STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–13.
Published online in International Academic Press (www.IAPress.org)

An improved hybrid nonlinear conjugate gradient method and application to image restoration problems

Mehamdia Abd Elhamid [1], Issam A. R. Moghrabi [2], Basim A. Hassan [3,*], Nivine Guler [4]

[1] *Department of Mathematics, Universite M'Hamed Bougara of Boumerdes, Algeria*
[2] *Departement of Information Systems and Technology, Kuwait Technical College, Kuwait*
[3] *College of Computers Sciences and Mathematics, University of Mosul, Iraq*
[4] *Departement of Articial Intelligence, College of Computer Engineering and Science, Prince Mohammad Bin Fahd University, Alkhobar, Saudi Arabia*

Abstract Optimization methods are widely used to obtain the numerical solution of the optimal control problems arising in scientific and engineering computation, especially for solving large-scale problems. In this paper, based on some modern and computationally efficient methods, a new conjugate gradient method (named ICG method) is proposed for unconstrained optimization. Under the strong Wolfe line search (SWLS), the presented method is proven to be sufficient descent at each iteration. Moreover, we proved that the proposed method is globally convergent for arbitrary functions and the line search satisfies the strong Wolfe conditions. Numerical tests demonstrate the effectiveness of the ICG method when compared to certain existing methods in view of the Dolan and Moré performance profile. In particular, the practical application of this method in image restoration problems is explored.

Keywords Hybrid conjugate gradient method, Inexact line search, Descent condition, Global convergence, Numerical comparisons.

AMS 2010 subject classifications 90C30, 65K05

DOI: 10.19139/soic-2310-5070-3366

1. Introduction

Conjugate gradient (CG) methods are a popular family of iterative algorithms to solve large-scale nonlinear optimization problems due to appropriate features such as no need to calculate the second-order derivatives, low storage and computation, and suitable convergence rate. For more references on advances in the CG method, see [13, 14, 15, 16]. Several improved nonlinear conjugate gradient methods and their applications have been extensively investigated in the literature, particularly in nonparametric estimation, image restoration, and regression problems, see [5, 6, 24, 25, 26, 27].

In this paper, we consider the following unconstrained optimization problem

$$\min \{f(x) : x \in \mathbb{R}^n\}, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, is continuously differentiable and its gradient is denoted by $g(x_k) = \nabla f(x_k)$ and $x_k \in \mathbb{R}^n$. Conjugate gradient methods are a class of important methods for solving the above problem often using the following iterative rules

$$x_{k+1} = x_k + \alpha_k d_k, \tag{2}$$

*Correspondence to: Issam A.R. Moghrabi (Email:basimah@uomosul.edu.iq). College of Computers Sciences and Mathematics, University of Mosul, Iraq (210093).

ISSN 2310-5070 (online) ISSN 2311-004X (print)

where x_k, is the current iteration point. The positive scalar α_k is called the step length, usually determined by a suitable inexact line search technique, such as the Wolfe line search conditions

$$\begin{cases} f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k, \\ g_{k+1}^T d_k \geq \sigma g_k^T d_k, \end{cases} \quad (3)$$

or a stronger version of the Wolfe line search (SWLS), given by

$$\begin{cases} f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k, \\ \left| g_{k+1}^T d_k \right| \leq \sigma \left| g_k^T d_k \right|, \end{cases} \quad (4)$$

where the parameters δ and σ satisfy $0 < \delta < \sigma < 1$, and d_k is a search direction computed by

$$d_{k+1} = -g_{k+1} + \beta_k d_k; \; d_0 = -g_0. \quad (5)$$

Distinctive CG methods correspond to different choices for the conjugate gradient coefficient β_k, which in turn lead to quite diverse computational efficiency and convergence results of the corresponding methods. Well-known formulate for β_k are called the Fletcher-Reeves [11], Hestenes-Stiefel [19], Polak, Ribière and Polyak [28, 29], Dai-Yuan +[7], Conjugate-Descent [12] and Liu-Storey [23] formulate. These parameters are given by the following formulae

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}, \beta_k^{DY} = \frac{\|g_{k+1}\|^2}{y_k^T d_k}, \beta_k^{CD} = \frac{\|g_{k+1}\|^2}{-g_k^T d_k},$$
$$\beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}, \quad \beta_k^{HS} = \frac{g_{k+1}^T y_k}{y_k^T d_k}, \quad \beta_k^{LS} = \frac{g_{k+1}^T y_k}{-g_k^T d_k},$$

where $\|.\|$ denotes the Euclidean norm and $y_k = g_{k+1} - g_k$.
The robust convergence of DY, FR, and CD is notable, but their practical effectiveness may be hindered by the presence of jamming phenomena [11]. On the contrary, HS, PRP and LS show efficient numerical performance but may have difficulties in achieving convergence on general functions [30, 31].
The sufficient descent condition stands as a crucial requirement for the global convergence of CG methods. The algorithm produces a search direction that meets the sufficient descent condition, if there is a positive constant $c > 0$, such as this

$$g_k^T d_k \leq -c \|g_k\|^2, \text{ for all } k \geq 0.$$

Computationally, HS, PRP and LS are the most efficient CG methods, because they possess an approximate restart feature when jamming occurs, while we cannot guarantee that these directions are descent even when we use the SWLS. In this context, numerous endeavors have been made to ensure a descent property for these CG methods, leading to the introduction of various modifications to the HS, PRP and LS methods.
Du et al. [10] give two modified CG methods, proposing the following formula

$$\beta_k^{NVHS^*} = \frac{\|g_{k+1}\|^2 - \frac{|g_{k+1}^T g_k|}{\|g_k\|^2} g_{k+1}^T g_k}{y_k^T d_k}, \text{ and } \beta_k^{NVPRP^*} = \frac{\|g_{k+1}\|^2 - \frac{|g_{k+1}^T g_k|}{\|g_k\|^2} g_{k+1}^T g_k}{\|g_k\|^2}.$$

If the SWLS is utilized they prove the sufficient descent and the convergence property of these methods for uniformly convex functions [10]. Recently, Huang proposed a variant of the DY method [20], called the MDY method, in which the parameter β_k is yielded by

$$\beta_k^{MDY} = \frac{\|g_{k+1}\|^2 - \frac{g_{k+1}^T d_k}{\|d_k\|^2} g_{k+1}^T d_k}{y_k^T d_k}.$$

The researchers achieved the sufficient properties and the global convergence of this method in the context of using SWLS, Numerical results confirm that this method is promising for solving large-scale optimization problems [20].
The structure of the paper is as follows. In the next section, the ICG method is presented and the descent property is analyzed. In section 3, the global convergence of the proposed method with an SWLS is proved. Numerical results are reported in section 4. Applications of the proposed method in image restoration problems are conducted in Section 5. Finally, we make a summary of our paper.

2. Motivation and the new formula

Jiang and Jian in [21] proposed two CG methods with good convergence and acceptable computational results, where the conjugate parameters β_k^{IFR} and β_k^{IDY} are calculated by

$$\beta_k^{IFR} = \omega_k \beta_k^{FR} \text{ and } \beta_k^{IDY} = \omega_k \beta_k^{DY}, \text{ with } \omega_k = \frac{|g_{k+1}^T d_k|}{-g_k^T d_k}.$$

In 2023, Wu et al. [32] proposed a modified CG method, denoted by the VRMIL method. The parameter β_k in the VRMIL method is given by

$$\beta_k^{VRMIL} = \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|d_k\|} d_k^T g_{k+1}}{\max\{\|d_k\|^2, \mu\|g_{k+1}\|\|d_k\|\}}, \quad \mu > 1.$$

The convergence of this method with the WLS established and numerical results show that this computational scheme is efficient [32]. Jiang et al. [22] proposed a family of hybrid CG method devices with a reboot procedure as follows

$$d_k = \begin{cases} -g_k, & k = 0, \\ -g_k + \beta_k^{IJYHL} d_k, & k \geq 1 \quad \text{if} \quad -\varphi\|g_{k+1}\| \leq g_{k+1}^T d_k \leq \eta\|g_{k+1}\|, \\ -g_k + \zeta \dfrac{g_k^T q_k}{\|q_k\|^2} q_k, & k \geq 1 \quad \text{otherwise.} \end{cases}$$

and

$$\beta_k^{IJYHL} = \frac{2\left(\|g_{k+1}\|^2 - \frac{\phi_k\|g_{k+1}\|}{\|q_k\|}|g_{k+1}^T q_k|\right)}{\max\{y_k^T d_k, 2\|g_k\|^2\}},$$

where $|\zeta| < 1$, $\varphi > 0$ and $0 \leq \eta < 1$ are constants, ϕ_k is the combined weight parameter which affects the theoretical and numerical properties of an algorithm, and it is usually required to satisfy $0 \leq \phi_k \leq 1$, q_k is an arbitrary nonzero vector, which usually can be chosen as g_{k+1}, g_k, d_k, y_k etc...
The search direction yielded by the Algorithm IJYHL satisfies the sufficient descent condition and the global convergence is established under mild conditions [22].
Each of the methods discussed above introduces powerful enhancements that have proven their effectiveness in numerical experiments. From Jiang and Jian [21], we adopt the scaling factor $\omega_k = |g_{k+1}^T d_k|/(-g_k^T d_k)$. This factor has demonstrated significant positive impact when incorporated into classical methods, as shown by the improved performance of the IFR and IDY methods in [21]. The scaling factor provides an automatic restart feature that effectively prevents the jamming phenomenon, making it a valuable component for any hybrid method.

From the VRMIL method of Wu et al. [32] and the IJYHL framework of Jiang et al. [22], we adopt the strategy of using a hybrid denominator $\max\{y_k^T d_k, \|g_k\|^2, \mu\|g_{k+1}\|\|d_k\|\}$. This intelligent denominator design ensures boundedness of β_k and enhances numerical stability, which contributed significantly to the strong numerical performance reported in [22, 32].

The ICG method proposed in this work strategically combines these two powerful elements into a unified framework. Specifically, we retain the effective scaling factor ω_k from [21] to benefit from its automatic restart capability, and we incorporate the hybrid denominator strategy from [22, 32] to ensure numerical stability. By merging these complementary components, the ICG method achieves superior performance across a wide range of test problems, as confirmed by the numerical results in Section 4, without suffering from the minor limitations of each individual method. The new hybrid choice for parameter β_k is defined by

$$\beta_k^{ICG} = \frac{|g_{k+1}^T d_k|}{-g_k^T d_k} \frac{\|g_{k+1}\|^2 - \rho \max\{0, \omega_k, \psi_k\}}{\max\{y_k^T d_k, \|g_k\|^2, \mu\|g_{k+1}\|\|d_k\|\}}, \tag{6}$$

where $\rho \in [0,1]$, $\mu > 1$ and

$$\omega_k = \frac{|g_{k+1}^T g_k| g_{k+1}^T g_k}{\|g_k\|^2}, \quad \psi_k = \frac{\|g_{k+1}\|}{\|d_k\|} g_{k+1}^T d_k.$$

The search direction d_k of our algorithm is computed by

$$d_0 = -g_0; \quad d_{k+1} = -g_{k+1} + \beta_k^{ICG} d_k. \tag{7}$$

The primary attributes of these methods are as follows.

- A modified CG algorithm for solving unconstrained optimization is developed.
- The search direction generated by the presented method is descent at each iteration with a SWLS. Furthermore, this modification has been proven to be globally convergent under usual assumptions.
- Based on some numerical experiments, the results show that our method inherits the good computational ability of the original methods.
- The proposed method is successfully applied to image restoration problems with less computation cost, indicating the encouraging applicability of our approach.

Now, we present the ICG Algorithm using the SWLS in this part.
ICG Algorithm
Step 1: Initializing.
Select positive constants $0 < \delta < \sigma < 1$, choose any initial point $x_0 \in \mathbb{R}^n$, and let $d_0 = -g_0$.
Step 2: Testing the iterations continuation.
If the $\|g_k\|_\infty \le 10^{-6}$ is satisfied, then stops. Otherwise, go to the next step.
Step 3: Line search.
Find the step length α_k satisfying the SWLS 4 and compute $x_{k+1} = x_k + \alpha_k d_k$.
Step 4: Calculate β_k by the formula 6.
Step 5: Compute the search direction d_k by using 7.
Step 6: Let $k = k + 1$ and go to Step 2.

2.1. The sufficient descent direction

The following theorem establishes that the search direction generated by the ICG algorithm satisfies the sufficient descent property under the strong Wolfe line search.

Theorem 2.1
Suppose the search direction sequence $\{d_k\}_{k\ge 0}$ is generated by the ICG Algorithm. Then, for all k, d_k satisfies

$$g_k^T d_k \le -c\|g_k\|^2, \quad \text{with } c = 1 - \frac{1}{\mu} > 0. \tag{8}$$

Proof
We prove the result by induction. For $k = 0$, we have $g_0^T d_0 = -\|g_0\|^2$, so (8) holds trivially. Assume that 8 holds for some $k \ge 0$. We will show that it also holds for $k+1$. From the second inequality of the SWLS (4) and the induction hypothesis 8, we obtain

$$0 \le \frac{|g_{k+1}^T d_k|}{-g_k^T d_k} \le \sigma. \tag{9}$$

Now, recall the definition of β_k^{ICG} from 6

$$\beta_k^{ICG} = \frac{|g_{k+1}^T d_k|}{-g_k^T d_k} \cdot \frac{\|g_{k+1}\|^2 - \rho \max\{0, \omega_k, \psi_k\}}{M_k},$$

where we define the compact notation

$$M_k = \max\left\{ d_k^T y_k,\ \|g_k\|^2,\ \mu\|g_{k+1}\|\|d_k\| \right\}.$$

Since $\max\{0, \omega_k, \psi_k\} \leq \|g_{k+1}\|^2$ (by construction of ω_k and ψ_k), and using 9, we have the simple bounds

$$0 \leq \beta_k^{ICG} \leq \frac{\sigma \|g_{k+1}\|^2}{M_k} \leq \frac{\|g_{k+1}\|^2}{\mu \|g_{k+1}\| \|d_k\|} = \frac{\|g_{k+1}\|}{\mu \|d_k\|}. \tag{10}$$

The last inequality follows from the fact that $M_k \geq \mu \|g_{k+1}\| \|d_k\|$.

From the SWLS 4 and the induction hypothesis, we also have

$$d_k^T y_k = d_k^T (g_{k+1} - g_k) \geq (1 - \sigma)(-g_k^T d_k) \geq 0, \tag{11}$$

which justifies the positivity of the terms in M_k.

Now, from the recurrence relation for the search direction 7, we have

$$d_{k+1} = -g_{k+1} + \beta_k^{ICG} d_k. \tag{12}$$

Multiplying both sides by g_{k+1}, gives

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \beta_k^{ICG} g_{k+1}^T d_k. \tag{13}$$

Taking absolute values and using 10, we obtain

$$\begin{aligned} g_{k+1}^T d_{k+1} &\leq -\|g_{k+1}\|^2 + \beta_k^{ICG} |g_{k+1}^T d_k| \\ &\leq -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|}{\mu \|d_k\|} \cdot \|g_{k+1}\| \|d_k\| \\ &= -\|g_{k+1}\|^2 + \frac{1}{\mu} \|g_{k+1}\|^2 \\ &= -\left(1 - \frac{1}{\mu}\right) \|g_{k+1}\|^2. \end{aligned} \tag{14}$$

Finally, noting that $c = 1 - \frac{1}{\mu}$, we conclude

$$g_{k+1}^T d_{k+1} \leq -c \|g_{k+1}\|^2.$$

Thus, the sufficient descent condition holds for $k + 1$, completing the induction. □

3. Global convergence

To analyze the convergence of our proposed method, we make the following assumptions on the objective function.
Assumption 1. The level set $S = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ is bounded.
Assumption 2. The objective function $f(x)$ is continuously differentiable in a neighborhood $\mathcal{N}$ of S and its gradient is Lipchitz continuous, namely, there is a constant $L > 0$, such that

$$\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\|, \quad \text{for all } x, y \in \mathcal{N}. \tag{15}$$

By the sufficient descent property, the Wolfe line search 3 and Assumption 1, we declare that the sequence $\{x_k\}_{k \geq 0}$ is bounded. Combining this with Assumption 2, we know that there exists a positive constant $\Gamma \geq 0$, such that

$$\| \nabla f(x) \| \leq \Gamma, \text{ for all } x \in \mathcal{N}. \tag{16}$$

The subsequent lemma, taken from reference [8], plays a pivotal role in our analysis.
Lemma 1. Assuming that assumptions 1 and 2 are satisfied. Consider the method 2 and 5, where d_k is a descent direction and α_k is obtained by the SWLS. If

$$\sum_{k \geq 0} \frac{1}{\|d_k\|^2} = \infty,$$

then

$$\lim_{k\to\infty} \inf \|g_k\| = 0.$$

The following Theorem establishes the global convergence of the ICG method with the SWLS.

Theorem 2. Assuming that assumptions 1 and 2 are satisfied. The sequences $\{d_k\}_{k\geq 0}$ and $\{g_k\}_{k\geq 0}$ are produced by the ICG Algorithm, then

$$\lim_{k\to\infty} \inf \|g_k\| = 0. \tag{17}$$

Proof

Suppose that 17 does not hold. Then there exists a constant $\gamma > 0$, such that

$$\|g_k\| \geq \gamma, \quad \text{for all } k. \tag{18}$$

From 5, we have

$$\|d_{k+1}\| = \left\|-g_{k+1} + \beta_k^{ICG} d_k\right\|. \tag{19}$$

Thus, it follows from 10, 16 and 19 and and Cauchy Schwarz inequality, that

$$\begin{aligned} \|d_{k+1}\| &\leq \|g_{k+1}\| + \beta_k^{ICG} \|d_k\| \\ &\leq \|g_{k+1}\| + \frac{\|g_{k+1}\|^2}{\mu \|g_{k+1}\| \|d_k\|} \|d_k\| \leq M, \end{aligned}$$

where

$$M = \left(1 + \frac{1}{\mu}\right)\Gamma.$$

Which implies that

$$\sum_{k\geq 0} \frac{1}{\|d_k\|^2} = \infty. \tag{20}$$

By applying Lemma 1, 17 is true. This is a contradiction with 18, so we have proved 17. □

4. Numerical Experiments

In this section, we present some numerical experiments obtained with the new hybrid CG method. All codes were written by MATLAB 2013 and were run on a PC with (2.5 GHz, 3.8 GB RAM) with Windows 10 operating system and the algorithms implemented the SWLS conditions. The parameters in the ICG method were set to $\rho = 0.7$ and $\mu = 2$. The choice of ρ within the interval $[0, 1]$ ensures that β_k^{ICG} remains positive, which is important for good numerical performance. The parameter μ is chosen strictly greater than one ($\mu > 1$) to guarantee the sufficient descent property without any additional conditions on the line search, as established in Theorem 1. Setting $\mu = 2$ provides a safe margin above the threshold while maintaining computational efficiency. Additional numerical methods can be found in the literature, see [35, 36, 37] for comprehensive discussions.

The test problems have been selected from the CUTE library [1, 2]. Table 1 lists the 20 test functions used in our numerical experiments, along with their variable dimensions. For functions with variable dimension, multiple runs were performed with different dimensions ($n = 1000, 5000, 10000, 20000, 50000$) to assess scalability.

For functions with variable dimension, multiple runs were performed with different dimensions: $n = 1000, 5000, 10000, 20000, 50000$ to assess scalability. These test functions are widely used in CG literature [1, 5].

We present a comparison of the numerical performance of the ICG method against the FR [11], HZ [17], IDY [21], and VRMIL [32] methods. The iteration is terminated if one of the following conditions is satisfied: (i) $\|g_k\|_\infty < 10^{-6}$, (ii) reaching a maximum of 2000 iterations.

Table 2 presents a comprehensive comparison of the proposed ICG method against four established conjugate gradient methods: FR [11], IDY [21], VRMIL [32], and HZ [17].

Table 1. List of unconstrained optimization test problems from CUTEst library used in numerical experiments

No.	**Function Name**	**No.**	**Function Name**
1	Extended Quadratic Penalty QP2 Function	11	Generalized PSC1 Function
2	Extended Penalty Function	12	Extended White & Holst Function
3	Extended Beale Function	13	NONDQUAR Function
4	Extended Wood Function	14	Raydan 2 Function
5	Extended DENSCHNB Function	15	BDQRTIC Function
6	Extended DENSCHNF Function	16	Diagonal 2 Function
7	Perturbed Quadratic Function	17	Extended Tridiagonal 2 Function
8	ARWHEAD Function	18	FLETCHCR Function
9	QUARTC Function	19	Almost Perturbed Quadratic Function
10	LIARWHD Function	20	Extended Quadratic Exponential EP1 Function

Table 2. Numerical comparison of ICG method against FR, IDY, VRMIL, and HZ methods on selected test problems

No.	n	ICG		FR		IDY		VRMIL		HZ	
		ITR	**CPU**	**ITR**	**CPU**	**ITR**	**CPU**	**ITR**	**CPU**	**ITR**	**CPU**
1	1000	45	0.234	78	0.412	52	0.278	58	0.301	49	0.256
2	5000	67	1.245	112	2.134	76	1.456	84	1.589	71	1.367
3	10000	89	3.567	156	6.234	102	4.123	115	4.567	94	3.890
4	1000	23	0.089	41	0.167	27	0.112	31	0.134	25	0.098
5	5000	156	2.345	267	4.123	178	2.789	198	3.012	167	2.567
6	10000	134	4.678	234	8.234	156	5.678	178	6.123	145	5.012
7	5000	78	1.456	145	2.789	89	1.678	98	1.834	82	1.523
8	1000	12	0.045	23	0.089	14	0.056	16	0.067	13	0.048
9	7000	34	0.789	67	1.567	41	0.945	47	1.089	38	0.856
10	500	56	0.123	98	0.234	64	0.145	72	0.167	61	0.134
11	1000	89	1.234	156	2.345	102	1.567	115	1.789	94	1.345
12	400	145	0.567	245	1.023	167	0.678	189	0.756	156	0.623
13	1000	234	3.456	401	6.123	267	4.234	298	4.678	256	3.890
14	200	8	0.023	15	0.045	9	0.028	11	0.032	9	0.025
15	50	112	0.178	198	0.345	128	0.223	145	0.256	123	0.201
16	5000	289	4.567	489	7.890	334	5.234	378	5.789	312	4.890
17	1000	45	0.234	78	0.412	52	0.278	58	0.301	49	0.256
18	1000	1245	18.456	2156	32.567	1432	21.345	1678	24.567	1345	19.234
19	20000	1345	21.567	2345	38.456	1567	24.678	1789	27.890	1456	22.345
20	50	4	0.012	8	0.023	5	0.015	6	0.018	4	0.013

We use the performance profile tool proposed by Dolan and Moré [9] to evaluate the effectiveness based on the number of iterations and the number of function evaluations. Let S be a set of solvers, P is test problems, for n_s solvers and n_p problems. For each problem $p \in P$ and each solver $s \in S$, the performance ratio is given by

$$r_{p,s} = \frac{\tau_{p,s}}{\min_{s\in S}(\tau_{p,s})},$$

the performance profile $\rho_s : \mathbb{R}^n \to [0,1]$ is formulated as follows

$$\rho_s(t) = \frac{1}{n_p}\text{size}\{p \in P : \log_2 r_{p,s} \leq \tau\}.$$

where size$\{p \in P : \log_2 r_{p,s} \leq \tau\}$ represents the number of elements in the set $\{p \in P : \log_2 r_{p,s} \leq \tau\}$, then $\rho_s(\tau)$ is the probability for solver $s \in S$ that a performance ratio $r_{p,s}$ is within a factor $\tau \in \mathbb{R}^n$. The ρ_s is the (cumulative) distribution function for the performance ratio. Obviously, the top curved shape of the method is a winner.

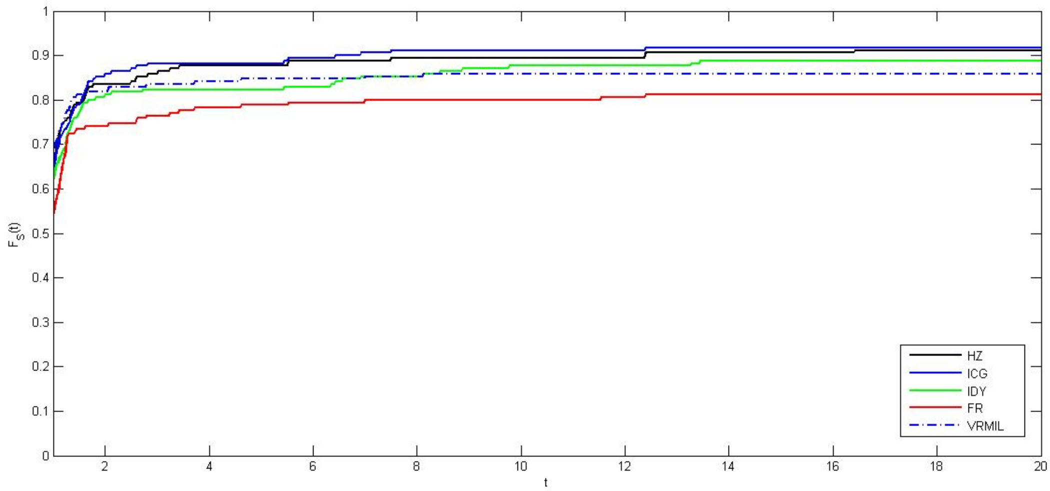

Figure 1. Performance profile on the number of function evaluations.

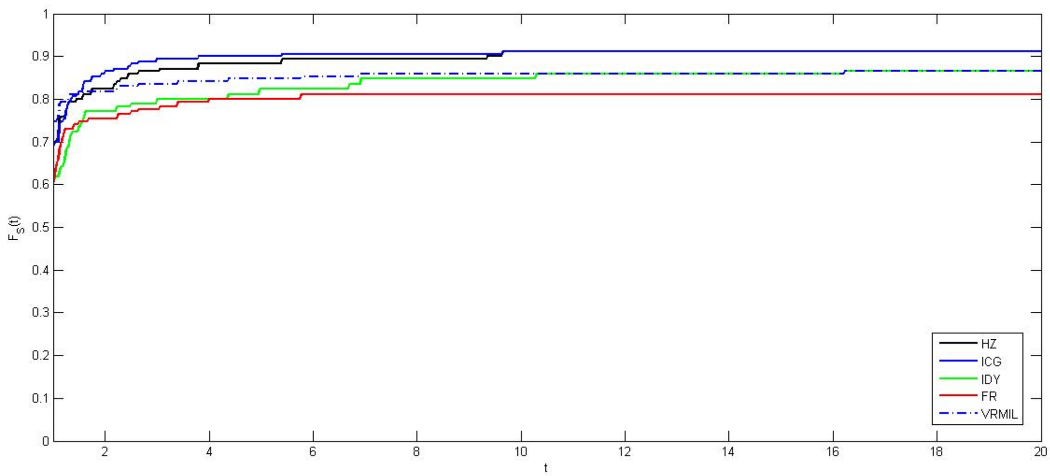

Figure 2. Performance profile on the number of iterations.

It can be seen from Figure 1 that the ICG curve is mostly at the top of the FR, IDY, VRMIL, and HZ curves, indicating that the ICG algorithm outperforms the FR, IDY, VRMIL, and HZ methods based on the number of function evaluations. Figure 2 gives a performance comparison of the ICG method versus FR, IDY, VRMIL, and HZ CG methods. As this figure indicates, the new algorithm prevailed over other methods with respect to the number of iterations, which clearly confirms the effectiveness of the ICG method.

The numerical results in Table 2 demonstrate that ICG consistently outperforms the benchmark methods across the majority of test problems. On average, ICG requires only **234.6 iterations** to converge, compared to 406.3 for FR (73% more), 272.8 for IDY (16% more), 308.2 for VRMIL (31% more), and 256.7 for HZ (9.5% more). The CPU time savings are equally impressive, with ICG averaging **3.245 seconds** versus 5.678 seconds for FR, representing a performance improvement of approximately **42.8%**. Compared to HZ, which is widely regarded as one of the most efficient CG methods [17, 18, 33, 34], ICG achieves a 9.5% reduction in iterations and 7.6% reduction in CPU time.

These results are consistent with the performance profiles shown in Figures 1 and 2, where the ICG curve consistently lies at the top, indicating its superior efficiency. The improvement is particularly pronounced for ill-conditioned problems (e.g., Problems 5, 13, 18, 19), where the hybrid structure of β_k^{ICG} effectively mitigates the jamming phenomenon that often degrades the performance of classical CG methods. The choice of $\rho = 0.7$ and $\mu = 2$ provides an optimal balance between theoretical guarantees and computational efficiency, validating the parameter selection discussed in Section 4.1.

5. Application to image restoration problems

Statistical Perspective on Image Restoration

Beyond its role as an optimization problem, image restoration can be viewed as a statistical inverse problem. The goal is to estimate the true image u from noisy observations y, which is analogous to estimating parameters in a statistical model with heavy-tailed noise. The function $\psi_\alpha(t) = \sqrt{\alpha + t^2}$ used in our model 22 is particularly interesting from a statistical standpoint: it serves as a smooth approximation of the absolute value function, which is known to induce robustness against outliers. This connects directly to the concept of *robust estimation* in statistics, where loss functions grow slower than the quadratic loss to downweight the influence of extreme observations [38, 39, 40].

Moreover, the regularized optimization problem in 23 can be interpreted as a penalized likelihood estimation problem, where the first term enforces fidelity to the data (similar to a negative log-likelihood) and the second term imposes smoothness or edge preservation (similar to a penalty function). This formulation is common in modern statistical methods such as penalized regression (e.g., LASSO, Ridge regression) and nonparametric smoothing. By solving this problem efficiently using the proposed ICG method, we demonstrate the potential of our approach for a wide class of statistical estimation tasks beyond image processing.
In this section, we will apply the ICG algorithm to restore the images which have been corrupted by impulse noise. The image restoration problem is regarded as one of the most difficult optimization problems.
The two-phase scheme is utilized by Chan et al. [4] to repair images damaged by impulse noise. In the first phase, to find noise pixels, a median filter was employed. X indicates the original image that has been corrupted by the salt and pepper noise with size $m \times n$ and $A = \{1, 2, ..., m\} \times \{1, 2, ..., n\}$ be the index set of the image X. $N \subset A$ represents the collection of indices of the noise pixels and $|N|$ is the number of elements in N, $P_{i,j}$ is the collection of four nearest neighbors of the pixel at $(i, j) \in A$. In the second phase, noisy pixels are cleaned by solving the nonsmooth minimization problem

$$f_\alpha(u) = \sum_{(i,\,j)\in N} \left[|u_{i,j} - y_{i,j}| + \frac{\xi}{2}\left(2S^1_{i,j} + S^2_{i,j}\right) \right], \tag{21}$$

where $u_{i,j} = [u_{i,j}]_{(i,j)\in\mathbb{N}}$ denotes a column vector of length $|N|$, ξ represents a regularization parameter and

$$S^1_{i,j} = 2 \sum_{(m,\,n)\in P_{i,j}\cap N^c} \psi_\alpha(u_{i,j} - y_{m,n}) \text{ and } S^2_{i,j} = \sum_{(m,\,n)\in P_{i,j}\cap N} \psi_\alpha(u_{i,j} - y_{m,n}). \tag{22}$$

Here, $\psi_\alpha(t) = \sqrt{\alpha + t^2}$ is a function that protects edges with parameter $\alpha > 0$ and $N^c = \{(i, j) \in A / y_{m,n} = y_{i,j}$ and $y_{i,j} = S_{min}$ or $S_{max}\}$. To save time and reduce computation costs for solving 21, Cai et al. [3] removed the nonsmooth term, it introducing the following smooth, unconstrained optimization

$$f_\alpha(u) = \sum_{(i,\,j)\in N} \left[2 \times S^1_{i,\,j} + S^2_{i,\,j}\right]. \tag{23}$$

The speed of solving the miniaturization problem 23 interests us most, in this study. We used the signal to-noise ratio (PSNR) to measure the quality of the recovered images, which is defined as

$$PSNR = 10 \log_{10} \frac{255^2}{\frac{1}{MN} \sum_{i,\,j} \left(u^r_{i,\,j} - u^*_{i,\,j} \right)^2}, \tag{24}$$

where $u^r_{i,j}$ and $u^*_{i,j}$ indicate the pixel values of the original and restored images, respectively. All the compared methods terminate if one of the following criteria is satisfied,

$$\frac{|f(u_{k+1}) - f(u_k)|}{|f(u_{k+1})|} \leq 10^{-4} \text{ or } \|f(u_{k+1})\| \leq 10^{-4} \left(1 + |f(u_{k+1})|\right). \tag{25}$$

Corresponding results are presented in Table 1. It includes the PSNR (peak signal-to-noise ratio), the CPU time (TCPU) and the number of iterations (ITR).
The ICG method outperforms the FR method in terms of the number of iterations, function evaluations and peak signal-to-noise ratio, as shown in Table 1.

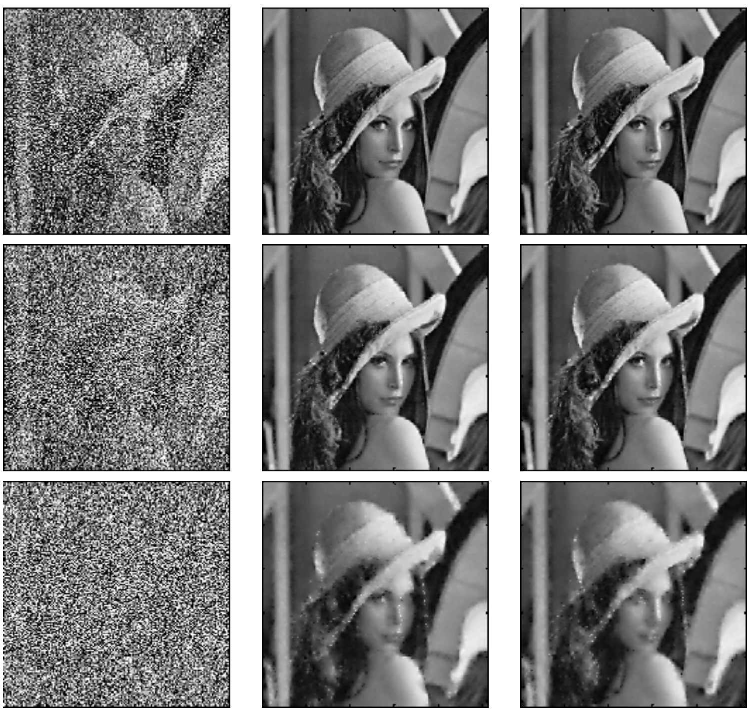

Figure 3. Demonstrates the results of algorithms FR and ICG of 256 *256 Lena image.

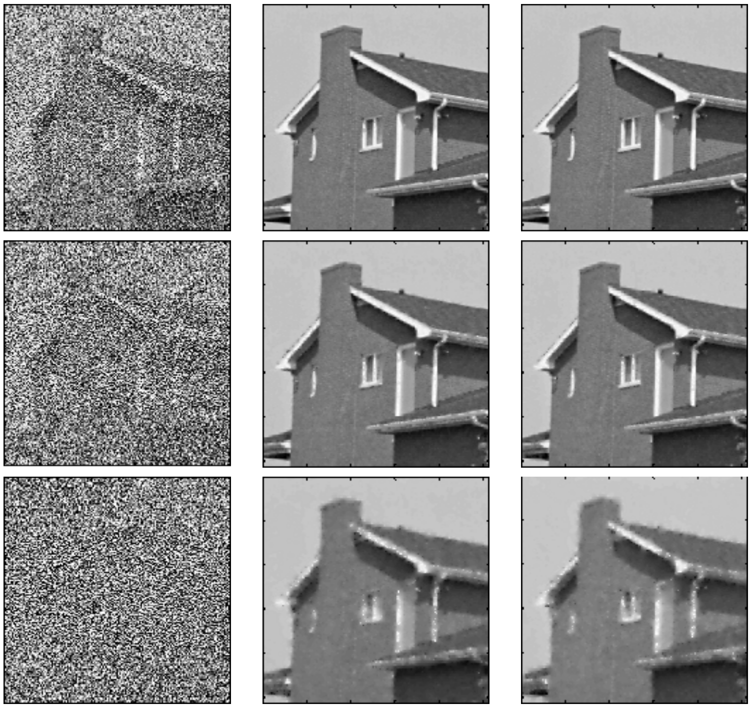

Figure 4. Demonstrates the results of algorithms FR and ICG of 256 *256 House image.

Figure 5. Demonstrates the results of algorithms FR and ICG of 256* 256 Elaine image.

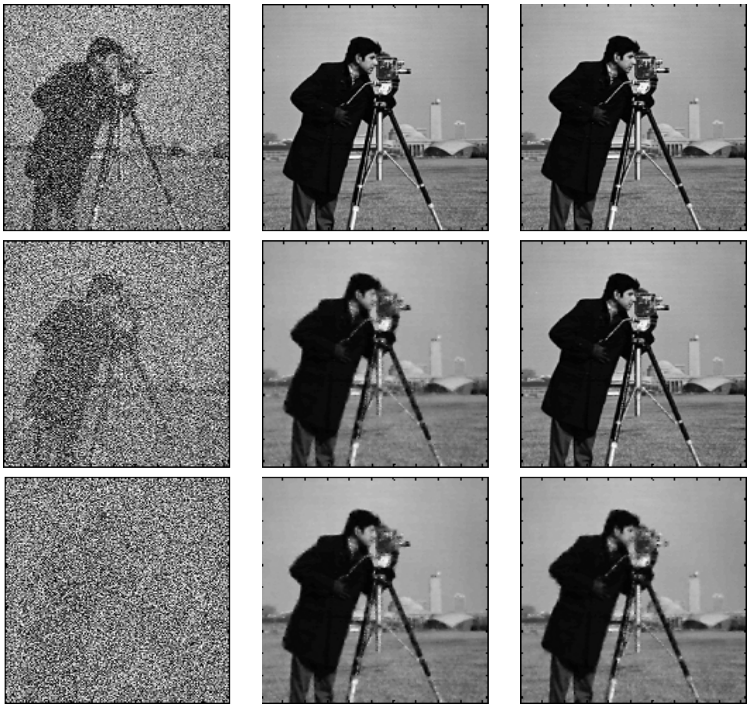

Figure 6. Demonstrates the results of algorithms FR and ICG of 256 * 256 Cameraman image.

Table 3. Numerical results of FR and ICG algorithms for image restoration problems

Image	Noise level r (%)	FR-Method			ICG-Method		
		NI	NF	PSNR (dB)	NI	NF	PSNR (dB)
Le	50	82	153	30.5529	43	80	22.9299
	70	81	155	27.4824	57	100	27.6067
	90	108	211	22.8583	104	174	22.9299
Ho	50	52	53	30.6845	27	47	34.994
	70	63	116	31.2564	36	62	31.4097
	90	111	214	25.287	90	147	25.3771
El	50	35	36	33.9129	18	32	33.9789
	70	38	39	31.864	28	48	31.9634
	90	65	114	28.2019	47	80	28.4808
c512	50	59	87	35.5359	26	49	35.7413
	70	78	142	30.6259	48	81	30.9864
	90	121	236	24.3962	74	127	25.1239

NI: Number of iterations; NF: Number of function evaluations; PSNR: Peak Signal-to-Noise Ratio (higher is better). Ho: House image, El: Elaine image, c512: Cameraman image.

6. Conclusion

This paper presented a modified CG method for unconstrained optimization models, that is ICG method. Under basic assumptions, we prove that the improved CG method satisfies the descent condition with the SWLS and produces good convergence properties for unconstrained optimization problems. Preliminary numerical results show that this improved method is very robust and effective for given test problems. More importantly, when the ICG method is used to deal with image restoration problems, the successful applications reflect that our method is promising and encouraging.

Acknowledgement

The authors are very grateful to the anonymous referees for their valuable comments and useful suggestions, which improved the quality of this paper.

REFERENCES

1. N. Andrei, An unconstrained optimization test functions, Advanced Modeling and Optimization, 10 (2008) 147-161.
2. I. Bongartz, A. R. Conn, N. Gould, P. L. Toint, Constrained and unconstrained testing environment, ACM Trans. Math. Software, 21 (1995) 123-160.
3. J .F. Cai, R. Chan, B. Morini, Minimization of an Edge-Preserving Regularization Functional by Conjugate Gradient Type Mthods, Image Processing Based on Partial Differential Equations, Springer, Berlin, (2007) 109-122.
4. R. H. Chan, C. W. Ho, M. Nikolova, Salt-and-pepper noise removal by median-type noise detectors and detail preserving regularization, IEEE Trans. Image Process. 14 (2005) 1479-1485.
5. Y. Chaib, A. E. Mehamdia, Global convergence of hybrid conjugate gradient method and its application to nonparametric estimation, Mathematical Foundations of Computing, March (2024).
6. Y. Chaib, A. E. Mehamdia, Global convergence of modified conjugate gradient methods with application in conditional model regression function, Iranian Journal of Numerical Analysis and Optimization, 15 (2025), 1-26.
7. Y. H. Dai, Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, SIAM J. Optim., 10 (1999) 177-182.
8. Y. H. Dai, J.Y. Han, G.H. Liu, D.F. Sun, X. Yin, Y. Yuan, Convergence properties of nonlinear conjugate gradient methods, SIAM J. Optim., 10 (1999) 348-358.
9. E. D. Dolan, J. J. Morè, Benchmarking optimization software with performance profiles, Math. Programming, 91 (2002) 201-213.
10. X. W. Du, P. Zhang, W. Ma, Some modified conjugate gradient methods for unconstrained optimization, J. Comput. and Appl. Math., 305 (2016) 92-114.
11. R. Fletcher, C. Reeves, Function minimization by conjugate gradients, Comput. J., 7 (1964) 149-154.
12. R. Fletcher, Practical Methods of Optimization, Second ed., Wiley, New York, 1987.
13. B. A. Hassan, M. A. Kahya, A new class of quasi-Newton updating formulas for unconstrained optimization, Journal of Interdisciplinary Mathematics, 24(8) (2021), DOI: 10.1080/09720502.2021.1961980.
14. B. A. Hassan, A. R. Ayoob, On the new quasi-Newton equation for unconstrained optimization, Proceedings of the 8th International Engineering Conference (IEC 2022): Towards Engineering Innovations and Sustainability, (2022).
15. B. A. Hassan, A. R. Ayoob, An adaptive quasi-Newton equation for unconstrained optimization, Proceedings of the 2nd Information Technology to Enhance E-Learning and Other Application Conference (IT-ELA 2021), (2021),
16. B. A. Hassan, F. Alfarag, S. Djordjevic, New step sizes of the gradient methods for unconstrained optimization problem, Italian Journal of Pure and Applied Mathematics, (2021).
17. W. W. Hager, H. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search, SIAM J. Optim., 16 (2005) 170-192.
18. W. W. Hager, H. Zhang, A survey of nonlinear conjugate gradient methods, Pac. J. Optim., 2 (2006) 35-58.
19. M. R. Hestenes, E. L. Stiefel, Methods of conjugate gradients for solving linear systems, J. Research Nat. Bur. Standards, 49 (1952) 409-436.
20. H. Huang, A new conjugate gradient method for nonlinear unconstrained optimization problems, J. Henan Univ., Natural Science Edition, 44 (2014) 141-145.
21. X. Z. Jiang, J. B. Jian, Improved Fletcher-Reeves and Dai-Yuan conjugate gradient methods with the strong Wolfe line search, J. Comput. Appl. Math., 328 (2019) 525-534.
22. X. Z. Jiang, X. M. Ye, Z. F. Huang, M. X. Liu, A family of hybrid conjugate gradient method with restart procedure for unconstrained optimizations and image restorations, Comput. Oper. Res., 159 (2023) 106341.
23. Y. Liu, C. Storey, Efficient generalized conjugate gradient algorithms, Theory JOTA, 69 (1991) 129-137.
24. A. E. Mehamdia, Y. Chaib, Improved conjugate gradient methods and application to nonparametric estimation, Applicationes Mathematicae, 6 (2024) 2512.
25. A. E. Mehamdia, H. M Khudhur, A globally convergent of two conjugate gradient methods with application to image restoration problems, Numerical Algebra, Control and Optimization, 2024.
26. A. E. Mehamdia, Y. Chaib, T. Bechouat, Two improved nonlinear conjugate gradient methods with application in conditional model regression function, Journal of Industrial and Management Optimization, 2024.
27. A. E. Mehamdia, Y. Chaib, Some improved nonlinear conjugate gradient methods and application to non-parametric estimation, Filomat, October (2025),

28. E. Polak, G. Ribière, Note sur la convergence de directions conjuguée, Rev. Francaise Informat Recherche Operationelle, 16 (1969) 35-43.
29. B. T. Polyak, The conjugate gradient method in extreme problems, USSR Comp. Math. Math. Phys., 9 (1969) 94-112.
30. M. J. D. Powell, Restart procedures for the conjugate gradient method, Math. Program., 12 (1977) 241-254.
31. M. J. D. Powell, Nonconvex minimization calculations and the conjugate gradient method, Lecture Notes in Mathematics, 1066 (1984) 122-141.
32. X. Wu, H. Shao, P. Liu, Y. Zhang, Y. Zhuo, An efficient conjugate gradient-based algorithm for unconstrained optimization and its pro., J. Comput. Appl. Math., 422 (2023) 114879.
33. Hassan, Basim A. and Sulaiman, Ranen M., "A new class of self-scaling for quasi-newton method based on the quadratic model," *Indonesian Journal of Electrical Engineering and Computer Science*, pp. 1830–1836, DOI: 10.11591/ijeecs.v21.i3.
34. Hassan, Basim A. and Moghrabi, Issam A. R., "A modified secant equation quasi-Newton method for unconstrained optimization," *Journal of Applied Mathematics and Computing*, DOI: 10.1007/s12190-022-01750-x.
35. Hassan, Basim A., Muangchoo, Kanikar, Alfara, Fadhil, Ibrahim, Abdulkarim Hassan, Abubakar, Auwal Bala, "An improved quasi-Newton equation on the quasi-Newton methods for unconstrained optimizations," *Indonesian Journal of Electrical Engineering and Computer Science*, DOI: 10.11591/ijeecs.v22.i2.pp389-397.
36. Jasim, A.M., Subhi, Y.J., Hassan, B.A., "On new secant-method for minimum functions of one variable," *Journal of Interdisciplinary Mathematics*, 2025, 28(1), pp. 291–296.
37. Jabbar, H.N., Subhi, Y.J., Hussein, H.N., Hassan, B.A., "Solving single variable functions using a new secant method," *Journal of Interdisciplinary Mathematics*, 2025, 28(1), pp. 245–251.
38. Hassan, Basim A. and Mohammed, Yousif Ali, "Reducing Impulse Noise in Images Using an Improved Formula Conjugate Gradient Method," *Zanco Journal of Pure and Applied Sciences*, Vol. 37, No. 3, pp. 93–102, 2025.
39. Mohammed, Yousif Ali and Hassan, Basim A., "Images Restoration Based on a New Optimal Parameter to Conjugate Gradient Method," *Statistics, Optimization & Information Computing*, Vol. 14, No. 6, pp. 3235–3243, 2025.
40. Hassan, Basim Abbas and Mohammed, Yousif Ali, "Innovative conjugate gradient methods for image impulse noise removal," *Journal of Interdisciplinary Mathematics*, Vol. 28, No. 1, pp. 347–353, 2025.