A New Adaptive Regularization Parameter Selection Based on Expected Patch Log Likelihood

Jianwei Zhang[1], Ze Qin[1] and Shunfeng Wang[1, *]

Abstract: Digital images have been applied to various areas such as evidence in courts. However, it always suffers from noise by criminals. This type of computer network security has become a hot issue that can't be ignored. In this paper, we focus on noise removal so as to provide guarantees for computer network security. Firstly, we introduce a well-known denoising method called Expected Patch Log Likelihood (EPLL) with Gaussian Mixture Model as its prior. This method achieves exciting results in noise removal. However, there remain problems to be solved such as preserving the edge and meaningful details in image denoising, cause it considers a constant as regularization parameter so that we denoise with the same strength on the whole image. This leads to a problem that edges and meaningful details may be oversmoothed. Under the consideration of preserving edges of the image, we introduce a new adaptive parameter selection based on EPLL by the use of the image gradient and variance, which varies with different regions of the image. Moreover, we add a gradient fidelity term to relieve staircase effect and preserve more details. The experiment shows that our proposed method proves the effectiveness not only in vision but also on quantitative evaluation.

Keywords: Computer network security, image denoising, EPLL, adaptive parameter, edges.

1 Introduction

Computer network security is always a hot topic because it is closely related to our daily life. However, it often suffers from attack. For example, as we all know digital images can be used as a powerful evidence in juridical practice, but criminals may corrupt images to disrupt the line of sight. For solving this, many experts focus on the source of images because it plays an important role in image forensics practice. Investigators may track suspects form images that taken by particular cameras. The most popular image forensic approach is mainly based on sensor pattern noise [Lucas, Fridrich and Goljian (2013)]. The key to this approach is extracting sensor pattern noise by advanced denoising filters. Therefore noise removal becomes a hot issue in computer forensics, and then much effort has been paid to image denoising techniques to obtain the high quality image.

A large number of methods have been proposed on image restoration in the past few

[1] School of Math and Statistics, Nanjing University of Information Science and Technology, Nanjing, 210044, China.

[*] Corresponding Author: Shunfeng Wang. Email: wsfnuist@126.com.

decades. Sparse representation is rather well-known among them [Dong, Zhang, Shi et al. (2011); Elad and Aharon (2006)]. The main idea of this theory is to consider image patches as the linear combination of some atoms based on a dictionary. The overcomplete dictionary can be ready-made, or it can be trained from image patches. Then low-rank approximation methods show amazing results on denoising [Dong, Shi, Li et al. (2014); Dong, Shi and Li (2013)]. It can be divided into two categories: the nuclear norm minimization methods [Cai, Candes and Shen (2010)], and the low rank matrix factorization approaches. The total variation (TV) [Beck and Teboulle (2009); Zheng, Jeon, Zhang et al. (2015); Zhang, Yu, Zheng et al. (2016); Zheng, Ma, Yu et al. (2017)] has always been active in image denoising. Building a proper image model as prior is of great significance in image denoising. So mixture models attract attention in image denoising due to its robustness, especially Gaussian Mixture Model (GMM). Then many meaningful explorations were devoted to learn image priors by GMM [Wang and Morel (2012); Yu, Sapiro and Mallat (2012); Zoran and Weiss (2011); Zhang, Liu, Li et al. (2017)]. Inspired by this, some further studies have been proposed [Zheng, Zhou, Jeon et al. (2017)]. These methods have already successfully prove their reliability.

As a major feature of the image, edges are full of meaningful structure information. Preserving edges occupies an important position in image denoising. Therefore, many researches on edge-preserving have been applied to noise removal. Anisotropic diffusion is one of the most efficient algorithm. In Perona et al. [Perona and Malik (2002)], the first pioneering model of anisotropic diffusion came into the public's view. The original Perona-Malik diffusion model mainly considered on the gradient information of the image. The basic idea behind the method is that gradient is an ideal indication of the edge. In detail, the classical P-M model smooths strongly at the homogeneous region which always has low gradient and stops diffusion at the inter-region edge that has large gradient. The number of iterations is the key to the success of the denoising result since too much iterations will lead to the loss of edges. Inspired by this, many further studies have been proposed due to its great success [Chao and Tsai (2010); Ma, Shen, Zhang et al. (2017); Masiseli and Gao (2016); Xu, Jia, Shi et al. (2016)], now the anisotropic diffusion approach has become a hot topic in image restoration [Wang, Guo, Chen et al. (2013)], image segmentation, texture segmentation, image enhancement [Giboa, Sochen and Zeevi (2004); Dang, Gao, Wang et al. (2015)], image smoothing [Lions, Morel and Coll (1992)] and defect detection [Malarvel, Sethumadhavan, Bhagi et al. (2017)]. However, there still remain tiny problems such as the huge cost of iteration and staircase effect. For solving the former, Tebini et al. [Tebini, Mbarki, Seddik et al. (2016)] proposed a fast and efficient method based on the tangent sigmoid model to obtain a higher speed of convergence. For solving the latter, Guo et al. [Guo, Sun, Zhang et al. (2012)] proposed an adaptive P-M diffusion which combined the original P-M model heat equation. Edge-preserving methods aim to remove noise effectively and preserve inter-region edges at the same time. In this paper, we focus on the edge-preserving, and propose a new method of adaptive regularization parameter selection. Moreover, we add a gradient-fidelity term to relieve the staircase effect and preserve more details of image.

In this paper, we focus on the noise removal method used in computer forensics. Good quality denoising methods enable experts to extract sensor pattern noise better and keep the accuracy of camera forensics. The noise removal method aims to enhance the security of computer identification.

This paper is organized as follows: Section 2 reviews the GMM model, and then introduces original EPLL algorithm briefly. Section 3 introduces the proposed method. In Section 4, we show experiment results as the proof of effectiveness of the proposed method. Section 5 makes the conclusion and summarizes the paper.

2 Original EPLL model

2.1 Normal denoising model

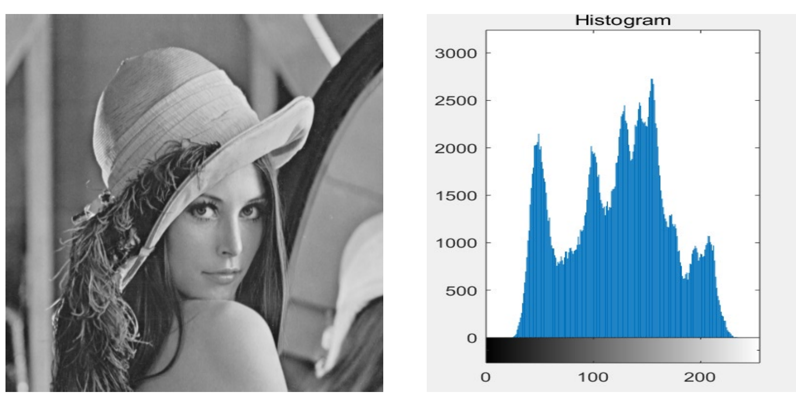
Assuming a noisy image $u_0 = u + v$ corrupted by additive white Gaussian noise. Our purpose is to achieve a pure image $\boldsymbol{u}$ from the observation. In this paper, we consider it as a Maximum a Posteriori (MAP) problem. MAP approach is based on Bayesian theory that maximizes the posteriori probability to obtain the clean image:

$$\arg\max_u p(u_0 \mid u)\, p(u) = \arg\min_u \left\{ \frac{\lambda}{2} \|u_0 - u\|^2 - \log p(u) \right\} \tag{1}$$

The first term is a data fidelity term, and the second term is the log of the prior. First of all, we need to learn prior knowledge. We consider the distribution $p(u)$ as the prior model which is independent of the data observation. Obviously, the distribution $p(u)$ directly affects the result, so how to build the prior model has become a hot issue. As we all know, image is a type of high-dimensional data with complex distribution, so directly learning the whole image prior is a giant challenge. To solve this, we can extract overlapping image patches with low dimension from the whole image, and naturally our goal is to make every patch is likely under our prior in statistics.

2.2 Gaussian mixture model (GMM)

Usually we can assume that one patch can be described by a specific distribution function. However, as is mentioned above, image is consisted of numerous independent patches that we can hardly describe it just by a single distribution.



a b

Figure 1: (a) Lena image (b) Gray distribution histogram of Lena image

As is shown in the Fig. 1, (b) represents gray distribution histogram of (a), abscissa represents the gray value, and ordinate represents the frequency of the gray value. Several peaks in (b) means one image can be considered as the superposition of several different Gaussian distributions, then naturally we call it Gaussian Mixture Model (GMM). GMM is always considered as the image prior instead of a single gaussian distribution.

In this paper, we train the GMM model by a set of clean natural image patches $D=\{a_1,a_2,\cdots a_n\}$ (in vectorized form). For a given patch a_i, we define the GMM distribution as the following:

$$p(a_i)=\sum_{i=1}^{K}\pi_k N\left(a_i\mid\mu_k,\Sigma_k\right) \tag{2}$$

where

$$N\left(a_i\mid\mu_k,\Sigma_k\right)=\left(2\pi\right)^{-\frac{d}{2}}\left|\Sigma_k\right|^{-\frac{1}{2}}\exp\left(-\frac{\left(a_i-\mu_k\right)^T\Sigma^{-1}\left(a_i-\mu_k\right)}{2}\right) \tag{3}$$

where K is the number of mixture components, π_k is the mixing weight for each mixture component, μ_k and Σ_k denote the corresponding mean vector and covariance matrix. Moreover, π_k denotes the priori probability, and $\sum_{k=1}^{K}\pi_k=1$. For notational simplicity, let $\Theta=\{\mu_k,\Sigma_k,\pi_k\}_{k=1}^{K}$ denote these parameters. Θ is learned by Expectation Maximization algorithm (EM).

In the E-step, we calculate the posterior probability for the component k as the following:

$$p\left(k\mid a_i,\Theta\right)=\frac{\pi_k N\left(a_i\mid\mu_k,\Sigma_k\right)}{\sum_{k=1}^{K}\pi_k N\left(a_i\mid u_k,\Sigma_k\right)} \tag{4}$$

Then, for all patches, we get the likelihood function as the following:

$$LL\left(D\right)=\sum_{i=1}^{N}\ln\left(\sum_{i=1}^{K}\pi_k N\left(a_i\mid\mu_k,\Sigma_k\right)\right) \tag{5}$$

In the M-step, to obtain parameters Θ, let $\frac{\partial LL\left(D\right)}{\partial\mu_k}=0$, then we get:

$$\sum_{i=1}^{N}\frac{\pi_k N\left(a_i\mid\mu_k,\Sigma_k\right)}{\sum_{k=1}^{K}\pi_k N\left(a_i\mid\mu_k,\Sigma_k\right)}\left(a_i-\mu_k\right)=0 \tag{6}$$

$$\mu_k=\frac{\sum_{i=1}^{N}p\left(k\mid a_i,\Theta\right)a_i}{\sum_{i=1}^{N}p\left(k\mid a_i,\Theta\right)} \tag{7}$$

Eq. (7) means the mean vector for each component can be estimated by weighted average method, the weight is the posterior probability.

Similarly, let $\frac{\partial LL(D)}{\partial \Sigma_k} = 0$, then we get:

$$\Sigma_k = \frac{\sum_{i=1}^{N} p(k \mid a_i, \Theta)(a_i - \mu_k)(a_i - \mu_k)^T}{\sum_{i=1}^{N} p(k \mid a_i, \Theta)} \tag{8}$$

For mixing weights π_k, considering the form of Lagrange, leading to the following:

$$LL(D) + \gamma\left(\sum_{k=1}^{K} \pi_k - 1\right) \tag{9}$$

subject to $\pi_k \geq 0$, $\sum_{k=1}^{K} \pi_k = 1$

where γ is the Lagrange multiplier.

Let the derivative of the Eq. (9) be $\pi_k = 0$, then

$$\sum_{i=1}^{N} \frac{N(a_i \mid \mu_k, \Sigma_k)}{\sum_{k=1}^{K} \pi_k N(a_i \mid \mu_k, \Sigma_k)} + \gamma = 0 \tag{10}$$

$$\pi_k = \frac{1}{N} \sum_{i=1}^{N} p(k \mid a_i, \Theta) \tag{11}$$

Eq. (11) means the mixing weight for each component is determined by the averaging posterior probability.

2.3 Expected patch log likelihood

EPLL [Zoran and Weiss (2011)] is an image restoration method that builds a model from patches to the whole image. The fundamental idea of EPLL is to make every patch have the highest log likelihood probability under our prior. EPLL is defined as:

$$EPLL(u) = \sum_{i} \log p(R_i u) \tag{12}$$

where R_i is an operator that extracts the patch from the whole image at the *i*-th position, obviously $R_i u$ is the extracted patch we need, λ denotes the regularization parameter and log means the likelihood of the *i*-th patch under the prior.

Our purpose is to obtain a pure image by minimizing the cost function as the following:

$$f(u \mid u_0) = \frac{\lambda}{2}\|u - u_0\|^2 - EPLL(u) \tag{13}$$

Similar to Eq. (1), Eq. (13) has two terms: the data fidelity term and the regularization term. We should point out that *EPLL (u)* denotes the total log probabilities of all overlapping patches, rather than the whole image. Overlapping patches means each pixel will appear several times in different patches, so each patch will be denoised independently and each pixel will be average according to its frequency.

3 Proposed method

Regularization parameter affects the quality of the restored image directly. As is shown in Chao et al. [Chao and Tsai (2010)], when it comes to image denoising, the adaptive regularization parameter always plays a central role. How to maintain a balance between preserving edges of the image and denoising remains a big problem for us to solve. So in this paper, a method with adaptive selection of regularization parameter for a better restored image is proposed as the following:

$$\left\{\frac{\lambda(x,y)}{2}\left\|u-u_0\right\|^2+\frac{\alpha(x,y)}{2}\left\|\nabla u-\nabla\left(G_\sigma * u_0\right)\right\|^2-\sum_i \log p\left(R_i u\right)\right\} \tag{14}$$

where the second term is the gradient fidelity term, ∇u denotes the image gradient, G_σ is a Gaussian filter operator, $\lambda(x,y)$ and $\alpha(x,y)$ are functions represents the parameter at the each pixel which related to the gradient of the image. They can be expressed as the following:

$$\lambda(x,y)=k_1\cdot\frac{1+\tanh\left(-\frac{\left|\nabla u(x,y)\right|\cdot\sigma_N^2(x,y)}{k_2}+\delta\right)}{1+\tanh(\delta)} \tag{15}$$

where k_1 and δ are constants, k_2 is the threshold value, is the gradient of pixel(x,y), and $\sigma_N^2(x,y)$ is the normalized variance at the pixel (x,y). We introduce the variance of the image because variance is less affected by noise compared to gradient. In order to keep the variance have the same weight as the gradient, normalize the variance by:

$$\sigma_N^2(x,y)=1+\frac{\sigma^2(x,y)-\min\sigma^2}{\max\sigma^2-\min\sigma^2}\cdot 254 \tag{16}$$

$$\alpha(x,y)=\frac{1}{1+\left(\frac{\nabla\left(G_\sigma * u_0(x,y)\right)}{k}\right)^{\eta(\nabla u(x,y))}} \tag{17}$$

where k_3 is still the threshold value, ∇u is the gradient of the image, $\eta(|\nabla u|)$ is set by：

$$\eta(\nabla u(x,y))=2+\frac{2}{1+0.5\cdot\left|\nabla u(x,y)^2\right|} \tag{18}$$

The regularization parameter function depends mainly on the gradient ∇u and variance of the image. $\lambda(x,y)$ is small when large gradient and variance are detected at edges of the image and $\alpha(x,y)$ is large so that edges will be preserved. $\lambda(x,y)$ is large when small gradient and variance are detected in the smooth regions so as to remove noise as much as possible. Meanwhile, $\alpha(x,y)$ is small, which aims to maintain the similar structure between the observation and the restored image.

Instead of solving the equation directly, we apply a method named "Half Quadratic Splitting algorithm" to the Eq. (14). A series of patches $\{z_i\}$ are introduced that each one refers to the corresponding overlapping patch $R_i u$ in the image, then we get the new cost function:

$$\min\left\{\frac{\lambda(x,y)}{2}\|u-u_0\|^2+\frac{\alpha(x,y)}{2}\|\nabla u-\nabla(G_\sigma * u_0)\|^2+\sum_i\left(\frac{\beta}{2}\|R_i u-z_i\|^2-\log p(z_i)\right)\right\} \tag{19}$$

where β is the penalty parameter, it is obvious that auxiliary variable $\{z_i\}$ equals to the patch $R_i u$ when β tends to be infinite.

For minimizing Eq. (19), firstly, we choose the Gaussian component that has the highest conditional mixing weight $k_{\max}$ for each patch $R_i u$, then optimize Eq. (19) by updating z_i and u alternately:

$$z_i^{n+1}=\left(\Sigma_{j\max}+\frac{1}{\beta}I\right)^{-1}\cdot\left(R_i u^n\Sigma_{j\max}+\frac{1}{\beta}u_{j\max}I\right) \tag{20}$$

$$\begin{aligned}u^{n+1}=u^n+\Delta t\Big[&\lambda(x,y)(u_0-u^n)-\sum_i\beta R_i^T(R_i u^n-z_i^n)\\&+\alpha(x,y)(u_{xx}^n+u_{yy}^n)-\alpha(x,y)\big(G_\sigma*(u_{0xx}+u_{0yy})\big)\Big]\end{aligned} \tag{21}$$

where Δt denotes the time step and I is the unit matrix.

Repeat such process for few iterations, usually 4 or 5. At each iteration, we fix the value of β. There are two normal ways for us to apply to the choice of β values. One option is to optimize the value by a series of training sets, the other way is to try to estimate the intensity of noise from the current image, then set β.

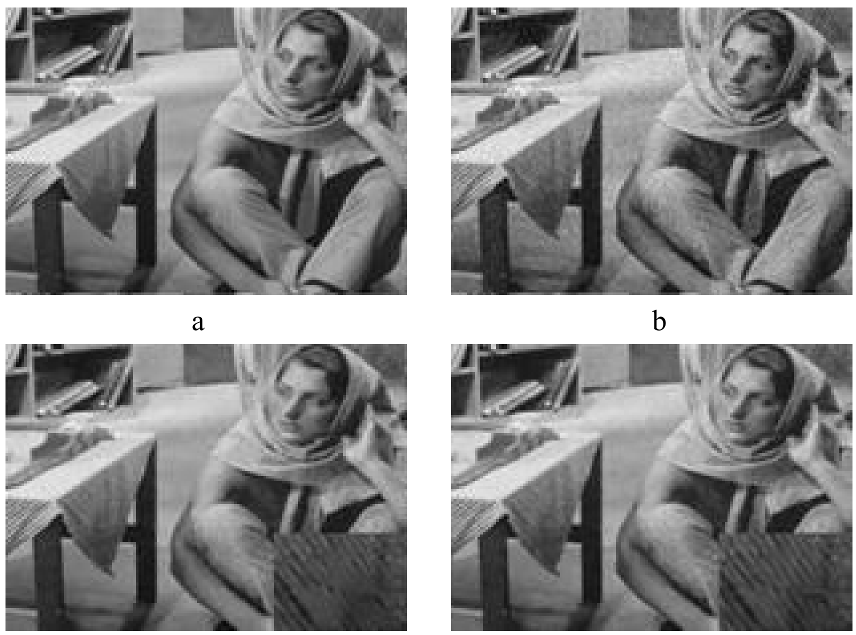
Generally speaking, the proposed method can be implemented as the following:

Table 1: Proposed algorithm

Input: corrupted image u_0, penalty parameter β, the time step Δt, regularization parameter functions $\lambda(x,y)$, $\alpha(x,y)$
Step Choose the most likely Gaussian mixing weights for each patch $R_i u$;
Calculate z_i^{n+1} using Eq. (20);
Pre-estimate image u^{n+1} by Eq. (21);
Repeat above steps for 4-5 times.

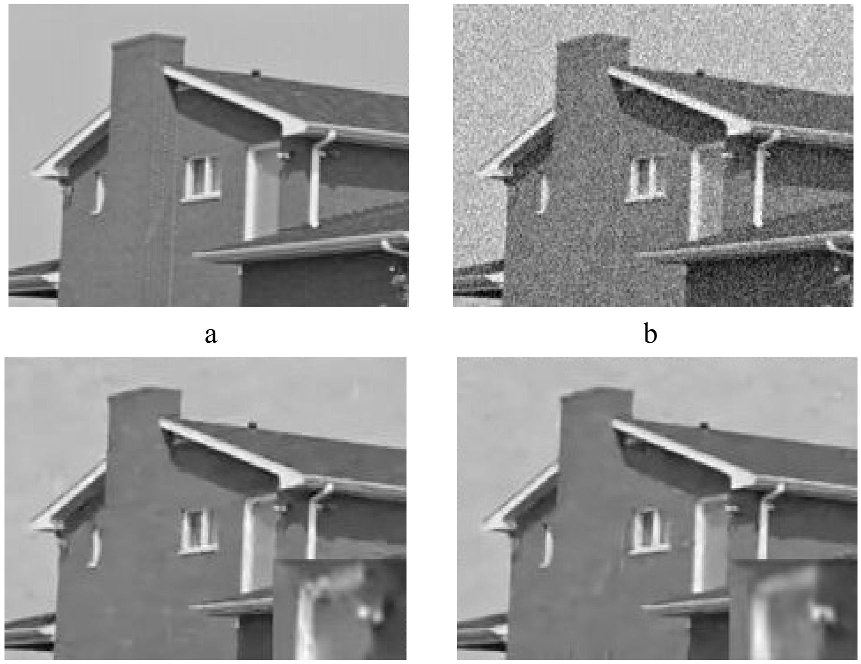
4 Implementation and experiment results

The set of 2×10^6 image patches used in our experiment are sampled from the Berkeley Segmentation Database (BSDS300). The GMM is trained with 200 mixture components by those patches. Adding all the patches with Gaussian noise of zero mean and a standard deviation $\sigma = 25$. Parameters in our proposed method are set as the following: the image patch size $\sqrt{L} = 8$, the penalty parameter $\beta = 1/\sigma^2 [1,2,3,4,5]$, $\Delta t = 0.0042$, the constants $k_1 = \frac{1}{\sigma^2}$, $k_2 = 20$, $k_3 = 70$, $\delta = 2$. The results are shown as the following:



c d

Figure 2: Denoising results on 'Barbara' image with a noise standard deviation σ=25. (a) Original image (b) Noisy image, (c) EPLL result (d), Proposed method



a b

c d

Figure 3. Denoising results on 'House' image with noise standard deviation σ=25. (a) Original image (b) Noisy image, (c) EPLL result (d), Proposed method

Table 2: PSNR results for test images with noise standard deviation σ=25

	EPLL	Proposed method
Barbara	28.55	28.72
House	32.15	32.41
Man	29.57	29.71
Hill	29.58	29.74
Lena	31.67	31.86

Table 3: PSNR results for test images with noise standard deviation σ=40

	EPLL	Proposed method
Barbara	26.01	29.29
House	29.93	30.21
Man	27.57	27.81
Hill	27.75	27.99
Lena	29.41	29.68

As a comparison, our proposed method preserves edges better and keeps more details of the image. As in shown in Figs. 2 and 3, the original EPLL method may denoise well, but it also leads to a problem that the loss of edges and details in some regions. Fig. 2(c) shows that details such textures are fuzzy in EPLL result, while clearer textures can be seen in proposed method. The gradient fidelity term proves its validity in preserving image details. Fig. 3(c) shows that the edge of house is barely visible in original EPLL result, while in Fig. 3(d) the edge is smoother and more real. This probably due to the fact that the new adaptive parameter can take a small number at edge areas so as to preserve edges better.

Tabs. 2 and 3 show numerical results between two methods on different images at different noise deviation. At both low and high noise deviation, various images can obtain higher PSNR by proposed method. We can see that the proposed method can achieve very competitive denoising performance. It is obvious that our method is superior to the original EPLL method.

5 Conclusions

Image is a carrier of abundant information, so directly learning image priors is a big challenge for us. GMM has proven its power in various areas due to its robustness, so image denoising based on GMM has become more and more active in the past decades. The EPLL algorithm is such a method that builds a model from the extracted patches to the whole image based on GMM. Classical P-M model has achieved satisfying results, but there is still weakness that it focuses only on the gradient. We should realize that details in image may involve variation information, therefore we should make full use of variation along with image gradient. In this paper, we construct an adaptive regularization parameter function based on the image gradient and variance. In addition, we add a gradient fidelity term to maintain the similar structure between the degraded image and the restored one. Our method performs well in image denoising, especially in edge and detail preserving, and shows an obvious advance compared with the original EPLL algorithm.

There is still much room for improvement. In the process of actual application, data tends to be symmetrical distribution from a statistical standpoint. GMM can be treated as deviation special distribution of the gaussian mixture distribution. Using a skew gaussian distribution as prior obviously will have stronger generalization ability. Moreover, we can consider the geometric characteristics of the image, so that the prior information dictionary more perfect and more accurate.

Acknowledgement: This paper is partly supported by the National Natural Science Foundation of China (GRANT No. 61672293).

References

Beck, A.; Teboulle, M. (2009): Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2419-2434.

Cai, J. F.; Candes, E. J.; Shen, Z. W. (2010): A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956-1982.

Chao, S. M.; Tsai, D. M. (2010): An improved anisotropic diffusion model for detail- and edge-preserving smoothing. *Pattern Recognition Letters*, vol. 31, no. 13, pp. 2012-2023.

Dang, C. Y.; Gao, J. M.; Wang, Z.; Chen, F. M.; Xiao, Y. L. (2015): Multi-step radiographic image enhancement conforming to weld defect segmentation. *IET Image Processing*, vol. 9, no. 11, pp. 943-950.

Dong, W. S.; Shi, G. M.; Li, X. (2013): Nonlocal image restoration with bilateral variance estimation: a low-rank approach. *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 700-711.

Dong, W. S.; Shi, G. M.; Li, X.; Ma, Y.; Huang, F. (2014): Compressive sensing via nonlocal low-rank regularization. *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3618-3632.

Dong, W. S.; Zhang, L.; Shi, G.; Li, X. (2011): Nonlocally centralized sparse representation for image restoration. *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1259-1266.

Elad, M.; Aharon, M. (2006): Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, vol. 15, no. 15, pp. 3736-3745.

Giboa, G.; Sochen, N.; Zeevi, Y. Y. (2004): Image enhancement and denoising by complex diffusion processes. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 26, no. 8, pp. 10201036.

Guo, Z. C.; Sun, J. B.; Zhang, D. Z.; Wu, B. Y. (2012): Adaptive perona-malik model based on the variable exponent for image denoising. *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 958-967.

Lions, P. L.; Morel, J. M.; Coll, T. (1992): Image selective smoothing and edge detection by nonlinear diffusion. *Siam Journal on Numerical Analysis*, vol. 29, no. 1, pp. 182-193.

Ma, X. S.; Shen, H. F.; Zhang, L. P.; Yang, J.; Zhang, H. Y. (2017): Adaptive anisotropic diffusion method for polarimetric SAR speckle filtering. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, vol. 8, no. 3, pp. 1041-1050.

Malarvel, M.; Sethumadhavan, G.; Bhagi, P. C. R. (2017): Anisotropic diffusion based denoising on X-radiography images to detect weld defects. *Digital Signal Processing*, vol. 68, pp. 112-126.

Masiseli, B. J.; Gao, H. J. (2016): Robust edge detector based on anisotropic diffusion-driven process. *Information Processing Letters*, vol. 116, no. 5, pp. 373-378.

Perona, P.; Malik, J. (2002): Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 12, no. 7, pp. 629-639.

Tebini, S.; Mbarki, Z.; Seddik, H.; Braiek, E. B. (2016): Rapid and efficient image restoration technique based on new adaptive anisotropic diffusion function. *Digital Signal Processing*, vol. 48, no. C, pp. 201-215.

Wang, Y. Q.; Guo, J. C.; Chen, W. F.; Zhang. W. X. (2013): Image denoising using modified Perona-Malik model based on directional Laplacian. *Signal Processing*, vol. 93, no. 9, pp. 2548-2558.

Wang, Y. Q.; Morel, J. M. (2012): SURE guided Gaussian mixture image denoising. *Siam Journal on Imaging Sciences*, vol. 6, no. 2, pp. 999-1034.

Xu, J. T.; Jia, Y. Y.; Shi, Z. F.; Pang, K. (2016): An improved anisotropic diffusion filter with semi-adaptive threshold for edge preservation. *Signal Processing*, vol. 119, pp. 80-91.

Yu, G. S.; Sapiro, G.; Mallat, S. G. (2012): Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity. *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2481-2499.

Zhang, J. W.; Liu, J.; Li, T.; Zheng, Y. H.; Wang, J. (2017): Gaussian mixture model learning based image denoising method with adaptive regularization parameters. *Multimedia Tools and Applications*, vol. 76, no. 9, pp. 11471-11483.

Zhang, J. W.; Yu, Q. Q.; Zheng, Y. H.; Zhang, H.; Wu, J. (2016): Regularization parameter selection for TV image denoising using spatially adaptive local spectral response. *Journal of Internet Technology*, vol. 17, no. 6, pp. 1117-1124.

Zheng, Y. H.; Jeon, B.; Zhang, J. W.; Chen, Y. J. (2015): Adaptively determining regularization parameters in non-local total variation for image denoising. *Electronics Letters*, vol. 5, no. 2, pp. 144-145.

Zheng, Y. H.; Ma, K.; Yu, Q. Q.; Zhang, J. W.; Wang, J. (2017): Regularization parameter selection for total variation model based on local spectral response. *Journal of Information Processing Systems*, vol. 13, no. 5, pp. 1168-1182.

Zheng, Y. H.; Zhou, X. Z.; Jeon, B.; Shen, J.; Zhang, H. (2017): Multi-scale patch prior learning for image denoising using Student's-t mixture model. *Journal of Internet Technology*, vol. 18, no. 7, pp. 1553-1560.

Zoran, D.; Weiss, Y. (2011): From learning models of natural image patches to whole image restoration. *IEEE International Conference on Computer Vision*, pp. 479-486.