Non-Convex Sparse Regularization Image Restoration via Atan Function

Zhijun Luo, Zhibin Zhu, Lirong Wang, and Yingying Li

***Abstract*—This study introduces an innovative approach within the sparse regularization framework, replacing the traditional total variation (TV) regularization with a non-convex regularizer based on the arctangent (Atan) function. The Atan-based regularization improves sparse representation and edge preservation through its non-convex properties, effectively overcoming the limitations of convex regularizers in detail reconstruction and artifact suppression. The model ensures overall convexity with careful parameter selection, thus maintaining guaranteed convergence during optimization. The alternating direction method of multipliers (ADMM) algorithm is employed to address the optimization challenges of the non-convex regularizer, demonstrating robust computational efficiency. Extensive experiments on image deblurring tasks show that the proposed method significantly outperforms traditional TV-based approaches in quantitative metrics and visual quality.**

***Index Terms*—Image restoration, Non-convex functions, Regularization, Sparse model**

I. Introduction

SPARSE regularization technology, a widely employed methodology in image processing, optimizes data representation and computational efficiency by exploiting the inherent sparsity of natural images. The fundamental principle of sparsity theory is based on approximating signal representations by minimizing non-zero coefficients in appropriate transform domains. This approach has established itself as a cornerstone technique with extensive applications across compressive sensing, signal reconstruction, and data compression [1]–[6]. In digital image processing, sparse approximation techniques have demonstrated remarkable success in critical tasks, including image compression, noise reduction, and restoration. By decomposing images into sparse linear combinations of basis functions within transformed domains, these techniques enable efficient processing and transmission while maintaining perceptual quality. Notably, sparse representation-based compression techniques achieve substantial file size reduction without compromising visual fidelity, effectively balancing storage efficiency and reconstruction accuracy. The image restoration problem under sparsity constraints can be mathematically formulated as an underdetermined linear system:

$$y := AX + e, \tag{1}$$

where X is the original image, A represents a linear operator, y is the observed image, and e represents noise. Generally speaking, the problem (1) is an ill-posed problem that poses difficulties in solving. The concept of total variation (TV) regularization was first introduced by Rudin, Osher, and Fatemi in their seminal work published in [7]. This model has since become a cornerstone in image processing and computer vision, typically expressed as

$$\min_X \frac{1}{2}\| AX - y\|^2 + \lambda\| X\|_{TV}, \tag{2}$$

where $\|\bullet\|$ represents the 2-norm, $\|X\|_{TV}$ represents the total variation norm of X, and $\lambda > 0$ is the regularization parameter. The basic idea of this model is to utilize the differences between pixels or edge information in the image to perform image denoising and preserve edges. T. Chan and J. Shen [8] proposed a mathematical model for image inpainting. In the TV model, the total variation of the image is used as a regularization term to promote fewer smooth regions and more edge information in the restored image. Both the data fitting and regularization terms in this model are convex functions, making the problem a convex optimization problem. Methods such as the primal-dual algorithm [9], split Bregman method [10-12], and alternating direction method of multipliers (ADMM) [13-15] can effectively solve these problems. Consequently, this model has gained widespread attention and application since its proposal.

Although TV regularization models excel in edge preservation, they may produce an unnatural "block effect" in smooth regions, also called the "stair casing effect." This effect renders areas that should transition smoothly into layered and discontinuous planes, resembling stairs. Non-convex models have garnered significant attention in recent years for their versatility in capturing diverse image features and circumventing issues such as stair effects in TV models. For instance, the nonlocal means (NLM) model, as proposed by K. Han [16] and Y. Liu [17], removes noise by computing the mean of local image regions, effectively preserving edge information and being adept at handling

Manuscript received June 1, 2024; revised June 8, 2025.

This work was supported in part by Hunan Provincial Natural Science Foundation of China (2023JJ50081and 2024JJ7268), and partly by the Scientific Research Fund of Hunan Provincial Education Department 22C0601.

Zhijun Luo is an associate professor of School of Mathematics and Finance, Hunan University of Humanities, Science and Technology, Loudi 417000, China (corresponding author; e-mail: ldlzj123@163.com).

Zhibin Zhu is a professor in the School of Mathematics and Computing Science, the Guangxi Colleges and Universities Key Laboratory of Data Analysis and Computation, Guilin University of Electronic Technology, Guilin 541004, China (e-mail: optimization_zhu@163.com).

Lirong Wang is a lecturer of School of Mathematics and Finance, Hunan University of Humanities, Science and Technology, Loudi 417000, China (e-mail: ldwlr1234@163.com).

Yingying Li is a teaching assistant of School of Mathematics and Finance, Hunan University of Humanities, Science and Technology, Loudi 417000, China (e-mail: yingyli@qq.com).

Gaussian and salt-and-pepper noise. In sparse representation models [18], images can be efficiently represented through linear combinations of a few essential elements. These models are extensively utilized in image compression, denoising, and super-resolution domains [19-22]. Low-rank models [23-25] posit that an image or sequence of images can be constituted by a limited number of fundamental images approximated by low-rank matrices. These models are extensively employed in background modeling and image completion [26-29]. L. Xu [30] introduced non-convex regularization terms, enhancing the TV model's ability to preserve details and mitigate stair effects. This category of models is delineated as follows

$$\min_X \frac{1}{2}\| AX - y\|^2 + \lambda P(X), . \tag{3}$$

where $P(X)$ is a non-convex function serves as an alternative to the TV norm. Non-convex functions, including the l_p norm [31-32], smoothly clipped absolute deviation (SCAD) [33-34], and minimax concave penalty (MCP) [35-36], serve as alternatives to the TV norm. These functions can more effectively accommodate the specific properties or structures of data. Consequently, these methods have emerged as a focal point in research, replacing traditional TV regularization terms [37].

This article further explores the application of non-convex regularization techniques in image processing. We introduce the arctangent (Atan) function as a replacement for the TV norm, resulting in the following model:

$$\min_X \frac{1}{2}\| AX - y\|^2 + \lambda \| X\|_{\text{Atan}TV}. \tag{4}$$

The notation $\|X\|_{\text{Atan}TV}$ denotes a non-convex regularization term crafted with the Atan function, as detailed in Section 2. The Atan function preserves the overall convexity of the objective function while reducing the penalty on significant coefficients as much as possible. This property of preserving convexity mirrors that of the MCP function. Numerical experiments indicate that the Atan function excels in terms of fitting performance.

The organization of this article is as follows: Section 2 revisits the definition and properties of the Atan function and introduces a non-convex model for the Atan function; Section 3 outlines the specific solution steps for the ADMM algorithm; Section 4 validates the approach through numerical experiments, demonstrating the effectiveness of both the model and the algorithm; Section 5 offers a concise summary of the content of this chapter.

II. Atan Regularization Model

This section begins by revisiting the definition and properties of Atan penalty functions and extends these concepts to multivariate functions. Subsequently, a non-convex sparse model incorporating the Atan penalty function is introduced. By integrating the difference-of-convex (DC) framework, we reformulate the non-convex problem as the difference between two convex functions to find the solution. This decomposition is facilitated by the DC algorithm, which iteratively resolves convex subproblems to progressively approximate the global optimum.

Definition 2.1 The Atan penalty function is defined as follows:

$$\phi_{\text{Atan}}(x;a) = \begin{cases} |x|, & a = 0 \\ \frac{2}{a\sqrt{3}}\left(\arctan\left(\frac{1+2a|x|}{\sqrt{3}}\right) - \frac{\pi}{6}\right), & a > 0 \end{cases} \tag{5}$$

where $a \geq 0$ is the parameter.

Property 2.1 If the parameter $a \geq 0$ satisfies $0 < a < 1/\lambda$, then

1) The function $\theta(x) = \lambda\phi_{Atan}(x;a) + \frac{1}{2}(x-t)^2$ is strictly convex;

2) The proximity operator corresponding to $\phi_{\text{Atan}}(x;a)$ is

$$prox_{\phi}(t;\lambda,a) = \arg\min_x \left\{ \lambda\phi_{\text{Atan}}(x;a) + \frac{1}{2}(x-t)^2 \right\}, \tag{6}$$

which is a continuous nonlinear function, where $0 < \lambda$ is called the proximity parameter. The proximity operator is an effective tool for solving sparse problems; its specific calculations can be found in [39].

Within the difference of convex (DC) framework, consider another function that corresponds to it

$$\psi_{\text{Atan}}(x) = |x| - \phi_{\text{Atan}}(x;a). \tag{7}$$

Figs 1-3 show the images of $\phi_{\text{Atan}}(x;a)$, $\psi_{\text{Atan}}(x)$, and $prox_{\phi}(t;\lambda,a)$, respectively. It can be seen that $\psi_{\text{Atan}}(x)$ is a convex function, which is crucial for establishing new non-convex models.

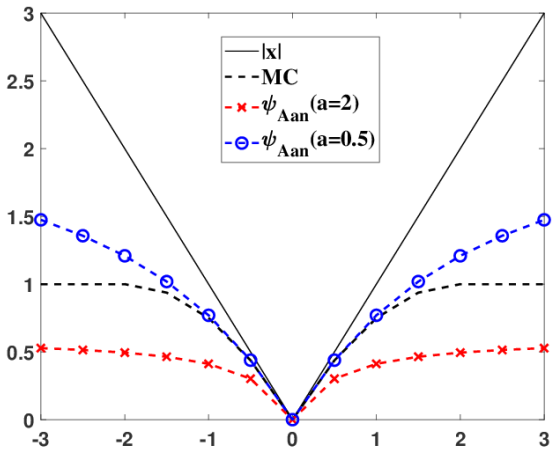


Fig. 1. The graph of $\phi_{\text{Atan}}(x;a)$.

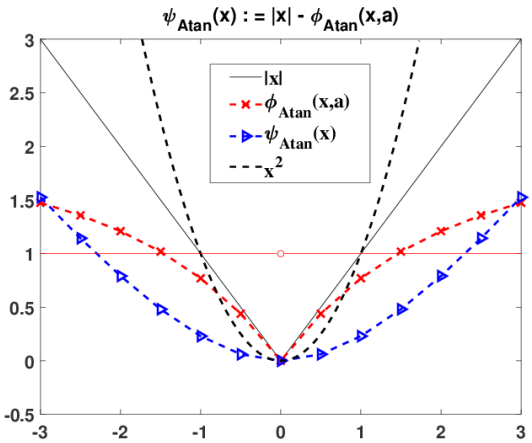


Fig. 2. The graph of $\psi_{\text{Atan}}(x)$.

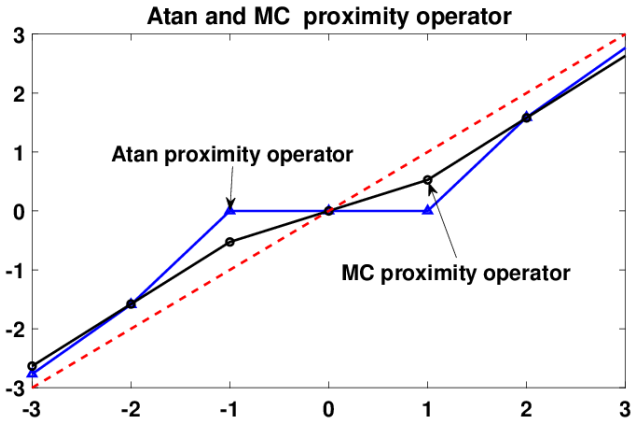


Fig. 3. The graph of $prox_{\phi}(t;\lambda,a)$.

The vector generalization of functions $\phi_{\text{Atan}}(x;a)$ and $\psi_{\text{Atan}}(x)$ yields the following expression:

$$\Psi_a(\mathbf{v})=\sum_{i=1}^{N}\psi_{\text{Atan}}(v_i),\quad \Phi_a(\mathbf{v})=\sum_{i=1}^{N}\phi_{\text{Atan}}(v_i),\quad \mathbf{v}\in \mathrm{R}^{\mathrm{N}},$$

where $\psi_{\text{Atan}}(v_i)$ and $\phi_{\text{Atan}}(v_i)$ are defined by (7). Furthermore, the following definition is obtained:

Definition 2.2 Define the AtanTV regularization term $\|X\|_{\text{Atan}TV}: R^N \to R$ for $X \in R^N$ as follows

$$\| X\|_{\text{Atan}TV}=\Phi_a(\mathbf{D}X)=\| \mathbf{D}X\|_1-\Psi_a(\mathbf{D}X), \quad (8)$$

where $\mathbf{D}$ is the gradient operator. In model (2), we consider using $\|X\|_{\text{Atan}TV}$ instead of the TV regularization term and obtain the following non-convex sparse model

$$\min_X \frac{1}{2}\| AX-y\|^2+\lambda\| X\|_{\text{Atan}TV}. \quad (9)$$

Lemma 2.1 Let $\lambda>0$ and $a\geq 0$. Define $\Phi(\bullet): R^N \to R$,

$$\Phi(X)=\lambda\| X\|_{\text{Atan}TV}+\frac{1}{2}\| y-AX\|^2. \quad (10)$$

If $\lambda a\mathbf{D}^T\mathbf{D}\preceq A^TA$, then Φ is convex.

Proof: Using (8), we write Φ as

$$\begin{aligned}\Phi(X)&=\lambda\| X\|_{\text{Atan}TV}+\frac{1}{2}\| y-AX\|^2\\&=\lambda\Phi_a(\mathbf{D}X)+\frac{1}{2}\left[\| \mathbf{y}\|^2+\| AX\|^2-2y^TAX\right]\\&=\left[\lambda\| \mathbf{D}X\|_1+\frac{1}{2}\| y\|^2-y^TAX\right]+\left[\frac{1}{2}\| AX\|^2-\lambda\Psi_a(\mathbf{D}X)\right].\end{aligned}$$

To prove that Φ is a convex function, it is sufficient to show that $\Psi(X)=\frac{1}{2}\| AX\|^2-\lambda\Psi_a(\mathbf{D}X)$ is convex. Using (8), we get

$$\Psi(X)=\underbrace{\frac{1}{2}(\| AX\|^2-\lambda a\| \mathbf{D}X\|^2)}_{\Psi_1}+\underbrace{\lambda(\frac{a}{2}\| \mathbf{D}X\|^2-\Psi_a(\mathbf{D}X))}_{\Psi_2}.$$

Based on the condition $\lambda a\mathbf{D}^T\mathbf{D}\preceq A^TA$, Ψ_1 is a convex function. According to Proposition 1 in reference [39], Ψ_2 is also a convex function. Thus, the proposition is established.

III. Numerical Algorithm

In this section, to effectively solve the problem (9), we first replace $\mathbf{D}X$ with an auxiliary variable z, and then transform the original optimization problem into an equation constraint form

$$\begin{aligned}&\min_X \ \lambda\Phi_a(z)\frac{1}{2}\| AX-y\|^2,\\&s.t. \quad \mathrm{z}=\mathbf{D}X\end{aligned} \quad (11)$$

The augmented Lagrangian function of (11) is

$$\mathcal{L}(X,z,\mathbf{w})=\lambda\Phi_a(z)+\frac{1}{2}\| y-\mathbf{A}X\|^2-\mathbf{w}^{\mathbf{T}}(z-\mathbf{D}X)+\frac{\rho}{2}\| \mathrm{z}-\mathbf{D}X\|^2,$$

where

$$\Phi_a(z)=\|z\|_1-\Psi_a(z),$$

$\mathbf{w}$ is the Lagrange multiplier and ρ is the penalty parameter. According to the traditional ADMM algorithm framework, the solution of (11) can be expressed as

$$\begin{cases}X^{k+1}=\operatorname{argmin}\mathcal{L}(X,z^k,\mathbf{w}^k)\\\quad=\min_X\{\frac{1}{2}\| y-\mathbf{A}X\|^2+\mathbf{w}^{kT}\mathbf{D}X+\frac{\rho}{2}\| \mathrm{z}^k-\mathbf{D}X\|^2\},\\z^{k+1}=\operatorname{argmin}\mathcal{L}(X^{k+1},z,\mathbf{w}^k)\\\quad=\min_z\{\lambda\Phi_a(z)-\mathbf{w}^{k\mathbf{T}}z+\frac{\rho}{2}\| \mathrm{z}-\mathbf{D}X^{k+1}\|^2\},\\\mathbf{w}^{k+1}=\mathbf{w}^k-(\mathbf{D}X^{k+1}-z^{k+1}).\end{cases}$$

Now consider the calculation for each subproblem. The X – subproblem can be derived from the first-order optimality condition.

$$X^{k+1}=(\rho\mathbf{D}^{\mathbf{T}}\mathbf{D}+\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-1}\mathbf{H}_k,$$

where $\mathbf{H}_k=\boldsymbol{\rho}\mathbf{D}^{\mathbf{T}}z^k+\mathbf{A}^{\mathbf{T}}y-\mathbf{D}^{\mathbf{T}}\mathbf{w}^k$.

For the $z-$subproblem,

$$z^{k+1}=\min_z\{\lambda\Phi_a(z)-\mathbf{w}^{k\mathbf{T}}z+\frac{\rho}{2}\| \mathrm{z}-\mathbf{D}X^{k+1}\|^2\}.$$

We expand and organize the quadratic terms in the above equation, omitting the constant to obtain

$$z^{k+1}=\min_z\{\Phi_a(z)+\frac{\rho}{2}\| \mathrm{z}-(\mathbf{D}X^{k+1}+\frac{\mathbf{w}^k}{\rho})\|^2\}.$$

Hence, we obtain the explicit solution of the above equation from the proximity operator as

$$z^{k+1}=prox_{\Phi_a,\rho}(\mathbf{D}X^{k+1}+\frac{\mathbf{w}^k}{\rho}).$$

The ADMM algorithm for solving problem (10) is provided below. Due to the inclusion of the AtanTV regularization term in the model, this algorithm is referred to as the AtanTV method.

Algorithm 3.1 (AtanTV):

Step 0: Initialization

Input $\rho,\lambda,a>0$ and tolerance $\varepsilon>0$.

Given $(X,z,\mathbf{w})=(X^0,z^0,\mathbf{w}^0)$, let $k:=0$;

Step 1: Compute

$$\begin{cases}X^{k+1}=(\rho\mathbf{D}^{\mathbf{T}}\mathbf{D}+\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-1}\mathbf{H}_k,\\z^{k+1}=prox_{\Phi_a,\rho}(\mathbf{t}^k),\\\mathbf{w}^{k+1}=\mathbf{w}^k-(\mathbf{D}X^{k+1}-z^{k+1}),\end{cases}$$

where $\mathbf{t}^k=\mathbf{D}X^{k+1}+\frac{\mathbf{w}^k}{\rho}$.

Step 2: Update

Set $\mathbf{w}^{k+1}=(X^{k+1},z^{k+1},\mathbf{w}^{k+1})$, if $\|\mathbf{w}^{k}-\mathbf{w}^{k+1}\|_2^2\le\varepsilon$, **STOP; otherwise, let** $k:=k+1$, go back to **Step 1**.

Lemma 2.1 demonstrates that, with suitable parameters, the proposed AtanTV model preserves the convexity of the objective function. Therefore, the algorithm introduced in this section can be theoretically analyzed using the framework described in [36] or through similar proofs found in [39].

IV. Experimental Results

This section will utilize the newly proposed AtanTV model for image deblurring experiments and compare it with the classical TV model [15] and the MCTV model [36]. In model (9), the matrix A represents the convolution operator K. For the experiments, a set of commonly used classic images of various sizes was selected, as shown in Fig. 4. From left to right, the images included are the "Cameraman" and "MRI" images of size 256×256, the "Lena" image of size 512×512, and the "Boat" image of size 1024×1024. All simulations were performed using Matlab R2015a on a PC with an Intel Core i5 CPU @ 2.20 GHz and 8 GB of RAM.

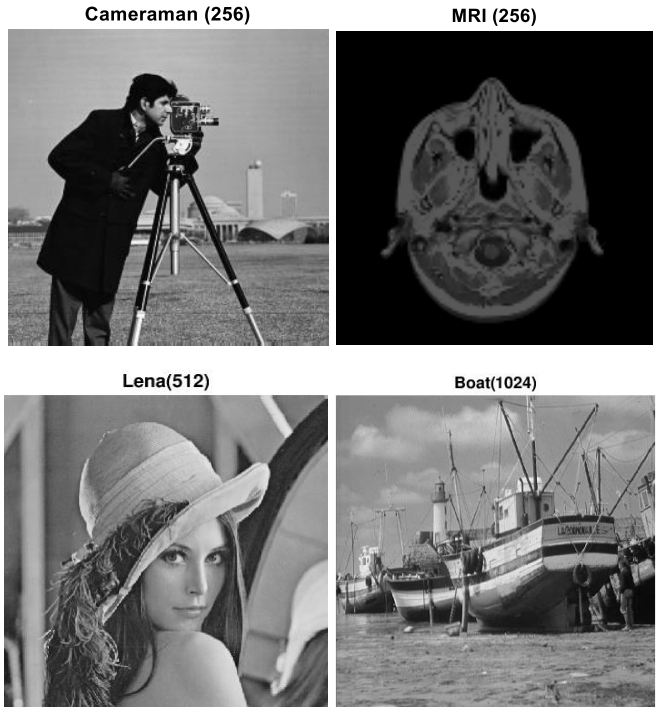


Fig. 4. Experimental images

We utilized the fspecial function from the Matlab toolbox in the perturbation simulation. For the experiments, different parameter settings were applied to images of various sizes to enhance the generalizability of the results. To evaluate the experimental outcomes of other models, we employed PSNR, RE, and SSIM metrics to measure the image reconstruction quality.

A. AtanTV Model Parameter Testing

In applying non-convex sparse models, parameter selection remains a critical unresolved issue that significantly influences the results. This article introduces a model named AtanTV, with the penalty parameter a ranging from 0 to $1/\lambda$, where λ is the regularization parameter. To explore the impact of various parameter combinations on the model's outcomes, we systematically tested multiple values for a and λ using the default settings of the fspecial ('Gaussian') function in Matlab for the image "Cameraman (256)" as presented in Table I. The experimental findings reveal a high sensitivity of parameter selection to both the model's convergence rate and the final solution's quality.

Additionally, Fig.5 visually depicts the relationship between λ and PSNR values. Further analysis indicates that the penalty parameter a exhibits relative stability in image deblurring applications. Specifically, when λ is set to 0.01 or 0.001, the difference between the minimum and maximum PSNR values does not exceed 8%. Therefore, careful parameter selection can enhance the model's stability and significantly improve the accuracy of the solution. This highlights the importance of fine-tuning hyperparameters to achieve optimal performance.

TABLE I
PSNR VALUES UNDER DIFFERENT PARAMETERS

λ	a	PSNR
10	0.1	32.774 dB
	0.01	29.252 dB
	0.001	23.574 dB
1	1	46.985 dB
	0.1	46.169 dB
	0.01	44.543dB
	0.001	41.654 dB
0.1	10	47.079dB
	1	46.999 dB
	0.1	47.058 dB
	0.01	47.000dB
	0.001	46.901 dB
0.01	100	47.052 dB
	10	47.029 dB
	1	47.050 dB
	0.1	**47.105 dB**
	0.01	47.070 dB
	0.001	47.050 dB
0.001	1000	47.027 dB
	100	47.022 dB
	10	47.030 dB
	1	47.049 dB
	0.1	47.043 dB
	0.01	47.041 dB
	0.001	47.027 dB

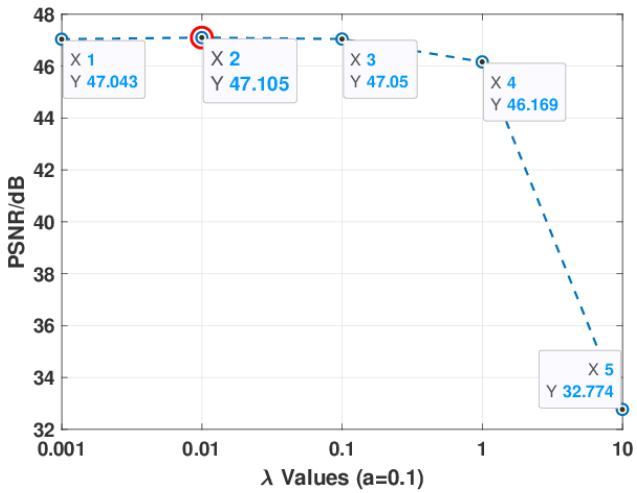


Fig. 5. λ $(a=0.1)$ vs. PSNR Curve

Taking into account the model types discussed subsequently, as well as the results presented in Table I and Fig. 5, we have chosen parameters $\lambda = 0.001, \rho = 1.5$, and $a = 0.1$ for the AtanTV model. The number of iterations was fixed at 250.

B. Evaluation of the New Method Amid Different Blur Disturbances

This sub-section explores the performance of the proposed new method in image restoration under three common types of blur disturbances. Fig. 6 illustrates the restoration results of the 256×256 "Cameraman" image under various blur perturbations using the AtanTV method. Table II presents the SSIM and CPU values for each method. The findings show that the AtanTV method excels in effectively managing Gaussian noise.

TABLE II
CPU TIME AND SSIM VALUES

Blur type	SSIM	CPU/s
('Gaussian',[5 5], 1)	1.0000	3.319
('motion', 9, 0)	0.7994.	4.174
('average',3)	0.8919.	2.741

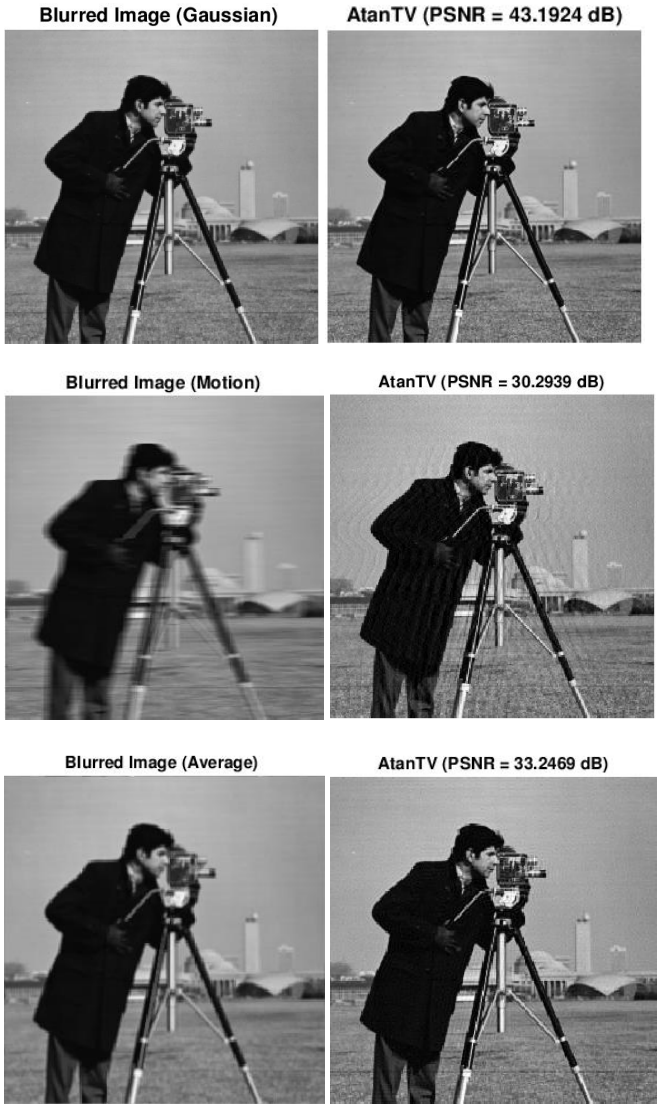


Fig. 6. AtanTV method recovery results

C. Performance Comparison among Different Models

In this subsection, we compare our method with existing TV and MCTV methods by conducting four sets of tests to evaluate the image reconstruction performance on images of varying sizes under different levels of blur interference. Fig. 7 shows the reconstruction results of the three methods.

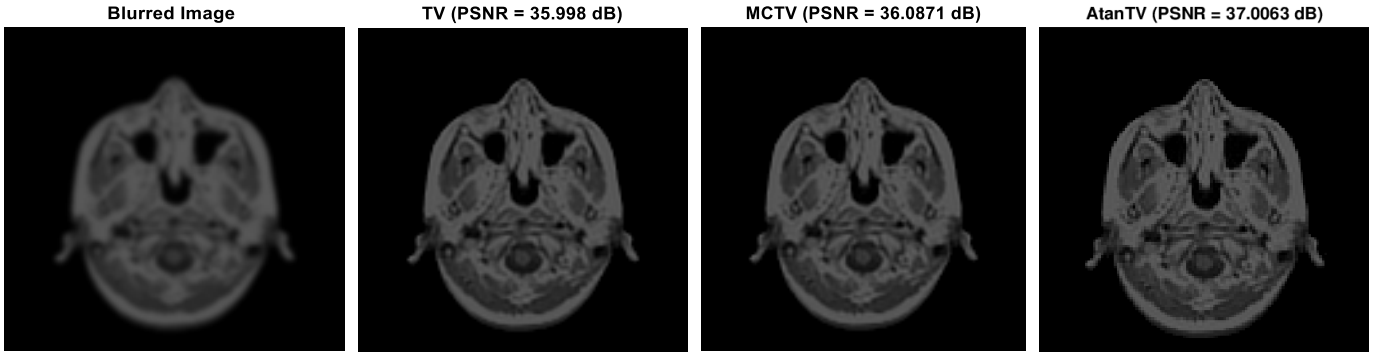
The deblurring performance of three methods (TV, MCTV, and AtanTV) on MRI (256) images is illustrated in Figs. 7(a2)-(a4) and (b2)-(b4). The findings reveal that AtanTV significantly outperforms TV and MCTV in restoring images under both mild Gaussian blur ('Gaussian,' [5 5], 1) and severe Gaussian blur ('Gaussian,' [10 10], 10). This indicates that AtanTV not only enhances image clarity but also preserves essential structural details often compromised by standard techniques.

Fig. 7 (c2)-(c4) shows the deblurring results for Lena (512) images subjected to heavier Gaussian blur, comparing the efficacy of TV, MCTV, and AtanTV. These results highlight the distinct approaches and impacts of each method on Gaussian blur. The AtanTV method, integrating TV regularization with non-convex function information, effectively restores fine image features and structures. This conclusion is further corroborated by the results in Fig. 7 (d2)-(d4).

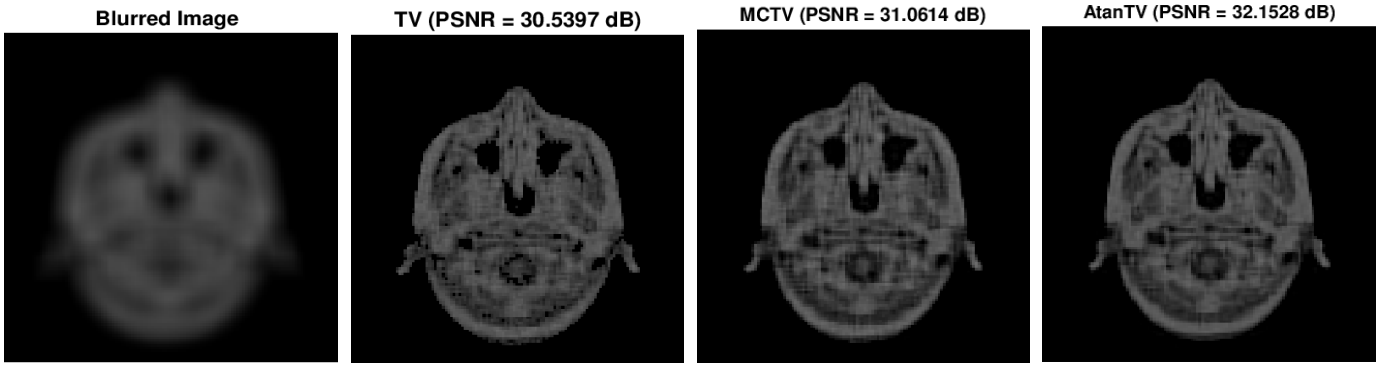
Table III illustrates the performance of four methods (TV, MCTV, AtanTV) across different images and blur settings, as assessed by relative error (RE%) and peak signal-to-noise ratio (PSNR/dB). For the MRI image (256x256) under Gaussian blur conditions ([5 5], 1) and ([10 10], 10), the AtanTV method achieves relative errors (RE%) of 9.06% and 15.82%, with PSNR values of 37.01dB and 32.15dB, respectively, outperforming both TV and MCTV methods. For the Lena image (512), AtanTV achieves an RE% of 5.04% and a PSNR of 31.76 dB, outperforming TV (RE% of 5.35%, PSNR of 30.83 dB) and MCTV (RE% of 5.29%, PSNR of 31.12 dB). Similarly, for the Boat image (1024x1024), AtanTV achieves a RE% of 4.53% and a PSNR of 32.27dB, surpassing TV and MCTV. In summary, the AtanTV method consistently delivers lower relative errors and higher PSNR across various deblurring scenarios, underscoring its efficacy in image deblurring applications.

TABLE III
NUMERICAL RESULTS

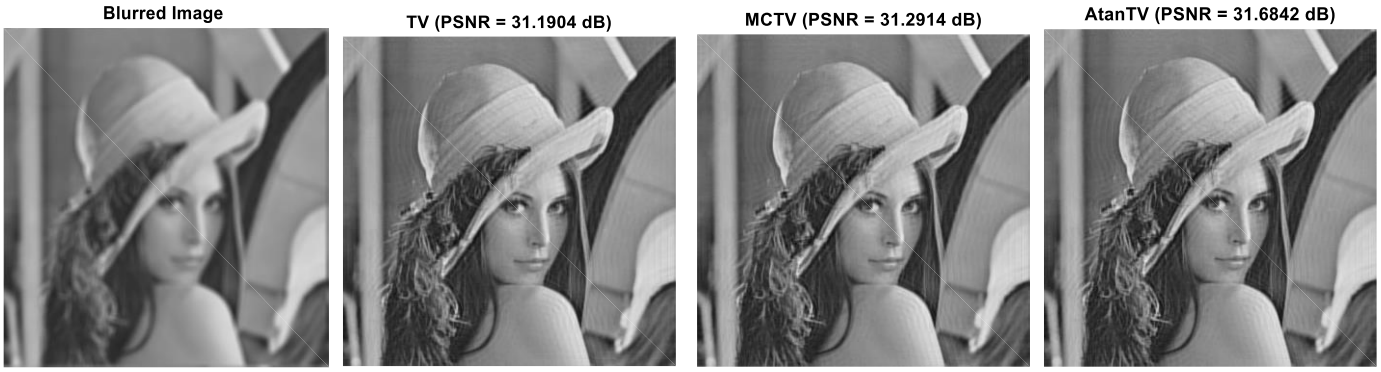
Image	Blur type	Method	RE%	PSNR/dB
MRI (256)	('Gaussian', [5 5], 1)	TV	10.29	36.00
		MCTV	10.18	36.08
		AtanTV	**9.06**	**37.01**
MRI (256)	('Gaussian', [10 10], 10)	TV	19.52	30.54
		MCTV	18.29	31.06
		AtanTV	**15.82**	**32.15**
Lena (512)	('Gaussian', [10 10], 10)	TV	5.35	31.19
		MCTV	5.29	31.29
		AtanTV	**5.04**	**31.68**
Boat (1024)	('Gaussian', [15 15], 15)	TV	4.89	31.64
		MCTV	4.81	31.77
		AtanTV	**4.53**	**32.27**



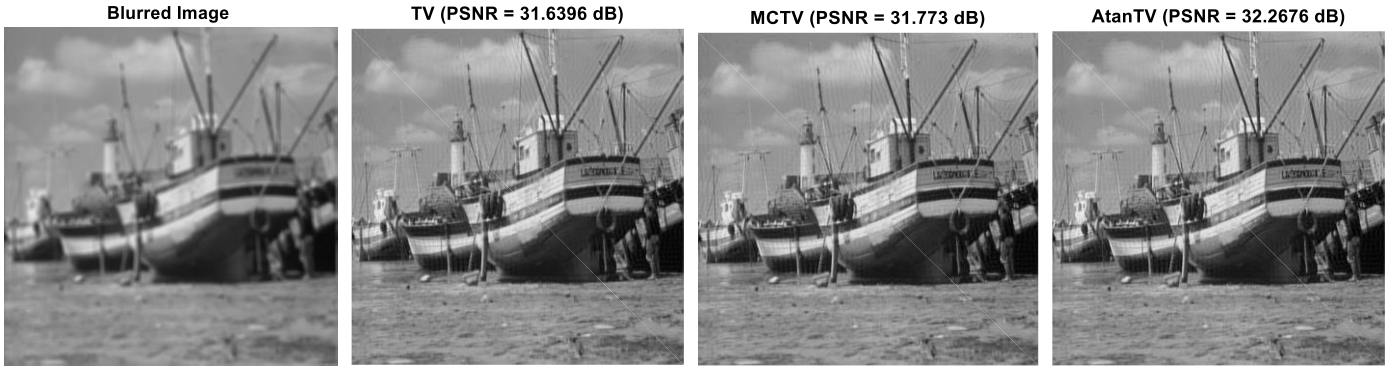
(**a1**) (**a2**) (**a3**) (**a4**)



(**b1**) (**b2**) (**b3**) (**b4**)



(**c1**) (**c2**) (**c3**) (**c4**)



(**d1**) (**d2**) (**d3**) (**d4**)

Fig. 7. Images restored by TV, MCTV, AtanTV models

V. Conclusion

This paper introduces a sparse, non-convex regularization model based on the arctangent function, which, under suitable parameters, facilitates convex optimization solutions for non-convex models. It effectively mitigates blur in image deblurring, enhancing clarity while retaining essential features and reducing noise. The model's versatility extends beyond image deblurring, with potential applications in dynamic imaging, machine learning, and signal processing.

Acknowledgment

The authors thank Prof. Ivan Selesnick for providing his Matlab code and making it freely available for download on his homepage (http://eeweb.poly.edu/iselesni/index.html).

References

[1] M. Aharon, M. Elad, and A. Bruckstein, "MK-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311-4322, 2006.

[2] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 460-489, 2005.

[3] J. Wright, D. Nowak, M. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Transactions on Signal Processing*, vol. 57, no.7, pp. 2479-2493, 2009.

[4] L. Wang, S. M. Deni, and Z. Zahid, "A LogTV nonconvex regularization model for magnetic resonance imaging," *Engineering Letters*, vol. 31, no.2, pp. 702-711, 2023.

[5] J. He, Y. Liu, and W. Lv, "A modified generalized relaxed splitting preconditioner for generalized saddle point problems," *IAENG International Journal of Computer Science*, vol. 50, no. 1, pp. 282-286, 2023.

[6] A. Mahdaoui, A. Ouahabi, and M. Moulay, "Image denoising using a compressive sensing approach based on regularization constraints," *Sensors*, vol. 22, no. 6, pp. 2199, 2022.

[7] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259-268, 1992.

[8] J. Shen and T. Chan, "Mathematical models for local nontexture inpaintings," *SIAM Journal on Applied Mathematics*, vol. 62, no. 3, pp. 1019-1043, 2002.

[9] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, pp. 120-145, 2011.

[10] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for l_1-minimization with applications to compressed sensing," *SIAM Journal on Imaging Sciences*, vol. 1, no. 1, pp. 143-168, 2008.

[11] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323-343, 2009.

[12] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends ® in Optimization*, vol. 1, no. 3, pp. 127-239, 2014.

[13] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 681-695, 2010.

[14] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1-122, 2011.

[15] J. Yang, Y. Zhang, and W. Yin, "A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 288-297, 2010.

[16] K. Han, Y. Wang, H. Chen, X. Chen, and J. Guo, "A survey on vision transformer," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 87-110, 2022.

[17] Y. Liu, J. Wang, X. Chen, Y. Guo, and Q. Peng, "A robust and fast nonlocal means algorithm for image denoising," *Journal of Computer Science and Technology*, vol. 23, no. 2, pp. 270–279, 2008.

[18] M. Elad, J. Starck, P. Querre, and D. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Applied and Computational Harmonic Analysis*, vol. 19, no. 3, pp. 340-358, 2005.

[19] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Transactions on Image Processing*, vol. 17, no. 1, pp. 53-69, 2007.

[20] C. Guillemot and O. Le Meur, "Image inpainting: Overview and recent advances," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 127-144, 2013.

[21] X. Zhang, D. Zhai, T. Li, Y. Zhou, and Y. Lin, "Image inpainting based on deep learning: A review," *Information Fusion*, vol. 90, no. 3, pp. 74-94, 2023.

[22] Y. Liu, S. Dutta, A. W. K. Kong, and C. Yeo, "An image inpainting approach to short-term load forecasting," *IEEE Transactions on Power Systems*, vol. 38, no. 1, pp. 177-187, 2022.

[23] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 210-227, 2008.

[24] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation: A low-rank approach," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 700-711, 2012.

[25] Y. Chen, Y. Guo, Y. Wang, D. Wang, C. Peng, and G. He, "Denoising of hyperspectral images using non-convex low-rank matrix approximation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 9, pp. 5366-5380, 2017.

[26] C. Geng, S. Huang, and S. Chen, "Recent advances in open set recognition: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 10, pp. 3614-3631, 2020.

[27] J. Peng, W. Sun, H. Li, W. Li, X. Meng, C. Ge, and Q. Du, "Low-rank and sparse representation for hyperspectral image processing: A review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 10, no. 1, pp. 10-43, 2021.

[28] S. Liao, S. Fu, Y. Li, and H. Han, "Image inpainting using non-convex low rank decomposition and multidirectional search," *Applied Mathematics and Computation*, vol. 452, no. 9, pp. 128048, 2023.

[29] U. S. Kamilov, C. A. Bouman, G. T. Buzzard, and B. Wohlberg, "Plug-and-play methods for integrating physical and learned models in computational imaging: theory, algorithms, and applications," *IEEE Signal Processing Magazine*, vol. 40, no. 1, pp. 85-97, 2023.

[30] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via L_0 gradient minimization," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 6, pp. 1-12, 2011.

[31] J. Huang, J. Horowitz, and S. Ma, "Asymptotic properties of bridge estimators in sparse high-dimensional regression models," *The Annals of Statistics*, vol. 36, no. 2, pp. 587-613, 2008.

[32] Z. Xu, X. Chang, F. Xu, and H. Zhang, "$L_{1/2}$ Regularization: A thresholding representation theory and a fast solver," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 7, pp. 1013-1027, 2012.

[33] J. Fan, and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," *Journal of the American Statistical Association*, vol. 96, no. 456, pp. 1348-1360, 2001.

[34] Z. Luo, Z. Zhu, and B. Zhang, "An SCADTV nonconvex regularization approach for magnetic resonance imaging," *IAENG International Journal of Computer Science*, vol. 48, no. 4, pp. 1013-1020, 2021.

[35] C. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *The Annals of Statistics*, vol. 38, no. 2, pp. 894-942, 2010.

[36] I. Selesnick, A. Lanza, S. Morigi, and F. Sgallari, "Non-convex total variation regularization for convex denoising of signals," *Journal of Mathematical Imaging and Vision*, vol. 62, no. 2, pp. 825-841, 2020.

[37] G. Zhu, X. Lv, L. Jiang, X. Sun, and B. Fang, "Non-convex regularization for convex image smoothing," *Signal Processing*, vol. 205, no. 4, pp. 108862, 2023.

[38] T. Yang, X. Yi, J. Wu, Y. Yuan, D. Wu, and Z. Meng, "A survey of distributed optimization," *Annual Reviews in Control*, vol. 47, no. 6, pp. 278-305, 2019.

[39] Y. Ding and I. Selesnick, "Artifact-free wavelet denoising: non-convex sparse regularization, convex optimization," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1364-1368, 2015.