A HYBRID EDGE-PRESERVING IMAGE SMOOTHING SCHEME FOR NOISE REMOVAL

Jinghong Zheng and Zhengguo Li

Signal Processing Department, Institute for Infocomm Research, 1 Fusionopolis Way, Singapore

ABSTRACT

In this paper, we propose a new image denoising scheme that is an integration of a content-adaptive guided filter and a collaborative Wiener filter. The proposed scheme consists of two steps. First a content-adaptive guided filter, which smoothes image based on spatial similarity within a local window, is applied. The content-adaptive guided filter can efficiently preserve edges while smoothing noise. A preliminary estimation of noise-free image can be obtained by the content-adaptive guided filter. In the second step, a patch-grouping based collaborative Wiener filter is adopted to exploit non-local similarity, and outputs final denoised image. Compared to the state-of-the-art denoising scheme, BM3D, the proposed method is more efficient in computation. Moreover, simulation results have shown that the proposed method can achieve comparable PSNR values and better visual quality on denoising of textural images.

1. INTRODUCTION

Denoising is a widely studied topic in the filed of image processing and computational photography. Numerous denoising schemes have been proposed. Local filters are used in many denoising schemes [1] [2] [3]. The pixel similarity with a small local window is analyzed by the local filters. Small pixel variations that may be caused by noise are smoothed while large intensity changes are preserved. The local filters are very efficient in computation, but the performance is restricted by the limited information in local areas. Details are often smoothed by the local filters as well as noise. To overcome this limitation, a content adaptive guided filter was introduced [4]. An edge-aware weight is included to preserve sharp edges. A joint bilateral filter scheme was proposed in [5], which uses the temporal correlation between images to distinguish noise and details. Then the details can be kept after denoising. Besides the local pixel correlation, the global pixel correlation is also utilized for denoising [6] [7]. Steering kernel transform is applied to achieve a sparse representation. By preserving high-magnitude transform coefficients that represent the true signals and discarding the small-magnitude coefficients, most noise signals can be removed. But sometimes a fixed 2-D transform cannot efficiently attain a sparse representation of certain image structure [11].

Compared to pixel correlation, structure correlation is more robust to noise. Patch-based non-local schemes were proposed to investigate the structure correlations in a noise image [8] [9]. The patch-based non-local schemes are based on the assumption that structure/content would be repeated in an image. Linear combining similar structure/content can obtain an approximation of true signal. In the BM3D [8], similar patches are grouped together. Then a 3-D transform is applied on grouped patches to achieve a sparse representation through local correlation within single patch and non-local correlation among similar patches. The similar structure of grouped patches brings the edge information to low frequency in transform domain. A filter is applied in transform domain to shrink high frequency coefficients. Thus the details are preserved while noise removal. On the other hand, the performances of the patch-based non-local schemes are heavily dependent on the patch similarity [11] [12]. If there are not enough similar patches or patches are wrongly matched, the signals would not be sparsely represented in transform domain. Details will be smoothed. To prevent the disturb of noise and find truly matched patches, transform coefficients [8] and the structural similarity index [12] are adopted to measure the similarity between patches. Another concern of the BM3D is its high computational cost.

In this paper, we proposed a new denosing scheme to overcome this limitation of patch-based non-local denoising schemes. The proposed method is based on an observation that the patch grouping in a textural noise image may be not accurate. Due to the complex structure and noise, the patches that are not similar may be grouped together. Thus the proposed scheme avoids the patch grouping in a noise image. In the first step, a content-adaptive guided filter proposed in [4] is applied on the input noise image. The content-adaptive guided filter includes an edge-aware weight that can efficiently preserve the edges. A preliminary estimation of the true signal can be obtained by the content-adaptive guided filter. Then in the second step, a patch-grouping based 3-D collaborative Wiener filter is adopted to exploit the non-local similarity. To find truly matched patches, the patch grouping is conducted based on the preliminary estimation that is outputted by the first step. As there is less noise in the preliminary estimation, the similar patches can be well grouped. A 3-D transform is applied on grouped patches. Then a Wiener filter, of which coefficients are derived from the preliminary estimation, is used to shrink the transform coefficients. Through inverse transform of the shrunk coefficients, a denoised image can be generated. Similar as the BM3D, the proposed algorithm also adopts a hybrid structure to achieve a sparse representation. Both the local pixel similarity and the non-local structure similarity are utilized in the proposed algorithm. As the content-adaptive guided filter has good edge-preserving performance, the proposed denoising scheme is especially suitable for denoising textural images. Furthermore, the proposed method is more efficient in computation than the BM3D, as the time-consuming 3-D collaborative hard-threshold filter in the BM3D is replaced by the content-adaptive guided filter. Compared to the 3-D collaborative hard-threshold filter, the time cost of the content-adaptive filter is negligible. The total processing time of denoising on black-and-white images can be saved about 49% by using the proposed algorithm.

The rest of this paper is organized as follows. The denoising scheme, BM3D, is reviewed in Section 2. The proposed algorithm is presented in Section 3. Some simulation results are displayed in

Section 4 to prove the performance of proposed algorithm. Finally, the conclusion remarks are given in Section 5.

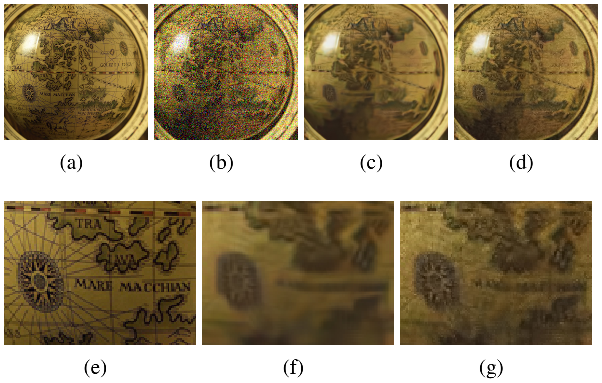

Fig. 1: Visual comparison of denoising results. (a) original image; (b) noise image ($\sigma^2 = 0.03$); (c) result of the BM3D; (d) result of the proposed scheme; (e, f, g) enlarged parts of (a, c, d) respectively.

2. IMAGE DENOISING BY 3-D TRANSFORM-DOMAIN COLLABORATIVE FILTER

The proposed algorithm can be considered as a modification of the BM3D. Thus a brief review of the BM3D is presented in this section.

The BM3D has two steps: a patch-grouping based 3-D collaborative hard-threshold filter, and a patch-grouping based 3-D collaborative Wiener filter. In the BM3D, a noise image Z is divided into fixed-size blocks. Let the currently processing block being denoted as Z_B. Through block matching within a noise image, a set of blocks that are similar to the Z_B can be found. Then a 3-D collaborative hard-threshold filter is applied on the grouped blocks. An array of denoised blocks $\mathbf{Z}_B^{ht}$ can be yielded as:

$$\mathbf{Z}_B^{ht} = \Gamma^{-1}(\gamma(\Gamma(\mathbf{Z}_B))), \tag{1}$$

where $\mathbf{Z}_B$ denotes a 3-D array of grouped blocks, and Γ represents a 3-D transform. γ is a hard-threshold operator. A basic estimate of the denoised image Z^{ht} can be obtained.

Then in the second step of the BM3D, the block matching is conducted based on the denoised image Z^{ht}. For a block Z_B^{ht}, a group of blocks $\bar{\mathbf{Z}}_B^{ht}$ can found through block matching. Then a 3-D collaborative Wiener filter is applied on the corresponding blocks of input noise image. The denoised blocks $\mathbf{Z}_B^{wie}$ can be estimated as:

$$\mathbf{Z}_B^{wie} = \Gamma^{-1}(\mathbf{W}(\Gamma(\bar{\mathbf{Z}}_B))), \tag{2}$$

where $\bar{\mathbf{Z}}_B$ is a block group that contains the corresponding blocks with $\bar{\mathbf{Z}}_B^{ht}$, and $\mathbf{W}$ is Wiener shrinkage coefficients. It is computed from the 3-D transform coefficients of $\mathbf{Z}_B^{ht}$ as:

$$\mathbf{W} = \frac{|\Gamma(\bar{\mathbf{Z}}_B^{ht})|^2}{|\Gamma(\bar{\mathbf{Z}}_B^{ht})|^2 + \sigma^2}, \tag{3}$$

where σ is the noise variance.

In the BM3D, grouping similar blocks is conducted in both of the first step and the second step. By applying a 3-D collaborative filter on the grouped blocks, the local correlation within a block and the non-local correlation among similar blocks are used to achieve a sparse representation. The first step generates a basic estimate of a true image. Then this basic estimate is used as a reference image in the second step to improve the reliability of block matching and compute the Wiener coefficients. Thus the quality of the basic estimate is critical for the performance. Improving the quality of the basic estimate can further improve the quality of final output image. A challenging for the first step of the BM3D is the block matching [11]. With the interference of noise, the block matching in textural image will be not reliable. In this case, the quality of the basic estimate would be dropped. Thus we propose a content-adaptive guided filter which mainly relies on the local information to generate the basic estimate. The unreliable block-matching in a textural noise image is avoided in the proposed method.

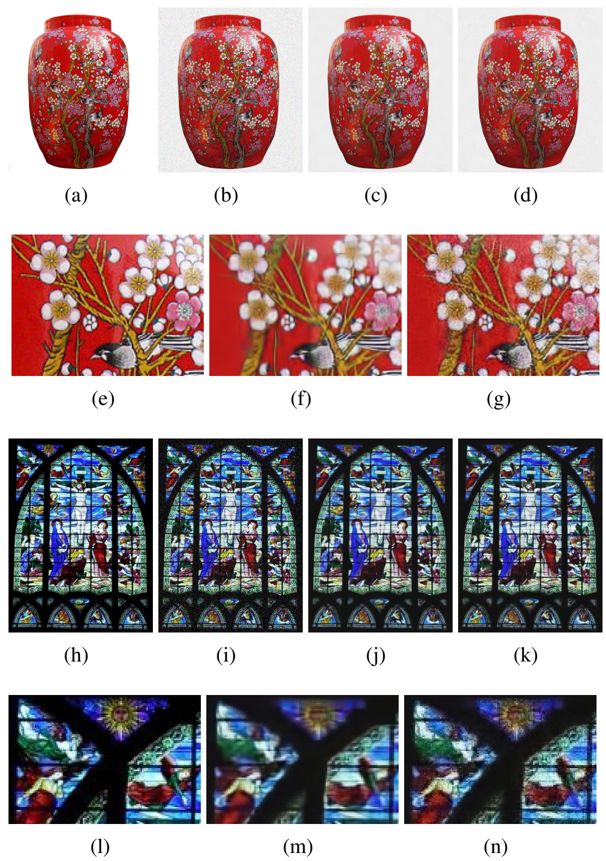

Fig. 2: Visual comparison of denoising results. (a, h) original images; (b, i) noise images; (c, j) results of the BM3D; (d, k) results of the proposed scheme; (e, f, g) enlarged parts of (a, c, d) respectively; (l, m, n) enlarged parts of (h, j, k) respectively.

3. DENOISING BY CONTENT-ADAPTIVE GUIDED FILTER AND 3-D COLLABORATIVE WIENER FILTER

The proposed denoising algorithm uses a hybrid approach to exploit the local and non-local correlations for denoising. There are two parts in the proposed method: a content-adaptive guided filter [4] and a 3-D collaborative Wiener filter [8].

3.1. Content-adaptive guided filter

The content-adaptive guided filter used in the first step is based on the observation that the human visual system (HVS) is more sensitive to pixels at edges than those pixels in smooth regions. Thus the edges should be preserved during noise smoothing [4].

A denoised image $\hat{Z}$ can be expressed as a linear transform of the noise image Z in a local window Ω [1] [4]:

$$\hat{Z}(p) = a_{p'}Z(p) + b_{p'}, \forall p \in \Omega(p'), \tag{4}$$

where $a_{p'}$ and $b_{p'}$ are two constants in Ω.

To determine the linear coefficients $(a_{p'}, b_{p'})$, a cost function on noise image Z and denoised image $\hat{Z}$ is defined as

$$\sum_{p\in\Omega(p')} [(a_{p'}Z(p) + b_{p'} - Z(p))^2 + \frac{\lambda}{\Phi_Z(p')} a_{p'}^2]. \tag{5}$$

Here, the edge-aware weight Φ_Z is computed by the local variances σ_Z^2 of 3×3 windows of all pixels as follows:

$$\Phi_Z(p') = \frac{1}{N}\sum_{p=1}^{N} \frac{(\frac{\sigma_Z^2(p')+\nu_1}{\mu_Z^2(p')+\nu_2})^\zeta}{(\frac{\sigma_Z^2(p)+\nu_1}{\mu_Z^2(p)+\nu_2})^\zeta}, \tag{6}$$

where N is the total pixel number, and μ is mean pixel value. ν_1, ν_2 and ζ are three constants. The value of ν_1 is $(0.001{*}L)^2$ with L being the dynamic range of input image. The value of ν_2 is 10^{-9}. The value of ζ is selected as 0.75 in this paper.

A denoised image $\hat{Z}$ can be obtained by the content-adaptive guided filter as:

$$\hat{Z}(p) = \bar{a}_p Z(p) + \bar{b}_p, \tag{7}$$

where $\bar{a}_p$ and $\bar{b}_p$ are average values of $a_{p'}$ and $b_{p'}$ among all overlapped windows that a pixel p is involved.

Compared to the hard-threshold collaborative filter used in the first step of BM3D, the content-adaptive filter can achieve better edge preserving performance than the hard-threshold collaborative filter for textural images. In the hard-threshold collaborative filter, edge preserving is achieved through non-local block correlation. Blocks with similar structure are grouped together. The repeated edge information can be easily distinguished from random noise in transform domain and then be preserved. But in a textual image, it is challenging to group truly matched blocks due to the interference of noise. In this case, the sparse representation between edge and noise will not be available in transform domain. The hard-threshold collaborative filter will smooth details as well as noise. The content-adaptive guided filter uses the local information to remove noise. Due to the edge-aware weight, the details can be efficiently preserved. As a trade-off, the noise near edges will be preserved as well. But the noise would be masked by the nearby edges and is not easily to be perceived.

Moreover, the content-adaptive guided filter is more efficient in computation. The block matching in the hard-threshold collaborative filter is time-consuming, especially when the search range is large. The 3D transform and inverse transform are complex and require large memory to process. On the other hand, the complexity of the content-adaptive guided filter is $O(N)$ for an image with N pixels. Because of the box filter in [1], the window size does not effect the computational cost of the content-adaptive guided filter.

Table 1: PSNR values of the denoised textural images

Image	σ^2	BM3D	Proposed Method
Globe	0.02	25.93	26.07
	0.03	25.00	25.28
	0.04	24.40	24.55
	0.05	23.82	23.83
Vase	0.02	24.03	24.26
	0.03	22.44	22.74
	0.04	21.29	21.63
	0.05	20.43	20.78
Window	0.02	24.49	24.48
	0.03	22.87	22.97
	0.04	21.67	21.86
	0.05	20.75	21.02

3.2. Collaborative Wiener Filter

In the second step of the proposed algorithm, a 3-D collaborative Wiener filter, which is presented in the BM3D [8], is applied to exploit the non-local image correlation.

The collaborative Wiener filter exploits non-local image correlation through block matching. The block matching is conducted on the output image of content-adaptive guided filter, $\hat{Z}$, to group the similar blocks. For a block in the image $\hat{Z}$, which is denoted as $\hat{Z}_B$, a group of similar blocks $\hat{\mathbf{Z}}_B$ can be formed through block matching. As the noise in the image $\hat{Z}$ is attenuated, it is more possible for block matching to find truly matched blocks. Then a block group $\mathbf{Z}_B$ in input image Z, which contains the same position blocks as $\hat{\mathbf{Z}}_B$, can be formed correspondingly. Similar to the BM3D, a 3-D transform is applied on $\mathbf{Z}_B$ to achieve a sparse representation in transform domain. A Wiener filter is used to shrink the 3-D transform coefficients.

If blocks $\hat{\mathbf{Z}}_B$ are flat, most of their transform coefficients will be small magnitude. The corresponding Wiener coefficients $\mathbf{W}$ will be close to 0. Then the coefficients of $\mathbf{Z}_B$ will be shrunk to 0 by the Wiener filter. The denoised blocks $\hat{\mathbf{Z}}_B^{wie}$ will be smooth. On the other hand, if the details in $\hat{\mathbf{Z}}_B$ are preserved, the corresponding Wiener coefficient will approximate to 1. Then the details in input image will be preserved after noise removal. As the content-adaptive guided filter used in the proposed algorithm has better edge preserving performance than the collaborative hard-threshold filter of the BM3D, the image $\hat{Z}$ has more detail information than Z^{ht}. Then the Wiener coefficients derived from $\hat{Z}$ will preserve more high frequency coefficients. The final denoised image $\hat{Z}^{wie}$ produced by the proposed scheme will contain more details.

4. SIMULATION RESULTS

In this section, a few images are tested to compare the proposed algorithm with the state-of-the-art denoising scheme, BM3D [8]. Both of the methods are implemented in C programming. The BM3D code is got from [13], which is provided by its authors. The simulation is carried on Dell Precision T7400 with Intel Quad Core

CPU 3.2 GHz and 4GB of RAM. The noise images are generated through artificially adding noise on clean images. Then the noise variance σ is available. In simulation, the parameters of the BM3D are set as the default values in [8]. The window size of content-adaptive guided filter is 3×3. The denoised results on textural images with $\sigma^2 = 0.03$ are presented in Figs. 1 and 2 for visual comparison. As shown in Fig. 1, the proposed method can produce a sharper image than the BM3D. In Fig. 1(f) that is the result of BM3D, the lines and edges are blurred while the details in Fig. 1(g) are visible. The PSNR values of the denoised images are presented in Table 1. It can be seen that the proposed algorithm can achieve slightly higher PSNR values than the BM3D for textural images.

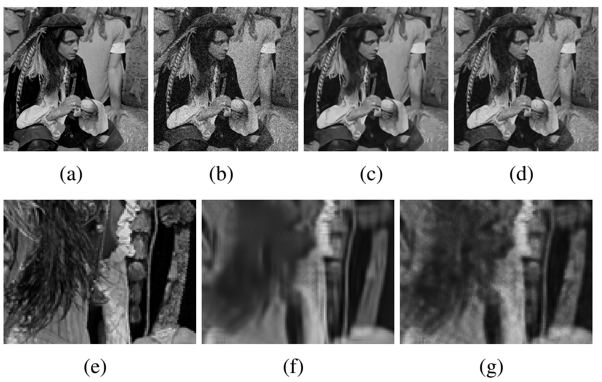

Fig. 3: Denoising results of "Man". (a) original image; (b) noise image; (c) result of the BM3D (PSNR = 24.70dB); (d) result of the proposed scheme (PSNR = 24.65dB); (e, f, g) enlarged parts of (a, c, d) respectively.

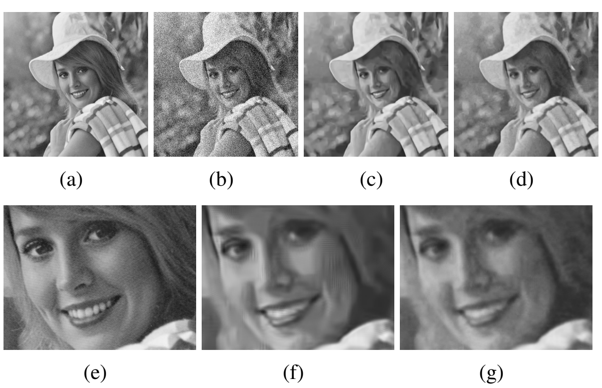

Fig. 4: Denoising results of "Elaine". (a) original image; (b) noise image; (c) result of the BM3D (PSNR = 29.15dB); (d) result of the proposed scheme (PSNR = 28.89dB); (e, f, g) enlarged parts of (a, c, d) respectively.

Besides the textural images in Figs. 1 and 2, three classical images that were tested by many denoising schemes are used to evaluate the performance of proposed scheme, as shown in Figs. 3, 4 and 5. The overall visual qualities of the images that are generated by the BM3D and the proposed scheme are very close. When parts of the images are zoomed in, it can be seen that the results of the proposed algorithm can present more small details. On the other hand, the BM3D can achieve better performance on denoising of flat areas. For example, the girl's face and hair are clearer in Fig. 5(d), which is the result of the proposed algorithm. But the background in Fig. 5(d) is not as smooth as the result of the BM3D, Fig. 5(f).

Fig. 5: Denoising results of "Girl". (a) original image; (b) noise image; (c) result of the BM3D (PSNR = 26.92dB); (d) result of the proposed scheme (PSNR = 26.90dB); (e, f, g) enlarged parts of (a, c, d) respectively.

It is worth noting that the proposed scheme can achieve comparable denoising performance to the BM3D but takes much less processing time. BM3D takes $15.21s$ to process the black-and-white image in Fig. 4, whose resolution is 512×512. The first step of the BM3D costs $7.64s$, and the second step costs $7.57s$. For the same image, the content-adaptive guided filter only takes $0.12s$ to denoise. Compare to the first step of the BM3D, the computational time of content-adaptive filter is negligible. The total cost time of the proposed scheme is $7.76s$. Thus the proposed scheme can save up to 49% processing time for the denoising of black-and-white images. As the processing time of both methods are linear with resolution, the bigger the image, the more time saved by the proposed scheme.

5. CONCLUSIONS

A new denoising scheme which includes a content-adaptive guided filter and a 3-D collaborative Wiener filter is proposed. The content-adaptive guided filter is first applied on a noise image to attenuate noise through exploiting pixel correlations in a local window. A preliminarily denoised image can be generated by the content-adaptive filter. Then a 3-D collaborative Wiener filter is applied on the preliminarily denoised image to exploit the non-local structural similarity. Thus both the local and non-local information is used by the proposed scheme to remove noise. As the preliminarily denoised image has less noise and plenty details, the 3-D collaborative Wiener filter can find truly matched blocks through block matching and then achieve a sparse representation in transform domain. The coefficients of the Wiener filter, which are derived from the preliminarily denoised image, can better preserve fine details. Simulation results show that the proposed scheme can achieve comparable performance with the BM3D but with less complexity.

References

[1] K. He, J. Sun, and X. Tang, “Guided image filtering,” IEEE Trans. on Pattern Analysis and Machine Learning, vol. 35, no. 6, pp. 1397-1409, Jun. 2013.

[2] Z. G. Li, J. H. Zheng, Z. J. Zhu, S. Q. Wu, and S. Rahardja, “A bilateral filter in gradient domain,” In IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, pp. 1113-1116, Kyoto, Japan, Mar. 2012.

[3] F. Durand and J. Dorsey, “Fast bilateral filtering for the display of high-dynamic-range images,” ACM Trans. on Graphics, vol.21, no.3, pp.257-266, Aug. 2002.

[4] Z. G. Li, J. H. Zheng, and Z. Zhu, “Content adaptive guided image filtering,” In IEEE Int. Conf. on Multimedia & Expo, Jul. 2014.

[5] G. Petschnigg, M. Agrawala, H. Hoppe, R. Szeliski, M. Cohen, and K. Toyama, “Digital photography with flash and non-flash image pairs,” ACM Trans. on Graphics, vol. 22, no. 3, pp.1-9, Aug. 2004.

[6] S. Yang, W. Min, L. Zhao, and Z. Wang, “Image noise reduction via geometric multiscale ridgelet support vector transform and dictionary learning,” IEEE Trans. on Image Process., vol. 22, no. 11, pp. 4161-4169, Nov. 2013.

[7] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, “Image denoising using scale mixtures of Gaussian in the wavelet domain,” IEEE Trans. on Image Process., vol. 12, no. 11, pp. 1338-1351, Nov. 2003.

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

[9] A. Buades, B. Coll, and J. M. Morel, “A non-local algorithm for image denoising,” In IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, vol. 2, pp. 60-65, 2005.

[10] P. Chatterijee, and P. Milanfar, “Patch-based near-optimal image denoising,” IEEE Trans. on Image Process., vol. 21, no. 4, pp. 1635-1649, Apr. 2012.

[11] H. Talebi, and P. Milanfar, “Global image denoising,” IEEE Trans. on Image Process., vol. 23, no. 2, pp. 755-768, Feb. 2014.

[12] A. Rehman, and Z. Wang, “SSIM-based non-local means image denoising,”, In IEEE Int. Conf. on Image Processing, pp.217-220, 2011.

[13] The BM3D webpage: http://www.cs.tut.fi/ foi/GCF-BM3D/.