A Novel Patch-based Image Denoising Algorithm using Finite Radon Transform for Good Visual Quality

Yun-Xia Liu, Ngai-Fong Law and Wan-Chi Siu
The Hong Kong Polytechnic University, Kowloon, Hong Kong
E-mail: 09902337r@polyu.edu.hk Tel: +852-27666202

***Abstract*—Patch-based denoising methods have recently emerged due to its good denoising performance. In this paper, based on analysis of the optimal over-complete patch aggregation, we highlight the importance of a local transform for good image features representation. A finite Radon transform (FRAT) based two-stage over-complete image denoising algorithm is then proposed for obtaining good visual quality of denoised images. Experimental results demonstrate good performance in that the denoised images obtained by the proposed method are less influenced by artifacts.**

I. INTRODUCTION

Noise is always un-avoidable during the process of image acquisition and transmission. Image denoising is an important pre-processing procedure to relieve human observers from the annoying noisy images. The rapid development of high-resolution imaging devices further imposes challenge on image denoising algorithms: to have high denoised quality, but at low computational complexity.

Dating back from the famous VisualShrink thresholding method [1], discrete wavelet transform (DWT) based methods have been extensively studied. Compared with the decimated DWT approaches, over-complete transforms often lead to better denoising performance. An example is the Bayesian least square estimator with steerable pyramid domain Gaussian scale mixture model (BLS-GSM) [2]. With the consideration that human visual system is more sensitive to edges and textures, various multi-resolution geometrical analysis schemes (e.g. ridgelet[3-4], curvelet[5], contourlet[6]) are proposed for having sparse coefficients and localizing feature. Though higher peak signal-to-noise ratios (PSNR) were reported, these methods all suffered from serious ringing artifacts due to shrinkages in transform domain. On the other hand, over-complete denoising can be achieved in spatial domain. The introduction of "non-local" idea [7] brings extensive discussion about spatial domain methods. In the original paper of non-local mean (NLM), denoised pixel intensity is obtained as a weighted average of similar pixels in a local surrounding window. The BM3D[8] scheme can be regarded as the patch extension version of NLM, where the third dimension is formed by adaptive stacking of similar patches into a group. Denoising is carried out in the local 3D transform domain in which 2D DWT with 'bior' kernels and 1D Haar kernel are used. It is followed by hard thresholding and wiener filtering in the basic and final estimation stage respectively.

Based on analysis of the importance of the local 2D transform within the BM3D framework, we propose a two-stage patch-based denoising algorithm based on the finite Radon transform (FRAT). The FRAT is a non-separable near-orthogonal 2D transform which is good at preserving linear singularity. Local geometric features are approximated in basis of "lines" in the proposed algorithm as opposed to "points" in the BM3D. Experiment results show that the proposed method demonstrate good performance in terms of both subjective measure and objective measure of PSNR and the Structural SIMilarity index (SSIM) [9]. Denoised images are least influenced by artifacts.

This paper is organized as follows. Section 2 briefly reviews the framework of over-complete patch based denoising algorithm and the finite Radon transform. In section 3, we discuss the proposed patch-based algorithm. Section 4 presents some experimental results and the paper is concluded in section 5.

II. PATCH-BASED OVER-COMPLETE DENOISING

In this section, we first briefly review the state of art patch-based over complete denoising algorithms using BM3D algorithm as an example. Based on discussion of the optimal weighting of the overlapping patches, we highlight the importance of local 2D transform, which motivates us to consider the use of finite Radon transform as local 2D transform for better denosing performance.

A. Basic framework of BM3D

The BM3D algorithm is an image denoising strategy based on an enhanced sparse representation in local transform domain. It consists of three major procedures which are *grouping, collaborative filtering* and *aggregation,* respectively.

The enhancement of the sparsity is achieved by *grouping* similar 2-D image fragments (e.g., blocks) into 3-D data arrays which is termed as "groups." *Collaborative filtering* is developed to deal with these 3-D groups to obtain a local estimate of the uncontaminated image. It contains three successive steps: 3-D transformation of a group, shrinkage of the transform spectrum, and inverse 3-D transformation. *Aggregation* is an averaging procedure which combines redundant overlapping blocks for the final denoising output at each specific pixel position.

B.

As patches overlap with each other, the optimal aggregation problem emerges naturally when denoised patches are going to be mapped back to their original position. In [10], one-dimensional formulations are used. It was found that the optimal weight is inversely proportional to the number of retained coefficients after hard thresholding. It is quite appealing as optimal estimates specifically rely on the assumption that the utilized transforms provide sparse decompositions. In fact, the aggregation strategy used by BM3D is in agreement with this result.

However, the sparsity of local 3D coefficients in the BM3D framework is mainly due to the 1D haar wavelet along the third dimension within a group. This is verified as only fractional PSNR decrease is observed without noticeable visual quality degradation when we turn off the aggregation part of the BM3D algorithm.

The 2D separable DWT used for local patch graphical information presentation is effective for point singularities, but not linear singularities. This leads to smearing-mud-like artifacts when spatial domain aggregation is finished. It is especially serious when the image is suffered from high level noise. The requirement of a more suitable local 2D transform motivates us to consider the finite Radon transform for preserving geometrical structure.

C. Finite Radon transform

The finite Radon transform origins from the combinational algebra and is first introduced for image coding [11]. It is later found to be very effective at linear singularity representation by multi-resolution geometrical analysis researchers [4].

Denote $Z_p = \{0,1,\cdots,p-1\}$ and $Z_p^* = \{0,1,\cdots,p-1,p\}$, where p is a prime number. Note that Z_p is a finite field with modulo p operations. The FRAT [4] of a real function on the finite grid Z_p^2 is defined as

$$\begin{aligned} r_{k,l} &= \frac{1}{\sqrt{p}} \sum_{(i,j)\in L_{k,l}} f[i,j] \\ &= \frac{1}{\sqrt{p}} \sum_{i=0}^{p-1}\sum_{j=0}^{p-1} f[i,j]\delta_{L_{k,l}} \\ &= < f, \varphi_{k,l} >, \qquad k \in Z_p^*, l \in Z_p, \end{aligned} \tag{1}$$

where $L_{k,l}$ denotes FRAT lines on Z_p^2: $\begin{cases} L_{k,l} = \{(i,j): j = ki + l (\mathrm{mod}\, p), i \in Z_p\}, 0 \le k < p \\ L_{p,l} = \{(i,j): j \in Z_p\} \end{cases}$, $\varphi_{k,l}$ denotes the basis on Z_p^2 corresponding to $L_{k,l}$: $\delta_{L_{k,l}}[i,j] = \begin{cases} 1, & [i,j] \in L_{k,l} \\ 0, & [i,j] \notin L_{k,l} \end{cases}$, $\varphi_{k,l} = \frac{1}{\sqrt{p}}\delta_{L_{k,l}}$. For simplicity, we denote it as $T_{FRAT}(f) = r$.

Fig. 1 illustrates all FRAT lines when $p = 11$. One can easily observe that FRAT is suitable for line singularity presentation for its good energy concentration property. It is a near-orthogonal transform that a $p \times p$ image patch is mapped to a $p \times (p+1)$ coefficient matrix in the transform domain. Do constructed an optimal ordering for FRAT [4] by selecting the optimal normal vector from the normal vector parameter sets which all correspond to the same slice as inspired by the projection-slice theorem. We assume the optimal ordering FRAT is adopted in this paper.

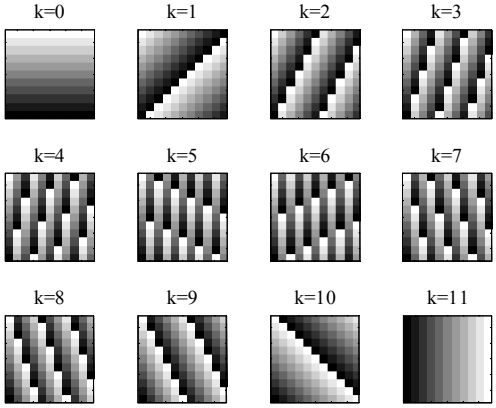

Fig. 1 FRAT lines. (p=11).

FRAT is an invertible transform that has perfect reconstruction. The inverse finite Radon transform (IFRAT) can be obtained by the finite back-projection (FBP) operator. For any coefficient matrix $r: \{r_{k,l}\}_{k \in Z_p^*, l \in Z_p}$ in the FRAT domain, we have

$$FRAT_r^{-1}[i,j] = FBP_r[i,j] = \frac{1}{\sqrt{p}} \sum_{(k,l)\in P_{i,j}} r_{k,l} \qquad (i,j) \in Z_p^2, \tag{2}$$

where $P_{i,j}$ denotes the sets of indexes that all go through a point $(i,j) \in Z_p^2$. Note that both FRAT and IFRAT have very low computational complexity that only summations are needed.

III. PROPOSED DENOISING ALGORITHM

In order to provide better presentation for line singularities in image and thus for better denoising visual quality, we propose a two-stage patch-based over-complete denoising algorithm using FRAT as the local 2D transform. Fig. 2 shows the flowchart of the algorithm consisting of the basic stage and the final stage. Each stage consists of three steps, namely l_2 -norm based patch grouping, local 3D transform with FRAT, and transform domain noise attenuation and aggregation, respectively.

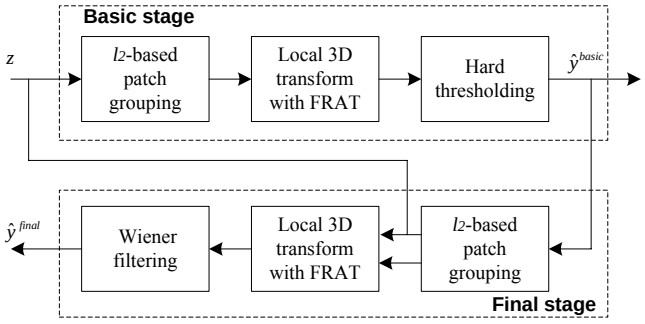

Fig. 2 Flowchart of the proposed algorithm.

The denoising problem of a noisy image $z: X \to R$ can be written as,

$$z(x) = y(x) + n(x),\ x \in X \tag{3}$$

where x denotes the 2-D spatial coordinate that belongs to the image domain X , y is the original image to be estimated, and n is Independent and identically distributed (i.i.d.) zero-mean Gaussian noise with variance σ_n^2 . With Z_x , we denote a local $N_1 \times N_1$ patch extracted from z , where x is the coordinate of the canter pixel and N_1 is the fixed patch size. The superscripts "*basic*" and "*final*" are used to distinguish variables in different stages. Common variables in the two stages are referred without superscript.

A. l_2 -norm based patch grouping

The strong mutual similarity within a group helps for better transform domain sparsity which is essential for success separation of noise and signal. Rather than using clustering methods that partition features in vector space, block matching is adopted in our algorithm for simplicity and effectiveness.

Pairwise similarity testing of two patches Z_a and Z_b is measured by l_2 norm defined as,

$$\|Z_a - Z_b\|_2^2 = \sum_{i=1}^{N_1 \times N_1} \left(z_{a_i} - z_{b_i}\right)^2 \tag{4}$$

as it can be easily shown that

$$E\left[\|Z_a - Z_b\|_2^2\right] = E\left[\|Y_a - Y_b\|_2^2\right] + 2\sigma_n^2 \tag{5}$$

It means statistically, the l_2 distance between two noisy patches maintains the same as that between the clean patches, up to a constant term $2\sigma_n^2$. It thus can be used as a robust criterion for selecting similar patches.

In the basic stage, the l_2 distance for each of the reference patch Z_{x_R} and its nearby neighbors within a $N_s \times N_s$ search window is computed. The most similar N_2 patches with least l_2 distances are stacked to form a group $G_{x_R}^{basic}$ with a fixed size of $N_1 \times N_1 \times N_2$. Locations of selected patches $S_{x_R}^{basic}$ should be recorded as they have to be returned back after the noise attenuation.

In the final stage, the l_2 -norm based patch grouping is performed on the basic estimated image $\hat{y}^{basic}$, as we believe noise in $\hat{y}^{basic}$ has been greatly reduced which can thus be used to obtain a more reliable grouping result. After obtaining the locations of selected patches $S_{x_R}^{final}$, we should have two groups extracted from the basic and noisy images, namely pilot group $PG_{x_R}^{final}$ and noisy group $NG_{x_R}^{final}$, respectively. Wiener filtering will then be performed as described in section 3.C.

B. Local 3D transform with FRAT

Patch transforms is of vital importance for both preserving local image structure and efficient aggregation. Local 2D transform which provides efficient presentation of image singularities and gives sparse transform coefficients is expected to give better denoising results. The 2D and 1D transforms are different in that,

- 2D Transform: codes the local geometrical information of current patch. Better sparsity means better for signal and noise separation.
- 1D Transform: enhances sparsity by coding the similarities between similar patches within the group.

Though the authors of BM3D mentioned that the selection with different local 2D transforms has little influence on the denoising effect, this observation is valid only for 2D separable transforms. The use of FRAT as local 2D transform will give us a new perspective of graphical information representation.

The local 3D transform with FRAT step is same for the basic and final estimate stages. We first perform T_{FRAT} to all the patches selected in current group G_{x_R} . It is then followed by a 1D haar DWT to the FRAT coefficients along the third dimension. Denote this operator as T_{FRAT}^{3D} , the local 3D coefficient is obtained as,

$$C_{x_R} = T_{FRAT}^{3D}\left(G_{x_R}^{basic}\right) = T_{Haar}\left(T_{FRAT}\left(G_{x_R}\right)\right) \tag{6}$$

This 3D transform is applied to all groups in the basic and final stages (i.e., $G_{x_R}^{basic}$, $PG_{x_R}^{final}$ and $NG_{x_R}^{final}$), the corresponding coefficients are denoted as $C_{x_R}^{basic}$, $PC_{x_R}^{final}$ and $NC_{x_R}^{final}$, respectively .

C. Transform domain noise attenuation and aggregation

After the patch grouping and local 3D transform with FRAT, the obtained coefficient C_{x_R} is sparse with respect to the image. However, the additive white Gaussian noise (AWGN) in the local 3D transform domain is also AWGN.

In the basic estimate stage, we can easily remove the noise by thresholding, i.e.,

$$\hat{C}_{x_R}^{basic} = Thr_{\lambda_{3D}\sigma_n}\left(C_{x_R}^{basic}\right) \tag{7}$$

where $Thr_{\lambda_{3D}\cdot\sigma_n}$ denotes a hard thresholding operator with threshold $\lambda_{3D}\cdot\sigma_n$ .In the final estimate stage, denoised coefficients is obtained by point-wise wiener filtering in the 3D domain as

$$\hat{C}_{x_R}^{final} = \frac{\left|PC_{x_R}^{final}\right|^2}{\left|PC_{x_R}^{final}\right|^2 + \sigma_n^2} \cdot NC_{x_R}^{final} \tag{8}$$

Then, denoised image patches are transformed back to the spatial domain using the inverse FRAT and inverse Haar transform:

$$\hat{G}_{x_R} = T_{FRAT}^{3D^{-1}}\left(G_{x_R}\right) = T_{FRAT}^{-1}\left(T_{Haar}^{-1}\left(\hat{C}_{x_R}\right)\right) \tag{9}$$

to obtain the denoised image patches.

Different from the complicated aggregation step, no weighting is performed (in other words, equal weighting is used). The estimated image is obtained as

$$\hat{y}(x)=\frac{\sum_{x_R\in X}\sum_{x_m\in S_{x_R}}\hat{G}_{x_R}(x)}{\sum_{x_R\in X}\sum_{x_m\in S_{x_R}}\chi_{x_m}(x)},\qquad \forall x\in X \tag{10}$$

where $\chi_{x_m}: X\to\{0,1\}$ is an indicator to see whether current pixel locates in the selected patches S_{x_R}. This aggregation strategy works to both basic and final stages.

D. Implementation issues

➢ *Reference block sampling for speedup*

To perform efficient denoising, one pixel in every $N_{step}\times N_{step}$ pixels is used in the image grid. With careful selection of N_{step} value, the over-completeness of the algorithm guarantees that there is no performance degradation.

➢ *Parameter setting*

In the proposed method, the number of parameters is greatly reduced as compared with BM3D. Only five parameters are needed and their values are fixed in all of our experiments. They are $N_1=7$ for image size of 256×256 and $N_1=11$ for image size of 512×512, $N_2=16$, $N_s=39$, $\lambda_{3D}=2.7$ and $N_{step}=3$.

➢ *Computational complexity*

The number of operations per pixel is approximately equal to

$$3C_{T_{2D}}+\frac{2\left(N_1^2+N_2\right)N_s^2}{N_{step}^2}+\frac{3\left(N_2C_{T_{2D}}+N_1^2C_{T_{1D}}\right)N_s^2}{N_{step}^2} \tag{11}$$

where C_T denotes the number of multiplications required for a transform T. As only summation operations are required in the FRAT, $3N_1^2\left(1+\frac{N_2N_s^2}{N_{step}^2}\right)$ multiplications per pixel are saved as compared with BM3D.

IV. Experimental results

We test the effectiveness of our proposed algorithm with six benchmark denoising schemes. The BLS-GSM [2] algorithm is the best wavelet domain denoiser, while hard-thresholding of Curvelet [5] and Contourlet [6] coefficients are representatives of multi-resolution denoising algorithms. The non-local mean [7], optimal spatial adaptation (SA) [12] and BM3D [8] represents the state-of-the-art spatial domain denoising algorithms.

Table1 compares PSNR as well as SSIM results for "house" and "lena" at different noise levels. The last column indicates PSNR difference averaged at the four noise levels as compared with the proposed method.

TABLE I

PSNR (IN DB) AND SSIM COMPARISON OF DIFFERENT DENOISING ALGORITHMS.

Methods	Noise Standard Deviation σ_n				Averaged Δ
	10	20	30	40	
	House 256*256				
BLS-GSM[2]	35.34	32.37	30.56	29.24	**-0.86**
	(0.893)	(0.846)	(0.814)	(0.787)	**(-0.025)**
Curvelet[5]	34.42	31.15	29.24	27.79	**-2.09**
	(0.876)	(0.803)	(0.759)	(0.722)	**(-0.070)**
Contourlet[6]	34.57	31.98	29.92	28.46	**-1.51**
	(0.879)	(0.841)	(0.812)	(0.784)	**(-0.031)**
NLM[7]	35.01	32.26	27.23	19.69	**-4.19**
	(0.885)	(0.849)	(0.685)	(0.273)	**(-0.187)**
SA[12]	35.24	33.08	31.49	30.11	**-0.26**
	(0.884)	(0.856)	(0.833)	(0.810)	**(-0.014)**
BM3D[8]	36.73	33.79	32.12	30.77	0.61
	(0.921)	(0.872)	(0.848)	(0.830)	(0.008)
Proposed	35.87	33.16	31.59	30.34	0
	(0.911)	(0.863)	(0.841)	(0.823)	(0)
	Lena 512*512				
BLS-GSM[2]	35.62	32.67	30.89	29.62	0.21
	0.966	0.936	0.907	0.880	**(-0.002)**
Curvelet[5]	34.61	31.63	29.78	28.47	**-0.87**
	0.957	0.914	0.877	0.842	**(-0.026)**
Contourlet[6]	34.83	32.02	30.24	28.92	**-0.49**
	0.962	0.928	0.896	0.864	**(-0.011)**
NLM[7]	34.80	32.10	27.41	19.77	**-3.48**
	0.956	0.929	0.868	0.593	**(-0.005)**
SA[12]	35.05	32.53	30.84	29.56	**-0.01**
	0.965	0.935	0.904	0.872	**(-0.004)**
BM3D[8]	35.90	33.05	31.25	30.01	0.56
	0.969	0.940	0.912	0.889	(0.004)
Proposed	35.22	32.45	30.77	29.54	0
	0.966	0.936	0.909	0.884	(0)

From Table 1, we see that our proposed algorithm has good PSNR and SSIM results, consisting outperforms other algorithms except BM3D. Although PSNR is lower than BM3D, reconstructed images have better visual quality. This can be verified by the relative high SSIM values. In fact, it is not hard to understand why the proposed method has lower PSNR as compared with BM3D. The BM3D aims at point-wise optimal reconstruction of the noise corrupted signal that favors the point-wise PSNR calculation. On the other hand, our algorithm aims at better preservation of the local geometrical structure which will be more important for obtaining high quality image subjectively.

Fig.3 shows the denoised images ($\sigma_n=25$ for "Lena" image) by different methods. We see that all transform domain methods, i.e., BLS-GSM, curvelet, and contourlet, suffer from severe ringing-like artifacts. The pixel-wise processing mechanism of NLM, SA and BM3D results in smearing-mud-like artifacts. We can observe that the proposed method generates more smoothed denoising output that has the best visual quality. It is particularly obvious that the fake tumor resulted by artifacts at Lena's chin looks more natural using the proposed method.

Fig.3. Subjective quality comparison of denoised images different denoising methods. (a) Noisy image, PSNR =20.17 dB, SSIM = 0.600; (b) BLS-GSM, PSNR = 31.70dB, SSIM = 0.921; (c) Curvelet, PSNR =30.63 dB, SSIM = 0.895; (d) Contourlet, PSNR = 31.06dB, SSIM = 0.912 ; (e) NLM, PSNR = 30.44dB, SSIM = 0.910; (f) SA, PSNR = 31.62dB, SSIM = 0.920 ; (g) BM3D, PSNR = 32.06dB, SSIM = 0.926; (h) Proposed, PSNR = 31.53dB, SSIM = 0.922.

V. CONCLUSIONS

In this paper, we propose a two-stage patch based over complete denoising algorithm using finite Radon transform for preserving line singularities information. Experimental results show that the denoised images have the best visual quality as compared with those obtained by BM3D and other state-of-the-art denoising algorithms. Good PSNR and SSIM results are also obtained at low computations. Our future work is to develop a multi-resolution generalization of the algorithm. Meanwhile, how to reduce the "wrap-around" effects of FRAT to further improve the visual quality is the main focus of our further study.

ACKNOWLEDGMENT

This work is supported by the Centre for Multimedia Signal Processing, Department of Electronic and Information Engineering and the Hong Kong Polytechnic University. Liu Yun-Xia acknowledges the research studentships provided by the University.

REFERENCES

[1] D. Donoho and I. Johnstone.: Ideal spatial adaptation via wavelet shrinkage, Biometrika, Dec. Vol. 81, pp.425−455, 1994.

[2] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain, " IEEE Trans. Image Process., vol. 12, no. 11, pp. 1338–1351, Nov.2003.

[3] EJ Candes and D.L. Donoho: Ridgelets: a key to higher-dimensional intermittency? [J]. Philosophical Transactions of the Royal Society of London Series A, 357(1760), pp. 2495-2509, 1999.

[4] M.N. Do, M. Vetterli, "The finite ridgelet transform for image representation, " IEEE Trans. Image Process., vol. 12, no. 1, pp. 16-28, Jan.2003

[5] J.L. Starck, E. J. Candes, and D.L. Donoho, "The Curvelet Transform for Image Denoising, " IEEE Trans. Image Process., vol. 11, no. 6, pp. 670-684, Jun.2002

[6] A. L. da Cunha, J. Zhou, M. N. Do, "The Nonsubsampled Contourlet Transform: Theory, Design, and Applications," IEEE Trans. Image Process., vol. 15, no. 10, pp. 3089–3101, Oct. 2006.

[7] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one, " Multiscale Model. Simul., vol. 4, no. 2, pp. 490-530, Jul.2005.

[8] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3D transform-domain collaborative filtering. IEEE Trans. on Image Processing. vol.16, no.8, pp. 2080-2095, August 2007.

[9] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. "Image quality assessment: From error visibility to structural similarity, " IEEE Trans. Image Process., vol. 13, no. 4, pp. 600-612, Apr.2004.

[10] O. G. Guleryuz, "Weighted Averaging for Denoising With. Overcomplete Dictionaries, " IEEE Trans. on Image Process., vol. 16, no. 12, pp. 3020-3034, Dec.2007.

[11] G. Beylkin.: Discrete Radon transform, IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-35, no. 2, pp.162-172, 1987.

[12] C. Kervrann and J. Boulanger, "Optimal spatial adaptation for patch based image denoising, " IEEE Trans. Image Process., vol. 15, no. 10, pp. 2866-2878, Oct.2006.