

Research Article

A Descent Conjugate Gradient Algorithm for Optimization Problems and Its Applications in Image Restoration and Compression Sensing

Junyue Cao^{1,2} and Jinzhao Wu^{3,4} 

¹Chengdu Institute of Computer Application Chinese Academy of Sciences, Chengdu, China

²University of the Chinese Academy of Sciences, Beijing, China

³The School of Computer and Electronic Information, Guangxi University, Nanning 530004, China

⁴Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis, Guangxi University for Nationalities, Nanning 530006, China

Correspondence should be addressed to Jinzhao Wu; tony_gx@aliyun.com

Received 7 July 2020; Accepted 25 August 2020; Published 29 September 2020

Guest Editor: Wenjie Liu

Copyright © 2020 Junyue Cao and Jinzhao Wu. 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.

It is well known that the nonlinear conjugate gradient algorithm is one of the effective algorithms for optimization problems since it has low storage and simple structure properties. This motivates us to make a further study to design a modified conjugate gradient formula for the optimization model, and this proposed conjugate gradient algorithm possesses several properties: (1) the search direction possesses not only the gradient value but also the function value; (2) the presented direction has both the sufficient descent property and the trust region feature; (3) the proposed algorithm has the global convergence for nonconvex functions; (4) the experiment is done for the image restoration problems and compression sensing to prove the performance of the new algorithm.

1. Introduction

Consider the following model defined by

$$\min\{f(x) \mid x \in R^n\}, \quad (1)$$

where $f: R^n \rightarrow R$ is a continuous function. The above problem (1) has many practical applied fields, such as economics, biology, and engineering. It is well known that the nonlinear conjugate gradient (CG) method is one of the most effective methods for (1). The CG algorithm has the following iterative formula with

$$x_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, 2, \dots, \quad (2)$$

where α_k denotes the steplength, x_k is the k th iterative point, and d_k is the search direction designed by

$$d_k = \begin{cases} -g_k + \beta_k d_{k-1}, & \text{if } k \geq 1, \\ -g_k, & \text{if } k = 0, \end{cases} \quad (3)$$

where $g_k = \nabla f(x_k)$ is the gradient and β_k is a scalar which determines different CG algorithms ([1–7], etc.), where the Polak–Ribière–Polak (PRP) formula [6, 7] is one of the well-known nonlinear CG formulas with

$$\beta_k^{\text{PRP}} = \frac{g_k^T (g_k - g_{k-1})}{\|g_{k-1}\|^2}, \quad (4)$$

where $g_{k-1} = \nabla f(x_{k-1})$ and $\|\cdot\|$ is the Euclidean norm. The PRP method has been studied by many scholars, and many results are obtained (see [7–12], etc.) since the PRP algorithm has the superior numerical performance but for convergence. At present, under the weak Wolfe–Powell (WWP) inexact line search and for nonconvex functions, the global convergence of the PRP algorithm is still open, and it

is one of the well-known open problems in optimization fields. Based on the PRP formula, many modified nonlinear CG formulas are done ([13–16], etc.) because many scholars want to use the perfect numerical attitude of it. Recently, Yuan et al. [17] open up a new way by modifying the WWP line search technique and partly proved the global convergence of the PRP algorithm. Further results are obtained (see [18–20], etc.) by this technique. It has been proved that the nonlinear CG algorithms can be used in nonlinear equations, nonsmooth optimization, and image restoration problems (see [21–24], etc.).

We all know that the sufficient descent property designed by

$$d_k^T g_k \leq -c \|g_k\|^2, \quad c > 0, \quad (5)$$

plays an important role for convergence analysis in CG methods (see [13, 14, 24], etc.), where $c > 0$ is a constant. There is another crucial condition about the scalar $\beta_k \geq 0$ that has been pointed out by Powell [10] and further emphasized in the global convergence [11, 12]. Thus, under the assumption of the sufficient descent condition and the WWP technique, a modified PRP formula $\beta_k^{\text{PRP}^+} = \max\{0, \beta_k^{\text{PRP}}\}$ is presented by Gilbert and Nocedal [13], and its global convergence for nonconvex functions is established. All of these observations tell us that both property (5) and $\beta_k \geq 0$ are very important in the CG algorithms. To get one of the conditions or both of them, many scholars made a further study and got many interesting results. Yu [25] presented a modified PRP nonlinear CG formula designed by

$$\beta_k^{\text{mPRP1}} = \frac{g_{k+1}^T y_k}{\|g_k\|} - \mu \frac{\|y_k\| \|g_{k+1}^T d_k\|}{\|g_k\|^4}, \quad (6)$$

where $\mu > (1/4)$ is a positive constant and $y_k = g_{k+1} - g_k$, which has property (5) with $c = 1 - (1/4\mu)$. Yuan [12] proposed a further formula defined by

$$\beta_k^{\text{mPRP2}} = \frac{g_{k+1}^T y_k}{\|g_k\|^2} - \min \left\{ \frac{g_{k+1}^T y_k}{\|g_k\|^2}, \mu \frac{\|y_k\| \|g_{k+1}^T d_k\|}{\|g_k\|^4} \right\}, \quad (7)$$

which possesses not only property (5) with $c = 1 - (1/4\mu)$ but also the scalar $\beta_k^{\text{mPRP2}} \geq 0$. To get a greater drop, a three-term FR CG formula is given by Zhang et al. [26]:

$$d_{k+1} = -\theta_k g_{k+1} + \beta_k^{\text{FR}} d_k, \quad (8)$$

$$\theta_k = \frac{d_k^T y_k}{\|g_k\|^2},$$

where it has (5) with $c = -1$. Dai and Tian [27] gave another CG direction designed by

$$d_{k+1} = \begin{cases} -\left(1 + \beta_k \frac{d_k^T g_{k+1}}{\|g_k\|^2}\right) g_{k+1} + \beta_k d_k, & \text{if } k \geq 0, \\ -g_k, & \text{if } k = 0, \end{cases} \quad (9)$$

which also possesses (5) with $c = -1$. The global convergence of the above CG method is proved by Dai and Tian [27] for $\beta_k = \beta_k^{\text{YuanPRP}}$ and $\beta_k = \beta_k^{\text{YuPRP}}$. For nonconvex functions and the effective Armijo line search, they did not analyze them. One the main reasons lies in the trust region feature. To overcome it, we [28] proposed a CG formula designed by

$$d_{k+1} = \begin{cases} -g_k, & \text{if } k = 0, \\ -g_{k+1} + \frac{\beta_k d_k - \beta_k \left(d_k^T g_{k+1} / \|g_{k+1}\|^2 \right) g_{k+1}}{\gamma_k}, & \text{if } k \geq 0, \end{cases} \quad (10)$$

where $\gamma_k = (\|\beta_k\| \|d_k\| / \|g_{k+1}\|)$, which possesses not only (5) with $c = -1$ but also the trust region property. It has been proved that the CG formula will have better numerical performance if it possesses not only the gradient value information but also the function value [29]. This motivates us to present a CG formula based on (10) designed by

$$d_{k+1} = \begin{cases} -g_k, & \text{if } k = 0, \\ -g_{k+1} + \frac{\beta_k^* d_k - \beta_k^* \left(d_k^T g_{k+1} / \|g_{k+1}\|^2 \right) g_{k+1}}{\gamma_k^*}, & \text{if } k \geq 0, \end{cases} \quad (11)$$

where $\gamma_k^* = (\|\beta_k^*\| \|d_k\| / \|g_{k+1}\|)$, $\beta_k^* = \beta_k^{\text{mPRP2}^*} = (g_{k+1}^T y_k^* / \|g_k\|^2) - \min\{(g_{k+1}^T y_k^* / \|g_k\|^2), \mu (\|y_k^*\| \|g_{k+1}^T d_k\| / \|g_k\|^4)\}$, and $y_k^* = y_k + \rho_k$ with $\rho_k = (\max\{\varrho_k, 0\} / \|s_k\|^2)$ and $\varrho_k = 2[f(x_k) - f(x_{k+1})] + [g_{k+1} + g_k]^T s_k$ and $s_k = x_{k+1} - x_k$. The new vector y_k^* [30] has been proved that it has some good properties in theory and experiment. Yuan et al. [29] use it in the CG formula and get some good results. These achievements inspire us to propose the new CG direction (11), and this paper possesses the following features:

- (i) The sufficient property and the trust region feature are obtained
- (ii) The new direction possesses not only the gradient value but also the function value
- (iii) The given algorithm has the global convergence under the Armijo line search for nonconvex functions
- (iv) The experiments for image restoration problems and compression sensing are done to test the performance of the new algorithm

The next section states the given algorithm. The convergence analysis is given in Section 3, and experiments are done in Section 4, respectively. The last section proposes one conclusion.

2. Algorithm

Based on the discussions of the above section, the CG algorithm is listed in Algorithm 1.

Initial step: given any initial point $x_0 \in R^n$ and positive constants $\epsilon \in (0, 1), \sigma_0 > 0, \delta \in (0, 1), \mu > (1/4), \sigma \in (0, 1)$, set $d_0 = -g_0 = -\nabla f(x_0)$ and $k := 0$.
 Step 1: stop if $\|g_k\| \leq \epsilon$ is true.
 Step 2: find $\alpha_k = \sigma_0 \sigma^{i_k}$ such that

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g(x_k)^T d_k,$$
 where $i_k = \min\{0, 1, 2, \dots\}$ satisfying the equation in Step 2 mentioned in Algorithm 1.
 Step 3: set $x_{k+1} = x_k + \alpha_k d_k$.
 Step 4: stop if $\|g_{k+1}\| \leq \epsilon$ holds.
 Step 5: compute d_k by (11).
 Step 6: set $k := k + 1$ and go to Step 2.

ALGORITHM 1: Three-term conjugate gradient algorithm.

Theorem 1. *The direction d_k is defined by (11); then there exists a positive $\beta > 0$ satisfying*

$$d_k^T g_k = -\|g_k\|^2, \quad \forall k \geq 0, \quad (12)$$

$$\|d_k\| \leq \beta \|g_k\|, \quad \forall k \geq 0. \quad (13)$$

Proof. By (11), we directly get (12) and (13) for $k = 0$ with $\beta = -1$. If $k \geq 0$, using (11) again, we have

$$\begin{aligned} g_{k+1}^T d_{k+1} &= g_{k+1}^T \left[-g_{k+1} + \frac{\beta_k^{mPRP2} d_k - \beta_k^{mPRP2} \left(d_k^T g_{k+1} / \|g_{k+1}\|^2 \right) g_k}{\gamma_k^*} \right] \\ &= -g_{k+1}^T g_{k+1} + \frac{\beta_k^{mPRP2} g_{k+1}^T d_k - \beta_k^{mPRP2} \left(d_k^T g_{k+1} / \|g_{k+1}\|^2 \right) g_{k+1}^T g_k}{\left(|\beta_k^{mPRP2}| \|d_k\| / \|g_{k+1}\| \right)} \\ &= -\|g_{k+1}\|^2 + \frac{\beta_k^{mPRP2} g_{k+1}^T d_k - \beta_k^{mPRP2} d_k^T g_{k+1}}{\left(|\beta_k^{mPRP2}| \|d_k\| / \|g_{k+1}\| \right)} \\ &= -\|g_{k+1}\|^2, \end{aligned} \quad (14)$$

then (12) is true. By (11) again, we can get

$$\begin{aligned} \|d_{k+1}\| &= \left\| -g_{k+1} + \frac{\beta_k^{mPRP2} d_k - \beta_k^{mPRP2} \frac{d_k^T g_{k+1}}{\|g_{k+1}\|^2} g_k}{\gamma_k^*} \right\| \\ &\leq \|g_{k+1}\| + \frac{|\beta_k^{mPRP2}| \|d_k\| + |\beta_k^{mPRP2}| \left(\|d_k\| \|g_{k+1}\| / \|g_{k+1}\|^2 \right) \|g_k\|}{\left(|\beta_k^{mPRP2}| \|d_k\| / \|g_{k+1}\| \right)} \\ &= \|g_{k+1}\| + \frac{2|\beta_k^{mPRP2}| \|d_k\|}{\left(|\beta_k^{mPRP2}| \|d_k\| / \|g_{k+1}\| \right)} \\ &= 3\|g_{k+1}\|, \end{aligned} \quad (15)$$

which implies that (13) holds by choosing $\beta \in [3, +\infty)$. We complete the proof.

Remark 1. The relation (13) is the so-called trust region feature, and the above theorem tells us that direction (11) has not only the sufficient descent property but also the trust region feature. Both these relations (12) and (13) will make the proof of the global convergence of Algorithm 1 be easy to be established.

3. Global Convergence

For the nonconvex functions, the global convergence of Algorithm 1 is established under the following assumptions.

Assumption 1. Assume that the function $f(x)$ has at least a stationary point x^* , namely, $\|g(x^*)\| = 0$ is true. Suppose that the level set defined by $L_0 = \{x \mid f(x) \leq f(x_0)\}$ is bounded.

Assumption 2. The function $f(x)$ is twice continuously differentiable and bounded below, and its $g(x)$ is Lipschitz continuous. We also assume that there exists a positive constant $L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\|, \quad x, y \in R^n. \quad (16)$$

Now, we prove the global convergence of Algorithm 1.

Theorem 2. *Let Assumption 1 be true. Then, we get*

$$\lim_{k \rightarrow \infty} \|g_k\| = 0. \quad (17)$$

Proof. Using (12) and the Step 2 of Algorithm 1, we obtain

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g(x_k)^T d_k < f_k - \delta \alpha_k \|g(x_k)\|^2, \quad (18)$$

which means that the sequence $\{f(x_k)\}$ is descent and the following relation

$$\delta \alpha_k \|g(x_k)\|^2 \leq f(x_k) - f(x_k + \alpha_k d_k), \quad (19)$$

is true. For $k = 0$ to ∞ , by summing the above inequalities and Assumption 1, we deduce that

$$\sum_{k=0}^{\infty} \delta \alpha_k \|g(x_k)\|^2 \leq f(x_0) - f_{\infty} < +\infty, \quad (20)$$

holds. Thus, we have

$$\lim_{k \rightarrow \infty} \alpha_k \|g(x_k)\|^2 = 0. \quad (21)$$

This implies that

$$\lim_{k \rightarrow \infty} \|g(x_k)\| = 0, \quad (22)$$

or

$$\lim_{k \rightarrow \infty} \alpha_k = 0. \quad (23)$$

Suppose that (22) holds, the proof of this theorem is complete. Assuming that (23) is true, we aim to get (17). Let the stepsize α_k satisfy the equation in Step 2 in Algorithm 1; for $\alpha_k^* = (\alpha_k/\sigma)$, we have

$$f(x_k + \alpha_k^* d_k) > f(x_k) + \delta \alpha_k^* d_k^T g(x_k). \quad (24)$$

By (12) and (13) and the well-known mean value theorem, we obtain

$$\begin{aligned} f(x_k + \alpha_k^* d_k) - f(x_k) &= \alpha_k^* d_k^T g(x_k) + (\alpha_k^*)^2 O(\|d_k\|^2) \\ &= -\frac{\alpha_k}{\sigma} \|g(x_k)\|^2 + \frac{\alpha_k^2}{\sigma^2} O(\|d_k\|^2) \\ &> \delta \alpha_k^* d_k^T g(x_k) \\ &= -\delta \frac{\alpha_k}{\sigma} \|g(x_k)\|^2, \end{aligned} \quad (25)$$

which implies that

$$\alpha_k > \frac{\sigma(1-\delta)\|g(x_k)\|^2}{O(\|d_k\|^2)} \geq \sigma(1-\delta)O\left(\frac{1}{\beta}\right), \quad (26)$$

is true. This is a contradiction to (23). Then, only relation (22) holds. We complete the proof.

Remark 2. We can see that the proof process of the global convergence is very simple since the defined direction (11) has not only the good sufficient descent property (12) but also the perfect trust region feature (13).

4. Numerical Results

The numerical experiments for image restoration problems and compression sensing will be done by Algorithm 1 and the normal PRP algorithm, respectively. All codes are run on a PC with an Intel (R) Core (TM) i7-7700T CPU @ 2.9 GHz, 16.00 GB of RAM, and the Windows 10 operating system and written by MATLAB r2014a. The parameters are chosen as $\sigma = 0.5$, $\sigma_0 = 0.1$, $\delta = 0.9$, and $\mu = 300$.

4.1. Image Restoration Problems. Setting x be the true image which has $M \times N$ pixels and $(i, j) \in A = \{1, 2, \dots, M\} \times \{1, 2, 3, \dots, N\}$. At a pixel location (i, j) , $x_{i,j}$ denotes the gray level of x . Then, defining a set N by

$$N := \{(i, j) \in A \mid \bar{\zeta}_{i,j} \neq \zeta_{i,j}, \zeta_{i,j} = s_{\min} \text{ or } s_{\max}\}, \quad (27)$$

which is the index set of the noise candidates. Suppose that ζ is the observed noisy image of x corrupted by salt-and-pepper noise, we let $\phi_{i,j} = \{(i, j-1), (i, j+1), (i-1, j), (i+1, j)\}$ be the neighborhood of (i, j) . By applying an adaptive median filter to the noisy image y , $\bar{\zeta}$ is defined by the image obtained. s_{\max} denotes the maximum of a noisy pixel, and s_{\min} denotes the minimum of a noisy pixel. The following conclusions can be obtained: (i) if $(i, j) \in N$, then $\zeta_{i,j}$ must



FIGURE 1: Restoration of Barbara, man, Baboon, and Lena by Algorithm 1 and the PRP algorithm. From left to right: a noisy image with 20% salt-and-pepper noise and restorations obtained by minimizing z with the PRP algorithm and Algorithm 1.

be restored. A pixel (i, j) is identified as uncorrupted, and its original value is kept, which means that $w_{i,j}^* = \zeta_{i,j}$ with the element $w_{i,j}^*$ of the denoised image w by the two-phase method. (ii) If $(i, j) \notin N$ holds, $w_{m,n}^* = \zeta_{m,n}$ is stetted and $\zeta_{m,n}$

is restored if $(m, n) \in \phi_{i,j} \cap N$. Chan et al. [31] presented the new function f^α and minimized it for the restored images without a nonsmooth term, which has the following form:

$$f^\alpha(w) = \sum_{(i,j) \in N} \left\{ \sum_{(m,n) \in \phi_{i,j} \setminus N} \psi^\alpha(w_{i,j} - \zeta_{m,n}) + \frac{1}{2} \sum_{(m,n) \in \phi_{i,j} \cap N} \psi^\alpha(w_{i,j} - \zeta_{m,n}) \right\}, \quad (28)$$



FIGURE 2: Restoration of Barbara, man, Baboon, and Lena by Algorithm 1 and the PRP algorithm. From left to right: a noisy image with 40% salt-and-pepper noise and restorations obtained by minimizing z with the PRP algorithm and Algorithm 1.

where α is a constant and ψ^α is an even edge-preserving potential function. The numerical performance of f^α is noteworthy [32, 33].

We choose Barbara (512×512), man (256×256), Baboon (512×512), and Lena (256×256) as the tested images. The well-known PRP CG algorithm (PRP algorithm) is also done to compare with Algorithm 1. The detailed performances are listed in Figures 1 and 2.

Figures 1 and 2 tell us that these two algorithms (Algorithm 1 and the PRP algorithm) are successful to solve these image restoration problems, and the results are good.

To directly compare their performances, the restoration performance is assessed by applying the peak signal-to-noise ratio (PSNR) defined in [34–36] which is computed and listed in Table 1. From the value of Table 1, we can see that Algorithm 1 is competitive to the PRP algorithm since its PSNR value is less than that of the PRP algorithm.

4.2. Compressive Sensing. In this section, the following compressive sensing images are tested: Phantom (256×256), Fruits (256×256), and Boat (256×256). These three images are treated as 256 vectors, and the size of the

TABLE 1: PSNR Algorithm 1 and the PRP algorithm.

20% noise	Barbara	Man	Baboon	Lena	Average
Algorithm 1	31.115	38.0355	29.4393	41.0674	34.9143
PRP algorithm	31.1118	37.9583	29.4534	41.356	34.969
40% noise	Barbara	Man	Baboon	Lena	Average
Algorithm 1	27.5415	34.0063	25.8947	36.6496	31.0230
PRP algorithm	27.6153	34.5375	25.8571	36.701	31.1777

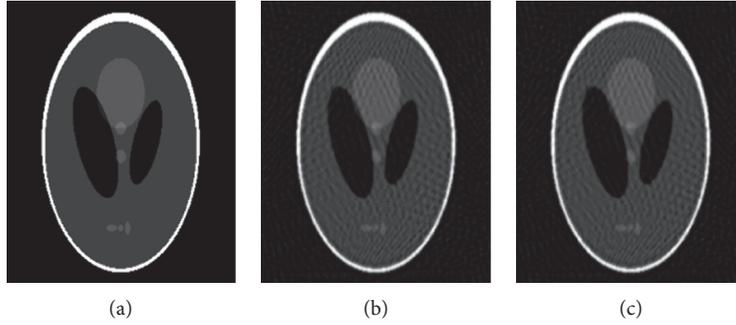


FIGURE 3: Phantom: (a) the general images, (b) the recovered images by Algorithm 1, and (c) the recovered images by the PRP algorithm.

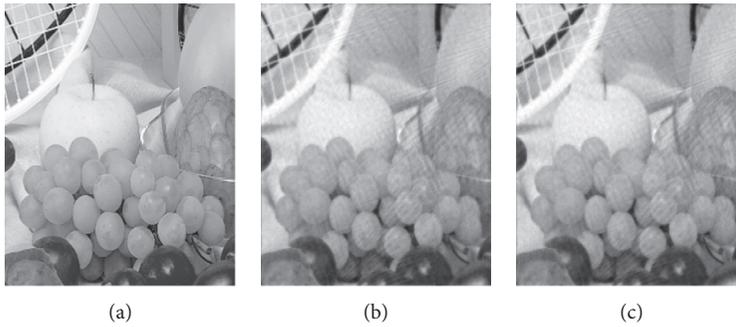


FIGURE 4: Fruits: (a) the general images, (b) the recovered images by Algorithm 1, and the (c) recovered images by the PRP algorithm.

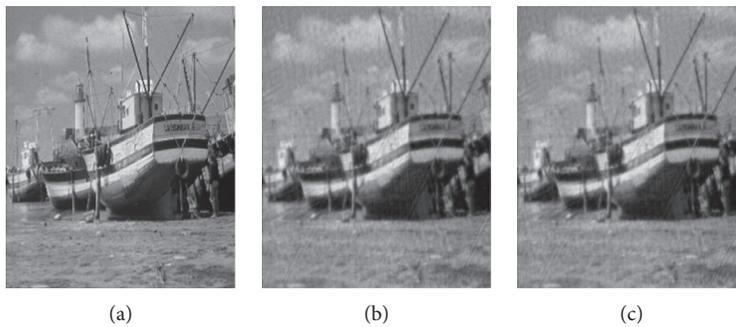


FIGURE 5: Boat: (a) the general images, (b) the recovered images by Algorithm 1, and (c) the recovered images by the PRP algorithm.

observation matrix is 100×256 . The so-called Fourier transform technology is used, and the measurement is the Fourier domain.

Figures 3–5 turn out that these two algorithms work well for these figures, and they can successfully solve them.

5. Conclusion

This paper, by designing a CG algorithm, studies the unconstrained optimization problems. The given method possesses not only the sufficient descent property but also

the trust region feature. The global convergence is proved by a simple way. The image restoration problems and compressive sensing problems are tested to show that the proposed algorithm is better than the normal algorithm. In the future, we will focus on the following aspects to be paid attention: (i) we believe there are many perfect CG algorithms which can be successfully used for image restoration problems and compressive sensing; (ii) more experiments will be done to test the performance of the new algorithm.

Data Availability

All data are included in the paper.

Conflicts of Interest

There are no potential conflicts of interest.

Acknowledgments

The authors would like to thank the support of the funds. This work was supported by the National Natural Science Foundation of China under Grant no. 61772006, the Science and Technology Program of Guangxi under Grant no. AB17129012, the Science and Technology Major Project of Guangxi under Grant no. AA17204096, the Special Fund for Scientific and Technological Bases and Talents of Guangxi under Grant no. 2016AD05050, and the Special Fund for Bagui Scholars of Guangxi.

References

- [1] Y. Dai and Y. Yuan, "A nonlinear conjugate gradient with a strong global convergence properties," *SIAM Journal on Optimization*, vol. 10, pp. 177–182, 2000.
- [2] R. Fletcher, *Practical Methods of Optimization*, John Wiley and Sons, New York, NY, USA, 2nd edition, 1987.
- [3] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, no. 2, pp. 149–154, 1964.
- [4] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, 1952.
- [5] Y. Liu and C. Storey, "Efficient generalized conjugate gradient algorithms, part 1: theory," *Journal of Optimization Theory and Applications*, vol. 69, no. 1, pp. 129–137, 1991.
- [6] B. T. Polak, "The conjugate gradient method in extreme problems," *Computational Mathematics and Mathematical Physics*, vol. 9, no. 4, pp. 94–112, 1969.
- [7] E. Polak and G. Ribière, "Note sur la convergence de méthodes de directions conjuguées," *Revue française d'informatique et de recherche opérationnelle. Série rouge*, vol. 3, no. 16, pp. 35–43, 1969.
- [8] Y. Dai, "Convergence properties of the BFGS algorithm," *SIAM Journal on Optimization*, vol. 13, no. 3, pp. 693–701, 2002.
- [9] Y. Dai, "Analysis of conjugate gradient methods," Ph.D. Thesis, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, Beijing, China, 1997.
- [10] M. J. D. Powell, "Nonconvex minimization calculations and the conjugate gradient method," *Lecture Notes in Mathematics*, vol. 1066, pp. 122–141, Springer-Verlag, Berlin, Germany, 1984.
- [11] M. J. D. Powell, "Convergence properties of algorithms for nonlinear optimization," *SIAM Review*, vol. 28, no. 4, pp. 487–500, 1986.
- [12] G. Yuan, "Modified nonlinear conjugate gradient methods with sufficient descent property for large-scale optimization problems," *Optimization Letters*, vol. 3, no. 1, pp. 11–21, 2009.
- [13] J. C. Gilbert and J. Nocedal, "Global convergence properties of conjugate gradient methods for optimization," *SIAM Journal on Optimization*, vol. 2, no. 1, pp. 21–42, 1992.
- [14] W. W. Hager and H. Zhang, "A new conjugate gradient method with guaranteed descent and an efficient line search," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 170–192, 2005.
- [15] W. W. Hager and H. Zhang, "Algorithm 851," *ACM Transactions on Mathematical Software*, vol. 32, no. 1, pp. 113–137, 2006.
- [16] Z. Wei, S. Yao, and L. Liu, "The convergence properties of some new conjugate gradient methods," *Applied Mathematics and Computation*, vol. 183, no. 2, pp. 1341–1350, 2006.
- [17] G. Yuan, Z. Wei, and X. Lu, "Global convergence of BFGS and PRP methods under a modified weak Wolfe-Powell line search," *Applied Mathematical Modelling*, vol. 47, pp. 811–825, 2017.
- [18] X. Li, S. Wang, Z. Jin, and H. Pham, "A conjugate gradient algorithm under Yuan-Wei-Lu line search technique for large-scale minimization optimization models," *Mathematical Problems in Engineering*, vol. 2018, Article ID 4729318, 11 pages, 2018.
- [19] G. Yuan, Z. Sheng, B. Wang, W. Hu, and C. Li, "The global convergence of a modified BFGS method for nonconvex functions," *Journal of Computational and Applied Mathematics*, vol. 327, pp. 274–294, 2018.
- [20] G. Yuan, Z. Wei, and Y. Yang, "The global convergence of the Polak-Ribière-Polyak conjugate gradient algorithm under inexact line search for nonconvex functions," *Journal of Computational and Applied Mathematics*, vol. 362, pp. 262–275, 2019.
- [21] J. Cao and J. Wu, "A conjugate gradient algorithm and its applications in image restoration," *Applied Numerical Mathematics*, vol. 152, pp. 243–252, 2020.
- [22] G. Yuan, T. Li, and W. Hu, "A conjugate gradient algorithm for large-scale nonlinear equations and image restoration problems," *Applied Numerical Mathematics*, vol. 147, pp. 129–141, 2020.
- [23] G. Yuan, J. Lu, and Z. Wang, "The PRP conjugate gradient algorithm with a modified WWP line search and its application in the image restoration problems," *Applied Numerical Mathematics*, vol. 152, pp. 1–11, 2020.
- [24] G. Yuan, Z. Meng, and Y. Li, "A modified Hestenes and Stiefel conjugate gradient algorithm for large-scale nonsmooth minimizations and nonlinear equations," *Journal of Optimization Theory and Applications*, vol. 168, no. 1, pp. 129–152, 2016.
- [25] G. Yu, "Nonlinear self-scaling conjugate gradient methods for large-scale optimization problems," Thesis of Doctors Degree, Sun Yat-Sen University, Guangzhou, China, 2007.
- [26] L. Zhang, W. Zhou, and D. Li, "Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search," *Numerische Mathematik*, vol. 104, no. 4, pp. 561–572, 2006.

- [27] Z.-F. Dai and B.-S. Tian, "Global convergence of some modified PRP nonlinear conjugate gradient methods," *Optimization Letters*, vol. 5, no. 4, pp. 615–630, 2011.
- [28] J. Cao and J. Wu, "A conjugate gradient algorithm and its applications in image restoration," *Applied Numerical Mathematics*, vol. 152, pp. 243–252, 2019.
- [29] G. Yuan, Z. Wei, and Q. Zhao, "A modified Polak-Ribière-Polyak conjugate gradient algorithm for large-scale optimization problems," *IIE Transactions*, vol. 46, no. 4, pp. 397–413, 2014.
- [30] G. Yuan and Z. Wei, "Convergence analysis of a modified BFGS method on convex minimizations," *Computational Optimization and Applications*, vol. 47, no. 2, pp. 237–255, 2010.
- [31] R. H. Chan, C. W. Ho, C. Y. Leung, and M. Nikolova, "Minimization of detail-preserving regularization functional by Newton's method with continuation," in *Proceedings of IEEE International Conference on Image Processing*, pp. 125–128, Genova, Italy, September 2005.
- [32] J. F. Cai, R. H. Chan, and B. Morini, "Minimization of an edge-preserving regularization functional by conjugate gradient types methods," in *Image Processing Based on Partial Differential Equations*, pp. 109–122, Springer, Berlin, Germany, 2007.
- [33] Y. Dong, R. H. Chan, and S. Xu, "A detection statistic for random-valued impulse noise," *IEEE Transactions on Image Processing*, vol. 16, no. 4, pp. 1112–1120, 2007.
- [34] A. Bovik, *Handbook of Image and Video Processing*, Academic, New York, NY, USA, 2000.
- [35] F. Rahpeymaii, K. Amini, T. Allahviranloo, and M. R. Malkhalifeh, "A new class of conjugate gradient methods for unconstrained smooth optimization and absolute value equations," *Calcolo*, vol. 56, 2019.
- [36] G. Yu, J. Huang, and Y. Zhou, "A descent spectral conjugate gradient method for impulse noise removal," *Applied Mathematics Letters*, vol. 23, no. 5, pp. 555–560, 2010.