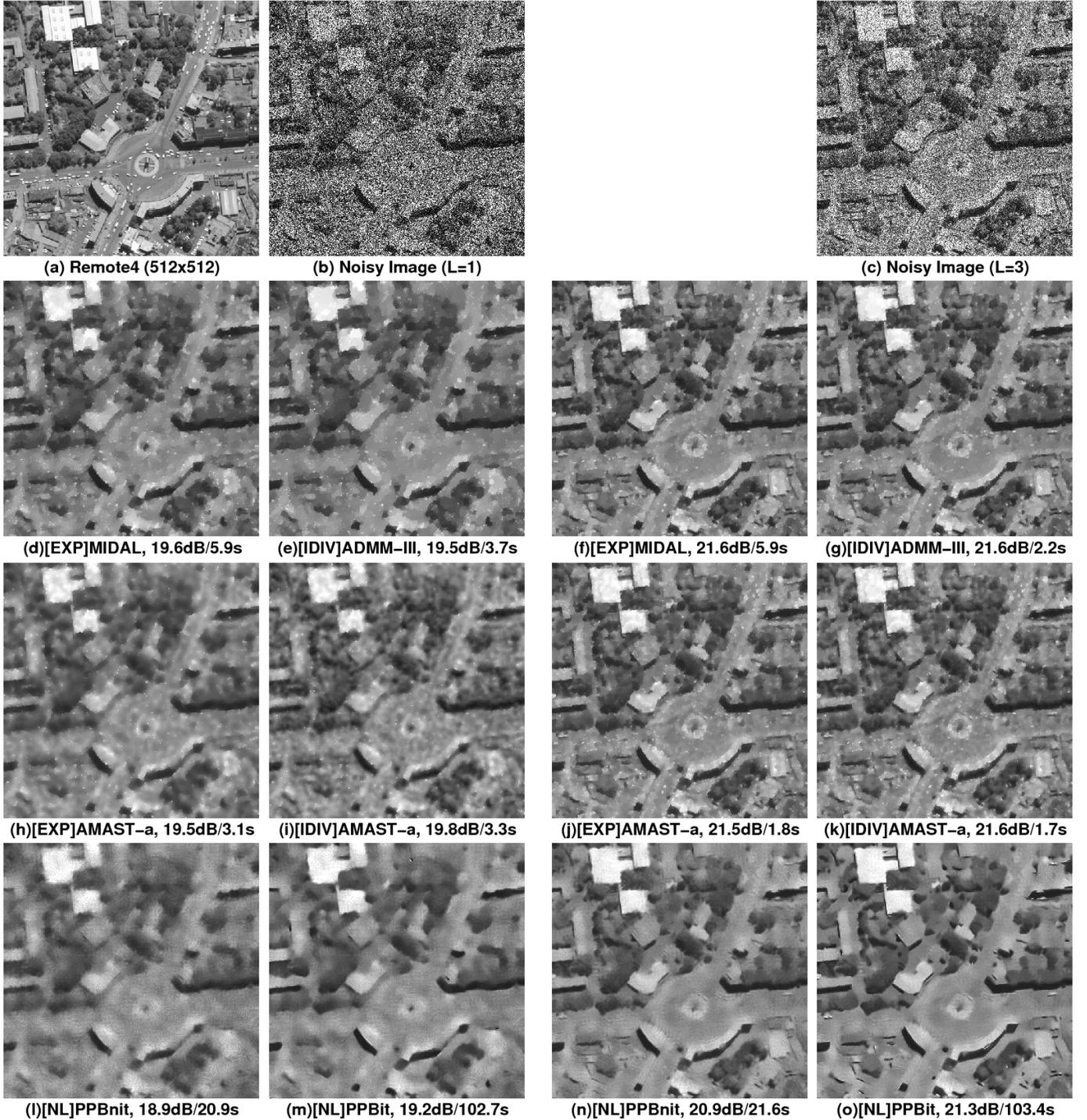


Fig. 6: Performance comparison of the AMAST-a with the MIDAL for the exponential model, with the ADMM-III for the I-divergence model, and the nonlocal means based PPB. The given numbers are PSNR/CPU time. The proposed method AMAST-a is faster than all the other algorithms. The PSNR of the AMAST-a is comparable to that of all the other algorithms. Since there is no periodic pattern, the PSNR of the PPB is less than that of the AMAST-a. As commented in [13], the PPB well preserves edges, but small and thin structures, such as cars on the road, are removed.

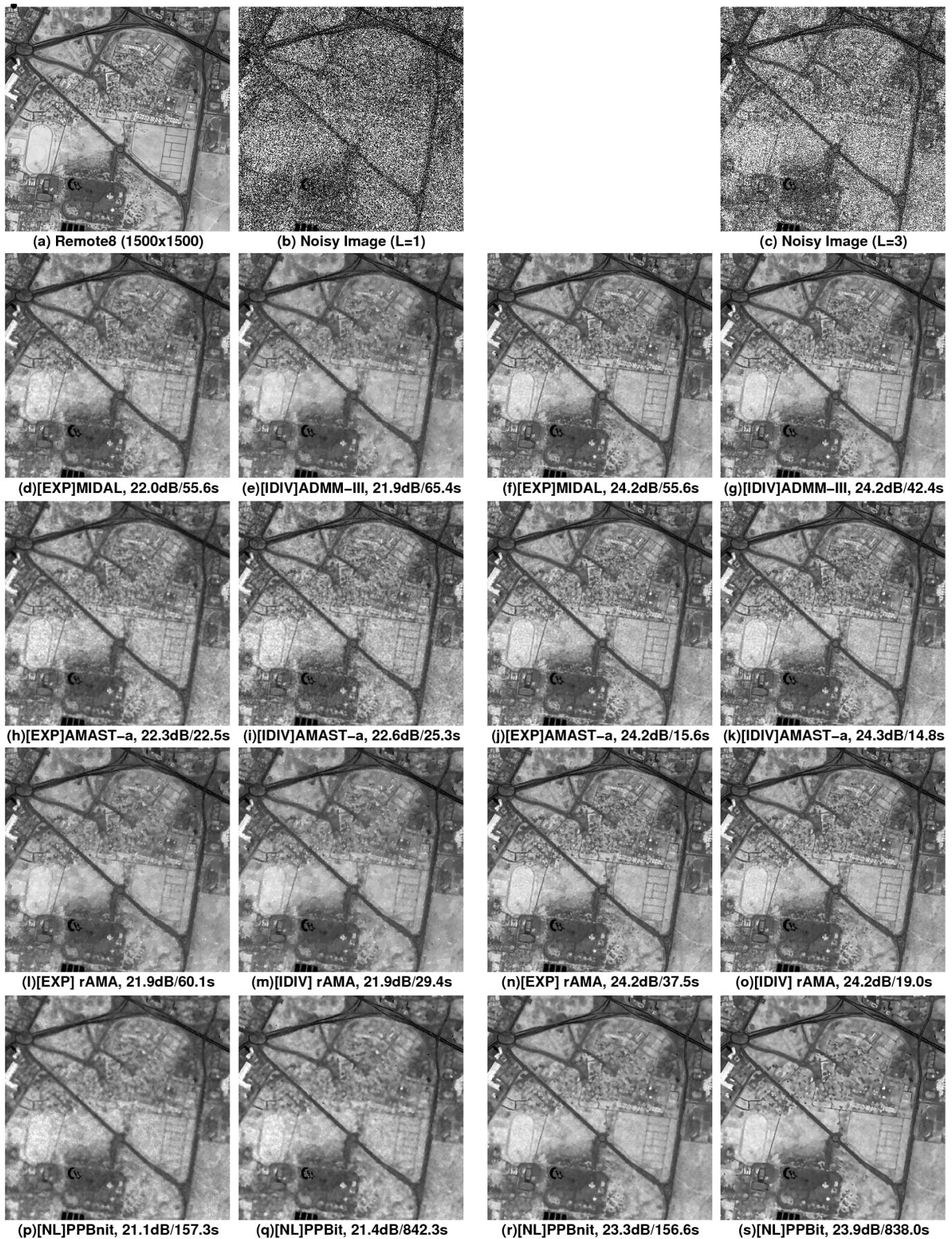


Fig. 7: Performance comparison of the AMAST-a and rAMA with the MIDAL for the exponential model, the ADMM-III for the I-divergence model, and the nonlocal means based PPB. The given numbers are PSNR/CPU time. For the exponential model, the proposed algorithm AMAST-a is at least 2.4 times faster than the MIDAL and rAMA with comparable PSNR. For the I-divergence model, the AMAST-a is the fastest algorithm with better PSNR.

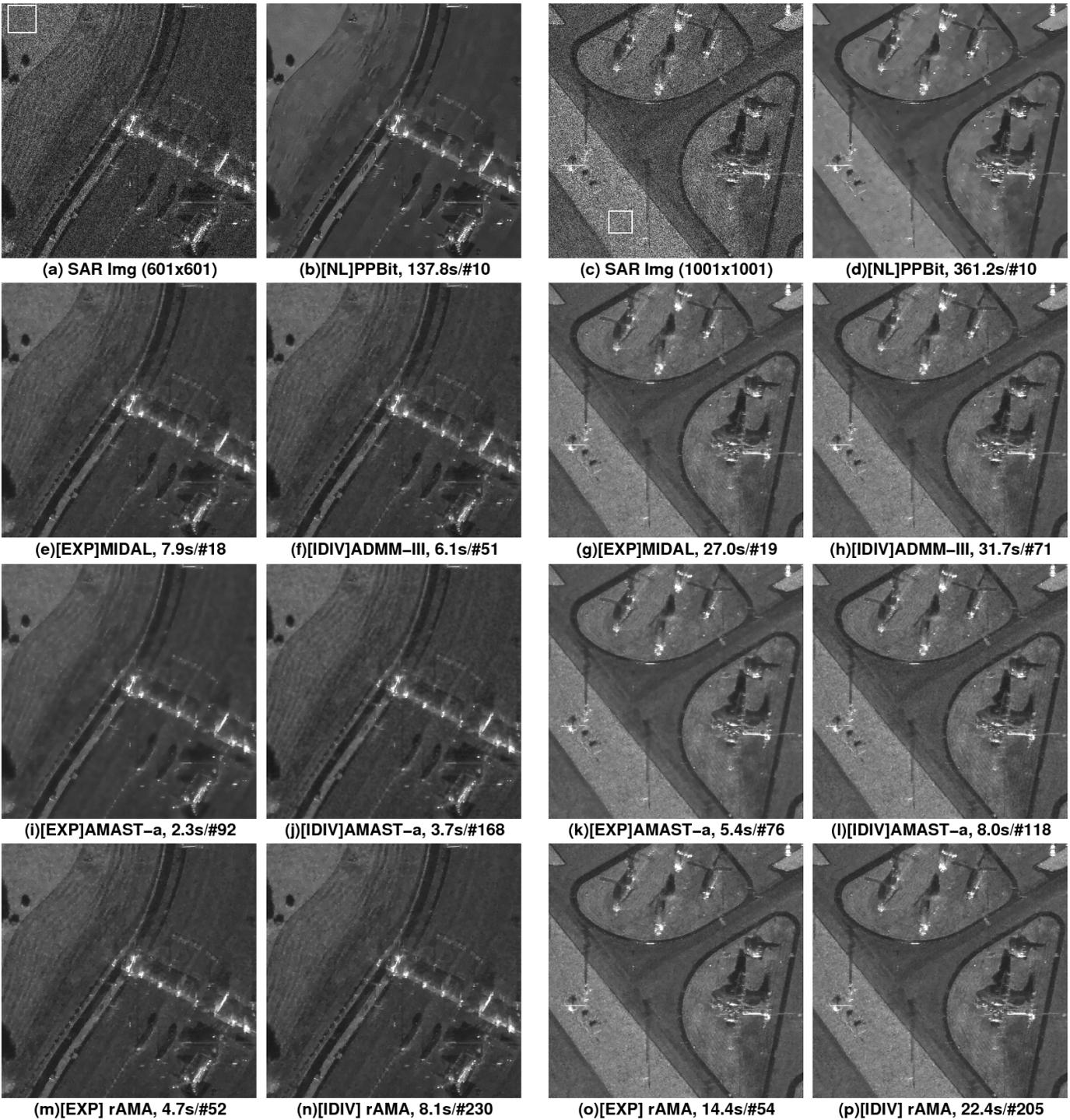


Fig. 8: Performance comparison of the AMAST-a with the MIDAL for the exponential model, with the ADMM-III for the ID-divergence model, and with the nonlocal means based PPBit. The given numbers are CPU time/#iterations. From homogeneous regions of real SAR images (white squares), we estimate the equivalent number of looks; $ENL=3.47$ for (a) and $ENL=3.51$ for (b). Therefore, we set $L=3.5$. Note that (a) and (c) are the absolute values of the complex raw data and the PDF of the speckle is theoretically the Rayleigh distribution. But we empirically find that the Gamma distribution with $L = 3.5$ does better approximate the true noise distribution.

The PPBit seems to remove the speckle in homogeneous regions well and preserve edges and small bright scatterers well. But it attenuates thin strip lines in (a). It seems that the visual quality of the AMAST-a is similar to that of the rAMA, MIDAL, and ADMM-III. The proposed algorithm AMAST-a is at least two times faster than the rAMA, MIDAL, and ADMM-III and at least 37 times faster than the PPBit. Note that, for better visualization, all images are displayed after square root transformation is applied.