On image restoration problems using new conjugate gradient methods

Basim A. Hassan, Haneen A. Alashoor
Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul, Mosul, Iraq

Abstract

The nonlinear conjugate gradient algorithm is one of the effective algorithms for optimization since it has low storage and simple structure properties. The coefficient conjugate is the basis of conjugate gradient algorithms with the desirable conjugate property. In this manuscript, we have derived a new second order information for the Hessian from objective function, which can give a new search direction. Based on new search direction, we have proposed the update formula interesting and nonlinear conjugate gradient method. Under wolfe line search and mild assumptions on objective function, the method possess sufficient descent property and are always globally convergent. Numerical results show that the method is effective and competitive to recover the original image from an image corrupted by impulse noise.

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

Keywords:
Conjugate gradient method
Convergence property
Descent property
Image restoration problems
New coefficient conjugate

1. INTRODUCTION

Images are often corrupted by impulse noise. Salt-and-pepper and random-valued impulse noise are removed using two-phase techniques. The first phase employs decision-based median filters to identify pixels that are prone to noise corruption (noise candidates). These noise candidates are restored in the second phase using a detail-preserving regularization approach that preserves edges and noise-free pixels. Let $X$ be the genuine image and $A = \{1,2,3, \ldots M\} \times \{1,2,3, \ldots N\}$ be the index set of $X$. The collection of indices of the noise pixels observed in the first phase is denoted by $N \subset A$. The noise pixels are then recovered in the second phase by (1):

$$F_\alpha(u) = \sum_{i,j \in N} \left[ |u_{i,j} - y_{i,j}| + \frac{\beta}{2} (2 \times S_{i,j}^1 + S_{i,j}^2) \right]$$

(1)

where $\beta$ is the regularization parameter,$S_{i,j}^1 = 2 \sum_{(m,n) \in P_{i,j} \cap N} \phi_\alpha(u_{i,j} - y_{m,n}), S_{i,j}^2 = \sum_{(m,n) \in P_{i,j} \cap N} \phi_\alpha(u_{i,j} - y_{m,n})$ and edge-preserving potential function is $\phi_\alpha = \sqrt{\alpha + x^2}, \alpha > 0$. Let $P_{i,j}$ be the set of four closest neighbors of the pixel at location $(i,j) \in A$, $y_{i,j}$ the image's observed pixel value at position $(i,j)$, and $u_{i,j} = \left[ u_{i,j} \right]_{(i,j) \in N}$ a lexicographically ordered column vector of length $\epsilon$. The number of elements in $N$ is represented by the letter $\epsilon. In actuality, the smooth data-fitting term is eliminated from the minimization, and only noisy pixels are restored. The smooth functional that results is as shown in (2).

$$F_\alpha(u) = \sum_{(i,j) \in N} \left( 2 \times S_{i,j}^1 + S_{i,j}^2 \right)$$

(2)

Journal homepage: http://ijeecs.iaescore.com
For more details see Xue et al. [1] and Yu et al. [2]. Our aim is to minimize a function of \( n \) variables by using iteration methods:

\[
\text{Min} f(u) , \ u \in \mathbb{R}^N
\]

(3)

where \( f \) is smooth and its gradient \( g_{k+1} \) is available. Many practical problems can be translated into the problems, see Dai and Wen [3] and Hajmohammadi [4].

One of the most widely remarkable and powerful used minimization functions is conjugate gradient methods. Begins the minimization process with an initial estimate \( x_0 \) and an initial search direction \( d_0 = -g_0 \).

As we noted that the iterative formula given by (4):

\[
u_{k+1} = u_k + \alpha_k d_k
\]

(4)

where \( \alpha_k \) is a step length generated by a suitable line search and \( d_k \) is the search direction, see [5]. Obviously, the approximate adequate step size associated to quadratic function \( f_k \) is (5).

\[
\alpha_k = -\frac{g_k^T d_k}{d_k^T g_k}
\]

(5)

For general nonlinear functions, it is necessary to use an iterative procedure [6]. We use the typical Wolfe criteria to determine the step length throughout our operation, as shown in (6) and (7):

\[
f(u_k + \alpha_k d_k) \leq f(u_k) + \delta \alpha_k g_k^T d_k
\]

(6)

\[
d_k^T g(u_k + \alpha_k d_k) \geq \sigma d_k^T g_k
\]

(7)

where \( 0 < \delta < \sigma < 1 \), Moghrabi [7]. In conjugate gradient methods, the search directions can be defined recursively:

\[
d_{k+1} = -g_{k+1} + \beta_k s_k
\]

(8)

where the coefficient \( \beta_k \) is so chosen that \( d_k \) and \( d_{k+1} \) must fulfill the conjugacy property. In practical computations promising, the CG methods include suggested by Fletcher and Reeves (FR) [8] and Dai and Yuan (DY) [9] is believed of the efficient methods, respectively given by:

\[
\beta_k^{FR} = \frac{g_k^T s_{k+1}}{s_k^T g_k}, \beta_k^{DY} = \frac{g_k^T g_{k+1}}{s_k^T g_k}
\]

(9)

or via other formulaes (e.g. see [4], [10]-[12]). Some of conjugate gradient (CG) approaches are numerically successful, while the others are conceptually successful.

Later, if the point \( x_{k+1} \) is close enough to a local minimizer \( u^* \) in pure Newton’s technique, the search direction of good to follow, that is, \( d_{k+1} = -Q^{-1}_{k+1} g_{k+1} \), is determined. Nazareth [13], inspired by this, rewrites in (5).

\[
- Q^{-1}_{k+1} g_{k+1} = -g_{k+1} + \beta_k s_k
\]

(10)

On this guideline, this work is improved to investigate ideas from Newton updates, in order to build definite new method for conjugate gradient methods. They proved to be really effective in practice and showed mature convergence properties. This idea was presented as a technical method in some manuscripts, for example, to maximize in the advantages of the original respective conjugate gradient methods [7], [14]-[16]. To introduce a new method, we will find the Hessian approximation of the minimum of a function \( f(u) \), which give a new search direction and choose the coefficient conjugate satisfies above relation.

2. OUR NEW COEFFICIENT CONJUGATE

A very important coefficient conjugate in minimization is that for conjugate gradient method. In order to derive a new coefficient conjugate, we exposed a second order Taylor series as:

\[
f(x) = f(x_{k+1}) + g_{k+1}^T (x - x_{k+1}) + \frac{1}{2} (x - x_{k+1})^T Q(x_{k+1}) (x - x_{k+1})
\]

(11)
can see derivative as (12).

\[ g_{k+1} = g_k + Q(x_{k+1})s_k \]  

(12)

A good second order curvature information, we obtained (13).

\[ s_k^T Q(x_{k+1})s_k = (f_{k+1} - f_k) - 3/2 s_k^T g_k \]  

(13)

The yielded matrix \( Q(x_{k+1}) \) can be as (14).

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

(14)

Putting \( Q(x_{k+1}) \) in (10) we get (15).

\[ \beta_k = \left( 1 - \frac{s_k^T s_k}{(f_{k+1} - f_k) - 3/2 s_k^T g_k} \right) g_{k+1}^T y_k \]  

(15)

In addition by (14) is gained (16).

\[ \beta_k = \left( 1 - \frac{s_k^T s_k}{(f_{k+1} - f_k) + 3/2 s_k^T g_k} \right) g_{k+1}^T y_k \]  

(16)

To ensure the sufficient descent condition, we will do some algebra manipulations on above formula, we get:

\[ \beta_k = \frac{1}{s_k^T y_k} \left( y_k - \alpha \frac{||y_k||^2}{s_k^T y_k} s_k \right)^T g_{k+1} \]  

(17)

where:

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

(18a)

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

(18b)

this formulas will assist us to proving that the direction is satisfied the sufficient descent property. We denote by binaryed neural networks (BNN) and block truncated-Newton (BTN) the methods defined by (17). We introduced the BNN and BTN methods by the following algorithm, which would be used for its convergence analysis. New algorithm (BNN and BTN Algorithms):

- Stage 1. Give initial point \( u_1 \). Set \( k = 1 \) and \( d_1 = -g_1 \). If \( ||g_1|| \leq 10^{-6} \), then stop.
- Stage 2. Evaluate \( \alpha_k \) satisfying the Wolfe conditions (6) and (7).
- Stage 3. Let \( u_{k+1} = u_k + \alpha_k d_k \). If \( ||g_{k+1}|| \leq 10^{-6} \), then stop.
- Stage 4. Evaluate \( \beta_k \) by the formulae (17) then generate \( d_{k+1} \) by (8).
- Stage 5. Set \( k = k + 1 \) and continue with step 2.

**Theorem 1.**

Let \( s_k, y_k, g_{k+1} \in R^n, \beta_k \in R \) and \( \beta_k^{BNN} = \frac{1}{s_k^T y_k} \left( y_k - \alpha \frac{||y_k||^2}{s_k^T y_k} s_k \right)^T g_{k+1} \). If \( s_k^T y_k \neq 0 \), then

\[ d_{k+1}^T g_{k+1} \leq - \left[ 1 - \frac{1}{4\alpha} \right] ||g_{k+1}||^2 \]

**Proof :**

Since \( d_0 = -g_0 \), we have \( g_0^T d_0 = -||g_0||^2 \), which satisfy (17). Multiplying (8) by \( g_{k+1} \), we have:

\[ d_{k+1}^T g_{k+1} = -||g_{k+1}||^2 + \left( \frac{g_{k+1}^T y_k}{s_k^T y_k} - \alpha \frac{||y_k||^2}{s_k^T y_k} \right) s_k^T g_{k+1} \]

(19)

Yielding:
On image restoration problems using new conjugate gradient methods

\[
d^T_{k+1}g_{k+1} = \frac{(\gamma_{k+1}y_k \Sigma_k^T \gamma_{k+1}) (\Sigma_k^T \gamma_{k+1}) - \beta g_{k+1}^T \Sigma_k^2 \gamma_{k+1}^2 - \mu ||y_k||^2 (\gamma_{k+1} \Sigma_k^T y_k)^2}{(\Sigma_k^T y_k)^2}
\]  \tag{20}

now, by applying:

\[
w = \frac{1}{\sqrt{2\sigma}} (\Sigma_k^T y_k) g_{k+1} \quad \text{and} \quad v = \sqrt{2\sigma} (g_{k+1}^T \Sigma_k) y_k
\]  \tag{21}

into the inequality:

\[
w^T v \leq \frac{1}{2} (||w||^2 + ||v||^2)
\]  \tag{22}

we can obtain:

\[
\begin{align*}
(g_{k+1}^T y_k) (\Sigma_k^T g_{k+1}) (\Sigma_k^T y_k) & \leq \frac{1}{2} \left( \frac{1}{\sqrt{2\sigma}} (\Sigma_k^T y_k)^2 ||g_{k+1}||^2 + 2\sigma (\Sigma_k^T g_{k+1})^2 ||y_k||^2 \right)
\end{align*}
\]  \tag{23}

therefore by (23) and (20), we have that:

\[
d^T_{k+1}g_{k+1} \leq \frac{1}{\sqrt{2\sigma}} (\Sigma_k^T y_k)^2 ||g_{k+1}||^2 + [\sigma - \sigma (\Sigma_k^T g_{k+1})^2 ||y_k||^2]
\]  \tag{24}

therefore, we get:

\[
d^T_{k+1}g_{k+1} \leq - \left[ 1 - \frac{1}{4\sigma} \right] ||g_{k+1}||^2
\]  \tag{25}

which completes our proof. In BTN method similar to the proof done in BNN.

3. CONVERGENCE ANALYSIS

To understand the global convergence theorem of the BNN technique, we must first understand the assumptions listed below. 1. The \( \Omega = \{ x \in \mathbb{R}^n \mid f(x) \leq f(x_1) \} \) is a bounded level set. 2. In some neighborhood \( \mathcal{N} \) of \( \Omega \), the gradient of function \( f \) is Lipschitz continuous, namely, there exists a constant \( L > 0 \) such that:

\[
||g(o) - g(\tau)|| \leq L \ ||o - \tau||, \forall \tau, o \in \mathcal{N}
\]  \tag{26}

We demonstrate the Dai et al. [17] theorem it is very important for deducing global convergence.

Lemma 1.

Let \( x_k \) be generated by (4), \( d_k \) satisfy descent property and \( \alpha_k \) be satisfy (6)-(7). If:

\[
\sum_{k=0}^{\infty} \frac{1}{\|d_{k+1}\|^2} = \infty
\]  \tag{27}

then:

\[
\lim_{k \to \infty} \inf ||g_{k+1}|| = 0
\]  \tag{28}

We introduced our main theorem of this paper.

Theorem 2.

Consider the method formed by (8), (17) and \( f(x) \) satisfies Assumption 1 and 2. If step size \( \alpha_k \) satisfies wolfe conditions (6) and (7), then we obtain:

\[
\lim_{k \to \infty} \inf ||g_k|| = 0
\]  \tag{29}

Proof:

We may deduce the following from (8) and the definition of \( \beta_k \) given by (17)

\[
||d_{k+1}|| = ||-g_{k+1} + \beta_k^{\text{BNN}} d_k|| \leq ||g_{k+1}|| +
\]
This relationship demonstrates that:
\[
\sum_{k=1}^{\infty} \frac{1}{||d_k||^2} \geq \left( \frac{a_k}{a_k + 1 + \alpha} \right) \sum_{k=1}^{\infty} 1 = \infty
\]

we obtain \( \lim_{k \to \infty} ||d_k|| = 0 \) from Lemma 1, which is identical to \( \lim_{k \to \infty} ||g_k|| = 0 \) for a uniformly convex function. In BTN way, one can find its proof similar to the proof done in BNN.

4. NUMERICAL RESULT

The goal of the results of applying the proposed BNN and FR methods to test images (Lena, House, Cameraman, and Elaine), will be reported in Table 1, to analyze the practical performance of Algorithms. The method was implemented in MATLAB on a PC-type computer with the stoping criterion:
\[
\frac{|f(u_k) - f(u_{k-1})|}{f(u_k)} \leq 10^{-4} \text{ and } ||f(u_k)|| \leq 10^{-4} (1 + |f(u_k)|)
\]

more details can be found in [18]-[24]. The examined pictures are presented in Table 1, and the comprehensive numerical results of our testing are reported in Table 1, where NI, NF, peak signal to noise ratio (PSNR) denotes the number of iterations, function evaluations, and PSNR, which is defined as:
\[
PSNR = 10 \log_{10} \frac{255^2}{\sum_{i,j} (u_{ij} - u_{ij}^*)^2}
\]

where pixel values for the restored and original images are denoted by \( u_{ij}^* \) and \( u_{ij} \), respectively.

<table>
<thead>
<tr>
<th>Image</th>
<th>Noise level r (%)</th>
<th>FR-Method</th>
<th>BNN-Method</th>
<th>BTN-Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NI</td>
<td>NF</td>
<td>PSNR (dB)</td>
<td>NI</td>
</tr>
<tr>
<td>Le</td>
<td>50</td>
<td>82</td>
<td>153</td>
<td>30.5529</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>81</td>
<td>155</td>
<td>27.4824</td>
</tr>
<tr>
<td>ho</td>
<td>90</td>
<td>108</td>
<td>211</td>
<td>22.8583</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>52</td>
<td>53</td>
<td>30.6845</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>63</td>
<td>116</td>
<td>31.2564</td>
</tr>
<tr>
<td>El</td>
<td>90</td>
<td>111</td>
<td>214</td>
<td>25.287</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>35</td>
<td>36</td>
<td>33.9129</td>
</tr>
<tr>
<td>c512</td>
<td>70</td>
<td>38</td>
<td>39</td>
<td>31.864</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>65</td>
<td>114</td>
<td>28.2019</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>59</td>
<td>87</td>
<td>35.5359</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>78</td>
<td>142</td>
<td>30.6259</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>121</td>
<td>236</td>
<td>24.3962</td>
</tr>
</tbody>
</table>

Figures 1-4 show the obtained results by denoised images. Figure 1(a), Figure 2(a), Figure 3(a), and Figure 4(a) are the images corrupted with 70% salt-and-pepper noise; Figure 1(b), Figure 2(b), Figure 3(b) and Figure 4(b) are results of FR method; Figure 1(c), Figure 2(c) and Figure 3(c), and Figure 4(c) are results of the BNN method; Figure 1(d), Figure 2(d), Figure 3(d) and Figure 4(d) are results of the BTN method. Take everything together, the numerical experiments show that the proposed method performs well for removing impulse noise images. The field of research work in optimization is wider as in [25]-[30].
Figure 1. Comparing images results of algorithms: (a) denoised images with 70% salt-and-pepper noise, (b) recovered images through FR, (c) and (d) restored images using BNN and BTN of 256 * 256

Figure 2. Comparing images results of algorithms: (a) denoised images with 70% salt-and-pepper noise, (b) recovered images through FR, (c) and (d) restored images using BNN and BTN of 256 * 256

Figure 3. Comparing images results of algorithms: (a) denoised images with 70% salt-and-pepper noise, (b) recovered images through FR, (c) and (d) restored images using BNN and BTN of 256 * 256

Figure 4. Comparing images results of algorithms: (a) denoised images with 70% salt-and-pepper noise, (b) recovered images through FR, (c) and (d) restored images using BNN and BTN of 256 * 256
5. CONCLUSION

Using the second order Taylor series to deriving a new coefficient conjugate for conjugate gradient method for image restoration problems. The algorithms exhibit global convergence and the necessary descent property holds. In numerical tests, the new method has made significant progress. It demonstrates that the novel conjugate gradient approach outperforms the conventional FR conjugate gradient method.

REFERENCES


On image restoration problems using new conjugate gradient methods (Basim A. Hassan)

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.

Haneen A. Alashoor is a Lecturer at Faculty of Basic Education/Sharqat from University of Tikrit. Received The Bachelors (B.S.C) and Master (M.SC) Degrees in The College of Computer Science and Mathematics/Department of Mathematics from the University of Mosul in 2013 and 2018 Respectively with specialization in optimization. To date, he has published more than 5 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 methods with application in signal recovery and image restoration. She can be contacted at email: haneen19921006@gmail.com.