Improvement of conjugate gradient methods for removing impulse noise images

Basim A. Hassan, Ali Ahmed A. Abdullah
Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul, Mosul, Iraq

ABSTRACT
Optimization problems occur in most disciplines like engineering, physics, mathematics, economics, administration, commerce, social sciences, and even politics. The conjugate coefficient is the cornerstone of conjugate gradient algorithms with the desired conjugate property. In this study, we discovered fresh second order information for the Hessian from the target function, which might lead to a new search direction. Based on a unique search direction, we proposed the update formula and nonlinear conjugate gradient technique. Under Wolfe line search and moderate objective function assumptions, the strategy has acceptable descent property and is always globally convergent. According to numerical results, the technique is successful and competitive in recovering the original picture from an image corrupted by impulsive noise.

This is an open access article under the CC BY-SA license.

1. INTRODUCTION
This paper presents an iterative technique for solving optimization problems using an edge-preserving regularization (EPR) objective function. In general, two equations explain impulse noise. The first equation uses an adaptive median filter (AMF) to detect pixel values that may be contaminated by impulse noise. Let \( X \) be the true picture and \( A = \{1,2,3,\ldots,M\} \times \{1,2,3,\ldots,N\} \) be its index set. Let \( N \subset A \) be the indices of the first phase noise pixels. Let \( P_{i,j} \) be the set of four closest neighbors of the pixel at \((i,j) \in A\), \( y_{i,j} \) be the observed pixel value at \((i,j)\), and \( u_{i,j} = [u_{i,j}]_{(i,j) \in N} \) be a lexicographically ordered column vector of length \( c \). \( N \) has \( c \) components. This is done by reducing the following functional:

\[
f_a(u) = \sum_{(i,j) \in N} \left[ u_{i,j} - y_{i,j} \right]^2 + \frac{\beta}{2} \left( 2 \times S_x^2 + S_y^2 \right), \tag{1}\]

where \( \beta \) is the regularization parameter, and \( S_x^2 = 2 \sum_{(m,n) \in P_{i,j} \cap N} \phi_a(u_{i,j} - y_{m,n}) \). The \( \phi_a = \sqrt{\alpha + x^2}, \alpha > 0 \) is an example of an edge-preserving potential. In reality, the non-smooth data-fitting term is unnecessary in the second phase, when only noisy pixels are restored, Yu et al. [2]. Thus, several optimization techniques may be extended to minimize the smooth edge-preserving regularization (EPR) functional:

\[
f_a(u) = \sum_{(i,j) \in N} \left[ 2 \times S_x^2 + S_y^2 \right] \tag{2}\]
Image restoration uses conjugate gradients, image restoration problems are expressed as:

$$\min f(u), \ u \in \mathbb{R}^n$$

(3)

where \( f: \mathbb{R}^n \rightarrow \mathbb{R} \) is smooth function. Gilbert and Nocedal [3]. A general conjugate gradient algorithm generates a sequence of iterates by the rule:

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

(4)

where the step size \( \alpha_k \) is positive and the directions \( d_k \) are computed using the updating formula:

$$d_{k+1} = -g_{k+1} + \beta_k d_k$$

(5)

which \( \beta_k \) is a scalar known as the conjugate gradient parameter. Different choices of \( \beta_k \) lead to various conjugate gradient methods, Andrei [4]. In this paper, we focus our attention on well-known method such as Fletcher and Reeves [5] given by:

$$\beta_{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$$

(6)

it is well known that the FR possess nice convergence properties. In the past decades, a variety of conjugate gradient methods are developed. There are some well known conjugate gradient methods, such as [6]-[10]. Other nonlinear conjugate gradient methods and their global convergence can be found in [11]. We can get the step-size \( \alpha_k \) using the exact line research:

$$\alpha_k = -\frac{g_k^T d_k}{d_k^T Q d_k}$$

(7)

see, [12]. Usually, in (2), the steplength \( \alpha_k \) is computed using the Wolfe line search conditions:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k$$

(8)

$$d_k^T g(x_k + \alpha_k d_k) \geq \sigma d_k^T g_k$$

(9)

where \( 0 < \delta < \sigma < 1 \), Wolfe [13], [14]. During the last decade, much effort has been devoted to develop new conjugate gradient methods which not only possess strong convergence properties but are also computationally superior to the classical methods. The most typical feature of conjugate gradient methods is conjugacy, namely, the search directions generated by (3) should possess the following conjugacy condition:

$$d_k^T Q d_k = 0$$

(10)

researchers focus their attention on conjugacy condition. One of the remarkable results is obtained in conjugate gradient methods. For a good reference for studies describing the latest CG coefficients with important result and various modifications from \( \beta_k \), Xue et al. [1]. For further references on the optimization methods, please refer to [15]-[18].

Using the conjugacy condition, we are now ready to give some new formulas of nonlinear conjugate gradient methods. Global convergence of these formulas have been established. Finally, some of the numerical results have been reported, which show the effectiveness of the new formula.

2. NEW CONJUGATE GRADIENT COEFFICIENT

In our paper, By adopting some idea that in the literature [19]. We first consider the second order Taylor series of \( f(x) \) as (11).

$$f(u) = f(u_{k+1}) + g_{k+1}^T(u - u_{k+1}) + \frac{1}{2}(u - u_{k+1})^T Q(u_k)(u - u_{k+1})$$

(11)

Finding the derivative yields:

$$g_{k+1} = g_k + Q(u_k)s_k$$

(12)

The parameter, \( \beta_k \), in the linear conjugate gradient method is given by:
\[
\beta_k = \frac{g_{k+1}^T Q s_k}{d_k^T Q s_k}
\]

(13)

where \( Q \) is Hessian matrix and where \( \beta_k \) is satisfies the conjugacy condition, Hassan and Sulaiman [19].

Now, we shall consider another expression of the denominator \( d_k^T Q s_k \). Using (11) and (7) in (11), we obtain:

\[
s_k^T Q (u_k) s_k = \frac{2}{3} s_k^T y_k + \frac{2}{3} (f_k - f_{k+1})
\]

(14)

which implies that:

\[
d_k^T Q (u_k) s_k = \frac{2}{3} d_k^T y_k + \frac{2}{3} (f_k - f_{k+1}) / \alpha_k
\]

(15)

from this we advance formula,

\[
\beta_k = \frac{g_{k+1}^T y_k}{2/3 d_k^T y_k + 2/3 (f_k - f_{k+1}) / \alpha_k}
\]

(16)

since \( f \) is quadratic model by using exact line search, then (16) reduces to:

\[
\beta_k = \frac{\|g_{k+1}\|^2}{2/3 d_k^T y_k + 2/3 (f_k - f_{k+1}) / \alpha_k}
\]

(17)

and

\[
\beta_k = \frac{\|g_{k+1}\|^2}{-2/3 d_k^T g_k + 2/3 (f_k - f_{k+1}) / \alpha_k}
\]

(18)

and

\[
\beta_k = \frac{\|g_{k+1}\|^2}{2/3 d_k^T g_k + 2/3 (f_k - f_{k+1}) / \alpha_k}
\]

(19)

our formula, so-called BA1, BA2 and BA3. They show that the new method globally convergent for general functions under some proper conditions. Below we present the BA algorithms:

- Stage 1: Set \( k = 1 \), then \( d_1 = -g_1 \) and select \( u_1 \).
- Stage 2: Test for Continuation of Iterations.
- Stage 3: Calculate the step length by using Wolfe line search (5) and (6).
- Stage 4: Compute \( d_{k+1} \) using (5).
- Stage 5: Compute \( u_{k+1} \) using (4).
- Stage 6: Set \( k = k+1 \), and go to Step 1.

**Theorem (2.2)**

The search direction defined by (3) with (17)-(19) satisfy:

\[
d_{k+1}^T g_{k+1} < 0 \quad \text{and} \quad d_{k+1}^T g_{k+1} = \beta_k d_k^T g_k
\]

(20)

**Proof:**

Clearly by (3), if \( k = 0 \) then \( g_0^T d_0 = -\|g_0\|^2 \) holds, let \( d_{k}^T g_k < 0 \) for all \( k \). Multiplying both sides of (3) with \( g_{k+1}^T \), we get:

\[
d_{k+1}^T g_{k+1} = -g_{k+1}^T g_{k+1} + \beta_k d_k^T g_{k+1}
\]

\[
= -\beta_k (2/3 d_k^T y_k + 2/3 (f_k - f_{k+1}) / \alpha_k) + \beta_k d_k^T g_{k+1}
\]

(21)

From (21), we obtain:

\[
d_{k+1}^T g_{k+1} = \beta_k [d_k^T g_{k+1} - (2/3 d_k^T y_k + 2/3 (f_k - f_{k+1}) / \alpha_k)]
\]

(22)

also, by using (17) and (21) we get:

\[
d_{k+1}^T g_{k+1} = \beta_k d_k^T g_k
\]

(23)

by the \( d_k^T g_k < 0 \) and (23), we can write:

\[
d_{k+1}^T g_{k+1} < 0
\]

(24)
hence $d_{k+1}^T g_{k+1} < 0$ and $d_{k+1}^T g_{k+1} = \beta_k d_k^T g_k$ holds, which completes the proof. Similarly, other methods can be proved.

3. CONVERGENCE ANALYSIS

In the above section, we have proved that the new methods has descent property that is independent of the line search and the function convexity. In this section, we will make use of this property to establish the global convergence for the new methods using a variety of line searches. Suppose that the objective function satisfies the following assumption.

Assumptions:
I) The level set $\mathcal{D} = \{ u \in \mathbb{R}^n : f(u) \leq f(u_t) \}$, is bounded.
II) Gradient is Lipschitz continuous, that is, for $L > 0$:

$$\|g(P) - g(O)\| \leq L \|P - O\|, \forall P, O \in \mathcal{A} \hspace{2cm} (25)$$

for more details see [20]-[22].

By having these assumptions, Zoutendijk [23] has proven the following Lemma.

Lemma (3.1):
Suppose Assumption are satisfied. In any iteration method if $\alpha_k$ is satisfied Wolfe line search, then:

$$\sum_{k=1}^{\infty} \left( \frac{g_k^T d_k}{\|d_k\|^2} \right) < \infty \hspace{2cm} (26)$$

Proof: For see Sulaiman and Hassan [24] and Zhang and Xu [25].

Theorem (3.2):
Suppose that Assumption and Lemma 1 holds. Then,

$$\liminf_{k \to \infty} \|g_k\| = 0 \hspace{2cm} (27)$$

Proof:
By induction, let that (27) is not true. Suppose that there exists $c_1 > 0$ such that $\|g_k\| \geq c_1$ for all $k \in n$. Squaring both sides of (8), we get:

$$\|d_{k+1}\|^2 + \|g_{k+1}\|^2 + 2d_{k+1}^T g_{k+1} = (\beta_k)^2 \|d_k\|^2 \hspace{2cm} (28)$$

using (23), yields:

$$\|d_{k+1}\|^2 = \frac{(d_{k+1}^T g_{k+1})^2}{(d_k^T g_k)^2} \|d_k\|^2 - 2d_{k+1}^T g_{k+1} - \|g_{k+1}\|^2 \hspace{2cm} (29)$$

divide by (29) by $(d_{k+1}^T g_{k+1})^2$, we get:

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} = \frac{\|d_k\|^2}{(d_k^T g_k)^2} - \frac{\|g_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} - \frac{2}{(d_{k+1}^T g_{k+1})^2} \hspace{2cm} (30)$$

using (3), (17), and (30), we obtain:

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} \leq \frac{\|d_k\|^2}{(d_k^T g_k)^2} - \left( \frac{\|g_{k+1}\|}{(d_{k+1}^T g_{k+1})} + \frac{1}{\|g_{k+1}\|^2} \right) + \frac{1}{\|g_{k+1}\|^2} \leq \frac{\|d_k\|^2}{(d_k^T g_k)^2} + \frac{1}{\|g_{k+1}\|^2} \hspace{2cm} (31)$$

hence,

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} \leq \sum_{i=1}^{k+1} \frac{1}{\|d_i\|^2} \hspace{2cm} (32)$$

therefore,

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} \leq \frac{k+1}{c_1^2} \hspace{2cm} (33)$$
from (33), we obtain:

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{||d_k||^2} = \infty$$  \hspace{1cm} (34)

based on Lemma 1, we get \( \liminf_{k \to \infty} \| g_k \| = 0 \) holds.

4. NUMERICAL RESULTS

The numerical findings in this section indicate the effectiveness of New in the reduction of salt-and-pepper impulse noise. New and FR methods are tested in our trials. MATLAB r2017a is used to write and execute all of the programs. The following are the stopping criteria for both methods (35).

$$\frac{|f(u_k) - f(u_{k-1})|}{|f(u_k)|} \leq 10^{-4} \text{ and } \|f(u_k)\| \leq 10^{-4}(1 + |f(u_k)|)$$  \hspace{1cm} (35)

Lena, House, Cameraman, and Elaine make up the test photos. PSNR (peak signal-to-noise ratio) is a quantitative metric that may be used to evaluate the quality of the restoration process:

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{\frac{1}{MN} \sum_{i,j} (u_{i,j}^r - u_{i,j}^*)^2}$$  \hspace{1cm} (36)

in this case, \( u_{i,j}^r \) and \( u_{i,j}^* \) represent the restored and original image's pixel values, respectively.

The number of iterations (NI) and the number of function evaluations (NF) needed for the whole denoising process, as well as the PSNR of the recovered picture, are reported in this paper. We can observe from Table 1 that the New technique is much quicker than the FR method for the vast majority of the test photographs. Furthermore, we see that the PSNR values obtained by the New and FR methods are fairly close. Conclusion. Table 1 shows that the suggested methods outperform the FR approach in terms of number of iterations, function evaluations, and peak signal to noise ratio when it comes to eliminating impulse noise from photos.

Figures 1-4 show the obtained results by denoised images. Figures (a1), (a2), (a3) and (a4) are the images corrupted with 70% salt-and-pepper noise; Figures (b1), (b2), (b3) and (b4) are results of FR method; Figures (c1), (c2), (c3) and (c4) are results of the BA1 method; Figures (d1), (d2), (d3) and (d4) are results of the BA2 method; and Figures (e1), (e2), (e3) and (e4) are results of the BA3 method.

Figure 1. Demonstrates the results of algorithms: (a1) denoised images with 70% salt-and-pepper noise, (b1) FR method, (c1) BA1 method, (d1) BA2 method and (e1) BA3 method of 256 * 256 Lena images

Figure 2. Demonstrates the results of algorithms: (a2) denoised images with 70% salt-and-pepper noise, (b2) FR method, (c2) BA1 method, (d2) BA2 method and (e2) BA3 method of 256 * 256 House image
demonstrate that the novel technique has a low computing cost and is effective at solving signal processing problems. The parameters of the new technique are generated from a quadratic model, which is described in this paper. The numerical findings and Table 1 support these conclusions.

Table 1. Numerical results of FR, BA1, BA2 and BA3 algorithms

| Image  | Noise level r (%) | NI   | NF   | PSNR (dB) | NI   | NF   | PSNR (dB) | NI   | NF   | PSNR (dB) | NI   | NF   | PSNR (dB) |
|--------|-------------------|------|------|-----------|------|------|-----------|------|------|-----------|------|------|-----------|
| Le     | 50                | 82   | 153  | 33.5529  | 42   | 89   | 30.7873  | 42   | 91   | 30.4787  | 44   | 93   | 30.6782  |
| Ho     | 50                | 52   | 53   | 30.6845  | 28   | 57   | 34.8651  | 32   | 67   | 34.7529  | 33   | 68   | 34.7004  |
| El     | 50                | 35   | 36   | 33.9129  | 23   | 43   | 33.895   | 21   | 40   | 33.875   | 25   | 47   | 33.8553  |
| c512   | 50                | 59   | 87   | 35.5359  | 26   | 56   | 35.3508  | 23   | 49   | 35.7124  | 30   | 61   | 35.4487  |

5. CONCLUSIONS

In this research, we offer novel CG approaches for minimizing the smooth regularization functional for impulse noise reduction, which we call smooth regularization functions. The parameters of the new technique are generated from a quadratic model, which is described in this paper. The numerical findings demonstrate that the novel technique has a low computing cost and is effective at solving signal processing difficulties.

REFERENCES

[1] W. Xue, J. Ren, X. Zheng, Z. Liu, and Y. Liang, “A new DTV conjugate gradient method and applications to image denoising,” IEEE Transactions on Information and Systems, vol. E101D, no. 12, pp. 2984–2990, Dec. 2010, doi: 10.1587/transinf.2010EDP7210.

[2] G. Yu, J. Huang, and Y. Zhou, “A descent spectral conjugate gradient method for impulse noise removal,” Applied Mathematics Letters, vol. 33, no. 5, pp. 555–560, May 2010, doi: 10.1016/j.aml.2010.01.010.

[3] 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, Feb. 1992, doi: 10.1137/0802005.

[4] N. Andrei, “Open problems in nonlinear conjugate gradient algorithms for unconstrained optimization,” Bulletin of the Malaysian Mathematical Sciences Society, vol. 34, no. 2, pp. 319–330, 2011.

[5] R. Fletcher, “Function minimization by conjugate gradients,” The Computer Journal, vol. 7, no. 2, pp. 149–154, Feb. 1964, doi: 10.1093/comjnl/7.2.149.

[6] Y. H. Dai and Y. Yuan, “A nonlinear conjugate gradient method with a strong global convergence property,” SIAM Journal on Optimization, vol. 10, no. 1, pp. 177–182, Jan. 1999, doi: 10.1137/S1052623497318992.

[7] 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, p. 409, Dec. 1952, doi: 10.6028/jres.049.044.
BIOGRAPHIES OF AUTHORS

Basim A. Hassan is currently a Professor in Department of Mathematics, College of Computer Science and Mathematics, University of Mosul. He obtained his M.Sc and Ph.D degrees in Mathematics from the University of Mosul, in 2000 and 2010, respectively with specialization in optimization. To date, he has published more than 80 research paper in various international journals and conferences. He currently works on iterative methods. His research interest in applied mathematics, with a field of concentration of optimization include conjugate gradient, steepest descent methods, Broyden’s family and quasi-Newton methods with application in signal recovery and image restoration. He can be contacted at email: basimah@uomosul.edu.iq.

Ali Ahmed Abdullah received the bachelor's (B.Sc) and master's (M.Sc) degrees in mathematics from University of Mosul-college of Computer Science And Mathematics-Department Of Mathematics in 2008 and 2020 respectively. He currently works as a Teacher in the ministry education of Iraq-Nineveh directorate of education-Abo-Tamam high school. His Master’s Degree was in pure mathematics and he has several researches in the field of projective geometry. His most recent research and interest are the iterative methods for unconstrained optimization with Application. He can be contacted at email: ali2005ah@gmail.com.

[8] 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, Apr. 1991, doi: 10.1007/BF00940464.
[9] 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, May 1969, doi: 10.1051/m2an/196903R100351.
[10] B. T. Polyak, “The conjugate gradient method in extremal problems,” USSR Computational Mathematics and Mathematical Physics, vol. 9, no. 4, pp. 94–112, Jan. 1969, doi: 10.1016/0041-5553(69)90035-4.
[11] W. W. Hager and H. Zhang, “A survey of nonlinear conjugate gradient methods,” Pacific journal of Optimization, vol. 2, no. 1, pp. 35–58, 2006. [Online]. Available: http://www.ybook.co.jp/online2/opj/pdf/vol2/p35.html.
[12] J. Nocedal and S. J. Wright, Numerical optimization, 2nd ed. Springer New York, 2006.
[13] P. Wolfe, “Convergence conditions for ascent methods,” SIAM Review, vol. 11, no. 2, pp. 226–235, Apr. 1969, doi: 10.1137/1011036.
[14] P. Wolfe, “Convergence conditions for ascent methods. II: some corrections,” SIAM Review, vol. 13, no. 2, pp. 185–188, Apr. 1971, doi: 10.1137/1013035.
[15] I. M. Sulaiman, N. A. Bakar, M. Mamat, B. A. Hassan, M. Malik, and A. M. Ahmed, “A new hybrid conjugate gradient algorithm for optimization models and its application to regression analysis,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 23, no. 2, pp. 1100–1109, Aug. 2021, doi: 10.11591/ijeecs.v23.i2.pp1100-1109.
[16] J. Zhang and C. Xu, “Properties and numerical performance of quasi-Newton methods with modified quasi-Newton equations,” Journal of Computational and Applied Mathematics, vol. 137, no. 2, pp. 269–278, Dec. 2001, doi: 10.1016/S0774-0517(00)00713-5.
[17] B. A. Hassan and A. R. Ayooob, “An adaptive quasi-newton equation for unconstrained optimization,” in Proceedings of 2021 2nd Information Technology to Enhance E-Learning and other Application Conference, IT-ELA 2021, Dec. 2021, pp. 1–5, doi: 10.1109/IT-ELAS2201.2021.9773580.
[18] M. Malik, M. Mamat, S. S. Abas, I. M. Sulaiman, and S. Sukono, “Performance analysis of new spectral and hybrid conjugate gradient methods for solving unconstrained optimization problems,” IAENG International Journal of Computer Science, vol. 48, no. 1, 2021.
[19] B. A. Hassan and R. M. Sulaiman, “A new class of self-scaling for quasi-newton method based on the quadratic model,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 21, no. 3, pp. 1830–1836, Mar. 2021, doi: 10.11591/ijeecs.v21.i3.pp1830-1836.
[20] H. Isiduka and Y. Narushima, “Conjugate gradient methods using value of objective function for unconstrained optimization,” Optimization Letters, vol. 6, no. 5, pp. 941–955, Jun. 2012, doi: 10.1007/s11590-011-0324-0.
[21] Y. Dai, J. Han, G. Liu, D. Sun, H. Yin, and Y. X. Yuan, “Convergence properties of nonlinear conjugate gradient methods,” SIAM Journal on Optimization, vol. 10, no. 2, pp. 345–358, Jan. 1999, doi: 10.1137/10052623494268443.
[22] B. A. Hassan, K. Muangkan, F. Alfarag, A. H. Ibrahim, and A. B. Abubakar, “An improved quasi-Newton equation on the quasi-Newton methods for unconstrained optimizations,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 22, no. 2, p. 997, May 2021, doi: 10.11591/ijeecs.v22.i2.pp997-1005.
[23] G. Zoutendijk, “Nonlinear programming, computational methods,” Integer and Nonlinear Programming, pp. 37–86, 1970.
[24] R. M. Sulaiman and B. A. Hassan, “Using a new coefficient conjugate gradient method for solving unconstrained optimization problems,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 27, no. 3, pp. 1635–1641, 2022, doi: 10.11591/ijeecs.v27.i3.pp1642-1648.
[25] P. Stanimirovic and M. Miladinovic, “Accelerated gradient descent methods with line search,” Numer. Algorithms, vol. 54, pp. 503–520, Aug. 2010, doi: 10.1007/s11075-009-9350-8.