

Research Article

Measurement Matrix Optimization via Mutual Coherence Minimization for Compressively Sensed Signals Reconstruction

Ziran Wei ^{1,2,3} Jianlin Zhang ^{1,3} Zhiyong Xu ^{1,3} and Yong Liu²

¹Key Laboratory of Optical Engineering, Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, China

²School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China

³University of Chinese Academy of Sciences, Beijing 100049, China

Correspondence should be addressed to Ziran Wei; weiziran15@mails.ucas.ac.cn

Received 15 June 2020; Revised 13 August 2020; Accepted 24 August 2020; Published 29 September 2020

Academic Editor: Krzysztof Okarma

Copyright © 2020 Ziran Wei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For signals reconstruction based on compressive sensing, to reconstruct signals of higher accuracy with lower compression rates, it is required that there is a smaller mutual coherence between the measurement matrix and the sparsifying matrix. Mutual coherence between the measurement matrix and sparsifying matrix can be expressed indirectly by the property of the Gram matrix. On the basis of the Gram matrix, a new optimization algorithm of acquiring a measurement matrix has been proposed in this paper. Firstly, a new mathematical model is designed and a new method of initializing measurement matrix is adopted to optimize the measurement matrix. Then, the loss function of the new algorithm model is solved by the gradient projection-based method of Gram matrix approximating an identity matrix. Finally, the optimized measurement matrix is generated by minimizing mutual coherence between measurement matrix and sparsifying matrix. Compared with the conventional measurement matrices and the traditional optimization methods, the proposed new algorithm effectively improves the performance of optimized measurement matrices in reconstructing one-dimensional sparse signals and two-dimensional image signals that are not sparse. The superior performance of the proposed method in this paper has been fully tested and verified by a large number of experiments.

1. Introduction

The theory of compressive sensing (compressed sensing, CS) was proposed by Donoho et al. in 2006. The main idea of compressed sensing is to combine signal sampling and signal compression with the premise that the original signal is sparse or can be sparsely represented [1]. The sampling of compressed sensing completes the compression by reducing signal dimensions without the intermediate stage of Nyquist Sampling [2]. Then, the original signal is reconstructed directly with less sampling number than Nyquist Sampling through the corresponding signal reconstruction algorithms, which saves transmission and storage costs and reduces computational complexity and energy consumption [1, 3–5]. Because compressed sensing can sample signals with a lower frequency than that required by Shannon–Nyquist Theorem, compressed sensing improves the compressibility and

recoverability for signals [3, 6, 7]. These properties make CS have a wide range of applications in many fields of signals processing [2, 8]. Some typical CS-based applications include single-pixel imaging [9–13], recovery of images and video [14], wireless image sensor networks [15], applications in classification problem [16, 17], and some biomedical signals processing fields [4, 18–21], such as nuclear magnetic resonance imaging (MRI) [22–24] and electrocardiographic (ECG) signals processing [25, 26].

In the sampling process of CS, the original discrete signal is denoted as x , $x \in C^n$. The measured or sampled signal y is obtained by the measurement matrix Φ with a size $m \times n$, $m \ll n$. So, we have the following:

$$y = \Phi x + e, \quad (1)$$

where e is the Gaussian white noise with variance σ^2 .

The original signal x is reconstructed from y . Because the dimensions of y are much smaller than those of x , an underdetermined equation must be solved. To solve the underdetermined equation, it is required that x should be sparse or can be represented sparsely [6, 27]. The sparsity level of discrete signal is usually expressed by the L0 norm $\|x\|_0$, and $\|x\|_0$ denotes the number of nonzero elements in x .

According to equation (1), we can get the following constrained optimization expression by adding $\|x\|_0$:

$$\begin{aligned} \hat{x} &= \arg \min_x \|x\|_0 \\ \text{s.t. } &\|\Phi x - y\|_2^2 \leq \varepsilon, \end{aligned} \quad (2)$$

where ε is a nonnegative real parameter.

We use the orthogonal sparsifying basis Ψ to represent x sparsely, $\Psi = [\psi_1, \psi_2, \psi_3, \dots, \psi_n]$. Thus,

$$x = \sum_1^n s_i \psi_i = \Psi s, \quad (3)$$

where $s = [s_1, s_2, s_3, \dots, s_n]^T$, $[\]^T$ is the transpose operator and s is the sparse coefficients vector. In this paper, the discrete wavelet transform (DWT) is used as the sparsifying basis for its good sparsifying performance [28, 29].

So in order to obtain s , a new expression is obtained according to equation (3):

$$\begin{aligned} \hat{s} &= \arg \min_s \|s\|_0 \\ \text{s.t. } &\|\Phi \Psi s - y\|_2^2 \leq \varepsilon, \end{aligned} \quad (4)$$

where \hat{s} is the estimated solution of s , and the estimation of the original signal is acquired as $\hat{x} = \Psi \hat{s}$. $\Theta = \Phi \Psi$ is called the equivalent dictionary. When x is a sparse signal, Ψ can be considered as an identity matrix of size $n \times n$ and Θ is almost equivalent to Φ .

One can know that the sparsity level of the original signal, the design of the measurement matrix and the signal reconstruction algorithm are the three main parts of CS theory [30]. Some recent works have been done to optimize CS theory, and there are different optimization approaches, technical principles, and application scenarios in these works [31–33], such as adapted compressed sensing and optimized projections for compressed sensing, which further enriches and optimizes the compressed sensing theory and approaches. For CS-based signals reconstruction algorithms, greedy pursuit algorithms [34–38], minimum norm optimization algorithms [39–41] and iterative threshold algorithms [42, 43] are usually used. The orthogonal matching pursuit (OMP) algorithm is one of greedy pursuit algorithms and has stable performance in reconstructing signals because of its simplicity and efficiency to get an approximated solution [36–38]. Thus, the OMP algorithm is used to reconstruct signals in this paper.

It has been proved that the lower mutual coherence between the measurement matrix and the sparsifying matrix is usually helpful to ensure the successful reconstruction of the signal [44], and random measurement matrices may be incoherence to most fixed orthogonal sparsifying matrices with large probability [3, 45]. Commonly used random

measurement matrices include 0-1 binary sparse random matrix, Gaussian random matrix, Bernoulli random matrix, and part Hadamard random matrix. In addition to the conventional random measurement matrices mentioned above, well optimized measurement matrices can often lead to a good performance of signals reconstruction [46–50], and measurement matrices are also usually optimized for some specific application scenarios [51–53]. For example, in the literature [47], Elad proposed a very classic algorithm that tried to minimize the average mutual coherence between the measurement matrix and sparsifying basis by the method of iteratively updating the measurement matrix when the sparsifying basis is fixed. Although Elad's method effectively improved the performance of CS-based signals reconstruction, the accuracy of signals reconstruction was not enough and the method was time-consuming for it requires a lot of iterations. In Wang's work [48], Wang proposed a weighted measurement matrix method that improved the restricted isometry constant (RIC) of the measurement matrix based on singular value decomposition (SVD). Experimental results showed that the signals reconstruction results using the optimized matrices of Wang's method are better than those by direct reconstruction for matrices [48].

Although these works have been contributed to optimizing measurement matrices, there are some limitations that usually require some specific experimental conditions and experimental objects in the previous works. Thus, based on the original works, it is necessary to develop a better optimization algorithm of generating the measurement matrix adaptive to a corresponding fixed sparsifying basis and improving the performance of measurement matrix with higher signals reconstruction accuracy under different measurement numbers and different signal sparsity levels, and the works of generating new optimized measurement matrices are contributed in this paper. Different from the conventional measurement matrices and the works in the literature [47] and [48], the main works of this paper are that we contribute a new measurement matrix optimization algorithm based on the proposed new mathematical model and the new measurement matrix initialization method. In one of our previous research works [54], the optimized measurement matrices were just fixed as the binary measurement matrices and were applied to our single-pixel camera only for the two-dimensional signals imaging. Owing to some previous research conclusions, we focus on seeking the optimal measurement matrices that are not only for the binary measurement matrices and the two-dimensional image signals, but also for adaptive measurement matrices and one-dimensional sparse signals and the medical image, and our works are further conducted under different signal sparsity levels and different compression rates.

The remainder of this paper is organized as follows: Section 2 introduces the basic theory and the related background of the measurement matrix. In Section 3, our new algorithm is proposed, the new algorithm model is analyzed and the detailed algorithm process is introduced. In Section 4, lots of experiments were done to test the

performance of our method in reconstructing one-dimensional sparse signals and two-dimensional image signals. In Section 5, analysis and discussion are presented about our algorithm and the experimental results. Finally, Section 6 summarizes the work of this paper.

2. Measurement Matrix Formulation

From the above theoretical introduction of CS, one can know that the performance of the measurement matrix directly influences the results of signals reconstruction, and the measurement matrix is the important intermediate link between signal sparsifying and signal reconstruction. Candès have contributed pioneering works to the theory of compressed sensing. They have published a series of important papers about the conditions that measurement matrices should meet and the relationship between the number of measuring signals and the signal sparseness [33, 55, 56]. The literature [30] proves that if the measurement matrix and the orthogonal sparsifying basis are as incoherent (orthogonal) as possible, the measurement matrix would have better performance with a large probability.

Generally, the sparsifying basis Ψ is known and fixed, so the mutual incoherence between the measurement matrix and sparsifying basis can be minimized as much as possible by designing and optimizing the measurement matrix. Also, the intuitive understanding is presented that if Φ and Ψ are incoherent or uncorrelated, sampling from original signal would contain more new information that has not already been represented by the known basis Ψ . Thus, the performance of the measurement matrix depends largely on the mutual coherence between the measurement matrix and the sparsifying matrix.

According to the related mathematical theory [47, 57], the incoherence between the measurement matrix and the sparsifying matrix can be indirectly transformed into the problem that Gram matrix approximates an identity matrix as closely as possible. Gram matrix is defined as follows:

$$\text{Gram} := \Theta^H \Theta, \quad \Theta = \Phi \Psi, \quad (5)$$

where Θ^H is the conjugate transpose operator of Θ .

The incoherence denotes the orthogonal property between the measurement matrix and the sparsifying matrix. According to the definition of the Gram matrix, minimizing the mutual coherence between Φ and Ψ is equivalent to minimizing the absolute off-diagonal elements in the corresponding Gram matrix and making the Gram matrix as close as possible to the identity matrix [57]. The closer the Gram matrix approximates the identity matrix, the smaller the mutual coherence between Φ and Ψ is. Thus, constructing a more optimized measurement matrix is equivalent to solving the following optimization problem:

$$\hat{\Phi} = \arg \min_{\Phi} \|(\Phi \Psi)^H \Phi \Psi - I\|_2^2, \quad (6)$$

where I is an identity matrix with the size $n \times n$. Thus, minimizing equation (6) can effectively and indirectly minimize mutual coherence. In the literature [47], Elad's work is just based on the model (6) via the minimization of

the average of the off-diagonal entries in the Gram matrix to seek an optimized measurement matrix.

3. Measurement Matrix Optimization

In this section, we give the mathematical definition of mutual coherence between the measurement matrix and the sparsifying matrix, and a new gradient projection based optimization algorithm is proposed to generate the new optimized measurement matrix.

3.1. Mathematical Definition of Mutual Coherence. On the basis of the Gram matrix, in [47], Elad gave a definition of the mutual coherence between Φ and Ψ by using the equivalent dictionary Θ , $\Theta = \Phi \Psi$, $\Theta = [\theta_1, \theta_2, \theta_3, \dots, \theta_n]$. Then the mutual coherence $\mu(\Theta)$ is defined as follows:

$$\mu(\Theta) := \max_{i \neq j, 1 \leq i, j \leq n} \left\{ \frac{\theta_i^T \theta_j}{\|\theta_i\|_2 \|\theta_j\|_2} \right\}. \quad (7)$$

Furthermore, from the perspective of between measurement matrix and sparsifying matrix, for the measurement matrix Φ with the size $m \times n$, $\Phi = [\phi_1, \phi_2, \phi_3, \dots, \phi_m]^T$, according to the orthogonal property between Φ and Ψ , another possible measurement value of the mutual coherence between Φ and Ψ is expressed as follows:

$$\mu(\Phi, \Psi) := \sqrt{n} \max_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}} |\langle \phi_i, \psi_j \rangle|, \quad (8)$$

$\langle \phi_i, \psi_j \rangle$ is the vector multiplication operator. Literature [58] gives the corresponding definition of mutual coherence, which plays an important role in reconstructing signals using the OMP algorithm, so the OMP algorithm is used as the signal reconstruction algorithm in our experiments of Section 4.

Consequently, equation (8) can denote the mutual coherence between Φ and Ψ , when the mutual coherence between Φ and Ψ is kept as small as possible in the sampling process of measurement matrix, which may ensure that more new information that has not been represented by the known sparsifying matrix Ψ can be sampled. That is to say, the incoherence property can ensure that original signals are reconstructed successfully with high probability in CS. If the measurement matrix is designed or optimized by minimizing equation (6), most sparse signals can be successfully recovered with the optimized measurement matrix.

3.2. New Function Model. One can know that, to solve and minimize equation (6), shrinking the absolute off-diagonal entries of the Gram matrix is usually time-consuming [47], and the accuracy of the measurement matrix obtained by equation (6) is also not enough. Thus, based on the original model (6), we come up with a new idea to optimize the original equation (6) by adding a constraint term of approximate equivalent dictionary $\Theta \approx \hat{\Phi} \Psi$ to ensure higher

accuracy of the measurement matrix, which is expressed as follows:

$$\begin{aligned} \widehat{\Phi} &= \arg \min_{\Phi} \|\Theta^H \Theta - I\|_2^2 \\ \text{s.t. } \|\Theta - \widehat{\Phi}\Psi\|_2^2 &\leq \zeta, \end{aligned} \quad (9)$$

where ζ is the upper bound of this term of approximate equivalent dictionary and $\widehat{\Phi}$ is an estimated approximate solution of Φ . Equation (9) makes full use of the properties of the Gram matrix and equivalent dictionary $\Theta = \Phi\Psi$.

We can also update equation (9) into another form by exchanging the penalty term and transferring the focus of equation (9) to the approximate term of the equivalent dictionary $\|\Theta - \widehat{\Phi}\Psi\|_2^2$, so the following expression is obtained:

$$\begin{aligned} \widehat{\Phi} &= \arg \min_{\Phi} \|\Theta - \widehat{\Phi}\Psi\|_2^2 \\ \text{s.t. } \|\Theta^H \Theta - I\|_2^2 &\leq \xi, \end{aligned} \quad (10)$$

where ξ is the upper bound of error of Gram matrix. Equation (10) is a convex constrained optimization problem. In order to solve equation (10), we need to define a new objective function or loss function to convert the constrained optimization problem to an unconstrained optimization problem. Therefore, we design a new loss function model based on equation (10) by Lagrangian method, and the original equation (9) is further transformed to the following:

$$\widehat{\Phi} = \arg \min_{\Phi} F(\widehat{\Phi}) = \left(\frac{1}{2}\right) \|\Theta - \widehat{\Phi}\Psi\|_2^2 + \tau \|(\widehat{\Phi}\Psi)^T \widehat{\Phi}\Psi - I\|_2^2, \quad (11)$$

so that the constraint becomes a penalty. For a proper choice of τ , the two problems (10) and (11) are equivalent, where τ is a weight parameter chosen empirically balancing the two items of equation (11). $F(\widehat{\Phi})$ is our proposed objective function (loss function) that requires to be solved for the minimum.

Our proposed new algorithm model (11) is a convex unconstrained optimization problem. The regularization penalty term $\|\Theta - \widehat{\Phi}\Psi\|_2^2$ denotes the approximate solution of $\Theta = \Phi\Psi$ and makes it more accurate to solve the measurement matrix by balancing $\|\Theta - \widehat{\Phi}\Psi\|_2^2$ and $\|(\widehat{\Phi}\Psi)^T \widehat{\Phi}\Psi - I\|_2^2$ adjusting the value of τ . The new algorithm model (11) not only guarantees the smaller mutual coherence between the measurement matrix and the sparsifying matrix by the cost $\|(\widehat{\Phi}\Psi)^T \widehat{\Phi}\Psi - I\|_2^2$, but also ensures higher accuracy of the measurement matrix by solving the regularization penalty term $\|\Theta - \widehat{\Phi}\Psi\|_2^2$. The regularization penalty term $\|\Theta - \widehat{\Phi}\Psi\|_2^2$ plays an important role in \ new cost function (11) and the optimized measurement matrix generated by simultaneously minimizing the two terms of $\|(\widehat{\Phi}\Psi)^T \widehat{\Phi}\Psi - I\|_2^2$ and $\|\Theta - \widehat{\Phi}\Psi\|_2^2$ in new model (11) can guarantee the signal reconstruction with higher accuracy, which is the advantage of our proposed method of optimizing traditional cost function (6) by adding a new

regularization penalty term. The new objective function will be solved by the gradient projection based strategy.

3.3. Gradient Projection. For the two special terms in our new model (11), the descent direction of this loss function is not just the gradient direction, but the direction of gradient projection. The new loss function of model (11) can better ensure that our algorithm approximates an optimal solution. According to the analysis of gradient descent direction, an iterative gradient projection based method is used to seek the minimum value of loss function $F(\widehat{\Phi})$. Firstly, the gradient $\nabla F(\widehat{\Phi})$ of $F(\widehat{\Phi})$ is expressed by the derivative of $F(\widehat{\Phi})$ on $\widehat{\Phi}$, as follows:

$$\nabla F(\widehat{\Phi}) \equiv \frac{\partial F(\widehat{\Phi})}{\partial \widehat{\Phi}} = (\widehat{\Phi}\Psi - \Theta)\Psi + 4\tau \widehat{\Phi}\Psi \left((\widehat{\Phi}\Psi)^T \widehat{\Phi}\Psi - I \right) \Psi^T. \quad (12)$$

In the process of solving $F(\widehat{\Phi})$, $\widehat{\Phi}$ is updated continually at each iteration. That is, from the k th iteration to the $k+1$ -th iteration, $\Phi^{(k)}$ is updated to $\Phi^{(k+1)}$. To complete the update, first we choose a positive scalar parameter $\alpha^{(k)}$, $\alpha^{(k)} > 0$, and define the following equation:

$$w^{(k)} = \left(\Phi^{(k)} - \alpha^{(k)} \nabla F(\Phi^{(k)}) \right)_+, \quad (13)$$

where $\nabla F(\Phi^{(k)})$ is the gradient of $F(\widehat{\Phi})$ at the k th iteration, and $\Phi^{(k)}$ is the estimation solution of Φ at the k th iteration. $(\phi)_+$ is the positive-part operator defined as $(\phi)_+ = \max\{0, \phi\}$. At each iteration update of $\Phi^{(k)}$, our algorithm searches along the negative gradient direction $-\nabla F(\Phi^{(k)})$, projecting onto the nonnegative orthant and conducting a backtracking line search until a sufficient decrease is attained in $F(\Phi^{(k)})$.

Then the second scalar parameter $\lambda^{(k)}$ is defined, $\lambda^{(k)} \in [0, 1]$. So, the iteration from $\Phi^{(k)}$ to iterate $\Phi^{(k+1)}$ is expressed as follows:

$$\Phi^{(k+1)} = \Phi^{(k)} + \lambda^{(k)} (w^{(k)} - \Phi^{(k)}), \quad (14)$$

where

$$\delta^{(k)} = w^{(k)} - \Phi^{(k)}. \quad (15)$$

Since loss function (11) is quadratic, the line search parameter $\lambda^{(k)}$ in equation (14) can be calculated simply using the following closed-form expression:

$$\lambda^{(k)} = \text{mid} \left\{ 0, \frac{(\delta^{(k)})^T \nabla F(\Phi^{(k)})}{(\delta^{(k)})^T B \delta^{(k)}}, 1 \right\}, \quad (16)$$

where $B = \begin{bmatrix} \Psi\Psi^T & -\Psi\Psi^T \\ \Psi\Psi^T & \Psi\Psi^T \end{bmatrix}$, and when $(\delta^{(k)})^T B \delta^{(k)} = 0$, we set $\lambda^{(k)} = 1$.

To acquire the value of $\alpha^{(k)}$, the parameter $\gamma^{(k)}$ is defined as follows:

$$\gamma^{(k)} = (\delta^{(k)})^T B \delta^{(k)}, \quad (17)$$

then $\alpha^{(k)}$ is obtained by the following:

$$\alpha^{(k)} = \text{mid} \left\{ \alpha_{\min}, \frac{\|\delta^{(k)}\|_2^2}{\gamma^{(k)}}, \alpha_{\max} \right\}, \quad (18)$$

where α_{\min} and α_{\max} are lower bound and upper bound of $\alpha^{(k)}$ respectively, if $\gamma^{(k)} = 0$, $\alpha^{(k+1)} = \alpha_{\max}$.

The use of parameter $\lambda^{(k)}$ and parameter $\alpha^{(k)}$ reduces the possibility that $F(\hat{\Phi})$ may increase at some time in the iteration process, and improves the efficiency of our algorithm in searching the global optimal solution of Φ .

3.4. New Initialization. Since local minimum solutions are likely to haunt us in the iterations of the algorithm, it is usually time-consuming to solve this new cost function (11) directly by the gradient projection strategy. Therefore, in order to reduce time-consuming and improve efficiency in algorithm operation, a wise initialization of Φ could be of great worth and initializing Φ is very necessary before the algorithm is executed, which directly determines the algorithm operation direction and influences the algorithm operation results. Therefore, based on the two terms' characteristics of objective function (11), we put forward a new method of initializing Φ .

First, we initialize the equivalent dictionary Θ by the pseudo-inverse $(\Theta^H)^\perp$ of Θ^H with the ideal assumptions that $\Theta = \Phi\Psi$ and Gram matrix is equivalent to an identity matrix $\Theta^H\Theta = I = \Theta^H(\Theta^H)^{-1}$, so

$$\Theta\Theta^H(\Theta^H)^{-1} = \Theta. \quad (19)$$

Then, we suppose that the pseudo-inverse is equal to the true-inverse operation $(\Theta^H)^{-1} = (\Theta^H)^\perp$ in an ideal situation, so

$$(\Theta^H)^\perp = (\Theta^H)^{-1} = (\Theta\Theta^H)^{-1}\Theta. \quad (20)$$

According to $\Theta^H\Theta = I$, we can get $\Theta = (\Theta^H)^{-1} = (\Theta^H)^\perp$ ($\Theta^H\Theta = \Theta^H(\Theta^H)^{-1}$). Based on equation (10) and $\Theta = \Phi\Psi$, so the initial $\Theta^{(0)}$ of Θ is expressed as follows:

$$\Theta^{(0)} = (\Phi\Psi(\Phi\Psi)^H)^\perp\Phi\Psi. \quad (21)$$

Finally, we can get the initialized measurement matrix $\Phi^{(0)}$ from $\Theta^{(0)} = \Phi^{(0)}\Psi$, so $\Phi^{(0)}$ is expressed as follows:

$$\Phi^{(0)} = (\Phi\Psi(\Phi\Psi)^H)^{-1}\Phi, \quad (22)$$

the Φ is unknown in equation (22) and we use the way of random initialization to initialize the Φ of equation (22). Because the random measurement matrices may be incoherent to most fixed orthogonal sparsifying matrices with large probability as analyzed as described in the introduction part, we use the Gaussian random matrix as the Φ of equation (22) to obtain the initialized measurement matrix $\Phi^{(0)}$.

The above method of initializing Φ is our proposed new initialization based on the ideal Gram matrix. Our initialization takes full advantage of the coupling of the two approximate terms in the loss function (11) assuming the two

ideal situations of $\Theta = \Phi\Psi$ and $\Theta^H\Theta = I$ simultaneously, which conforms to the characteristics of the objective function and also enables our algorithm to apply fewer number of iterations. This new initialization method can make the loss function $F(\Phi)$ approximate the global optimal solution with higher efficiency and larger probability.

Consequently, $\Theta^{(0)}$ obtained by our initialization and the fixed Ψ are as the inputs of our algorithm to train a more optimized measurement matrix Φ , and the risk of local minima and the instability the solution of Φ falls into is ameliorated by the fact that the initialization solution of Φ has already been in the neighborhood of a desirable attraction basin of global optimum, and the global optimum is almost equivalent to the ideal solutions of $\Theta = \Phi\Psi$ and $(\Phi\Psi)^T\Phi\Psi = I$. In the process of our algorithm operation, the last $\Phi^{(k)}$ is the estimation of the ideal solutions of $\Theta = \Phi\Psi$ and $(\Phi\Psi)^T\Phi\Psi = I$ based on the $\Phi^{(0)}$ of equation (22), and the coupling of $\|\Theta - \hat{\Phi}\Psi\|_2^2$ and $\|(\hat{\Phi}\Psi)^T\hat{\Phi}\Psi - I\|_2^2$ approximating zero simultaneously is also the goal our algorithm search.

3.5. Pseudocodes for Solving Measurement Matrix. The specific and detailed steps of our proposed algorithm by gradient projection for solving problem (11) to get an optimized measurement matrix are presented in Algorithm 1.

In our proposed algorithm, after the measurement matrix is initialized, and then the measurement matrix is trained and optimized by making full use of the information of the known sparsifying matrix. After a large number of experiments (the experiments in Section 4), we tested some values for these parameters in Algorithm 1 and found empirically that these parameters perform relatively well, so τ is set to 0.35, ξ is set to 0.01, α_{\min} is set to 1×10^{-3} , and α_{\max} is set to 1×10^3 .

It is worth noting that, because there may be both positive and negative elements in the measurement matrix Φ . In the iteration process of our algorithm, in order to facilitate the algorithm operation, we divide Φ into a positive part U and a negative part V , then Φ can be expressed as follows:

$$\Phi = U - V, \quad u_{ij} \geq 0, v_{ij} \geq 0, \quad (23)$$

where $u_{ij} = (\phi_{ij})_+$ and $v_{ij} = (-\phi_{ij})_+$ for all $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$. u_{ij} , v_{ij} and ϕ_{ij} are the elements in the matrices U , V and Φ , respectively.

Finally, after acquiring the optimized solution of Φ from the above algorithm, the columns of Φ need to be regularized to get the final optimized measurement matrix and make the experimental results more balanced and more stable. The regularization is expressed as follows:

$$\Phi(:, j) \leftarrow \frac{\Phi(:, j)}{\|\Phi(:, j)\|_2}. \quad (24)$$

According to the above analysis, the measurement matrix obtained by our optimization algorithm may have lower coherence with the specific sparsifying basis theoretically. Next, we will test the performance of the optimized

The algorithm process.

- (1) Initialization: the number of iterations $k, k = 0$; set $\Theta^{(0)}$ and $\Phi^{(0)}$ by equations (21) and (22); choose parameters α_{\min} and α_{\max} , $\alpha^{(0)} \in [\alpha_{\min}, \alpha_{\max}]$.
- (2) Calculate the value of $\delta^{(k)}$ by equation (15).
- (3) Find the parameter $\lambda^{(k)}$ on the interval $\lambda^{(k)} \in [0, 1]$ by equation (16).
- (4) The line backtracking search to minimize $F(\Phi^{(k+1)})$: $\Phi^{(k+1)} = \Phi^{(k)} + \lambda^{(k)}\delta^{(k)}$.
- (5) Update $\alpha^{(k)}$ by equations (17) and (18).
- (6) When the algorithm performs convergence and meets the termination condition: $(\|\nabla F(z^{(k)})\|_2 / \|z^{(k)}\|_2) \leq \xi$
Stop the iteration, and $\Phi^{(k+1)}$ is the final approximate solution of Φ , ξ is the threshold parameter of termination; otherwise, set $k = k + 1$ and return to 2 until the termination condition is met.

ALGORITHM 1: The pseudocodes of solving new measurement matrix.

measurement matrix by specific experiments compared with the conventional random measurement matrices and the classic optimization methods in literature [47] and [48].

4. Experiments

In the section, to test the performance of the optimized measurement matrices after adding the new regularization penalty term to the traditional cost function (6), we conducted the numerical experiments that provided the distribution of the off-diagonal elements of the Gram matrix obtained using different measurement matrices and different methods. In addition, more experiments were conducted to reconstruct one-dimensional sparse signals and two-dimensional nonsparse images using different measurement matrices and different methods. All the experiments were operated in the environment of MATLAB 2017b on a standard PC with a 2.20 GHz Intel Core i5 CPU and 4 GB memory, running on the Windows 10 operating system.

4.1. Off-Diagonal Elements' Distribution of Gram Matrix.

According to equation (6), making the Gram matrix as close as possible to the identity matrix equivalently minimizes the mutual coherence between Φ and Ψ . To test the incoherence performance of the measurement matrix optimized by our proposed method, we conducted the experiments at $m = 256$ and $n = 512$ that provided the distribution of the off-diagonal elements of the Gram matrix obtained using different measurement matrices and different methods, as shown in Figure 1. The different measurement matrices include the four conventional random measurement matrices that are 0-1 binary sparse matrix, Gaussian matrix, Bernoulli matrix, part Hadamard matrix, and the different methods are respectively the Elad's method in literature [47] and our methods before and after optimizing the traditional method of the cost function (6) by adding the regularization penalty term. The method of directly solving the equation (6) by the gradient projection strategy without the regularization penalty term is denoted as "No Regularization (NoReg)" in this paper.

In the experiments, in order to facilitate the statistics of the number of off-diagonal elements, the absolute value of off-diagonal elements was rounded to two significant digits

with the accuracy of 0.01, which means that there are 100 evenly distributed points from 0 to 1 and the number of off-diagonal elements is counted at every point. The known square DWT (discrete wavelet transform) matrix was given as the sparsifying dictionary Ψ . Because the measurement matrix was optimized indirectly based on singular value decomposition in Wang's method, we cannot directly get the optimized measurement matrix of Wang's method. Thus, the off-diagonal elements' distribution of the Gram matrix of Wang's method is not provided in Figure 1. The method of No Regularization is given as a comparison to demonstrate the important role of the added regularization penalty term $\|\Theta - \Phi\Psi\|_2^2$ in the new cost function (11). Because the absolute value of off-diagonal elements of random matrices is too small, Figure 1 is divided into two subfigures (a) and (b) for displaying the curves of all the compared matrices. In Figure 1(a), the curves of yellow, blue, green and cyan have very low values and can hardly display normally, while the curves of black, red, and magenta are too narrow to display well in Figure 1(b). Therefore, in order to display the curves of all the colors, two subfigures of Figure 1 have to be plotted with two different scales.

According to the results of Figure 1, the distributions of the off-diagonal elements of the Gram matrix of four random measurement matrices are close and similar. Because the random matrices had not been optimized via mutual coherence minimization, the distributions of the off-diagonal elements of the Gram matrix of the four random measurement matrices are far worse than that of No Regularization, Elad's method and our method. Furthermore, our method outperforms the No Regularization and Elad's method in increasing the frequency of the off-diagonal elements with low absolute value and reducing the frequency of large absolute values via mutual coherence minimization, better attempting to make the columns of $\Theta\Psi$ as close to orthogonal as possible. Thus, the experimental results of Figure 1 show that the added regularization penalty term $\|\Theta - \Phi\Psi\|_2^2$ in the new cost function (11) plays a crucial role in improving the performance of our proposed algorithm and optimizing the traditional cost function (6), and demonstrate the effectiveness and superiority of our algorithm based on gradient projection in minimizing the mutual coherence between Φ and Ψ .

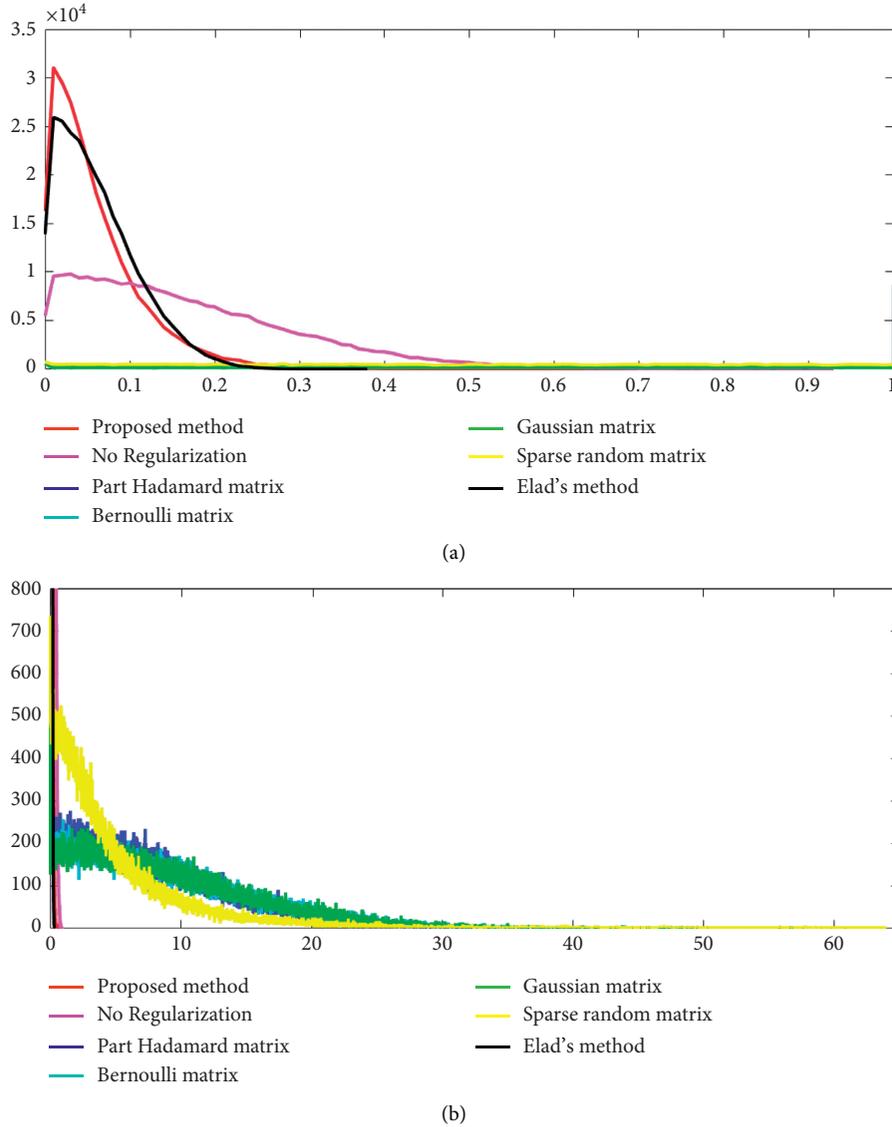


FIGURE 1: Distribution of the off-diagonal elements of the Gram matrix obtained using different measurement matrices and different methods at $m = 256$ and $n = 512$, when Ψ is a DWT matrix.

The following experiments were conducted to test the performance of our proposed algorithm in practical signals reconstruction, OMP algorithm was used as the CS-based signals reconstruction algorithm. The measurement value y was generated by equation (1) with the white Gaussian noise of variance $\sigma^2 = 10^{-4}$. One-dimensional sparse signals and two-dimensional nonsparse images were reconstructed by the new optimized measurement matrices generated by our method, compared with the four conventional random measurement matrices, the method of No Regularization, the Elad's method [47] and the Wang's method [48].

4.2. One-Dimensional Sparse Signals. For the reconstruction of one-dimensional sparse signals, the mean square error (MSE) is used to evaluate the accuracy of reconstructed one-dimensional signals. The smaller MSE is, the higher the

reconstruction accuracy of one-dimensional signals is MSE is defined as follows:

$$\text{MSE} = \left(\frac{1}{n}\right) \|\hat{x} - x\|_2^2, \quad (25)$$

where \hat{x} is the estimation of x .

In the experiments, since the one-dimensional sparse signals have already been sparse, there is no need to sparsify the original signals again, so the Ψ in the objective function (11) and equation (22) is replaced as an identity matrix of dimension n . $n = 512$, $m = 256$, $K = 100$ mean that the original one-dimensional sparse signal x of length 512 contains 100 randomly placed +1 and -1 spikes, and the other elements of x are zero. The measurement number m is 256 and the compression rate is 0.5. For one-dimensional sparse signals, the iteration number in the OMP algorithm is set to K . The results of reconstructing one-dimensional

sparse signals by different measurement matrices and different optimization methods at the compression rate 0.5 ($m=256$, $n=512$) are shown in Figure 2.

As shown in Figure 2, the optimized measurement matrix by our proposed method has the best performance in reconstructing one-dimensional sparse signals, and the recovered signals by the optimized measurement matrix have the lowest MSE and are closer to the original sparse signals, compared with the four conventional measurement matrices, the method of No Regularization, Elad's method and Wang's method, because the four conventional measurement matrices, the initialization matrix in our method, the method of No Regularization, the Elad's measurement matrix and the Wang's measurement matrix are all generated randomly. In order to avoid the randomness of experimental results as much as possible, more experiments should be done to verify if the measurement matrices have stable performance. Thus, five times experiments for each sparsity level were done in reconstructing one-dimensional sparse signals, and then we took the average of five experiments' MSEs as the final results. According to the experimental results, the evolution of the MSE versus the sparsity level K under different measurement matrices is plotted when the compression rate is 0.5 ($m=256$, $n=512$), as shown in Figure 3.

As shown in Figure 3, the MSEs of the reconstructed signals by our method are lower than the other four conventional measurement matrices, the method of No Regularization, the Elad's and Wang's methods at different sparsity level K , and our method can reconstruct the sparse signal with higher precision. In addition, to test our algorithm's performance of processing larger data, experiments were done to reconstruct larger sparse signals of length $n=1024$ and the sparsity level K is set to 128. In the same way, every result of MSE reported is an average over 5 experiments. The relationship curves between MSE and the measurement number m under different measurement matrices and different optimization methods are shown in Figure 4.

As shown in Figure 4, compared with other measurement matrices, the method of No Regularization, the Elad's and Wang's methods, our method also achieves the best performance with the lowest MSEs at different measurement numbers m . Thus, according to the above experimental results, it has been verified that the new measurement matrices generated by our optimization method have better performance in reconstructing one-dimensional sparse signals of different data lengths and different sparsity levels under different measurement numbers.

4.3. Two-Dimensional Nonsparse Images. In this subsection, we conduct the following experiments to test the performance of our method in reconstructing two-dimensional images. We choose eight representative images that are not sparse with the size of 512×512 , including nature scenes, persons, animals, detailed and texture images, respectively, as shown in Figure 5. In order to enhance the efficiency of

image reconstruction and reduce the computational complexity, we reconstruct the images by measuring and reconstructing each column of original images with the size of 512×1 separately at the compression rate 0.25 ($m=128$, $n=512$), and the 512 reconstructed columns were added back to a full reconstructed image of size 512×512 . The reconstruction method of column-by-column scanning was also used in Wang's method [48], and this approach not only reduces problem size but also eliminates the obvious block mosaics that the 2D pattern of block-based scanning produces in the reconstructed images, as described in the literature [59].

For two-dimensional nonsparse images, the iteration number in the OMP algorithm is set to the measurement number m . The peak signal to noise ratio (PSNR) and the structural similarity (SSIM) are used as the evaluation indexes of image quality, which are defined as follows:

$$\text{PSNR (dB)} = 20 \log \frac{255}{1/HW \sum_{i=1}^H \sum_{j=1}^W (\hat{x}(i, j) - x(i, j))^2},$$

$$\text{SSIM} = \frac{(2\mu_x \mu_{\hat{x}} + c_1)(2\sigma_{x\hat{x}} + c_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + c_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + c_2)}, \quad (26)$$

where x is the original image and \hat{x} is the reconstructed image. H and W are the height and the width of the image, respectively. μ_x and $\mu_{\hat{x}}$ are, respectively, the mean values of x and \hat{x} . σ_x^2 , $\sigma_{\hat{x}}^2$, and $\sigma_{x\hat{x}}$ are the variance and covariance values of x and \hat{x} , respectively. c_1 and c_2 are the constants that maintain stability, $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$. L is the dynamic range of pixel greyscale, $k_1 = 0.01$, $k_2 = 0.03$.

It is worth noting that, in the reconstruction of one-dimensional sparse signals, the sparsifying basis Ψ is an identity matrix. However, for the reconstruction of two-dimensional nonsparse images, since the images are not sparse, the square DWT matrix with 5-level wavelet is used as the sparsifying basis Ψ in our algorithm. The performance of the optimized measurement matrices in reconstructing images will be analyzed according to the following experimental results.

Table 1 shows the PSNRs and SSIMs of eight reconstructed images by different measurement matrices and different optimization methods at the compression rate of 0.25 ($m=128$). All the data in Table 1 were obtained by taking the average of five experiment runs' results and Table 2 shows the specific statistical data of each experiment run's results in reconstructing image *Building*, and the means and variances of PSNRs and SSIMs are calculated.

From Table 2, one can know that, according to the five experiment runs, the variances of PSNRs and SSIMs are very small, and the experimental results of every run are stable with little fluctuation. Moreover, for the experimental results of each run, our method has obvious advantages and effectively improve the PSNRs and SSIMs of reconstructed images. Therefore, five experiment runs can objectively reflect the actual performance of our algorithm and can prove that our method has stable and superior performance.

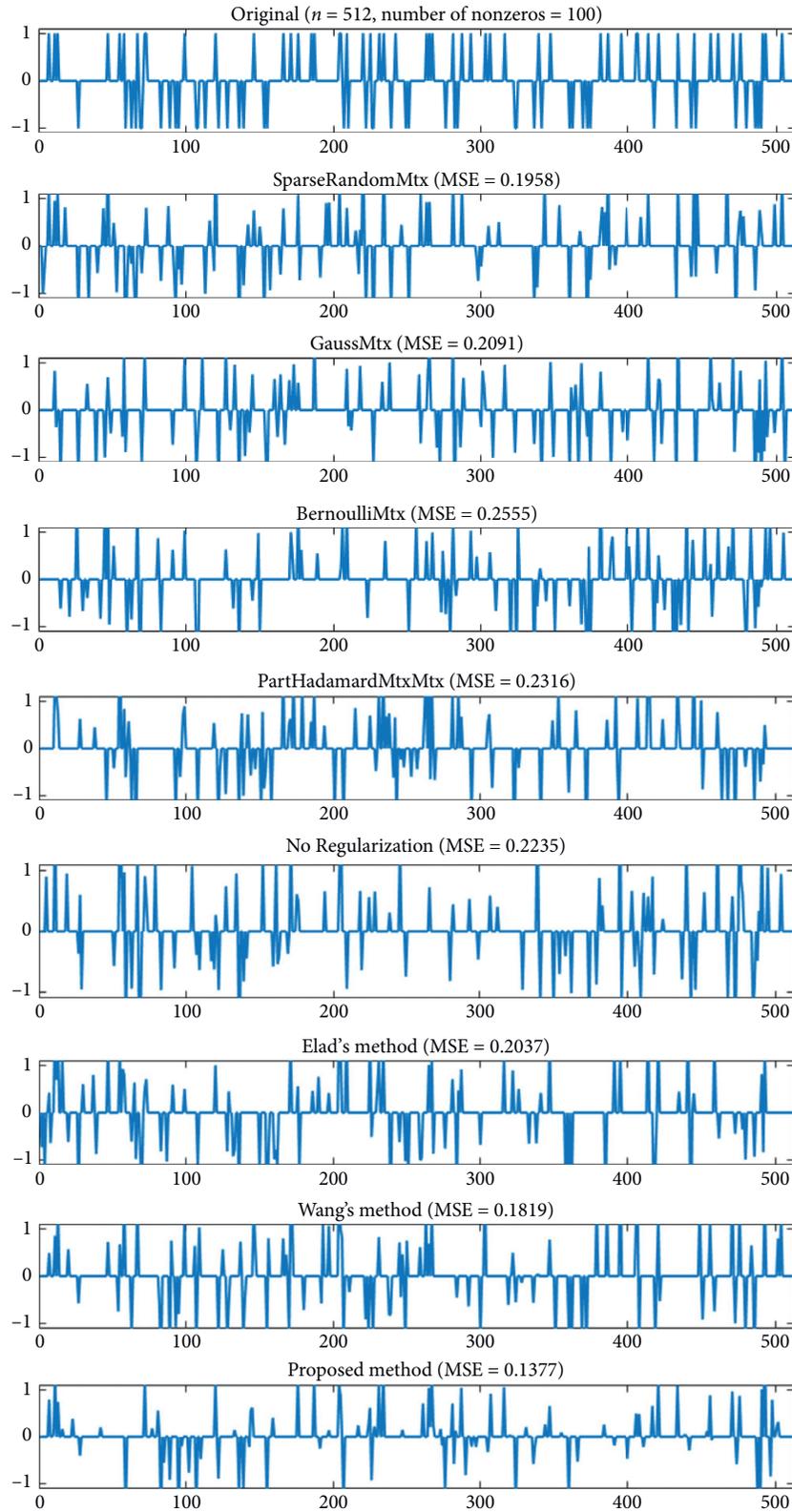


FIGURE 2: Results of reconstructing one-dimensional sparse signals by different measurement matrices and different methods at $m = 256$ and $n = 512$.

In order to visually show the improvement of the reconstructed images by our method, Figure 6 shows the actual reconstructed images of *Building* in one of five

experiment runs under different measurement matrices and different methods, and the local areas of image *Building* marked by red box are amplified.

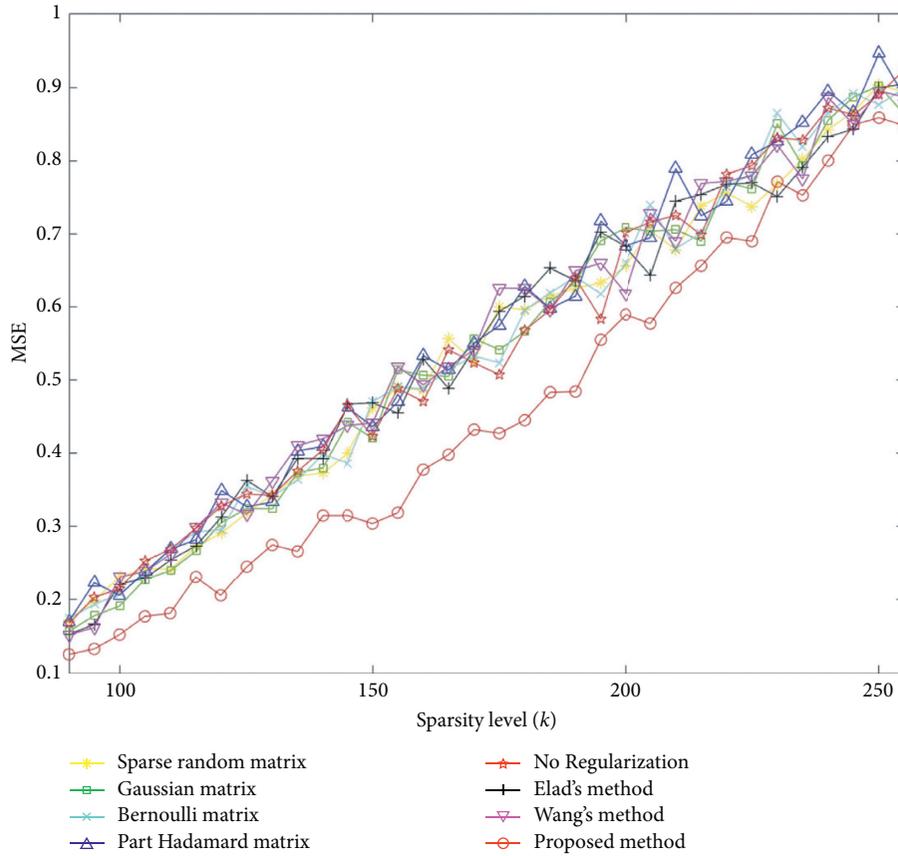


FIGURE 3: The relationships between MSE and sparsity level K in reconstructing one-dimensional sparse signals using different measurement matrices and different methods at $m = 256, n = 512$.

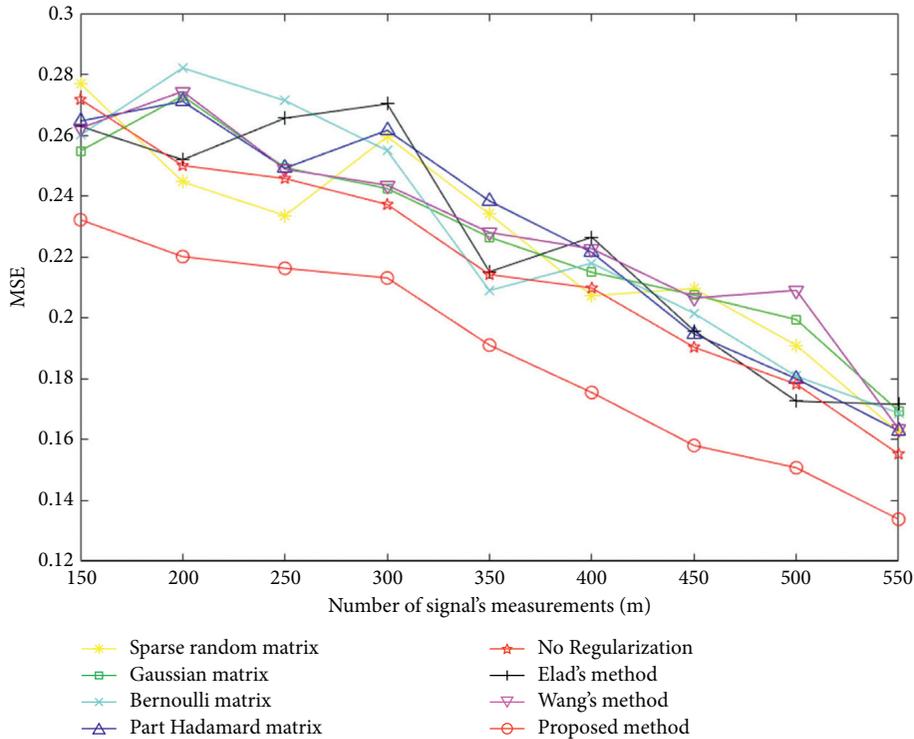


FIGURE 4: The relationships between MSE and measurement number (m) in reconstructing one-dimensional sparse signals using different measurement matrices and different methods at $n = 1024, K = 128$.



FIGURE 5: Eight representative images of (a) Peppers, (b) Boats, (c) Building (d) Cameraman, (e) Barbara, (f) Mandrill, (g) Goldhill, and (h) Fingerprint.

TABLE 1: PSNRs/SSIMs of reconstructed images by different measurement matrices at $m = 128$.

Images	Sparse random	Gaussian	Bernoulli	Part Hadamard	NoReg	Elad's	Wang's	Optimized matrix
Peppers	21.2336/0.33377	21.2234/0.3352	20.9420/0.32786	21.4817/0.34782	21.4106/0.33903	21.4965/0.34668	22.3604/0.4139	24.2672/0.48002
Boats	20.1211/0.32113	20.0598/0.31796	20.1037/0.32355	20.4870/0.34079	21.1489/0.32162	20.3158/0.33196	21.1873/0.3976	23.3487/0.48302
Building	24.0236/0.43797	24.4098/0.44937	24.0987/0.44321	24.7301/0.47496	24.2669/0.44778	24.8264/0.47467	26.1096/0.57190	27.2533/0.58487
Cameraman	21.8334/0.43157	21.5152/0.41978	21.6693/0.41437	22.0669/0.46524	21.6392/0.42525	22.1898/0.45114	22.7035/0.52447	25.8632/0.62734
Barbara	20.1602/0.42315	20.4094/0.43275	20.0468/0.41262	20.5244/0.44034	20.3292/0.43005	20.4563/0.4436	21.8251/0.5315	23.6868/0.6042
Mandrill	16.4434/0.20460	16.4298/0.20306	16.4234/0.20594	16.6225/0.21432	16.6581/0.21474	16.4741/0.20784	16.9468/0.2518	18.1533/0.30182
Goldhill	20.2144/0.29227	20.1329/0.28863	20.1679/0.29005	20.5636/0.31045	20.3922/0.30052	20.4659/0.30194	21.2499/0.37623	23.4258/0.46419
Fingerprint	15.6476/0.25184	15.5595/0.24256	15.5394/0.24149	15.7223/0.25411	15.6447/0.24709	15.6734/0.25524	17.0946/0.38497	17.8569/0.43580

From the experimental results of Tables 1 and 2 and Figure 6, in accordance with the quantitative results, one can know that our method generally has a better performance and the reconstructed image by our method is more visually pleasing at the compression rate of 0.25 ($m=128$), compared with the four conventional measurement matrices, the method of No Regularization, the Elad's, and Wang's methods. The PSNRs increase by about 1 to 3 dB, and the SSIMs increase by about 0.01 to 0.2 as shown by bold in Tables 1 and 2. In addition, from the reconstructed images in Figure 6, the reconstructed images by the four conventional measurement matrices have serious noises and are fuzzy in the image contours and edges. The reconstructed images by the method of No Regularization and the Elad's and Wang's methods also have some distortion and much noise, while the reconstructed images by our method have obviously better image reconstruction accuracy and better visual quality.

In order to further demonstrate the effectiveness of our method in reconstructing nonsparse images at a higher sampling rate, a person image Cameraman is reconstructed at the compression rate of 0.5 ($m = 256$). Figure 7 shows the image reconstruction results of Cameraman at $m = 256$ under different measurement matrices and different methods, and the local detailed areas of the image Cameraman marked by red box are also amplified.

As shown in Figure 7, compared with the four conventional measurement matrices, the image Cameraman reconstructed by our method is closer to the original image and has much higher image reconstruction quality and higher accuracy with lower noise and fewer visual artifacts. Also, there are higher contrast and richer details, especially in the enlarged area of the image Cameraman. Compared with the method of No Regularization, the Elad's and Wang's methods, our proposed method also enhances the PSNR and SSIM of the reconstructed image and reduces the

TABLE 2: The statistical PSNRs/SSIMs of five experiment runs' results in reconstructing image *Building* by different measurement matrices at $m = 128$.

Times	Sparse random	Gaussian	Bernoulli	Part Hadamard	NoReg	Elad's	Wang's	Optimized matrix
1	24.2901/ 0.44622	24.3995/ 0.44873	24.0885/ 0.43630	24.6580/ 0.47171	24.5528/ 0.46479	24.7666/ 0.47067	26.3642/ 0.58124	26.9759/ 0.54479
2	23.5791/ 0.41832	24.4172/ 0.45150	24.1424/ 0.44034	24.3653/ 0.45435	24.5539/ 0.45525	24.8598/ 0.47935	25.9501/ 0.56853	27.0524/ 0.58232
3	24.0906/ 0.44499	24.5570/ 0.45856	24.0547/ 0.44028	24.8548/ 0.47995	24.1901/ 0.44146	24.9355/ 0.47619	25.9409/ 0.56286	27.2789/ 0.58696
4	24.0453/ 0.43432	24.0803/ 0.43536	24.3196/ 0.44327	24.6074/ 0.46906	23.8855/ 0.43135	24.9349/ 0.48395	26.2562/ 0.56896	27.2887/ 0.59844
5	24.1129/ 0.44600	24.5950/ 0.45270	23.8883/ 0.45586	25.1650/ 0.49973	24.1520/ 0.44606	24.6354/ 0.46319	26.0366/ 0.57790	27.6707/ 0.61185
Means	24.0236/ 0.43797	24.4098/ 0.44937	24.0987/ 0.44321	24.7301/ 0.47496	24.2669/ 0.44778	24.8264/ 0.47467	26.1096/ 0.57190	27.2533/ 0.58487
Variances	0.0704/ 1.4532e-4	0.0412/ 7.4193e-5	0.0243/ 5.6149e-5	0.0895/ 2.7715e-4	0.0822/ 1.6450e-4	0.0162/ 6.4530e-5	0.0364/ 5.6239e-5	0.0734/ 6.3234e-4

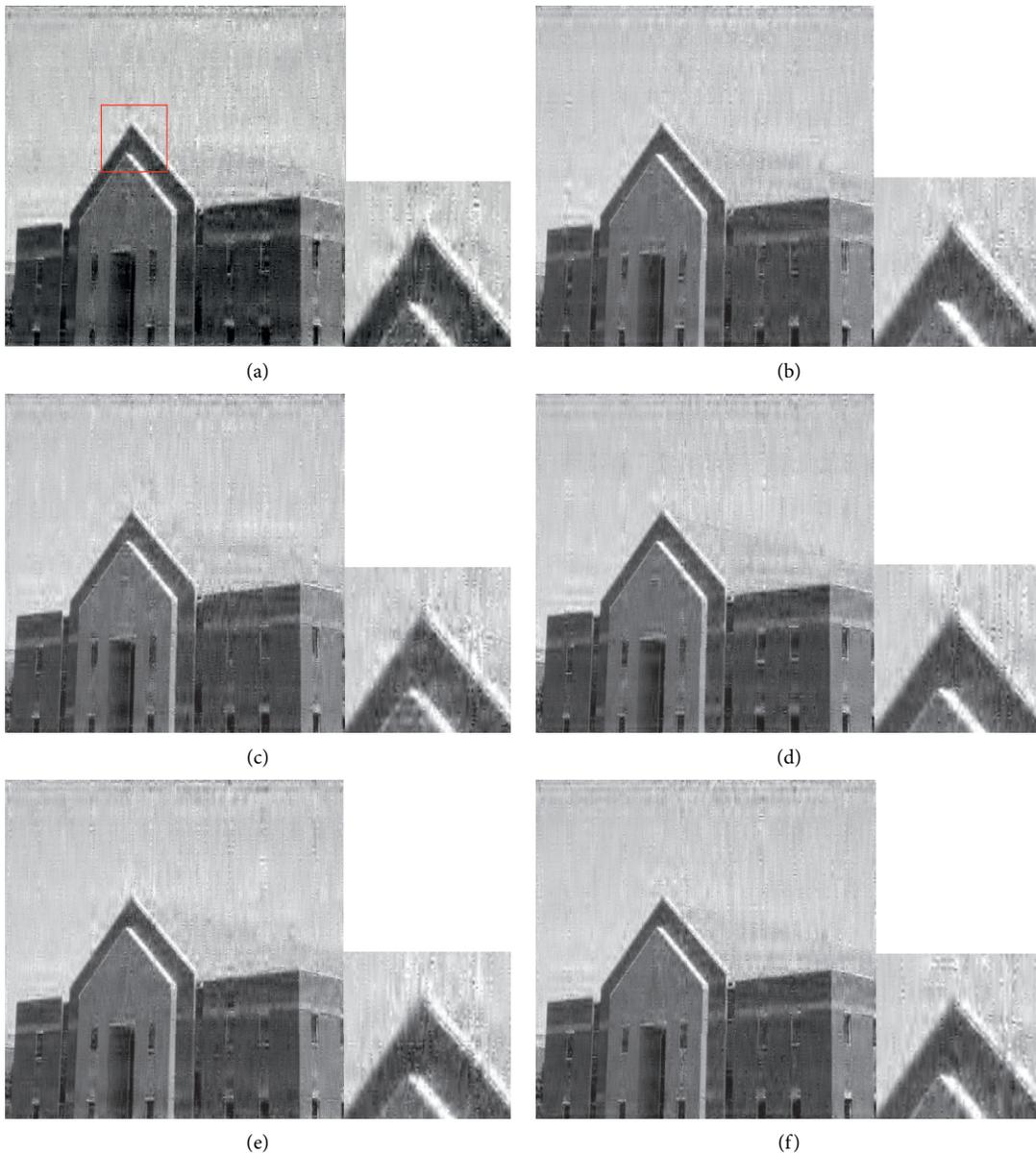


FIGURE 6: Continued.

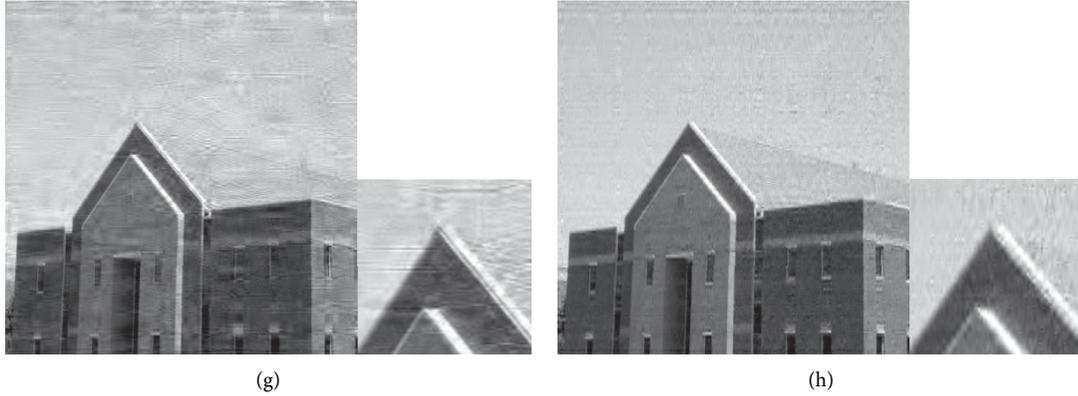


FIGURE 6: The reconstructed images Building by measurement matrices of (a) sparse random (PSNR: 24.0906 dB, SSIM: 0.44499), (b) Gaussian (PSNR: 24.5570 dB, SSIM: 0.45856), (c) Bernoulli (PSNR: 24.0547 dB, SSIM: 0.44028), (d) part Hadamard (PSNR: 24.8548 dB, SSIM: 0.47995), (e) No Regularization (PSNR: 24.1901 dB, SSIM: 0.44146), (f) Elad's method (PSNR: 24.9355 dB, SSIM: 0.47619), (g) Wang's method (PSNR: 25.9409 dB, SSIM: 0.56286), and (h) our proposed method (PSNR: 27.6707 dB, SSIM: 0.61185) at $m = 128$ and $n = 512$.

reconstruction errors, and the reconstructed image via our method is closer to the original image with better visibility of the details. The above experiments further show that our proposed algorithm is still superior to other conventional methods at a higher sampling rate of 0.5.

In addition, compressed sensing is widely applied in biomedical image processing, such as the image of magnetic resonance imaging (MRI). To verify that our method has universal performance for different types and different sizes of images and also has a good performance in reconstructing MRI images, we conducted the experiments on smaller data and reconstruct the image Brain acquired by MRI with the size of 128×128 using our method, the compression rate is 0.25 and 0.5 ($m = 32$, $m = 64$), and the DWT matrix with 3-level wavelet was used in this experiments. The experimental results are shown in Figures 8 and 9.

As shown in Figure 8, at the compression rate of 0.25, one can know that our method greatly improves the reconstruction quality of MRI image with PSNRs rising by about 3 to 4 dB and SSIMs rising by about 0.05 to 0.15, compared with the other seven methods. The other seven methods can hardly reconstruct the contour of the brain, while our method can reconstruct the basic contour and shape of the brain in the image Brain. In Figure 9, the image (h) is closer to the original image (a), and the other seven images have more visual artifacts and distortions. Figure 9(h) reconstructed by our optimized measurement matrix also has the highest PSNR and SSIM with more clarity and less noise, and our method creates the most visually comfortable results. Thus, these experiments reflect that our method also outperforms the other seven conventional methods in reconstructing MRI images of less size.

5. Analysis and Discussion

These above experiments have verified the superior performance of our new algorithm in optimizing measurement matrix by reconstructing one-dimensional sparse signals and two-dimensional nonspare images under different

sparsity levels and different measurement numbers. Our proposed method can make the elements distribution of the Gram matrix closer to the identity matrix than other methods. In addition, through comparing the experimental results of the new loss function (11) with and without the regularization penalty term, one can know that, after optimizing the traditional method of the cost function (6) by adding the regularization penalty term, our method has the superior performance in reconstructing higher precision signals and images than the traditional method of directly solving cost function (6) under different sparsity levels and different measurement numbers. Thus, the new cost function model (11) with the added regularization penalty term plays a crucial role in our proposed algorithm and makes the measurement matrix Φ have smaller coherence with Ψ , which demonstrates the effectiveness of adding the regularization penalty term in the traditional cost function.

Our optimized measurement matrices perform well in extracting information from original signals and restoring information from sampled signals so that our new method can reconstruct original signals of higher precision with higher probability. The main contributions of this paper are as follows:

- (i) The new algorithm model (11) of the cost function is proposed for seeking optimal measurement matrix that would be least coherent to the sparsifying basis, in theory, ensuring the accuracy of solving measurement matrix by adding the regularization penalty term;
- (ii) A new measurement matrix initialization method is designed based on the characteristics of our proposed new algorithm model (11). The new initialization method makes the new model more solvable by reducing the risk of local minima and instability the solution of Φ may fall into;
- (iii) Due to the coupling of measurement matrix and sparsifying basis, it is usually time-consuming to solve the model (6). Therefore, our method provides a good idea for solving the new mathematical



FIGURE 7: The reconstructed images Cameraman by measurement matrices of (a) sparse random (PSNR: 29.9159 dB, SSIM: 0.78685), (b) Gaussian (PSNR: 29.4204 dB, SSIM: 0.77486), (c) Bernoulli (PSNR: 29.3394 dB, SSIM: 0.77426), (d) part Hadamard (PSNR: 31.0353 dB, SSIM: 0.82327), (e) No Regularization (PSNR: 29.8907 dB, SSIM: 0.78744), (f) Elad's method (PSNR: 31.1338 dB, SSIM: 0.8232), (g) Wang's method (PSNR: 30.6554 dB, SSIM: 0.82174), and (h) our proposed method (PSNR: 32.0550 dB, SSIM: 0.83511) at $m = 256$ and $n = 512$.

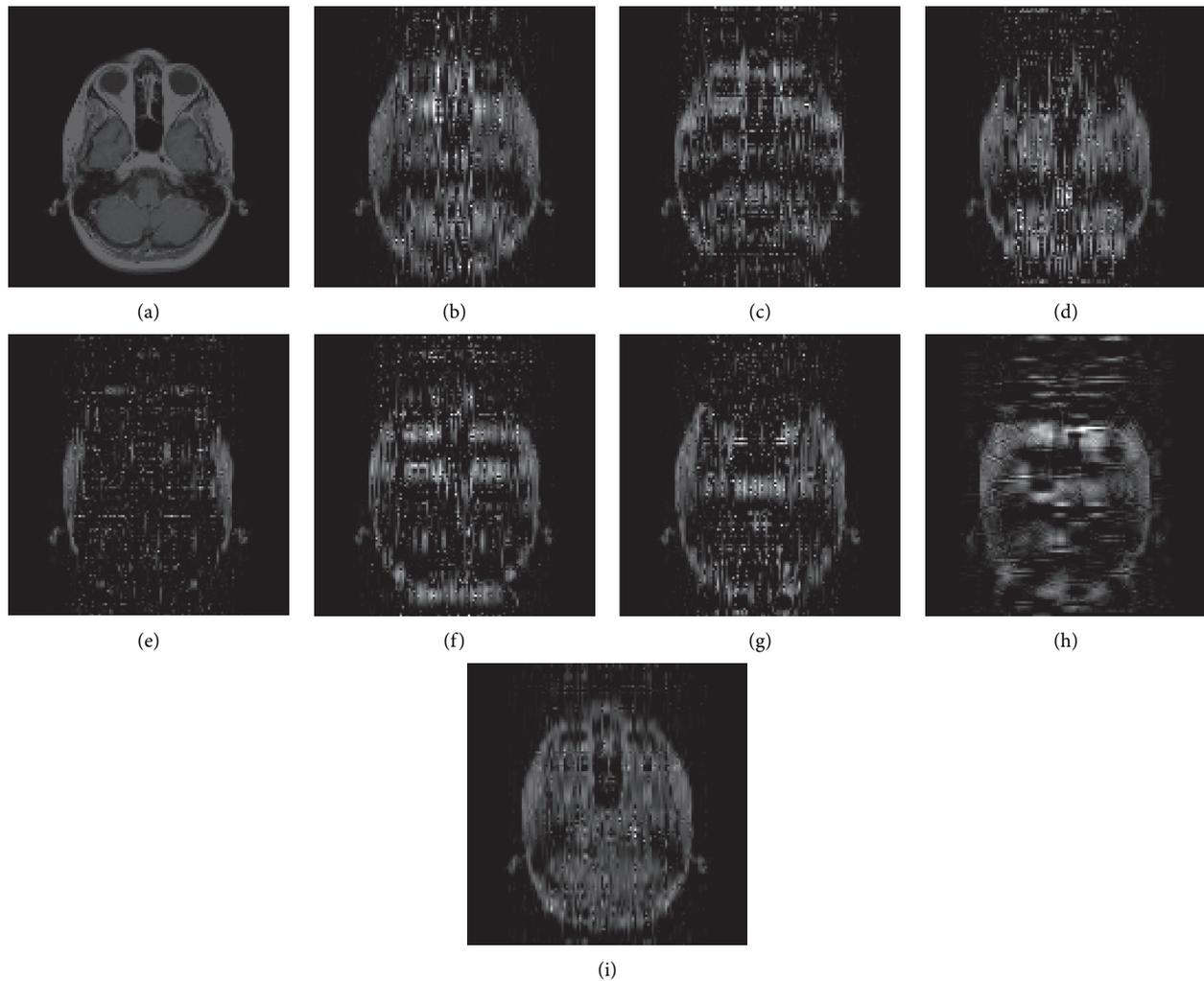


FIGURE 8: (a) Original MRI image Brain, reconstructed images by measurement matrices of (b) sparse random (PSNR: 20.15 dB, SSIM: 0.523), (c) Gaussian (PSNR: 19.24 dB, SSIM: 0.448), (d) Bernoulli (PSNR: 19.36 dB, SSIM: 0.481), (e) part Hadamard (PSNR: 18.38 dB, SSIM: 0.443), (f) No Regularization (PSNR: 18.62 dB, SSIM: 0.438), (g) Elad's method (PSNR: 19.29 dB, SSIM: 0.480), (h) Wang's method (PSNR: 19.15 dB, SSIM: 0.407), and (i) our proposed method (PSNR: 23.83 dB, SSIM: 0.574) at $m = 32$ and $n = 128$.

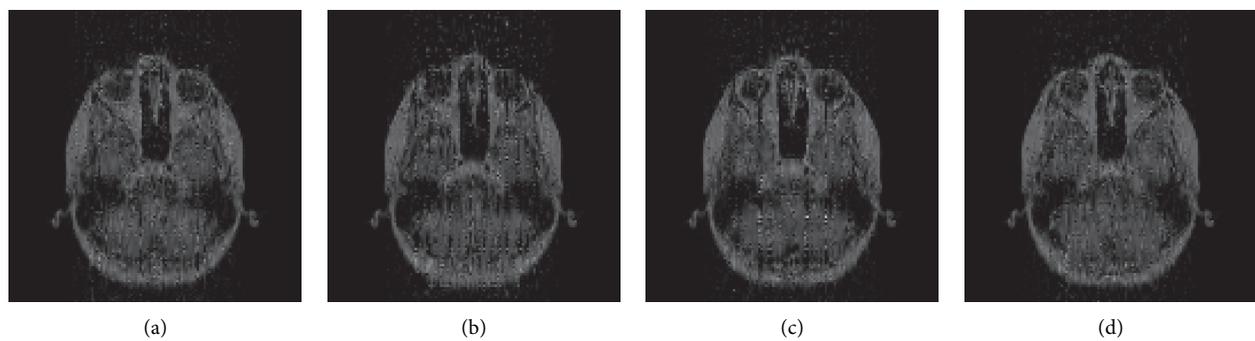


FIGURE 9: Continued.

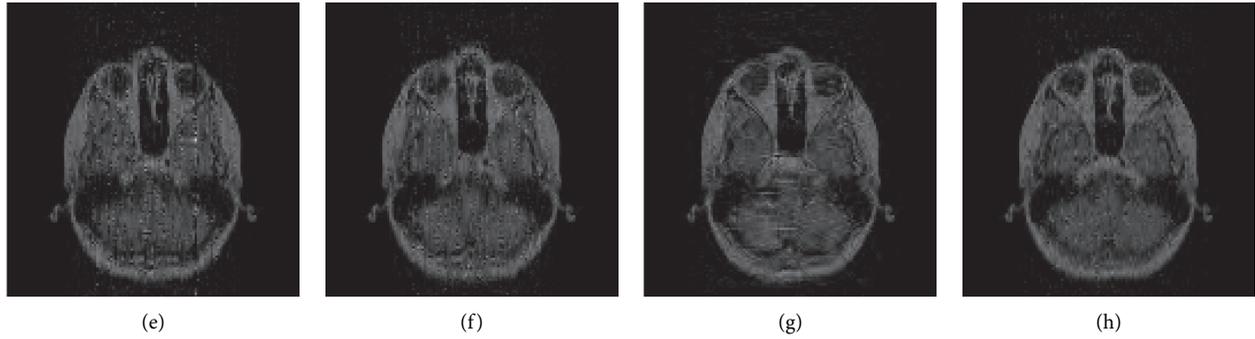


FIGURE 9: The reconstructed images Brain by measurement matrices of (a) sparse random (PSNR: 27.21 dB, SSIM: 0.708), (b) Gaussian (PSNR: 25.82 dB, SSIM: 0.679), (c) Bernoulli (PSNR: 26.54 dB, SSIM: 0.697), (d) part Hadamard (PSNR: 27.60 dB, SSIM: 0.722), (e) No Regularization (PSNR: 27.35 dB, SSIM: 0.720), (f) Elad's method (PSNR: 28.23 dB, SSIM: 0.747), (g) Wang's method (PSNR: 28.59 dB, SSIM: 0.729), and (h) our proposed method (PSNR: 30.03 dB, SSIM: 0.773) at $m = 64$ and $n = 128$.

optimization problem and constructing a more optimized measurement matrix;

- (iv) Based on the characteristics of our method, our proposed algorithm can find a solvable measurement matrix adaptive to the corresponding fixed sparsifying basis and have good flexibility and adaptability in reconstructing CS-based signals.

In addition, here are two important matters worth being discussed:

- (i) First, the amount of information the measurement matrix extracts or samples from original signals increases with the rise of measurements number, which makes it easier to recover the original signals for the reason that more valuable information can be obtained from original signals. This is why the improvement degree of PSNR and SSIM in reconstructing two-dimensional images by our method is smaller at a compression rate 0.5 than at a compression rate 0.25. Thus, the trade-off between compression rate and recovery accuracy has to be considered fully in the process of measuring and recovering signals.
- (ii) Second, small mutual coherence between the measurement matrix and sparsifying basis would ensure that original signals may be recovered successfully with high probability. The incoherence property is only a necessary condition and one of many factors (such as signal reconstruction algorithm, signal sparsifying transform, the original signal itself, and so on) that affect the final accuracy of signal reconstruction and the main work of this paper is to only optimize the measurement matrix, so the incoherence property cannot ensure that original signals are reconstructed successfully and precisely with a probability of 100% but just a large probability. Thus, the higher frequency of the off-diagonal elements with low absolute value in Figure 1 does not mean that the signal of higher accuracy will be reconstructed with a probability of 100% but just a

higher probability, and the results of signals reconstruction usually have some randomness, which is related to the original signal itself. Consequently, the experimental results of Figures 2–9 and Tables 1 and 2 may not necessarily and exactly conform to the distribution of Figure 1. Our proposed algorithm outperforms the traditional methods and can successfully reconstruct the signals with much higher probability in these experiments.

6. Conclusions

In this paper, the gradient projection based strategy is used to solve our proposed new algorithm model with the new method of initializing measurement matrix for compressively sensed signals reconstruction, which is a new idea for acquiring optimized measurement matrices via minimizing the mutual coherence between measurement matrix and sparsifying basis. According to the theoretical analysis of our method and the experimental results, we can conclude that our proposed measurement matrix optimization method has good performance and outperforms the conventional random measurement matrices and some existing optimization methods. Not only one-dimensional sparse signals but also two-dimensional nonspare images can be reconstructed by our optimized measurement matrices with less error, higher accuracy, and better quality. In addition, this new algorithm also has good performance in reconstructing the MRI image. Thus, our research by proposing the new algorithm in this paper may enlighten wide exploration in this direction of the field and has potential application value in the broader field of signal processing. It is of great worth to make further research in the incoherence property and the optimization algorithm of the measurement matrix in compressive sensing. This is also the focus of our future works to extend our research to broader applications.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Ziran Wei made significant contributions to this study regarding conception, methodology and analysis, and writing the manuscript. Jianlin Zhang contributed much to the research through his general guidance and advice and optimized the paper. Zhiyong Xu made research plans and managed the study project. Yong Liu gave lots of advice and helped in conducting experiments.

Acknowledgments

The authors are grateful to the Fraunhofer Institute for Computer Graphics Research IGD of Germany for full support. This research was funded by the National High Technology Research and Development Program of China (863 Program, grant no. G158207), the West Light Foundation for Innovative Talents (grant no. YA18K001), and the UCAS-Fraunhofer Joint Training Ph.D. Scholarship.

References

- [1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [2] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, "Compressive sensing: from theory to applications, a survey," *Journal of Communications and Networks*, vol. 15, no. 5, pp. 443–456, 2013.
- [3] E. Candès, "Compressive sampling," *Proceedings of the International Congress of Mathematicians*, vol. 3, pp. 1433–1452, Madrid, Spain, 2006.
- [4] Y. Wang, X. Li, K. Xu, F. Ren, and H. Yu, "Data-driven sampling matrix boolean optimization for energy-efficient biomedical signal acquisition by compressive sensing," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 11, no. 2, pp. 255–266, 2017.
- [5] B. Liu and Z. Zhang, "Quantized compressive sensing for low-power data compression and wireless telemonitoring," *IEEE Sensors Journal*, vol. 16, no. 23, pp. 8206–8213, 2016.
- [6] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.
- [7] F.-Y. Wu, K. Yang, and Z. Yang, "Compressed acquisition and denoising recovery of EMGdi signal in WSNs and iot," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 5, pp. 2210–2219, 2018.
- [8] H. Boche, R. Calderbank, G. Kutyniok, and J. Vybiral, *Compressed Sensing and its Applications*, Springer, Berlin, Germany, 2015.
- [9] M. F. Duarte, M. A. Davenport, D. Takhar et al., "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 83–91, 2008.
- [10] Z. Wei, J. Zhang, Z. Xu et al., "Improving the signal-to-noise ratio of super resolution imaging based on single pixel camera," *IEEE Photonics Journal*, vol. 11, no. 1, pp. 1–16, 2019.
- [11] A. C. Sankaranarayanan, C. Studer, and R. G. Baraniuk, "Csmuvi: video compressive sensing for spatial multiplexing cameras," in *Proceedings of IEEE International Conference on Computational Photography*, pp. 1–10, IEEE, Seattle, WA, USA, April 2012.
- [12] N. Radwell, K. J. Mitchell, G. M. Gibson, M. P. Edgar, R. Bowman, and M. J. Padgett, "Single-pixel infrared and visible microscope," *Optica*, vol. 1, no. 5, pp. 285–289, 2014.
- [13] M. P. Edgar, G. M. Gibson, R. W. Bowman et al., "Simultaneous real-time visible and infrared video with single-pixel detectors," *Sci. Rep.* vol. 5, p. 10669, 2015.
- [14] C. Chen, E. W. Tramel, and J. E. Fowler, "Compressed-sensing recovery of images and video using multihypothesis predictions," in *Proceedings of the 2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Pacific Grove, CA, USA, November 2011.
- [15] J. Zhang, Y. Yin, and Y. Yin, "Adaptive compressed sensing for wireless image sensor networks," *Multimedia Tools and Applications*, vol. 76, pp. 1–16, 2017.
- [16] Y. Gan, Z. Tong, and H. Chu, "Image classification with a deep network model based on compressive sensing," in *Proceedings of the International Conference on Signal Processing*, Hangzhou, China, October 2015.
- [17] D. Milioris, "Classification in twitter via compressive sensing," in *Proceedings of the Computer Communications Workshops*, IEEE, Hong Kong, China, May 2015.
- [18] G. J. J. Warmerdam, R. Vullings, L. Schmitt, J. O. E. H. Van Laar, and J. W. M. Bergmans, "Hierarchical probabilistic framework for fetal r-peak detection, using ecg waveform and heart rate information," *IEEE Transactions on Signal Processing*, vol. 66, no. 16, pp. 4388–4397, 2018.
- [19] H. Kim and S. Y. Chun, "Cancelable ecg biometrics using compressive sensing-generalized likelihood ratio test," *IEEE Access*, vol. 7, pp. 9232–9242, 2019.
- [20] S. K. Jain and B. Bhaumik, "An energy efficient ECG signal processor detecting cardiovascular diseases on smartphone," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 11, no. 2, pp. 314–323, 2017.
- [21] J. Zhang, J. Teng, and Y. Bai, "Improving sparse compressed sensing medical CT image reconstruction," *Automatic Control & Computer Sciences*, vol. 53, no. 3, pp. 281–289, 2019.
- [22] K. Haris, E. Hedström, F. Kording et al., "Free-breathing fetal cardiac MRI with Doppler ultrasound gating, compressed sensing, and motion compensation," *Journal of Magnetic Resonance Imaging*, vol. 51, no. 1, pp. 260–272, 2019.
- [23] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: the application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.
- [24] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, "*k-t* FOCUSS: a general compressed sensing framework for high resolution dynamic MRI," *Magnetic Resonance in Medicine*, vol. 61, no. 1, pp. 103–116, 2009.
- [25] J. Zhang, Z. Gu, Z. L. Yu, and Y. Li, "Energy-efficient ECG compression on wireless biosensors via minimal coherence sensing and weighted l1 minimization reconstruction," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 2, pp. 520–528, 2015.
- [26] D. Craven, B. McGinley, L. Kilmartin, M. Glavin, and E. Jones, "Energy-efficient compressed sensing for ambulatory ECG monitoring," *Computers in Biology and Medicine*, vol. 71, pp. 1–13, 2016.
- [27] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE*

- Transactions on Information Theory* vol. 52, no. 12, pp. 5406–5425, 2006.
- [28] Q. Wang, D. Li, and Y. Shen, “Intelligent nonconvex compressive sensing using prior information for image reconstruction by sparse representation,” *Neurocomputing*, vol. 224, pp. 71–81, 2017.
- [29] F.-Y. Wu, F. Tong, and Z. Yang, “EMGdi signal enhancement based on ICA decomposition and wavelet transform,” *Applied Soft Computing*, vol. 43, pp. 561–571, 2016.
- [30] R. Baraniuk, “Compressive sensing lecture notes,” *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 2007.
- [31] M. Mangia, R. F. Pareschi, and G. Setti, “Adapted compressed sensing: a game worth playing,” *IEEE Circuits and Systems Magazine*, vol. 20, no. 1, pp. 40–60, 2020.
- [32] N. Cleju, “Optimized projections for compressed sensing via rank-constrained nearest correlation matrix,” *Applied and Computational Harmonic Analysis*, vol. 36, no. 3, pp. 495–507, 2014.
- [33] J. Xu, Y. Pi, and Z. Cao, “Optimized projection matrix for compressive sensing,” *Eurasip Journal on Advances in Signal Processing*, vol. 2010, no. 1, p. 560349, 2010.
- [34] D. Needell and R. S. Vershynin, “Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 310–316, 2010.
- [35] J. Wang, S. Kwon, and B. Shim, “Generalized orthogonal matching pursuit,” *IEEE Transactions on Signal Processing*, vol. 60, pp. 6202–6216, 2012.
- [36] Z. Xu, “The performance of orthogonal multi-matching pursuit under the restricted isometry property,” *Journal of Computational Mathematics*, vol. 33, no. 5, pp. 495–516, 2015.
- [37] J. Wang and B. Shim, “Exact recovery of sparse signals using orthogonal matching pursuit: how many iterations do we need?” *IEEE Transactions on Signal Processing*, vol. 64, no. 16, pp. 4194–4202, 2016.
- [38] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, “Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition,” in *Proceedings of the 27th Annual Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, November 1993.
- [39] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.
- [40] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, “An interior-point method for large-scale ℓ_1 -regularized least squares,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007.
- [41] Z. Wei, J. Zhang, Z. Xu, Y. Huang, Y. Liu, and X. Fan, “Gradient projection with approximate L_0 norm minimization for sparse reconstruction in compressed sensing,” *Sensors*, vol. 18, no. 10, p. 3373, 2018.
- [42] T. Blumensath and M. E. Davies, “Iterative hard thresholding for compressed sensing,” *Applied & Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2008.
- [43] K. Bredies and D. A. Lorenz, “Iterative soft-thresholding converges linearly,” *Math Subject Classifications, Journal of Fourier Analysis and Applications*, vol. 14, pp. 813–837, 2007.
- [44] V. Abolghasemi, S. Ferdowsi, B. Makkiabadi et al., “On optimization of the measurement matrix for compressive sensing,” in *Proceedings of the 18th European Signal Processing Conference*, pp. 427–431, Aalborg, Denmark, August 2010.
- [45] E. J. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [46] C. Lu, H. Li, and Z. Lin, “Optimized projections for compressed sensing via direct mutual coherence minimization,” *Signal Processing*, vol. 15, pp. 45–55, 2015.
- [47] M. Elad, “Optimized projections for compressed sensing,” *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5695–5702, 2007.
- [48] Q. Wang and G. Qu, “Restricted isometry constant improvement based on a singular value decomposition-weighted measurement matrix for compressed sensing,” *Iet Communications*, vol. 11, no. 11, pp. 1706–1718, 2017.
- [49] V. Tiwari, P. P. Bansod, and A. Kumar, “Designing sparse sensing matrix for compressive sensing to reconstruct high resolution medical images,” *Cogent Engineering*, vol. 2, no. 1, p. 1017244, 2015.
- [50] A. Aziz, K. Singh, W. Osamy, and A. M. Khedr, “Optimising compressive sensing matrix using Chicken Swarm Optimization algorithm,” *IET Wireless Sensor Systems*, vol. 9, no. 5, pp. 306–312, 2019.
- [51] V. Abolghasemi, S. S. Ferdowsi, and S. Sanei, “A gradient-based alternating minimization approach for optimization of the measurement matrix in compressive sensing,” *Signal Processing*, vol. 92, no. 4, pp. 999–1009, 2012.
- [52] Y. J. Chen, Q. Zhang, Y. Luo et al., “Measurement matrix optimization for ISAR sparse imaging based on genetic algorithm,” *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1875–1879, 2016.
- [53] E. Candès and J. Romberg, “Sparsity and incoherence in compressive sampling,” *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007.
- [54] Z. Wei, Z. J. Zhang, and Y. Liu, “Optimization methods of compressively sensed image reconstruction based on single-pixel imaging,” *Applied Sciences*, vol. 10, no. 9, p. 3288, 2020.
- [55] S. Foucart and H. Rauhut, “A mathematical introduction to compressive sensing,” in *Applied and Numerical Harmonic Analysis*, J. J. Benedetto, Ed., Springer, New York, NY, USA, 2013.
- [56] D. L. Donoho and M. Elad, “Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization,” *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.
- [57] J. M. Duarte-Carvajalino, G. Sapiro, Sapiro et al., “Learning to sense sparse signals: simultaneous sensing matrix and sparsifying dictionary optimization,” *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1395–1408, 2009.
- [58] D. Donoho and Y. Tsaig, “Extensions of compressed sensing,” *Signal Processing*, vol. 86, no. 3, pp. 533–548, 2006.
- [59] J. E. Fowler, S. Mun, and E. W. Tramel, *Block-Based Compressed Sensing of Images and Video*, Now Publishers Inc., Boston, MA, USA, 2012.