RESEARCH ARTICLE
doi: 10.2306/scienceasia1513-1874.2019.45.092

ScienceAsia 45 (2019): **92–97**

An iterative method for impulse noise removal

Xuesha Wu[a,*], **Yun Ji**[b]

[a] School of Liberal and International Education, Chongqing College of Electronic Engineering, Chongqing 401331 China
[b] School of Electronics and Internet of Things, Chongqing College of Electronic Engineering, Chongqing 401331 China

*Corresponding author, e-mail: wuxuesha2013@126.com

Received 31 Oct 2018
Accepted 28 Feb 2019

ABSTRACT: Based on the conjugate gradient method, we propose a modified iterative method (MCG) for impulse noise removal based on the famous two-phase scheme. The noise candidates are firstly identified by the adaptive median filter, and then the MCG method recovers these noise candidates by minimizing an edge-preserving regularization function. A nice property of the MCG method is that the global convergence can be established without assuming that the objective function satisfies strong convexity under the Armijo-type line search. Numerical experiments show that the MCG method is effective to remove the impulse noise.

KEYWORDS: two-phase method, optimization, conjugate gradient method

MSC2010: 90C30 65K05

INTRODUCTION

In practice, images are often corrupted by impulse noise during image acquisition and transmission. Thus some algorithms for removing impulse noise have been intensively investigated. For example, the median filter and some of its variants[1,2] have been constructed based on the nonlinear digital filters[3]. These methods firstly locate possible noisy candidates, and then replace them with medians or their variants. Furthermore, these methods can detect the noisy pixels even at a high noise level. It is a pity that they cannot restore such pixels satisfactorily because they do not take into account local image features such as the possible presence of the edges. Subsequently, Nikolova[4] proposed a variational method for impulse noise removal, which can preserve details and the edges well, but the grey level of every pixel is changed including uncorrupted ones. Other types of methods have also been extensively studied[5–11].

In this paper we focus on the famous two-phase method for impulse noise removal proposed by Chan et al[11]. The noisy candidates are firstly detected by the adaptive median filter[1] or the adaptive centre-weighted median filter[12] for different types of noise, and then the detected noisy pixels are recovered by minimizing an objective function by some optimization methods. Chan et al gave some numerical results to show that the two-phase method is powerful even for a salt-and-pepper noise ratio as high as 90%. However, the given objective function is not smooth because it includes an ℓ_1 data-fitting term. This implies that it is costly to obtain the minimizer in the second phase. To reduce the cost of computing, Chan et al[13] showed that the non-smooth ℓ_1 data-fitting term can be deleted from the objective function, because only detected noisy pixels are restored in the minimizing process. Furthermore, some experiments show that the quality of the recovered images is not affected. Thus the objective function proposed in Ref. 11 can be reduced to the following function. Let X be the true image with M-by-N pixels, and $\mathscr{A} = \{1,2,3,\dots,M\} \times \{1,2,3,\dots,N\}$ be the index set of X. $\mathscr{N} \subset \mathscr{A}$ denotes the set of indices of the detected noisy pixels in the first phase. In the second phase, the detected noisy pixels can be restored by minimizing the following function.

$$\mathscr{F}_\alpha(u) = \sum_{(i,j)\in\mathscr{N}} \Bigg\{ \sum_{(m,n)\in\mathscr{V}_{i,j}\setminus\mathscr{N}} \varphi_\alpha(u_{i,j} - y_{m,n}) + \frac{1}{2} \sum_{(m,n)\in\mathscr{V}_{i,j}\cap\mathscr{N}} \varphi_\alpha(u_{i,j} - u_{m,n}) \Bigg\}, \quad (1)$$

where $\mathscr{V}_{i,j}$ is the set of the four closest neighbours of the pixels at the position $(i,j) \in \mathscr{A}$, $y_{m,n}$ is the observed pixel value of the image at the position (m,n), $u=[u_{i,j}]_{(i,j)\in\mathcal{N}}$ is a column vector of the length c ordered lexicographically in which c is the number of elements of $\mathcal{N}$, and $\varphi_\alpha(t)$ is an edge-preserving function which has a great influence on the features of $\mathscr{F}_\alpha(u)$.

The objective function $\mathscr{F}_\alpha$ is smooth if the selected edge-preserving function φ_α is smooth. This means that some first-order optimization algorithms can be used to recover the corrupted images by minimizing (1). Cai et al[14] used the conjugate gradient methods to recover the corrupted images without any line search. Yu et al[15] proposed a descent spectral conjugate gradient method for impulse noise removal. A favourite property of the proposed method is that the search direction generated is a descent direction at each iteration. Under the strong Wolfe line search, its global convergence could be established if $\mathscr{F}_\alpha$ is strongly convex. Liu et al[16] constructed a spectral gradient (SP) method to remove the impulse noise. The search direction generated by the SP method satisfies the sufficient descent property at each iteration, which is independent of any line search. Under the Armijo-type line search, the global convergence of the SP method is established for general smooth functions.

Hager et al[17] proposed a nonlinear conjugate gradient method where the directions are generated by the rule

$$d_0=-g_0,\quad d_k=-g_k+\beta_k^{\mathrm{HZ}}d_{k-1},\quad k\geqslant 1,$$

with

$$\beta_k^{\mathrm{HZ}}=\frac{1}{d_{k-1}^{\mathrm{T}}y_{k-1}}\left(y_{k-1}-2d_{k-1}\frac{\|y_{k-1}\|^2}{d_{k-1}^{\mathrm{T}}y_{k-1}}\right)^{\mathrm{T}}g_k.$$

Under the standard Wolfe line search, the global convergence of this method is established for strongly convex smooth functions. To prove the global convergence for general smooth functions, they set $\beta_k^{\mathrm{CG}}=\max\{\beta_k^{\mathrm{HZ}},\eta_k\}$, where

$$\eta_k=\frac{-1}{\|d_{k-1}\|\min\{\eta,\|g_{k-1}\|\}},\quad \eta>0.$$

This is the famous CG_DESCENT method, which is one of the most efficient conjugate gradient methods for solving unconstrained optimization problems. In this study, we are interested in the parameter β_k^{HZ}. Based on the parameter β_k^{HZ}, we propose a modified conjugate gradient (MCG) method for impulse noise removal. The MCG method inherits the nice property of the HZ method, i.e., the search direction generated by the MCG method always satisfies the sufficient descent property independent of any line search. In particular, its global convergence can be established for general smooth functions under the Armijo-type line search. Numerical experiments show that the proposed MCG method performs well for impulse noise removal.

ALGORITHM AND ITS GLOBAL CONVERGENCE

We start with the convergence theorem for general smooth function $f(x)$. Firstly, we give the modified conjugate gradient method.

Algorithm 1

Step 1: Give $x_0\in\mathbb{R}^n$, $\rho\in(0,1)$, $\tau>0$, $\delta\in(0,1)$, $t>1/4$, $\varepsilon>0$. Set $k=0$.

Step 2: If $\|g_k\|\leqslant\varepsilon$, stop.

Step 3: Define d_k by

$$d_k=\begin{cases}-g_k, & k=0,\\ -g_k+\beta_k d_{k-1}, & k\geqslant 1,\end{cases}\tag{2}$$

where g_k denotes the gradient of the objective function at the kth iterate and

$$\beta_k=\frac{g_k^{\mathrm{T}}y_{k-1}}{d_{k-1}^{\mathrm{T}}\gamma_{k-1}}-\frac{t\|y_{k-1}\|^2}{(d_{k-1}^{\mathrm{T}}\gamma_{k-1})^2}g_k^{\mathrm{T}}d_{k-1},\tag{3}$$

$\gamma_{k-1}=y_{k-1}+r_{k-1}d_{k-1}$, $y_{k-1}=g_k-g_{k-1}$, $r_{k-1}=1+\max\{0,-d_{k-1}^{\mathrm{T}}y_{k-1}/d_{k-1}^{\mathrm{T}}d_{k-1}\}$.

Step 4: Compute the step-size α_k.

Step 5: Let $x_{k+1}=x_k+\alpha_k d_k$, set $k:=k+1$, go to Step 2.

The following result indicates that the MCG method must satisfy the sufficient descent condition without any line search.

Lemma 1 *Let the sequences $\{d_k\}$ and $\{g_k\}$ be generated by the MCG method. Then*

$$g_k^{\mathrm{T}}d_k\leqslant-\left(1-\tfrac{1}{4t}\right)\|g_k\|^2,\ k\geqslant 0.\tag{4}$$

Proof: (4) holds for $k=0$. For $k\geqslant 1$, from (2) we have

$$g_k^T d_k=-\|g_k\|+\frac{(g_k^T y_{k-1})(g_k^T d_{k-1})}{d_{k-1}^T\gamma_{k-1}}-\frac{t\|y_{k-1}\|^2(g_k^T d_{k-1})^2}{(d_{k-1}^T\gamma_{k-1})^2}$$

$$
\begin{aligned}
&= \Big[-\|g_k\|(d_{k-1}^T \gamma_{k-1})^2 \\
&\quad + \tfrac{d_{k-1}^T \gamma_{k-1}}{\sqrt{2t}} g_k^T (\sqrt{2t} g_k^T d_{k-1}) y_{k-1} \\
&\quad - t\|y_{k-1}\|^2 (g_k^T d_{k-1})^2 \Big] / (d_{k-1}^T \gamma_{k-1})^2 \\
&\leqslant \Big[-\|g_k\|(d_{k-1}^T \gamma_{k-1})^2 + \tfrac{(d_{k-1}^T \gamma_{k-1})^2}{4t} \|g_k\|^2 \\
&\quad + t(g_k^T d_{k-1})^2 \|y_{k-1}\|^2 \\
&\quad - t\|y_{k-1}\|^2 (g_k^T d_{k-1})^2 \Big] / (d_{k-1}^T \gamma_{k-1})^2 \\
&= \quad -\big(1 - \tfrac{1}{4t}\big) \|g_k\|^2.
\end{aligned}
$$

□

To prove the global convergence of the MCG method, the following Armijo line search is needed, i.e., the step-size $\alpha_k = \max\{\rho^i \tau |g_k^T d_k| / \|d_k\|^2, i = 0, 1, 2, 3, \ldots\}$ satisfies

$$f(x_k + \alpha_k d_k) \leqslant f(x_k) - \delta \alpha_k^2 \|d_k\|^2, \quad k \geqslant 0. \tag{5}$$

The following result is easily obtained under the given line search.

Lemma 2 *Assume that the level set $\Gamma = \{x \in \mathbb{R}^n \mid f(x) \leqslant f(x_0)\}$ is bounded. Let the sequence $\{d_k\}$, $k \geqslant 0$, be generated by the MCG method, and the step-size α_k satisfies (5). Then*

$$\lim_{i \to \infty} \alpha_i \|d_i\| = 0. \tag{6}$$

Proof: By (5), we have

$$\sum_{i=0}^{k} \delta \alpha_i^2 \|d_i\|^2 \leqslant f(x_0) - f(x_{k+1}).$$

Since the level set Γ is bounded, then

$$\sum_{i=0}^{\infty} \delta \alpha_i^2 \|d_i\|^2 < \infty.$$

This indicates that the conclusion (6) holds. □

Theorem 1 *Assume that the level set $\Gamma = \{x \in \mathbb{R}^n \mid f(x) \leqslant f(x_0)\}$ is bounded. Let the sequence $\{g_k\}$, $k \geqslant 0$, be generated by the MCG method, and the step-size α_k satisfies (5). If g satisfies the Lipschitz condition, i.e., there exists a constant $L > 0$ such that*

$$\|g(x) - g(y)\| \leqslant L\|x - y\|, \quad \forall x, y \in \Gamma, \tag{7}$$

then we have

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

Proof: From the definition γ_{k-1} in the MCG algorithm, we have

$$
\begin{aligned}
d_{k-1}^{\mathrm{T}} \gamma_{k-1} &= d_{k-1}^{\mathrm{T}} y_{k-1} + r_{k-1} d_{k-1}^{\mathrm{T}} d_{k-1} \\
&\geqslant d_{k-1}^{\mathrm{T}} y_{k-1} + d_{k-1}^{\mathrm{T}} d_{k-1} - d_{k-1}^{\mathrm{T}} y_{k-1} \\
&\geqslant d_{k-1}^{\mathrm{T}} d_{k-1}.
\end{aligned} \tag{8}
$$

It follows from (4) that

$$\|d_k\| \geqslant \big(1 - \tfrac{1}{4t}\big) \|g_k\|,$$

then

$$\alpha_k \leqslant \frac{\tau \|g_k\|}{\|d_k\|} \leqslant \frac{\tau}{1 - \frac{1}{4t}} \triangleq q.$$

By (2), (7), and (8), it holds that

$$
\begin{aligned}
\|d_k\| &\leqslant \|g_k\| + \Big(\tfrac{|g_k^{\mathrm{T}} y_{k-1}|}{d_{k-1}^{\mathrm{T}} \gamma_{k-1}} + \tfrac{t\|y_{k-1}\|^2 |g_k^{\mathrm{T}} d_{k-1}|}{(d_{k-1}^{\mathrm{T}} \gamma_{k-1})^2} \Big) \|d_{k-1}\| \\
&\leqslant \|g_k\| + \\
&\quad \big(\tfrac{\alpha_{k-1} L \|g_k\| \|d_{k-1}\|}{\|d_{k-1}\|^2} + \tfrac{t \alpha_{k-1}^2 L^2 \|d_{k-1}\|^2 \|g_k\| \|d_{k-1}\|}{\|d_{k-1}\|^4} \big) \|d_{k-1}\| \\
&\leqslant (1 + \alpha_{k-1} L + t \alpha_{k-1}^2 L^2) \|g_k\| \\
&\leqslant (1 + qL + tq^2 L^2) \|g_k\|.
\end{aligned}
$$

Since the level set Γ is bounded and g satisfies the Lipschitz condition, then the sequence $\{g_k\}$ is bounded. Hence the sequence $\{d_k\}$ is bounded. The remaining proof is referred to the second part of Theorem 3.1 in Ref. 16. □

As pointed out by Cai et al[14], if φ_α is convex, continuously differentiable, and first-order Lipschitz continuous, then $\mathscr{F}_\alpha$ is continuously differentiable and first-order Lipschitz continuous, i.e., $\nabla \mathscr{F}_\alpha$ is Lipschitz continuous. Thus if we select the appropriate $\varphi_\alpha(t)$, the problem (1) is solved by the MCG method. Applying the theorem to the function $\mathscr{F}_\alpha$, from the result in Ref. 14 we have the following global convergence results.

Theorem 2 *Assume that $\varphi_\alpha(t)$ is even, convex, continuously differentiable, and strictly increasing in $|t|$. Let the sequence $\{u_k\}$, $k \geqslant 0$, be generated by the MCG method applied to $\mathscr{F}_\alpha$. If $\varphi_\alpha(t)$ is first-order Lipschitz continuous, then there exists a subsequence of $\{u_k\}$ converging to a global minimizer u^* of $\mathscr{F}_\alpha$.*

NUMERICAL EXPERIMENTS

In this section, some experimental results are given to show the performance of the MCG method in the context of impulse noise removal impulse noise. We select the salt-and-pepper impulse noise. The test images are 256×256 and 512×512 grey level images. We select the parameters in the MCG

Table 1 Performance of the MCG and SP methods for restoring the images with noise level 50%.

Image	Method	PSNR	Time (s)	Iter.†
Lena	MCG (t=1)	30.0952	2.7708	44.8
(256 × 256)	MCG (t=2)	30.1138	2.6402	44.4
	SP	29.9878	3.5500	44.6
Cameraman	MCG (t=1)	27.5288	3.7082	60.2
(256 × 256)	MCG (t=2)	27.5246	3.4110	59.4
	SP	27.5182	4.8000	61.0
Barbara	MCG (t=1)	26.4259	11.8686	41.4
(512 × 512)	MCG (t=2)	26.4213	10.9560	40.8
	SP	26.4032	16.4437	41.0
Banoon	MCG (t=1)	24.6085	12.0921	44.6
(512 × 512)	MCG (t=2)	24.5660	12.0572	44.6
	SP	24.5659	17.7937	45.0

† Iter.,the total number of iterations for the whole impulse noise removal process.

Table 2 Performance of the MCG and SP methods for restoring the images with noise level 70%.

Image	Method	PSNR	Time (s)	Iter.
Lena	MCG (t=1)	27.0811	4.0648	63.0
(256 × 256)	MCG (t=2)	27.1862	4.0128	63.2
	SP	27.1024	5.2313	64.8
Cameraman	MCG (t=1)	24.6753	5.2796	81.4
(256 × 256)	MCG (t=2)	24.7377	5.0870	83.2
	SP	24.6262	6.7031	81.6
Barbara	MCG (t=1)	24.5696	17.2668	54.6
(512 × 512)	MCG (t=2)	24.5549	16.1910	53.6
	SP	24.5489	26.9312	63.4
Banoon	MCG (t=1)	22.3741	17.7564	59.4
(512 × 512)	MCG (t=2)	22.3729	17.9706	60.4
	SP	22.3724	25.9437	59.8

Table 3 Performance of the MCG and SP methods for restoring the images with noise level 90%.

Image	Method	PSNR	Time (s)	Iter.
Lena	MCG (t=1)	22.7776	10.8462	119.8
(256 × 256)	MCG (t=2)	22.7800	11.3956	135.2
	SP	22.7748	12.8906	123.0
Cameraman	MCG (t=1)	21.2164	12.6844	152.2
(256 × 256)	MCG (t=2)	21.2021	11.7630	150.0
	SP	21.1851	16.7437	147.4
Barbara	MCG (t=1)	22.5561	46.8472	105.8
(512 × 512)	MCG (t=2)	22.5594	44.0846	107.6
	SP	22.5512	62.1031	106.0
Banoon	MCG (t=1)	20.3252	42.4632	104.6
(512 × 512)	MCG (t=2)	20.3118	43.7204	106.0
	SP	20.3075	59.9218	107.8

method as follows: $\delta = 0.5$, $\tau = \sqrt{99}/8$, $\rho = 0.5$, $t = 1, 2$, $\varepsilon = 10^{-4}$. In addition, we compare the performance of the MCG method with that of the SP method presented in Ref. 16. The SP method performs better than the Polak-Ribière conjugate gradient method which is the most effective among the methods considered in the numerical comparison in Ref. 14.

To solve the problem (1) by the MCG method, we select the Huber function[18] as the edge-preserving function, which is convex and first-order Lipschitz continuous. Its definition is as follows

$$\varphi_\alpha(t) = \begin{cases} \frac{t^2}{2\alpha}, & |t| \leqslant \alpha, \\ |t| - \frac{\alpha}{2}, & |t| > \alpha, \end{cases}$$

with $\alpha = 10$. The stopping criterion is

$$\frac{|\mathcal{F}(u_k) - \mathcal{F}(u_{k-1})|}{|\mathcal{F}(u_k)|} \leqslant 10^{-4}.$$

To assess the restoration performance, we use the peak signal-to-noise ratio PSNR[19], i.e.,

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

where $u^*_{i,j}$ and $x_{i,j}$ denote the pixel values of the restored image and the original image, respectively. To test the performance of both methods more fairly, the experiments are repeated for 5 different noise samples of each image, and the average of the 5 results is listed in Tables 1–3.

To a certain extent, Tables 1–3 indicate that, for the PSNR, the MCG method performs better than the SP method for most of the test images with different noise levels. Although the numbers of iterations of both methods are very similar, the MCG method is faster than the SP method, leading to about 25% saving in time. Fig. 1 lists the results restored by the methods from the corrupted images with noise level 70%. In a word, these results show that the MCG method can effectively restore the corrupted images with different noise levels.

CONCLUSIONS

In this study, we propose a modified conjugate gradient method based on the HZ method. Remarkableness of the proposed method is that its global convergence is established under the Armijo line search without assuming that the objective is strongly convex. We use the proposed method to remove the impulse noise in the two-phase method, and obtain efficient experiment results.

Fig. 1 Restoration of the images Lena, Cameraman, Barbara, and Banoon via the MCG(t=1), MCG(t=2), and SP methods. From left to right: the corrected image, the restorations obtained by the MCG(t=1), MCG(t=2), and SP methods, respectively.

REFERENCES

1. Hwang H, Haddad RA (1995) Adaptive median filters: new algorithms and results. *IEEE Trans Image Process* **4**, 499–502.
2. Huang TS, Yang GJ, Tang GY (1979) A fast two-dimensional median filtering algorithm. *IEEE Trans Acoust Speech Signal Process* **1**, 13–8.
3. Astola J, Kuosmanen P (1997) *Fundamentals of Nonlinear Digital Filtering*, CRC, Boca Raton.
4. Nikolova M (2004) A variational approach to remove outliers and impulse noise. *J Math Imaging Vision* **20**, 99–120.
5. Kimiaei M, Rostami M (2016) Impulse noise removal based on new hybrid conjugate gradient approach. *Kybernetika* **52**, 791–823.
6. Veerakumar T, Jagannath RPK, Subudhi BN, Esakkirajan S (2017) Impulse noise removal using adaptive radial basis function interpolation. *Circ Syst Signal Process* **36**, 1192–223.
7. Veerakumar T, Subudhi BN, Esakkirajan S, Pradhan PK (2017) Context model based edge preservation filter for impulse noise removal. *Expert Syst Appl* **88**, 29–44.
8. Chen QQ, Hung MH, Zou FM (2017) Effective and adaptive algorithm for pepper-and-salt noise removal. *IET Image Process* **11**, 709–16.
9. Gupta V, Chaurasia V, Shandilya M (2015) Random-valued impulse noise removal using adaptive dual threshold median filter. *J Vis Commun Image R* **26**, 296–304.
10. Esakkirajan S, Veerakumar T, Subramanyam AN, PremChand CH (2011) Removal of high density salt and pepper noise through modified decision based unsymmetric trimmed median filter. *IEEE Signal Process Lett* **18**, 287–90.
11. Chan RH, Ho CW, Nikolova M (2005) Salt-and-pepper noise removal by median-type noise detector and edge-preserving regularization. *IEEE Trans Image Process* **14**, 1479–85.
12. Chen T, Wu HR (2001) Adaptive impulse detection using center-weighted median filters. *IEEE Signal Process Lett* **8**, 1–3.
13. Chan RH, Ho CW, Leung CH, Nikolova M (2005) Minimization of detail-preserving regularization func-

tional by Newton's method with continuation. In: *Proceedings of IEEE International Conference on Image Processing*, Genova, Italy.
14. Cai JF, Chan RH, Morini B (2007) Minimization of an edge-preserving regularization functional by conjugate gradient types methods. In: Tai XC, Lie KA, Chan TF, Osher S (eds) *Image Processing Based on Partial Differential Equations*, Spring, Berlin, pp 109–22.
15. Yu GH, Huang JH, Zhou Y (2010) A descent spectral conjugate gradient method for impulse noise removal. *Appl Math Lett* **23**, 555–60.
16. Liu JK, Li SJ (2015) Spectral gradient method for impulse noise removal. *Optim Lett* **9**, 1341–51.
17. Hager WW, Zhang H (2005) A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM J Optim* **16**, 170–92.
18. Huber PJ (1973) Robust regression: Asymptotics, conjectures, and Monte Carlo. *Ann Statist* **1**, 799–821.
19. Bovik A (2000) *Handbook of Image and Video Processing*, Academic, New York.