EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 15, No. 4, 2022, 2011-2021
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

Efficient New Conjugate Gradient Methods for Removing Impulse Noise Images

Basim A. Hassan[1,*], Hameed M. Sadiq[1]

[1] *Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul, Iraq*

Abstract. In most applications, denoising image is fundamental to subsequent image processing operations. In this research, we derivation a new formula of conjugate gradient methods based on the quadratic model. The fact that the search direction created at each iteration of the proposed approach is descending and independent of the line search makes it interesting. The use of Wolfe conditions also determines the global convergence of the suggested approach. To prove the viability of the suggested approach, comparison tests on impulse noise reduction are given.

2020 Mathematics Subject Classifications: 90C30, 65K05, 49M37

Key Words and Phrases: Optimization, impulse noise, Conjugate gradient method, Global convergence

1. Introduction

Image denoising is a fundamental problem in image processing operations. One of its more investigated domains is image denoising which plays an adequate contribution in many applications. There are two main models to represent impulse noise [17]. One type of noise is known as salt-and-pepper noise, in which the noisy pixels can only accept the maximum and minimum pixel values possible within the dynamic range of the source image. The random-valued noise, which may have any random value between the maximum and minimum pixel values of dynamic range, is another type of impulsive noise [18]. One of the most significant issues in picture processing is the removal of above both noise. The average filter and its variations [7] may find the noisy pixels but return them badly when the noise ratio is large. These two approaches are quite common for this purpose. Unaltered gray levels exist in unharmed pixels. While the variational technique is capable of keeping the features and edges well, every pixel's gray level, including those that aren't damaged, is altered, see [19]. The recovered picture may also lose its details and become distorted.

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v15i4.4568

Email addresses: basimah@uomosul.edu.iq (B. A. Hassan),
hameedsadiq1975@gmail.com (H. M. Sadiq)

The two-phase method in [6] unify the advantages of both methods and we will give a brief description here. In the following, we let $x_{i,j}$ for $(i,j) \in A = 1,2,3,...,M \times 1,2,3,...,N$, be the gray level of a true M by N. Let the set of indices of the noisy pixels uncover in the first phase denote by N, where $N \subset A$. Let $u_{i,j} = [u_{i,j}]_{(i,j)\in N}$ be a column vector of length c ordered lexicographically (c is the number of elements of N), and $y_{i,j}$ denote the observed pixel value at position (i, j). Then, the second phase it to recover the noisy pixels by minimizing the following edge-preserving regularization function:

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

where $S^1_{i,j} = 2\sum_{(m,n)\in P_{i,j}\cap N^c} \varphi_\alpha(u_{i,j} - y_{m,n}), S^2_{i,j} = \sum_{(m,n)\in P_{i,j}\cap N} \varphi_\alpha(u_{i,j} - y_{m,n})$ and φ_α is an edge-preserving potential function having the parameter α. Examples of such $\varphi_\alpha(x)$ are: $\varphi_\alpha(x) = \sqrt{\alpha + x^2}, \alpha > 0$. However, because of the $|u_{i,j} - y_{i,j}|$ term, the functional of problem (1) is nonsmooth. It is commonly accepted that this nonsmooth term can separate from (1) because, on the one hand, it keeps the minimizer u close to the original picture y, ensuring that the original image's unaltered pixels are preserved. However, the two-phase approach simply cleans the noisy pixels, leaving the unharmed pixels unaltered, making issue (1) functional. Consequently, the nonsmooth term is not necessary. The word "data-fitting" in [4] should be removed, according to Cai et al. With this process, $f_\alpha(u)$ may be converted into a smooth function that can be effectively reduced. As a result, the objective function that we will reduce in this essay has the following shape:

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

Nowadays, conjugate gradient (CG) methods are regarded as popular and efficient algorithms to deal problems noise of image. In general, the method has the following form:

$$Minf(u); u \in R^n \tag{3}$$

where f is a continuously differentiable function. This problem may be effectively solved using an iterative technique at the $(k + 1)$ iteration by using the iteration form shown below:

$$u_{k+1} = u_k + \alpha_k d_k \tag{4}$$

where u_k is the current iterate point, d_k is a direction of f at u_k, and $\alpha_k > 0$ is step size obtained by a one-dimensional line search. step size α_k is obtained using several forms of line search, i.e., exact line search with quadratic model [20] as follows:

$$\alpha_k = -\frac{g_k^T d_k}{d_k^T Q d_k} \tag{5}$$

The strong Wolfe inexact line search is frequently taken into consideration in the convergence analysis implementation of nonlinear conjugate gradient techniques since exact line search for searching α_k is typically costly and impracticable. It seeks to identify a step size α_k that satisfies the two strong Wolfe requirements listed below [5], namely:

$$f(u_k + \alpha_k d_k) \leq f(u_k) + \delta \alpha_k g_k^T d_k \tag{6}$$

$$\left| g(u_k + \alpha_k d_k)^T d_k \right| \leq \sigma \left| g_k^T d_k \right| \tag{7}$$

where $0 < \delta < \sigma < 1$ are arbitrary constants and $g_k = \nabla f(x_k)$. The search direction d_k is computed by:

$$d_{k+1} = -g_{k+1} + \beta_k s_k \tag{8}$$

β_k is the conjugate gradient parameter that evaluates the performance and global convergence characteristics of several conjugate gradient technique types. The nonlinear conjugate gradient parameters include some well-known ones like the Fletcher and Reeves (FR)[10], conjugate descent (CD)[9], Dai and Yuan (DY) [8]. These parameters are given by the following formulae:

$$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}, \beta_k^{CD} = -\frac{g_{k+1}^T g_{k+1}}{d_k^T g_k} \; \beta_k^{DY} = \frac{g_{k+1}^T g_{k+1}}{d_k^T y_k}, \tag{9}$$

More information about other conjugate gradient methods [16].

Al-Baali [1] extended this result to an inexact line search and showed that the method generates sufficient descent direction under the strong Wolfe conditions using the constraint $\sigma < 1/2$. The global convergence properties of FR, DY, and CD methods with exalt line are strong, but they are prone to taking many short steps without making sufficient advancement to the minimum by Hager and Zhang [11].

To provide a broader context for developing conjugate gradient methods, Perry extended the classical conjugate condition to:

$$d_{k+1}^T y_k = -(H_{k+1} g_{k+1})^T y_k = -g_{k+1}(H_{k+1} y_k) = -g_{k+1}^T s_k = 0 \tag{10}$$

Many efforts have been made in few recent years to design new formulas for conjugate gradient method which are not only satisfied global convergence but also improve numerical performance for method. Remainder the conjugate gradient methods have many application in real life In our work, we found a new formula for conjugate gradient method with Wolfe–Powell generate a descent direction at each iteration in section 2 and the new formula for conjugate gradient method which is satisfied the global convergence in section 3 . In section 4, we present the numerical behavior of the method. The last section proposes one conclusion.

2. Propose new conjugate gradient method:

We will discuss the new parameter choice. Now using the quadratic formula for the objective function $f(x)$ we have:

$$f_{k+1} = f_k + s_k^T g_k + \frac{1}{2} s_k^T Q(u_k) s_k \tag{11}$$

where $Q(u_k)$ is the Hessian matrix of second-order derivatives. we derivative both sides of (11) for s_k, is presented as follows:

$$\nabla f_{k+1} = g_k + Q(u_k)s_k = 0 \tag{12}$$

Using (5) and (12) in (11), we have:

$$s_k^T Q(u_k)s_k = f_k - f_{k+1} + \frac{1}{2}s_k^T y_k \tag{13}$$

After some algebra from (11) and (13), as a result:

$$\beta_k = \frac{\frac{1}{2}g_{k+1}^T y_k + \frac{(f_k - f_{k+1})}{s_k^T y_k} g_{k+1}^T y_k}{d_k^T y_k}. \tag{14}$$

If exact line search is utilized, then β_k is such that:

$$\beta_k^{BV1} = \frac{\frac{1}{2}\|g_{k+1}\|^2 + \frac{(f_k - f_{k+1})}{s_k^T y_k}\|g_{k+1}\|^2}{d_k^T y_k}. \tag{15}$$

In particular, conclude:

$$\beta_k^{BV2} = \frac{\frac{1}{2}\|g_{k+1}\|^2 - \frac{(f_k - f_{k+1})}{s_k^T g_k}\|g_{k+1}\|^2}{-d_k^T g_k} \tag{16}$$

and

$$\beta_k^{BV3} = \frac{\frac{1}{2}\|g_{k+1}\|^2 + \frac{(f_k - f_{k+1})}{a_k\|g_k\|^2}\|g_{k+1}\|^2}{\|g_k\|^2}. \tag{17}$$

Thus, BV1, BV2 and BV3 are the new parameters of conjugate gradient.

Based on the above discussions, the presented algorithm is stated as follows:

Step 1: Given a starting point u_1. set $k = 0$ and $d_\circ = -g_\circ$.
Step 2: Compute β_k by (15), (16) and (17).
Step 3: Compute d_k by (8) and (15). If $\|g_k\| = 0$, then stop.
Step 4: Evaluate α_k satisfy the conditions (6) and (7).
Step 5: Update the new point by the recurrence expression (4).
Step 6: If $f(u_{k+1}) < f(u_k)$ and $\|g_k\| < \varepsilon$ then stop. otherwise go to Step 2 with $k = k+1$.

An important feature for any minimization algorithm is the descent or the sufficient descent property. The following theorem indicates that search direction d_k satisfies descent property of our algorithms.

Theorem 1. *In the algorithm (4), (8), (15), assume that α_k determined by the Wolfe line search (6)-(7) then the direction d_{k+1} given by (8) is a descent direction.*

Proof. If $k = 0$, then $d_1 = -g_1$, so $d_1^T g_1 = -\|g_1\|^2 < 0$. suppose that $d_k^T g_k < 0$ for all k. Multiply (8) by g_{k+1}^T, will give:

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \beta_k^{BV1} d_k^T g_{k+1} \tag{18}$$

Since

$$\|g_{k+1}\|^2 = \beta_k d_k^T y_k \tag{19}$$

Now, put (19) in (18) we obtain:

$$d_{k+1}^T g_{k+1} = -\beta_k d_k^T y_k + \beta_k d_k^T g_{k+1} = \beta_k[-d_k^T y_k + d_k^T g_{k+1}] = \beta_k d_k^T g_k < 0 \tag{20}$$

The proof is complete. The proof descent property of BV2 and BV3 is similar to proof BV1.

3. Convergence Analysis:

In order to establish the global convergence property of the method, we make the following standard assumptions for the objective function:

- For any initial point $x_1 \in R^n$, the level set $O = \{x \in R^n \backslash f(x) < f(x_1)\}$ is bounded.
- $f(x)$ is continuously differentiable in a neighborhood U of Ω, and its gradient $g(x)$ is Lipschitz continuous, namely, there exists a constant $L > 0$ such that:

$$\|g(x) - g(y)\| = L\|x - y\|, \forall x, y \in U. \tag{21}$$

To proceed, the well-known Zoutendijk condition [22] is reviewed in the following.

Lemma 1. *Suppose that Assumptions holds true. For any CG iterative algorithm defined by (4), where* d_k *is defined by (8), and the step-size* α_k *is obtained by the Wolfe line search. Then:*

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty. \tag{22}$$

In the following theorem, the convergence property of new Algorithm is proved.

Theorem 2. *Suppose the all assumptions holds true. Consider the sequence*$\{g_k\}$ *and* $\{d_k\}$ *generated by the proposed method, where* β_k *is given by (15), and* α_k *satisfies the Wolfe line search, then,*

$$\lim_{k \longrightarrow \infty} inf \|g_k\| = 0 \tag{23}$$

Proof. By contradiction, suppose that (23) is not correct. Therefore, there exists a constant $\varepsilon > 0$ such that:

$$\|g_{k+1}\| > \varepsilon \tag{24}$$

Upon squaring both sides of (8), we obtain:

$$\|d_{k+1}\|^2 + \|g_{k+1}\|^2 + 2d_{k+1}^T g_{k+1} = (\beta_k)^2 \|d_k\|^2 \tag{25}$$

Next, dividing both sides of the above inequality by $(g_{k+1}^T d_{k+1})^2$, we have:

$$\begin{aligned}\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} &= \frac{(\beta_k)^2\|d_k\|^2}{(d_{k+1}^T g_{k+1})^2} - \frac{2}{(d_{k+1}^T g_{k+1})^2} - \frac{\|g_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} \\ &= \frac{(\beta_k)^2\|d_k\|^2}{(d_{k+1}^T g_{k+1})^2} - \left(\frac{1}{\|g_{k+1}\|^2} + \frac{\|g_{k+1}\|}{(d_{k+1}^T g_{k+1})}\right)^2 + \frac{1}{\|g_{k+1}\|^2}\end{aligned} \tag{26}$$

which reduces to:

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} = \frac{(\beta_k)^2\|d_k\|^2}{(d_{k+1}^T g_{k+1})^2} + \frac{1}{\|g_{k+1}\|^2} \tag{27}$$

However, from (20) we have $\beta_k = \frac{d_{k+1}^T g_{k+1}}{d_k^T g_k}$, Substituting in (27), we obtain:

$$\begin{aligned}\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} &= \frac{(d_{k+1}^T g_{k+1})^2}{(d_k^T g_k)^2}\frac{\|d_k\|^2}{(d_{k+1}^T g_{k+1})^2} + \frac{1}{\|g_{k+1}\|^2} \\ &= \frac{\|d_k\|^2}{(d_k^T g_k)^2} + \frac{1}{\|g_{k+1}\|^2}\end{aligned} \tag{28}$$

Notice that $\|d_1\|^2 = -g_1^T d_1 = \|g_1\|^2$, which implies:

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} = \sum_{i=1}^{k+1} \frac{1}{\|g_i\|^2}. \tag{29}$$

Then we get from (29) and (24) that:

$$\frac{(d_k^T g_k)^2}{\|d_k\|^2} = \frac{\in^2}{k}. \tag{30}$$

Therefore:

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = \infty. \tag{31}$$

This result contradicts (22), The proof is completed. Global convergence property of BV2, BV3 algorithm are similar those of BV1 algorithm.

4. Numerical experiments

To illustrate the effectiveness of the suggested approach for salt-and-pepper impulse noise, we present some numerical results in this section. Table 1 lists the experimental outcomes. We report the number of iterations (NI), the number of function evaluations

(NOF) and the evaluation indexes used in the experiments were the PSNR (peak signal to noise ratio) see [3], which is defined as:

$$PSNR = 10log_{10}\frac{255^2}{\frac{1}{MN}\sum_{i,j}(u^r_{i,j} - u^*_{i,j})^2} \tag{32}$$

where $u^r_{i,j}$ and $u^*_{i,j}$ denote the pixel values of the restored image and the original image. We stop the iteration if the inequality:

$$\frac{|f(u_k) - f(u_{k-1})|}{|f(u_k)|} = 10^{-4}\ and\ \|f(u_k)\| = 10^{-4}(1 + |f(u_k)|) \tag{33}$$

are satisfied. In the experiment, a picture that has been lost or become hazy is recovered or recreated. Four different images are employed for the experiment, which includes, placeLena, House, Cameraman and Elaine by employing test images in [2],[21] results of experiment to images shown in Table (1). Images show that new Algorithm and FR Algorithm have good performance to solve the image restoration and it can successfully do this problem, for more details in this field see [12–15].

The numerical results show that for some situations, the suggested solution outperforms the FR method.

Table 1: Numerical results of FR, BV1, BV2 and BV3 algorithms.

Image	Noise level r (%)	FR-Method			BV1-Method			BV2-Method			BV3-Method		
		NI	NF	PSNR (dB)	NI	NF	PSNR (dB)	NI	NF	PSNR (dB)	NI	NF	PSNR (dB)
Le	50	82	153	30.5529	54.0	66.0	30.4969	40.0	84.0	30.5043	63.0	66.0	30.6177
	70	81	155	27.4824	53.0	65.0	27.6745	49.0	101.0	27.3327	60.0	63.0	27.4735
	90	108	211	22.8583	73.0	86.0	23.2218	55.0	109.0	22.5143	71.0	74.0	22.6418
Ho	50	52	53	30.6845	43.0	52.0	35.0342	29.0	59.0	34.5810	46.0	51.0	34.5704
	70	63	116	31.2564	41.0	50.0	30.5736	36.0	71.0	31.0401	40.0	44.0	30.8299
	90	111	214	25.287	58.0	64.0	25.0723	51.0	103.0	25.1160	63.0	67.0	25.0853
El	50	35	36	33.9129	28.0	33.0	33.8784	24.0	43.0	33.8835	34.0	35.0	33.8796
	70	38	39	31.864	42.0	48.0	31.8125	28.0	51.0	31.8465	39.0	40.0	31.8010
	90	65	114	28.2019	54.0	62.0	28.1780	38.0	70.0	28.0831	48.0	54.0	28.3744
c512	50	59	87	35.5359	45.0	54.0	35.3429	28.0	61.0	35.3759	44.0	52.0	35.2678
	70	78	142	30.6259	52.0	62.0	30.9240	35.0	75.0	30.7033	49.0	53.0	30.7929
	90	121	236	24.3962	69.0	81.0	24.7572	46.0	96.0	25.0248	67.0	73.0	24.8688

In terms of the number of iterations and function evaluations, as well as the peak signal to noise ratio, the recommended algorithms surpass the FR technique, as shown in Table (1).

5. Conclusions

We presented a powerful conjugate gradient technique. In addition to meeting the adequate descent criterion, the proposed approach is globally convergent. According to numerical findings, the approach operates well in practice and is superior than the widely used FR method. We also looked at our approach's aptitude for resolving several practical problems. In this manner, a typical issue from applications for image processing was taken into account. We demonstrated the acceptability of the image that our approach restored.

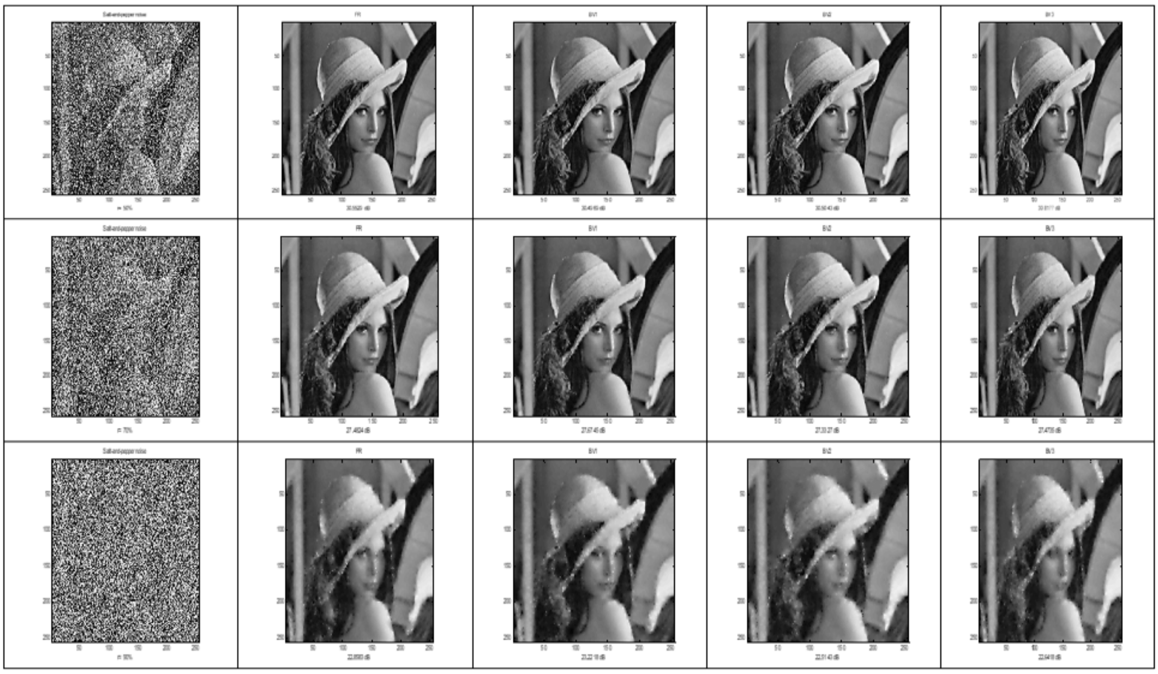

Figure 1: Demonstrates the results of algorithms FR, BV1, BV2 and BV3 of 256×256 Lena image.

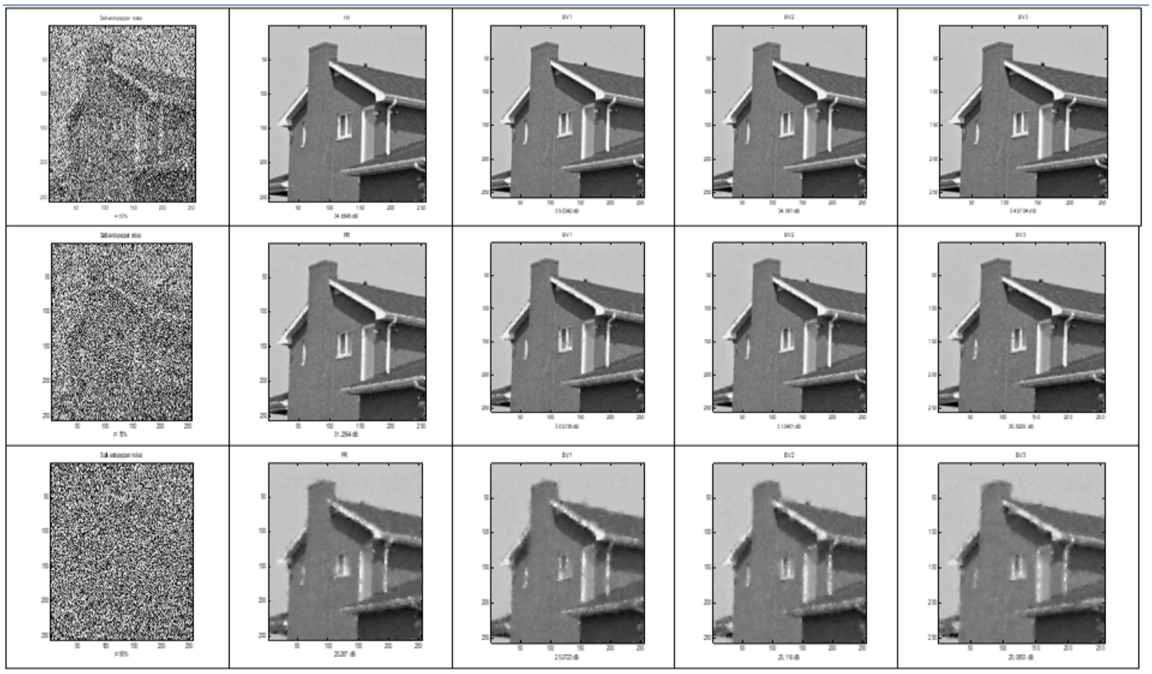

Figure 2: Demonstrates the results of algorithms FR, BV1, BV2 and BV3 of 256×256 House image.

Figure 3: Demonstrates the results of algorithms FR, BV1, BV2 and BV3 of 256×256 Elaine image.

Figure 4: Demonstrates the results of algorithms FR, BV1, BV2 and BV3 of 256×256 Cameraman image.

References

[1] Mehiddin Al-Baali. Descent property and global convergence of the fletcher—reeves method with inexact line search. *IMA Journal of Numerical Analysis*, 5(1):121–124, 1985.

[2] Neculai Andrei. An unconstrained optimization test functions collection. *Adv. Model. Optim*, 10(1):147–161, 2008.

[3] Alan C Bovik. *Handbook of image and video processing.* Academic press, 2010.

[4] Jian-Feng Cai, Raymond Chan, and Benedetta Morini. Minimization of an edge-preserving regularization functional by conjugate gradient type methods. In *Image Processing Based on Partial Differential Equations*, pages 109–122. Springer, 2007.

[5] Jian-Feng Cai, Raymond H Chan, and Mila Nikolova. Fast two-phase image deblurring under impulse noise. *Journal of Mathematical Imaging and Vision*, 36(1):46–53, 2010.

[6] Raymond H Chan, Chung-Wa Ho, and Mila Nikolova. Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. *IEEE Transactions on image processing*, 14(10):1479–1485, 2005.

[7] Tao Chen and Hong Ren Wu. Adaptive impulse detection using center-weighted median filters. *IEEE signal processing letters*, 8(1):1–3, 2001.

[8] Yu-Hong Dai and Yaxiang Yuan. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on optimization*, 10(1):177–182, 1999.

[9] John E Dennis Jr. Practical methods of optimization, vol. 1: Unconstrained optimization (r. fletcher), 1982.

[10] Reeves Fletcher and Colin M Reeves. Function minimization by conjugate gradients. *The computer journal*, 7(2):149–154, 1964.

[11] William W Hager and Hongchao Zhang. A survey of nonlinear conjugate gradient methods. *Pacific journal of Optimization*, 2(1):35–58, 2006.

[12] Basim A Hassan, Fadhil Alfarag, and Snezana Djordjevic. New step sizes of the gradient methods for unconstrained optimization problem. *Italian Journal of Pure and Applied Mathematics*, page 583.

[13] Basim A Hassan and Abdulrahman R Ayoob. An adaptive quasi-newton equation for unconstrained optimization. In *2021 2nd Information Technology To Enhance e-learning and Other Application (IT-ELA)*, pages 1–5. IEEE, 2021.

[14] Basim A Hassan and Abdulrahman R Ayoob. On the new quasi-newton equation for unconstrained optimization. In *2022 8th International Engineering Conference on Sustainable Technology and Development (IEC)*, pages 168–172. IEEE, 2022.

[15] Basim A Hassan and Ranen M. Sulaiman. Using a new type of formula conjugate on the gradient methods. *Indonesian Journal of Electrical Engineering and Computer Science*, 27(1):86–91, 2022.

[16] P Kaelo, P Mtagulwa, and MV Thuto. A globally convergent hybrid conjugate gradient method with strong wolfe conditions for unconstrained optimization. *Mathematical Sciences*, 14(1):1–9, 2020.

[17] Morteza Kimiaei and Majid Rostami. Impulse noise removal based on new hybrid conjugate gradient approach. *Kybernetika*, 52(5):791–823, 2016.

[18] Peter Mtagulwa and P Kaelo. A convergent modified hs-dy hybrid conjugate gradient method for unconstrained optimization problems. *Journal of Information and Optimization Sciences*, 40(1):97–113, 2019.

[19] Mila Nikolova. A variational approach to remove outliers and impulse noise. *Journal of Mathematical Imaging and Vision*, 20(1):99–120, 2004.

[20] Jorge Nocedal and Stephen J Wright. Numerical optimization, 2006.

[21] Gaohang Yu, Liqun Qi, Yimin Sun, and Yi Zhou. Impulse noise removal by a nonmonotone adaptive gradient method. *Signal processing*, 90(10):2891–2897, 2010.

[22] G Zoutendijk. Nonlinear programming, computational methods. *Integer and nonlinear programming*, pages 37–86, 1970.