Multidim Syst Sign Process
https://doi.org/10.1007/s11045-018-0577-1

A stochastic image denoising method based on adaptive patch-size

Liang Luo[1] · **Zhi-qin Zhao**[2] · **Xiao-ping Li**[1] · **Xiang-chu Feng**[3]

Received: 3 July 2017 / Revised: 3 April 2018 / Accepted: 12 April 2018

Abstract A new stochastic nonlocal denoising method based on adaptive patch-size is presented. The quality of restored image is improved by choosing the optimal nonlocal similar patch-size for each site of image individually. The method contains two phase. The first phase is to search the similar patches base on adaptive patch-size. The second phase is to design the denoising algorithm by making use of similar image patches obtained in the first step. The multiple clusters of similar patches for each pixel point are searched by using Markov-chain Monte Carlo sampling many times. Following, we adjust the patch-size according to the consistency of multiple clusters. This processing is repeated until we obtain the optimal patch-size and corresponding optimal patch cluster. We get the estimation of noise-free patch cluster by employing modified two-directional non-local method. Furthermore, the denoised image is obtained by using the method of superposition approach. The theoretical analysis and simulation results show that the method is feasible and effective.

Keywords Image denoising · Adaptive patch size · Markov-Chain Monte Carlo method · Two-directional non-local approximation

1 Introduction

Image is one of the most important information form. At present, the image processing technology has been widely used in remote sensing image scene classification (Cheng et al. 2018), saliency detection (Yao et al. 2017; Zhang et al. 2017a), object detection (Cheng et al. 2016) and image classification (Zhang et al. 2016, 2017c), etc. In all application areas, the

✉ Liang Luo
luoliang775@163.com

1 Department of Mathematics, School of Science, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

2 School of Science, Xi'an Shiyou University, Xi'an 710065, China

3 School of Mathematics and Statistics, Xidian University, Xi'an 710071, China

Published online: 19 April 2018

images with high quality is of great significance for subsequent analysis and treatment. But due to the limitation of resolution of the acquisition equipment, acquisition environment, the noises of equipments and communication conditions, images are polluted by noise in the process of acquisition, transmission, and record, which result in radiometric distortion and greatly reduces the interpretability of image target information. For the purpose of analysis and using of the original image, it is necessary for us to remove noise the original image first of all.

As before, the denoising method based on pixels is adopted to removal noising, such as median filtering, bilateral filtering, Tv (Rudin et al. 1992), regularity method about anisotropic diffusion (Wang and Feng 2008). These methods focus on calculating similarities between pixels to estimate the original image. In recent years, some scholars present the denoising methods based on image block, which use the spatial redundancy information of image. The advantage of this denoising approach is making use of the local structures information in the smoothing process, which over the method based on pixels. For example nonlocal mean denoising method (NLM), Block-matching and 3D-filtering algorithm (BM3D). A large number of experimental results show that the method based on image block can be more effectively in removing noise and keeping the structure of the original image compared with denoising method based on pixels. From the original nonlocal average method to the latest BM3D method, its basic idea is to searching the image patches corresponding to noise image block by making use of self-similarity of image, and then remove the noise and keep the original image structures.

There are two main ways of searching similar patch: one way is from large-scale database. For example, in literature (Levin et al. 2012), Levin etal got similar patches from database. However, we note that this searching strategy is not a practical way, required several days of computation, which get more kicks than halfpence. Another searching approach is based on the single image: the reference patches are come from the observed image. For example non-local mean algorithm (Buades et al. 2005), Block Method of 3-Dimension (BM3D) method (Dabov et al. 2007). These methods search the similar patches in whole image. This approach is very effective while it has relatively high computational cost especially when the size of the image is large. Buades etal presented a modified one (Buades et al. 2005) which searches self-similar patches only in a search window rather than over the whole image. Although it decreases the searching cost, it prevents farther patches form being selected.

Literature (Wong et al. 2011) adopted a stochastic approach via Markov Chain Monte Carlo (MCMC) sampling. The method adopts the Gaussian as proposal distribution in spatial domain to draw candidate patches and use similarity as criterion to accept or reject the candidates as the similar patches. It obviously found that this strategy not only break through the limitation of the search window, but also has less amount of calculation compared to the global search method. MCMC is an statistical method which is widely used in recent years. the basic idea of this method is to constructed the markov chain with stationary distribution which consider as the posterior distribution of parameter needed to be estimated, and make targeted calculation by utilizing samples has played a significant role in statistics, econometrics, physics and computing science over the last two decades. Compared with the ergodic search approach, this stochastic searching method by using MCMC is more quickly, furthermore is faster than method proposed in literature (Dabov et al. 2007).

In general, the method based on patch is more focus on various natural image priors such as gradient based prior, nonlocal self-similarity prior, and sparsity prior in process of searching similar patches for noise removal. However, these prior tend to smooth the detailed image textures, degrading the image visual quality. Little is known about their dependence on patch size.

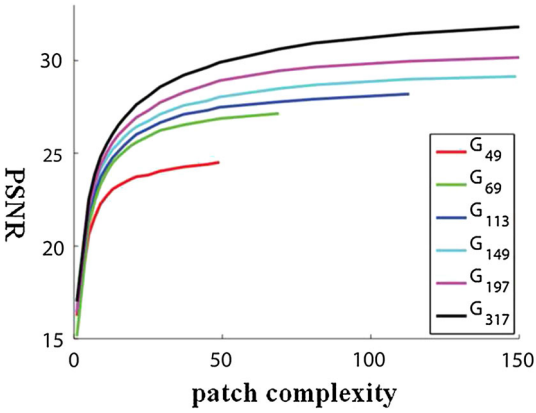

Fig. 1 The processing of searching similar patch based on adaptive patch size. For patch groups G_l of varying complexity, we present PSNR versus number of pixels d in window w_d, where $d = 1, 2, \ldots, l$. Higher curves correspond to smooth regions, which fatten at larger patch dimensions. Textured regions correspond to lower curves which not only run out of samples sooner, but also their curves fatten earlier

In fact, the similar patch size reflects the strength of smoothing in denoising process to some extent. Most of denoising method ignore the correlation between regional characteristics and similar block size, and consider that the patch size corresponding each pixel is fixed, namely taking the same smooth strength for high frequency detail and flat area. This lead to the over smooth or incompletely removal noise in implementations of denoising algorithm. So according to the different local characteristics, adaptively choosing appropriate smooth strength is one of the effective methods to improve denoising performance.

In other words, the idea of satisfaction strategy is imposing the strong smooth strength in flat regions, reducing smooth strength in detailed components, and at the same time giving consideration to both preserving the details of image and removing noise. The literature (Levin et al. 2012) analysis the relationship between the denoising gain and patch size requirements in a non parametric approach and present a law of diminishing return: when the patch size is increased, the difficulty of finding enough training data for an input noisy patch directly correlates with diminishing returns in denoising performance. That is, patch size directly influences the last effect of denoising (denoising error). When sample space is larger enough, increasing the patch-size indeed improves the performance of denoising algorithm. The experiment results (Fig. 1) demonstrate that patch based denoising can be improved mostly in flat areas and less in textured ones. It is worth stressing that sufficient sample is the premise condition for the results demonstrated in literature (Levin et al. 2012). Note that this premise condition is not practically applied in general algorithm due to its high computational complexity. As one coin has two sides, it also provide some practical and constructive guidelines on how we design a more practical image denoising method.

Through the another experiment, we found two phenomenon about the variation of patch size. On the one hand, the noise in the flat region is removed clean with relatively larger patch size while details in texture area is not kept better. On the other hand, when the patch size is small, there still have some noise leaving in the flat areas of image, but it can reduce noise effectively with a little loss of texture detail. The reason for this phenomenon is that statistical characteristics of smooth area is not estimated accurately which lead to noise still leave in the flat regions; For the texture region, structure and texture feature is obviously, so statistical characteristics is estimated accurately and maintain the image detail well.

Literature (Levin et al. 2012) also point out that in order to guarantee the accurate estimation of statistical character for the smooth regions and keep local structure components, the relatively larger patch size for the smooth regions is needed and relatively smaller patch size

is required for the texture regions when the similar patch is searched in a database with scale-fixed. In fact, texture regions are of complexity geometric structures, similar patch samples are dense enough and all statistic estimators provide consistent results for the small patch size. While for large patch size, similar patch sample density is insufficient, and each statistic estimator gives a very different result. In contrast, flat regions require a significantly larger sample size to search similar patch due to its simple local structure. The analysis mentioned above implied that there exist a close relationship between patch size and the performance gain, and told us that algorithm can adaptively adjust the patch size according to variance of image region. The guideline given in Levin et al. (2012) also provide a patch size prior as designing a denoising algorithm.

However, determining the attribution of region (flat area or texture region)is also a intractable problem. Levin et al firstly chose the part pixels as the tested point and find the sufficient clean similar patches of corresponding test pixel point from the massive databases, and then they divide the N clean samples into 10 groups, compute the non-parametric estimator on each group separately, and check if the variance of these estimators is much smaller than noise variance. For small patch size , samples are dense enough and all these estimators provide consistent results which indicate that it has consistency between similar block group. In this case, it easier to increase patch size for similar patches. For large patch size, sample density is insufficient, and each estimator gives a very different result which imply that the consistency of similar patch group has changed. In this case, it is unnecessary to increase the patch size. Finally, they obtain the optimal similar patch size for each pixel point and corresponding optimal similar patch group. In testing experiments, they use 1000 test pixels and 7×10^9 clean samples. At all considered noise levels, the adaptive approach significantly improves the fixed patch approach, by about 0.3–0.6 dB. At higher noise levels, it increasingly outperforms BM3D. However, we can find that time complexity of this method for determining the optimal patch size is high, required several days of computation. So this adaptive strategy is not practical.

After the similar patches being selected, in general, the weighted average method is adopted to estimate the noise-free image in Buades et al. (2005) and Dabov et al. (2007). It is worth to note that the weighted average method does not take full use of the 'similarity' between these patches and adverse to keep image structure. Different form the other average method, BM3D method integrate and aggregate to the similar patches by using collaborative filtering, and finally obtain the estimated image. Two phase:designing and creating database, data migrating, querying and analyzing data. Each step of this approach conclude: Grouping, collaborative filtering and aggregating. From the selection of similar patches to processing of similar patch groups, BM3D is more complex and more time-consuming than NLM algorithm. In Wong et al. (2011), authors presented an efficient two direction non local-based method (TDNL) (Zhang et al. 2013) to deal with the similar patches. Compared to the NLM algorithm, the TDNL model further exploits the two direction structure of similarity data. In addition, implementations of TDNL method are more convenient than BM3D since it has the closed-form solution.

Inspired by idea of prior learning for patch size in literature (Levin et al. 2012), In order to combine the quickness of MCMC strategy in searching similar patches with the adaptation in adjusting patch size, we present an novel stochastic non-local denoising method based on MCMC sampling and adaptive patch size basing stochastic sampling. The method contains two phase: searching the similar patches based adaptive patch size, and designing the denoising algorithm by making use of similar image patches obtained in the first step.

In order to make full use of the structure of similar patch and reduce the computational cost, in the first phase, the multiple clusters of similar patches for each pixel point are searched by

using Markov-chain Monte Carlo sampling many times. Following, we adjust the patch size according to the consistency of multiple clusters. This processing is repeated until we obtain the optimal patch size and corresponding optimal patch cluster. It is worth emphasizing that the similar patches are originate form the observed image instead of the massive database, which will reduce the costs dramatically. In the second steps, we proposed the modified TDNL model to estimate the noise-free patch cluster, namely, we get the initial estimation of noise free patch cluster by employing modified two direction approximation approach. Furthermore, the denoised image is obtained by using the method of superposition approach. TDNL method (Zhang et al. 2013) is better than the weighted average method proposed in Buades et al. (2005). Compared to the NLM algorithm, the TDNL model further exploits the two direction structure of similarity data.

Finally, the initial estimation of original patch is re-projected into the pixel space, and then the denoised image is obtained by using the primary superposition method. Compared to the constrained method and the way of singular value decomposition, the proposed strategy further reduce the computation complexity. We call the whole denoising processing as "An adaptive patch size stochastic nonlocal image denoising method (APSND)".

The proposed method ingeniously use MCMC strategy for searching similar patches corresponding pixel point, which easily adjust the patch size for each pixel point in different regions. Comparing with the fixed patch size approach, the adaptive one can better preserving the image details, guarantee the overall denoising performance. Comparing with the ergodic similar patch searching method both form database or single image, the proposed adaptive method based on MCMC sampling has the lower computation complexity. The main contributions of this paper are:

(1) The proposed algorithm adopt the appropriate random sampling strategy based on MCMC to search similar patches and adaptively chose the optimal patch size according to the variance of similar patch groups.
(2) The original image is approximated by the TDNL approximation method. The TDNL model further exploits the two-direction structure of similarity data, better retains the structure information of original image and achieves very competitive denoising performance.
(3) An effective image denoising algorithm based on adaptive patch size stochastic non-local technique is presented.

Numerical experiments illustrate that the proposed method achieves very competitive denoising performance to state-of-the-art methods, such as BM3D (Zhang et al. 2013), MCMCD (Wong et al. 2011). In addition, comparing with the Anat Levin's method proposed in Levin et al. (2012), our method gets twice the result with half the effort and significantly improves the fixed patch size approach. In contrast NLM method (Buades et al. 2005), the proposed algorithm keeps the integrity of edge and detail information and has the good denoising effect. Comparing with BM3D method, the proposed method can maintain the good visual quality.

Remainder of the paper is organized as follows: In Sect. 2, we briefly introduce theory of stochastic approach for searching similar patches based on adaptive patch size. Section 3 briefly reviews the TDNL strategy to similar patches. Section 4 gives image denoising algorithm proposed in this paper. Section 5 presents experimental results demonstrating the ability of the method of the proposed algorithm to restore artificially corrupted images with additive white Gaussian noise. We conclude the paper in Sect. 6.

2 The stochastic searching similar patches based on adaptive patch size

Let S be a discrete lattice upon which the image is defined. The relationship between the observed noise-contaminated image $y = \{y(s)|s \in S\}$, noise-free image $x = \{x(s)|s \in S\}$,and $n = \{n(s)|s \in S\}$ can be generally expressed as

$$y(s) = x(s) + n(s) \tag{1}$$

2.1 Algorithm of searching similar patch based on adaptive patch size

A new stochastic nonlocal denoising method based on adaptive patch size is presented. The quality of restored image is improved by choosing the optimal nonlocal similar patch size for each site of image individually. The method contains two phases. The first phase is to search the similar patches based adaptive patch size. The second phase is to design the denoising algorithm by making use of the similar image patches obtained in the first step.

The multiple clusters of similar patches for each pixel point are searched by using Markov-chain Monte Carlo sampling many times. Following, we adjust the patch size according to the consistency of multiple patch clusters and obtain group of similar patch corresponding the optimal patch size. This processing is repeated until we obtain the optimal patch size and corresponding optimal patch cluster. The process of adaptively stochastic searching similar patches based on MCMC is stated as follows:

The first step: given the initial patch size d, for each pixel point s_j, we search the corresponding similar neighborhood sequence $\Omega^d(s_j)$ using the MCMC sampling. In order to initialize the sequence of $\Omega^d(s_j)$, let $s_{j_0} = s_j$. According to the estimation method in Levin et al. (2012), we get sample point sequence $\Gamma(s_j) = \{s_{j_0}, \ldots, s\{j_\eta\}\}$ from a known distribution Q, where candidate sample s'_{jk} is drawn from a known instrumental distribution $Q(s'_{jk}|s_{jk-1})$. This processing is repeated N times, and then get N sample group sequences for each pixel point s_j. There are m samples in each group. At the pth iteration, we obtain the pth sample group with center at $s_{ji} \in \Gamma_p^d(s_j)$, patch size $d \times d$ denoted as $\Gamma_p^d(s_j) = \{s_{j_0}, \ldots, s\{j_\eta\}\}$. Finally, we get the N group sequences$\{\Omega_p^d\}_{p=1}^N$.

The second step: we compute the mean-variance $V^d(s_j) = variance(E^d(s_j))$ of sequence $\{\Omega_p^d\}_{p=1}^N$, where $E^d(s_j) = E\{\Omega_p^d(s_j)$ is the mean of sequence $\{\Omega_p^d\}_{p=1}^N$, V represents variance.

The third step: we adjust the optimal patch size according to the variance computed in the second step. we select the largest patch size as the optimal one for which the variance is still below threshold T, namely, $d = max\{d_i|V(\{E_p^{d_i}\}_{p=1}^N) < T\}$, which mean that there are enough similar patch in the observed image, and consistency of multiple groups remain unchanged. In this case, we increase the patch size d to d_1 and then return to the first step, continuously search the similar groups $\{\Omega_p^{d_1}\}_{p=1}^N$, and compute its mean-variance$V^{d_1}(s_j)$. If the variance upon threshold T, this indicates that the consistency has changed, in other word, there are not enough similar patches for the $s'_j s$ neighborhood. In this case, there are no need to increase the patch size, and only output the one similar patch group under the size d as the optimal patch group for $s'_j s$ neighborhood.

2.2 The analysis for algorithm of searching similar patch based on adaptive patch size

In the first step mentioned above, there are three main criterions in the selection of a suitable instrumental distribution for the purpose of denoising. First, distribution of sample points

should reflect the strong locality. Second, sampling under instrumental distribution is a global sampling. This proposed strategy means that the whole sampling process can be interpreted as a local sense of global sampling. With the above consideration in mind, for algorithm in this paper,a convenient choice for the instrumental distribution $Q(s'_{jk}|s_{jk-1})$ is a Gaussian distribution centered at s_{jk-1}, thus encouraging increased sampling of sites within close proximity while still allowing sites farther away be sampled, which utilize global information within the image. For proposed method, the Gaussian instrumental distribution $Q(s'_{jk}|s_{jk-1})$ can be defined as

$$Q(s'_{jk}|s_{jk-1)} = \frac{1}{\sqrt{2}\sigma_s} exp\left[-\left(\frac{s'_{jk}}{s_{jk-1}}\right)^2\right] \tag{2}$$

where $\sigma_s$represents the spatial variance of $Q(s'_{jk}|s_{jk-1})$.

To determine whether the sample candidate s'_{jk} is accepted as part of $\Omega_p^d(s_j)$ or not, the acceptance probability of s'_{jk} given s_{jk-1}

$$\alpha(s'_k|s_{k-1}) = min\left\{1, \frac{\varphi(s'_k|s_0)}{\varphi(s_{k-1}|s_0)}\right\} \tag{3}$$

where $\varphi(s'_k|s_0)$ is an objective function that estimates the likelihood between the neighborhoods of candidate point s'_k and initial point s_0.The acceptance probability in Eq. (3) suggests that the Markov chain then moves towards s'_k with acceptance probability $\alpha(s'_k|s_{k-1}) = min\left\{1, \frac{\varphi(s'_k|s_0)}{\varphi(s_{k-1}|s_0)}\right\}$, that is, the acceptance probability for s'_k is one when $\varphi(s'_k|s_0) > \varphi(s_{k-1}|s_0)$, Otherwise $[\varphi(s_{k-1}|s_0)]^{-1} \cdot \varphi(s'_k|s_0)$. The objective function in this paper is defined as

$$\varphi(s'_k|s_0) = \prod_j exp\left[-\frac{\left(y^{\rho_{s'_k}}(j)\right)^2 - (y^{\rho_{s_0}}(j))^2}{C}\right] \tag{4}$$

where $y^{\rho_{s'_k}}(j)$ and $y^{\rho_{s_0}}(j)$ are the intensities of the jth site in the local neighborhoods $\rho_{s'_k}$ and ρ_{s_0}, respectively, and C is a constant. It gives greater values to sites with similar local spatial intensity relationships, as they are more likely to belong to the same distribution as s. The sample candidate s'_k is then included into $\Omega_p^{d_i}(s)$ with a probability of $\gamma(s'_k|s_{k-1})$. This probabilistic acceptance step is realized in the proposed algorithm as follows. A value u is randomly sampled from a uniform distribution $U(0, 1)$. If $u > \gamma(s'_k|s_{k-1})$, then the sample candidate is discarded. However, if $u < \gamma(s'_k|s_{k-1})$, then s'_k is included into the sequence of samples $\Omega_p^{d_i}(s)$. The sampling process is repeated until the desired number of similar patches is achieved.

In the second stage, we give the criterion for adjusting the patch size according to the consistency of similar patch groups. The consistency of patch groups is measured by changeability of the mean-variance of similar patch groups. In fact, for different similar patch groups, we can easily compute means of these groups by weighted average way and obtain corresponding mean-variance of means. If there are enough similar patches for each pixels' neighborhood, the difference between the means of patch group is small. This also suggest that there is high consistency among patch groups. In this condition, we increase the patch size and continuously search the similar patch in the updated patch size. In contrast, when the patch size goes up to some extent, we could not get enough similar patches. This fact is

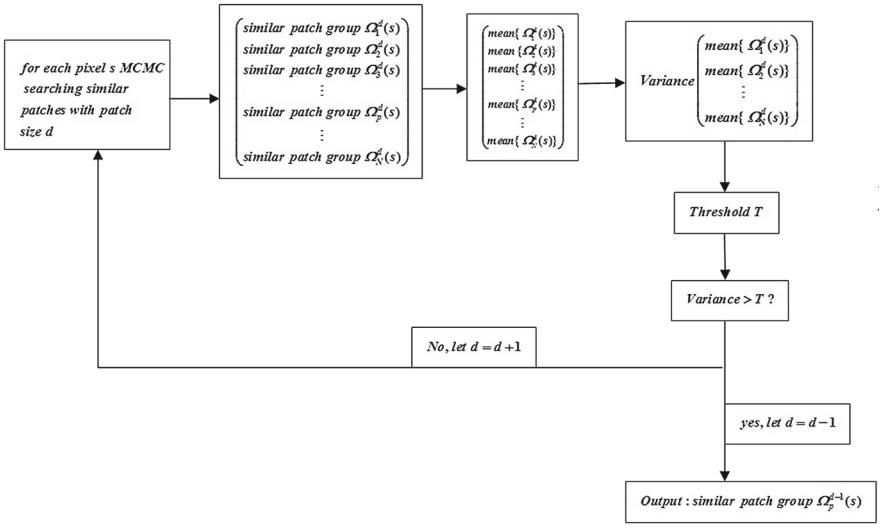

Fig. 2 The processing of searching similar patch based on adaptive patch size

supported by the obvious difference of mean. So we need to reduce patch size in a optimal situation.

Given the criterion stated in the second stage, we adjust the patch size and get the optimal patch size and corresponding optimal similar patch group, namely, let $d = max\{d_i | V(\{E_p^{d_i}(s_j)\}_{p=1}^N) < T\}$ be the optimal patch size for pixel s_j, let $\Omega(s_j) = \Omega_p^d(s_j)$ be the optimal patch group, where T is a threshold, which is defined as $\frac{1}{\sigma^2}$,σ^2 is the noise variance estimated by the method proposed in Levin et al. (2012). Figure 2 illustrates the above process in a simplified sequence flow diagram.

3 TDNL approximation of similar patches

Once the similar patches being selected, the weighted average method is used in Wang and Feng (2008), Levin et al. (2012) and Buades et al. (2005). However, this method does not take full use of the 'similarity' between these patches. In fact, stacking these non-local similar patches of s_j into a matrix, denoted by Y_{s_j}, we have $Y_{s_j} = X_{s_j} + N_{s_j}$, where X_{s_j} and N_{s_j} are the patch matrices of original and noise, respectively. Intuitively, there are similarity both columns and rows of the matrix. So the TDNL approximation method can be used to estimate X_{s_j} from Y_{s_j}.

In many cases, the average method is used to deal with the similar patches, while it does not take full use of the similarity between the patches. To avoid this, we employ the TDNL method proposed in Zhang et al. (2013) which fully use the two directional similarities between patches.

For ease of description, we introduce some denotations. Let C_s be the operation to extract the square patch centered at $x(s)$ and vectorize it as $C_s x$, D_s be operation to extract similar patches of $C_s x$ and rearrange them into a matrix D_s, Let R_s be the operation to extract the row of $D_s x$ corresponding to the central point of each patch arranged in $D_s x$ as $R_s x$, see Fig. 3.

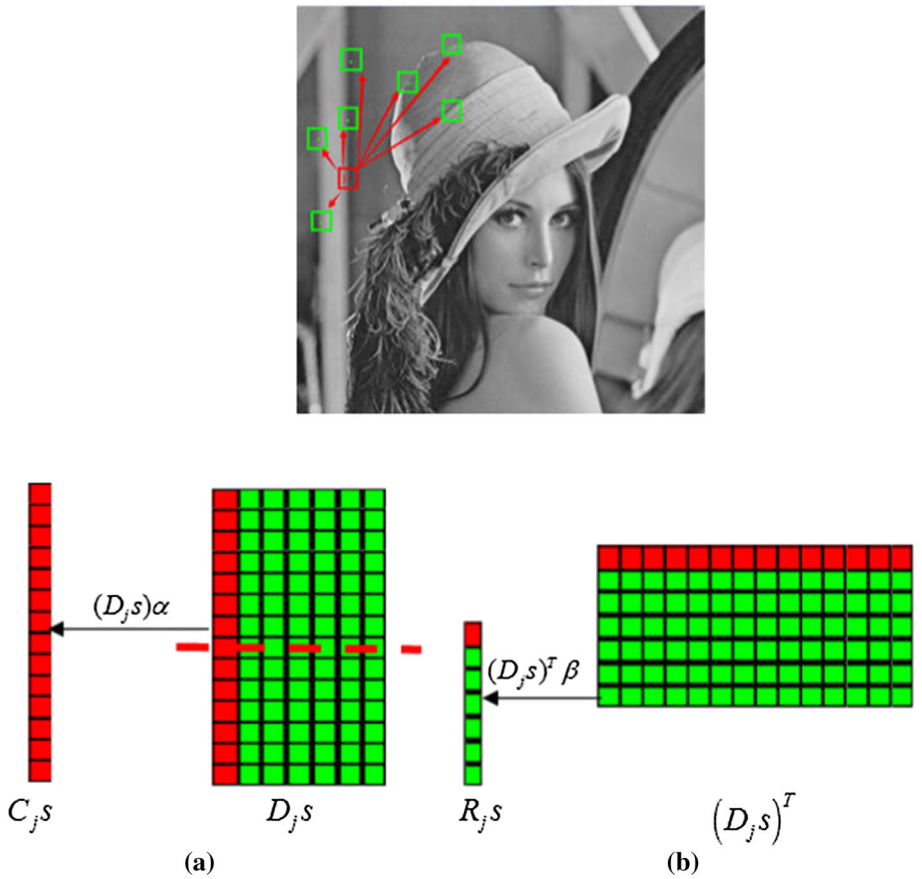

Fig. 3 The similarities between both column and rows of the in the self-representation style, **a** 'similarity' between the columns, **b** 'similarity' between the rows

In fact, once similar patches $D_s x$ are selected, we take columns of $D_s x$ as dictionary atoms to represent $C_s x$ instead of averaging them with the similarity measure as the weights. This can be formulated as

$$\alpha_s^* = \arg\min_{\alpha_s} \left\{ \|C_s x - (D_s x)\alpha_s\|_2^2 + \lambda_1 \|\alpha_s\|_\omega^\omega \right\} \tag{5}$$

where α_s is the coefficient vector and λ_1 is the regularizing parameter. The regularization term $\|\alpha_s\|_\omega^\omega$ is involved to avoid the trivial solution and ill-posedness of the problem. Zhang et al. (2017c) consider that $C_s x$ is almost equally coherent with each column of $D_s x$ and let $\omega = 2$. So the formulation is changed to the form

$$\alpha_s^* = \arg\min_{\alpha_s} \left\{ \|C_s x - (D_s x)\alpha_s\|_2^2 + \lambda_1 \|\alpha_s\|_2^2 \right\} \tag{6}$$

By row similarity, we mean that, the center pixels of all patches have identical AR coefficients (Wong et al. 2011), in terms of the matrix $D_s x$, the row consisted of the center pixels can be linearly represented by other rows through the so called nonlocal coherent AR. Sim-

ilarity between patches in $D_s x$ suggests that the AR coefficients for all patches should be identical. Similar to the model (6), Wong et al. (2011) give the following model to find the AR coefficients

$$\beta_s^* = \arg\min_{\beta_s} \left\{ \|R_s x - (D_s x)^T \beta_s\|_2^2 + \lambda_2 \|\beta_s\|_2^2 \right\} \tag{7}$$

where β_s is the AR coefficient vector and λ_2 is the regularizing parameter. The Eq. (7) can be viewed as a regularized nonlocal-AR model.

In order to take full use of the 'similarity' structures between the patches and get an ideal denoised image, we employ TDNL model (Wang and Feng 2008) in image.

$$\begin{aligned}(x, \{\alpha_s^*\}, \{\beta_s^*\}) = \arg\min_{x, \{\alpha_s\}, \{\beta_s\}} \Bigg\{ & \frac{\mu}{2}\|y - x\|_2^2 \\ & + \sum_s \begin{bmatrix} \|C_s x - (D_s x)\alpha_s\|_2^2 + \lambda_1 \|\alpha_s\|_2^2 \\ + \|R_s x - (D_s)^T \beta_s\|_2^2 + \lambda_2 \|\beta_s\|_2^2 \end{bmatrix} \Bigg\}\end{aligned} \tag{8}$$

It is worth stressing that the model (8) employed in this paper is different from the original TDNL proposed in Zhang et al. (2013). all patches are clustered, and then the TDNL is used to represent the patch. In this paper, we use TDNL more directly.

The first term of (8) globally forces the proximity between the noisy observation and the denoised image. The second term integrates similar patch priors to form global force to regularize the estimation. The TDNL model is closely related to regression model but directly takes the similarity data matrix as the dictionary. The key point is that it symmetrically exploits the similarities between both columns and rows of the $D_s x$ in the self-representation style, and does not impose the sparsity constraint.

We can perform Eq. (8) by alternatively minimizing according (Wong et al. 2011).This means we need to solve the following subproblems

$$\alpha_s^{k+1} = \arg\min_{\alpha_s} \left\{ \|C_s x^k - (D_s x^k)\alpha_s\|_2^2 + \lambda_1 \|\alpha_s\|_2^2 \right\} \tag{9}$$

$$\beta_s^{k+1} = \arg\min_{\beta_s} \left\{ \|R_s x^k - (D_s x^k)^T \beta_s\|_2^2 + \lambda_2 \|\beta_s\|_2^2 \right\} \tag{10}$$

$$\begin{aligned}(x^{k+1}) = \arg\min_{x} \Bigg\{ & \frac{\mu}{2}\|y - x\|_2^2 \\ & + \sum_s \begin{bmatrix} \|C_s x - (D_s x^k)\alpha_s\|_2^2 + \lambda_1 \|\alpha_s^{k+1}\|_2^2 \\ + \|R_s x - (D_s x^k)^T \beta_s\|_2^2 + \lambda_2 \|\beta_s^{k+1}\|_2^2 \end{bmatrix} \Bigg\}\end{aligned} \tag{11}$$

4 Denoising algorithm and analysis

4.1 Denoising algorithm

An adaptive patch size stochastic nonlocal image denoising method (APSND)
1 **input**:observation matrix observation matrix $y, m, \hat{x} = y$
2 **For**each $s_j \in S$
3 $d = 0$
4 **Do while** $V < T$

5 $d = d + 1$
6 **For** $p = 1$ *to* N
7 **For** $k = 1$ *to* m
8 $\Omega_p^d = s_j, s_{j0} = s_j$
9 Random sampling: $s'_{jk} \sim Q(s'_{jk}|s_{jk-1})$ is defined as Eq. (6)
10 Generate a random value l according to a uniform distribution $U(0, 1)$
11 Adjudicating: A compute acceptance probability $\alpha(s'_{jk}|s_{j,k-1})$ in Eq. (7)
12 B: If $\alpha < l$, include s'_{jk} into the sequence of samples $\Omega_p^d(s_j)$, let $\Omega_p^d(s_j) = \Omega_p^d(s_j) \cup \{s'_{jk}\}$, $k = k + 1$, $s_{jk} = s'_{jk}$, other wise, discard s'_{jk}, let $\Omega_p^d(s_j) = \Omega_p^d(s_j)$, $k = k$,$s_{jk} = s_{j,k-1}$
13 **End For**
14 $E_p^d = E\{\Omega_p^d(s_j)\}$
15 **End For**
16 $V = variance\{E_p^d\}_{p=1}^N$
17 **End Do**
18 Let $\Omega(s_j) = \Omega_N^d(s_j)$, stacking non-local similar patches $\Omega(s_j)$ into a matrix $D_j s$ and employing the TDNL strategy on $D_j s$ to get $\hat{x}(s_j)$
A: Employ the equation (9) and get the α_{s_j};
B: Employ the equation (10) and get the β_{s_j};
C: Employ the equation (11) and get the $\hat{x}_{s_j}$
19 Aggregate each patch to form the clean image $\hat{x}$=$\sum \hat{x}_{s_j}$
20 **End For**
21 **Output**:Computer the whole noise-free image estimate:$\hat{x}$

The whole process of algorithm is shown in Fig. 4.

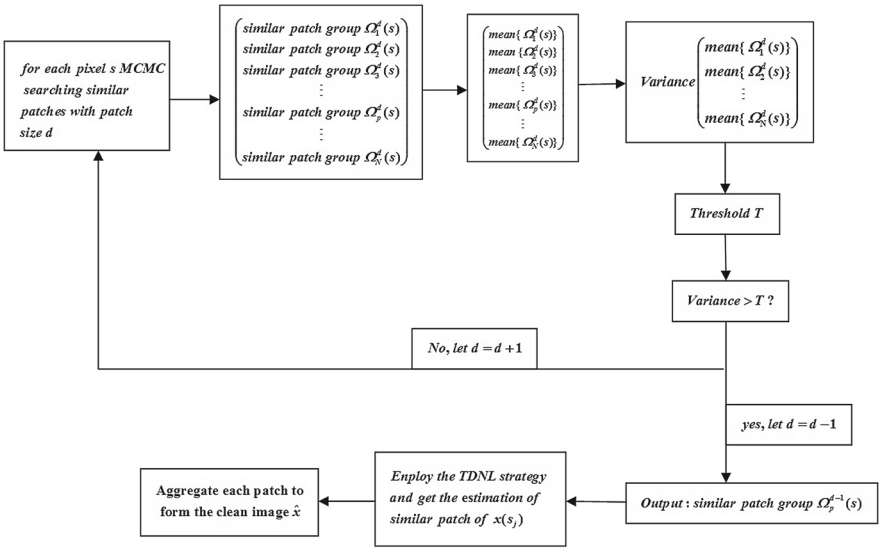

Fig. 4 The processing of APSND algorithm

4.2 Analysis of algorithm

Image self-similarity is considered among the patches over the whole image (Mahmoudi and Sapiro 2005; Orchard and Ebrahimi 2008). The well known NLM is very effective while it is computation-consuming in searching for similar patches especially for large image. Buades et al. presented a modified one (Buades et al. 2005) which searches for self-similar patches in a searching window rather than the whole image, to some extent the method lower the computation. However, for a searching window, the searching step still takes a lot of costs. Such as a searching window of size 41×41, the searching step will make 1600 times comparison for each pixel, which is a great burden. To alleviate it, literature (Wong et al. 2011) adopted a searching approach based on MCMC. In this method, we adopt MCMC strategy with adaptive patch size to select the similar patches. For each pixel point , the step of searching similar patches is repeated several times. In theory, the computing cost is higher than approach stated in literature (Wong et al. 2011), but it is still far lower than methods (Levin et al. 2012; Buades et al. 2005; Dabov et al. 2007). Noise of an image can be reduced by dealing with similar patches using many methods such as weighted average method (Buades et al. 2005), PCA approach (Zhang et al. 2010). However, the 1-norm regularization term based sparse representation method, or singular value decomposition based principle component analysis method is of higher computation complexity. In our paper, we adopt the TDNL technique, using the 2-norm regulation term, which gives closed form solutions of Eqs. (9)–(11) and reduce the time complexity. A brief conclusion is that in our method the computation complexity of searching step is lower than the NLM (Buades et al. 2005), litter higer than MCMCD (Wong et al. 2011),and the time complexity of the noise remove step is also lower compared with the sparse representation approach and the singular value decomposition based PCA method.

5 Experimental results

To evaluate the performance gains of the proposed image denoising algorithm, the four benchmark images (Fig. 5) named Lena, Barbara, peppers and house are used for the experiments. The noisy images are obtained by adding Gaussian white noise (noise standard deviations ranging from 20 to 50) to the noise free image. The proposed method is compared to state-of-the-art methods such as BM3D, NLM and literature (Wong et al. 2011) method under different noise levels to assess the overall performance of our method both quantitative indicators and visual quality.

The TDNL model can be viewed as the two-direction. The parameters λ_1, λ_2 are the regularizing parameter which have the relationship with the noise level and can affect the estimation of original image. Even the solution of Eq. (8) highlight the relation between TDNL model and Wiener filter. So This provide the cue for parameters λ_1, λ_2 choice. In paper (Zhang et al. 2013), author choose $\lambda_1 = c_1 \times \sigma_n^2$, $\lambda_2 = c_2 \times \sigma_n^2$, where c_1 and c_2 are both predefined positive constant. σ_n^2 is the noise variance which is estimated by using method proposed in John (1996) and Zhang and Wu (2008).

For quantitative comparison purposes, we adopt the PSNR and the Structural Similarity Index Metrix (SSIM) to test the denoising effect. The PSNR was computed according to the formula

$$PSNR = 10lg \frac{255^2}{\sum\sum \frac{(\hat{x}_{ij} - x_{ij})^2}{M}} \tag{12}$$

Fig. 5 Set of test images (from left to right): Lena, Barbara, peppers, hill

where x_{ij} is pixel point of original image, $\hat{x_{ij}}$ is the pixel point of denoised version, M is the total number of pixels in the image, i, j are the subscript of image pixel.

In addition to the PSNR, SSIM is also a brilliant metric for image quality assessment and has been used widely recent years.

$$SSIM(\hat{x}, x) = \frac{4\sigma_{\hat{x}P}\mu_{\hat{x}}\mu_x}{(\sigma_{\hat{x}}^2 + \sigma_x^2)(\mu_{\hat{x}}^2 + \mu_x^2)}, \tag{13}$$

where $\mu_{\hat{x}}$, μ_x, $\sigma_{\hat{x}}$, σ_x, $\sigma_{\hat{x}x}$ represent the mean, standard deviation and cross-correlation evaluations to $\hat{x}$, x, respectively. SSIM is an image assessment measure based on structural similarity index to evaluate the similarity of original image and the denoised image, higher value indicates higher similarity in structure.

Table 1 lists the PSNR and SSIM values of different methods on the four tested images. It is clear from the table that the BM3D has the highest PSNR value, which is because it sufficiently exploits the non-local redundancies in the image. The PSNR results of the proposed method are higher than MCMCD (Wong et al. 2011) and NLM which has the lowest PSNR value. In addition, let's focus on the SSIM value. From Table 1 we can see that BM3D again achieves the highest SSIM values except for the results of image "Lena",

Table 1 PSNR and SSIM results with respect to noise level. SSIM results are shown in brackets

Noise level	20	30	40	50
Lena 512 × 512				
BM3D (Dabov et al. 2007)	33.04 (0.87)	31.26 (0.84)	29.86 (0.81)	29.05 (0.78)
NLM (Buades et al. 2005)	29.13 (0.79)	27.30 (0.75)	26.18 (0.72)	26.51 (0.72)
MCMCD (Wong et al. 2011)	32.18 (0.86)	30.05 (0.82)	28.86 (0.70)	27.61 (0.75)
APSND	32.80 (0.87)	31.02 (0.84)	29.54 (0.82)	28.17 (0.79)
DnCNN (Zhang et al. 2017b)	**33.36** (**0.88**)	**31.57** (**0.85**)	**30.32** (**0.83**)	**29.38** (**0.81**)
Barbara 512 × 512				
BM3D (Dabov et al. 2007)	31.78 (0.90)	29.81 (0.86)	27.99 (0.82)	27.23 (0.76)
NLM (Buades et al. 2005)	29.23 (0.84)	26.75 (0.77)	25.10 (0.70)	23.96 (0.65)
MCMCD (Wong et al. 2011)	30.53 (0.89)	28.76 (0.82)	26.48 (0.75)	25.82 (0.71)
APSND	31.10 (0.89)	29.08 (0.85)	27.45 (0.81)	26.05 (0.76)
DnCNN (Zhang et al. 2017b)	31.09 (0.89)	28.92 (0.85)	27.25 (0.81)	26.22 (0.77)
Peppers 256 × 256				
BM3D (Dabov et al. 2007)	31.76 (0.88)	29.28 (0.85)	27.70 (0.81)	26.68 (0.77)
NLM (Buades et al. 2005)	28.77 (0.83)	26.53 (0.78)	24.85 (0.74)	24.43 (0.70)
MCMCD (Wong et al. 2011)	30.44 (0.86)	28.40 (0.81)	26.72 (0.78)	25.34 (0.72)
APSND	30.85 (0.88)	28.63 (0.85)	27.02 (0.81)	25.50 (0.78)
DnCNN (Zhang et al. 2017b)	**31.90** (**0.90**)	**29.91** (**0.86**)	**28.41** (**0.83**)	**27.12** (**0.81**)
House 256 × 256				
BM3D (Dabov et al. 2007)	33.77 (0.87)	32.09 (0.85)	30.65(0.83)	29.69 (0.82)
NLM (Buades et al. 2005)	32.00 (0.79)	29.37 (0.75)	27.35(0.72)	25.92 (0.70)
APSND	32.93 (0.86)	31.29 (0.84)	29.83 (0.82)	28.43 (0.80)
DnCNN (Zhang et al. 2017b)	**33.89** (**0.87**)	**32.24** (**0.85**)	**30.97** (**0.83**)	**30.00** (**0.82**)

The higher numerical indicators are given in bold

"peppers". The proposed method has outperformed PSNR, SSIM results as MCMCD, which indicates that our method has great advantages in preserving the structure information of the original image, the NLM denoising methods still has the lowest values in SSIM.

In addition,more recently many deep learning based image denoising methods have been proposed. We compare the proposed APSND method with state-of-the-art DnCNN method (Zhang et al. 2017b). Table 1 also lists the average PSNR and SSIM results of different methods for three general image dennoising tasks. As one can see, the DnCNN outperforms the other methods for Gaussian denoising, but it takes huge computational cost without the GPU implementation.

Figure 6 illustrates the variance of optimal patch size with the image region changes. There are many regional characteristic in " Barbara" image, such as edges, texture, details flat regions. So the optimal patch size for different regions also appear diverse situation.

Figure 7 demonstrates the relationship between the patch size and the noise level. From the experiment results, we can see that we need the similar patches with relatively large size when the noise level is high, on the contrary, we need the similar patches with small size. e.g., Fig. 7 shows that the optimal patch size is 3×3 for most of pixel points in the image at the low noise level $\sigma_n = 5$ and $\sigma_n = 10$, while at the higher noise level $\sigma_n = 40$, the

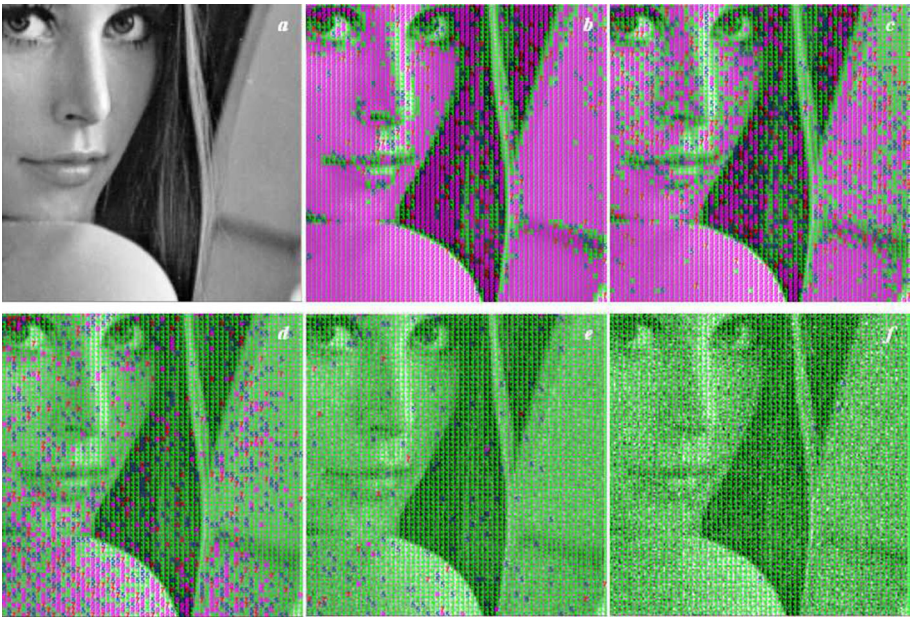

Fig. 6 Optimal patch size with different noise level for image "lena".Top left to right: original image, $\sigma = 0$, $\sigma = 5$, bottom left to right: $\sigma = 10$, $\sigma = 20$,$\sigma = 40$. Light green dot patch size is 3×3, dark green dot patch size is 5×5, red dot patch size is 7×7, rose dot patch size is 9×9 (Color figure online)

Fig. 7 optimal patch size with the image region changes. Light green dot patch size is 3×3, dark green dot patch size is 5×5, red dot patch size is 7×7, rose dot patch size is 9×9 (Color figure online)

Fig. 8 Denoising results on image lena by different strategies. Top left to right:Original image, fixed patch size $3 \times 3 + mean(PSNR : 24.88)$, fixed patch size $5 \times 5 + mean(PSNR : 26.41)$, Bottom left to right: fixed patch size $7 \times 7 + mean(PSNR : 25.68)$, adaptive patch size + mean ($PSNR : 26.64$) APSND ($PSNR : 31.02$)

optimal patch size is 9×9 for most ones. So for large noise the neighborhood definition is wider.

In order to demonstrate the superior of the APSND method compared to the fixed- patch size strategy, several comparison results about image "Lena" are given, Fig. 8b–d are results by using MCMC sampling method with different fixed patch size and taking the weighted average approach, respectively. Figure 8e is result by using MCMC sampling method with the adaptive patch-size and the weighted average approach. Figure 8f is result by proposed method in this paper, namely, by using MCMC sampling method with adaptive patch-size and taking TDNL method. From the comparison results, we can see that the weighted average denoising method based on the adaptive patch size is better than ones based on the fixed patch size. Furthermore, the APSND method is better than the weighted average denoising method (Levin et al. 2012) based on the adaptive patch-size.

In Figs. 9 and 10, we compare the visual quality of the denoised images by the competing algorithms. Figure 9 demonstrates that the APSND algorithms reconstructs more image details from the noisy observation. We observe that the results of our proposed method are better than that fixed patch size stochastic method for window size 3×3 and 5×5. Figure 10 shows images "Lena", "Barbara", "peppers", "house" contaminated by additive Gaussian noise with standard deviation $\sigma_n = 30$ and the corresponding denoised images using the tested method. As observed from the denoised image, the proposed method yields more satisfactory denoising results visually, NLM does not preserve fine details such as the stripe patterns in Barbara's clothes.

In order to further show the benefits of the APSND method in keeping visual quality, we crop and zoom-in the denoised results of Barbara in Fig. 11. It demonstrates that the proposed method reconstructs more image details from the noisy observation. Compared with BM3D,

Fig. 9 Denoising results on tested image by different methods. Left to right: Original image, Noised image BM3D, NLM, APSND method

the effect of reconstructing is very satisfying, especially in the region of face. At the same time, the APSND method also keep good texture details as well as BM3D.

Figure 12 demonstrates the visual comparison of adaptive versus fixed patch size denoising. Compared with the adaptive method, our method removes noise significantly as compared to the fixed-patch size ones. More interestingly, as can we seen in the highlighted window, when patch size is small, noise residuals are highly visible in the flat regions. With a large patch size, noise is visible around edges. Both edges and flat regions are handled properly by our approach the adaptive approach.

Figure 13 shows the zoomed-in regions of "peppers" image contaminated by additive Gaussian noise with $\sigma_n = 20$ and the corresponding denoised 3-D surface images using NLM approach, BM3D approach and the APSND approach. Comparing with the edge-preserving ability of these algorithms, we can see that APSND method can better keep the fine edge in images and provide quite a good visual quality. As can be seen in the highlighted window that the APSND method has a good opportunity to obtain clearer edges in the denoising output.

Figure 14 shows the "Lena" image contaminated by additive Gaussian noise with and the corresponding denoised edge images using the tested methods. Based on visual inspection, the APSND method provide better edge image quality due to having little outliers, smooth and continuous edge compared to the BM3D. Another quantitative indicator for evaluating behavior of proposed algorithm in preserving the edge is edge-strength-similarity-based image quality metric (ESSIM) proposed in Wang et al. (2004). For "Lena" image contam-

Fig. 10 Visual comparison of adaptive versus fixed patch size stochastic denoising comparison on sub-images clipped from image peppers, Barbara and Lena. (Left to right: fixed 3×3, fixed 5×5, Adaptive)

Fig. 11 Visual comparison of BM3D versus the proposed denoising comparison on sub-images clipped from image Barbara. (Left to right: Original image, BM3D, APSND method)

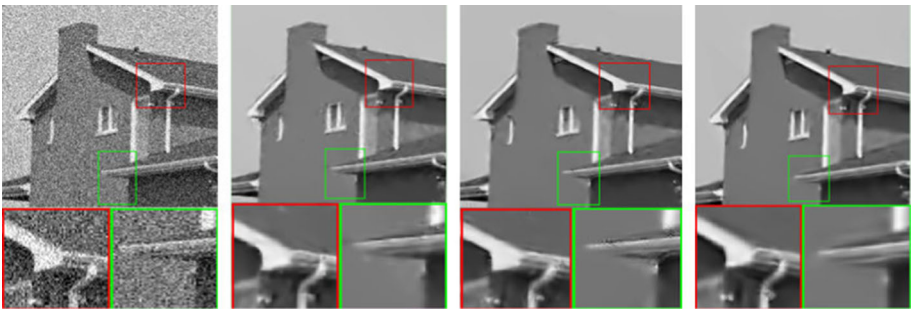

Fig. 12 Denoising results on image house by fixed patch size and adaptive patch size. (Left to right: Noised image, patch size 3×3, patch size 5×5, proposed approach)

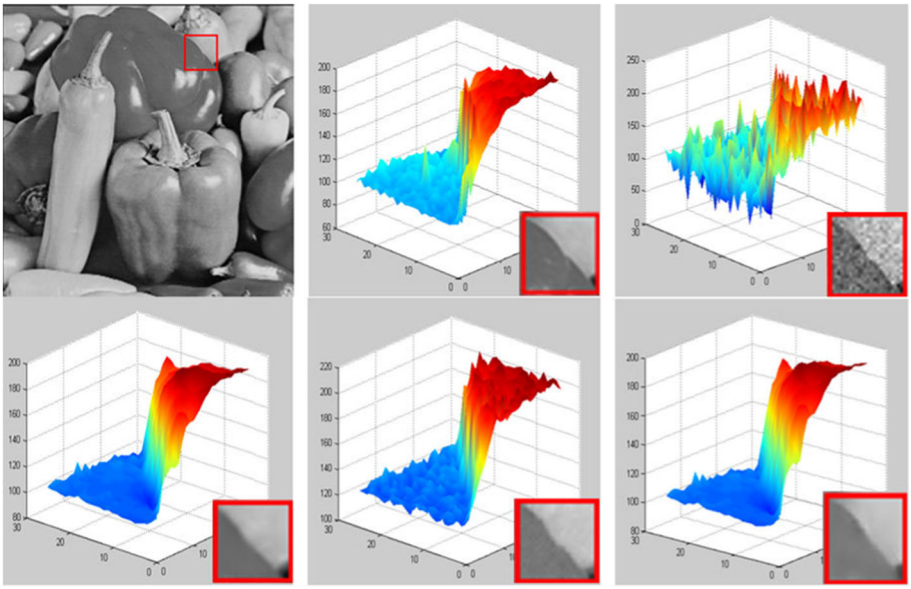

Fig. 13 Variance of optimal patch size with the image region changes

inated by additive Gaussian noise with, ESSIM values for corresponding denoised images using the BM3D, NLM, and APSND algorithm are 0.997, 0.994, 0.997, respectively. These results also indicate that proposed methods can filter noise while preserving image details. All these indicate that the APSND method better retain the detail information of the image itself. The results show that this method may reduce the image fuzziness, preserve the integrity of edge and detail information and has the good denoising effect.

6 Conclusions

In this study, a stochastic technique based on the adaptive patch size for image denoising (APSND) is given. We use the MCMC strategy to obtain the similar patch groups and adjust the optimal patch size according to the variance of these groups. Then the TDNL is employed

Fig. 14 "Lena" image contaminated by additive Gaussian noise with and the corresponding denoised edge images using the tested methods. Left to right: Original edge image, BM3D , proposed method

to estimate the noise-free image. Our experimental results demonstrate that the APSND algorithm not only achieves very competitive denoising performance in terms of PSNR, SSIM and visual quality, but also has lower computational complexity comparing the Anat Levin's work.

Despite we have high denoising quality,the APSND method typically suffer from some drawbacks. In fact,the proposed method is not effective for all image. For those image with having no significant changes in texture structure or with more repetitive texture in whole image, denoising effect is obvious. because with the increase of similar block size, we can find plenty of similar image blocks. for images with more texture and the texture structure changing significant, denoising effect is not ideal. because when a similar block size is small enough, similar patches can be easily found, but the denoising effect is not good due to keeping the structure of the image is not ideal in this case. with the large patch size,we can't find enough similar image block. This is drawback of the proposed method, also is working to improve later.

Acknowledgements This work is Supported by National Natural Science Foundation of China (NSFC 11601420) and Shanxi province education department fund item (No:16JK1708).

References

Buades, A., Coll, B., & Morel, J. M. (2005). A non-local algorithm for image denoising. *IEEE International Conference on Computer Vision and Pattern Recognition*, *2*, 20–25.

Cheng, G., Yang, C., Yao, X., Guo, L., & Han, J. (2018). When deep learning meets metric learning: remote sensing image scene classification via learning discriminative CNNs. *IEEE Transactions on Geoscience and Remote Sensing*,. https://doi.org/10.1109/TGRS.2017.2783902.

Cheng, G., Zhou, P., & Han, J. (2016). Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Transaction on Geoscience and Remote Sensing*, *54*(12), 7405–7415.

Dabov, K., Foi, V., & Katkovnik, V. (2007). Image denoising by sparse 3D transform-domain collaborative filtering. *IEEE Transaction on Image Processing*, *16*(8), 2080–2095.

John, I. (1996). Fast noise variance estimation. *Computer Vision and Image Understanding*, *64*(2), 300–302.

Levin, A., Nadler, B., Durand, F., Freeman, W. T. (2012). Patch Complexity, finite pixel correlations and optimal denoising. In *Lecture notes in computer science* (Vol. 7576, pp. 73–86).

Mahmoudi, M., & Sapiro, G. (2005). Fast image and video denoising via nonlocal means of similar neighborhoods. *IEEE Signal Processing Letters*, *12*(12), 839–842.

Orchard, J., Ebrahimi, M., Wong, A. (2008). Effcient nonlocal-means denoising using the SVD. In *Proceedings of IEEE international conference on image processing*.

Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, *60*, 259–268.

Wang, W., & Feng, X. (2008). Anisotropic diffusion with nonlinear structure tensor. *Multiscale Modeling and Simulation*, *7*(2), 963–977.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600612.

Wong, A., Mishra, A. K., Zhang, W., Fieguth, P. W., & Clausi, D. A. (2011). Stochastic image denoising based on Markov-Chain Monte Carlo sampling. *Signal Processing*, *91*(8), 2112–2120.

Yao, X., Han, J., Zhang, D., & Nie, F. (2017). Revisiting co-saliency detection: A novel approach based on two-stage multi-view spectral rotation co-clustering. *IEEE Transactions on Image Processing*, *26*(7), 3196–3209.

Zhang, D., Meng, D., & Han, J. (2017a). Co-saliency detection via a self-paced multiple-instance learning framework. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(5), 865–878.

Zhang, K., Zuo, W., Chen, Y., Meng, D., & Zhang, L. (2017b). Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, *26*(7), 3142–3155.

Zhang, L., Dong, W., Zhang, D., & Shi, G. (2010). Two-stage image denoising by principal component analysis with local pixel grouping. *Pattern Recognition*, *43*(4), 1531–1549.

Zhang, L., Yang, J., & Zhang, D. (2017c). Domain class consistency based transfer learning for image classification across domains. *Information Sciences*, *419*, 242–257.

Zhang, L., Zuo, W., & Zhang, D. (2016). LSDT: Latent sparse domain transfer learning for visual adaptation. *IEEE Transactions on Image Processing*, *25*(3), 1177–1191.

Zhang, X., Feng, X., & Wang, W. (2013). Two-direction nonlocal model for image denoising. *IEEE Transactions on Image Processing*, *22*(1), 408–412.

Zhang, X., & Wu, X. (2008). Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation. *IEEE Transactions on Image Processing*, *17*(6), 887–896.