CID: Combined Image Denoising in Spatial and Frequency Domains Using Web Images

Huanjing Yue[1*], Xiaoyan Sun[2], Jingyu Yang[1], Feng Wu[3]

[1] Tianjin University, Tianjin, China. {dayueer,yjy}@tju.edu.cn

[2] Microsoft Research Asia, Beijing, China. xysun@microsoft.com

[3] University of Science and Technology of China, Hefei, China. fengwu@ustc.edu.cn

Abstract

In this paper, we propose a novel two-step scheme to filter heavy noise from images with the assistance of retrieved Web images. There are two key technical contributions in our scheme. First, for every noisy image block, we build two three dimensional (3D) data cubes by using similar blocks in retrieved Web images and similar non-local blocks within the noisy image, respectively. To better use their correlations, we propose different denoising strategies. The denoising in the 3D cube built upon the retrieved images is performed as median filtering in the spatial domain, whereas the denoising in the other 3D cube is performed in the frequency domain. These two denoising results are then combined in the frequency domain to produce a denoising image. Second, to handle heavy noise, we further propose using the denoising image to improve image registration of the retrieved Web images, 3D cube building, and the estimation of filtering parameters in the frequency domain. Afterwards, the proposed denoising is performed on the noisy image again to generate the final denoising result. Our experimental results show that when the noise is high, the proposed scheme is better than BM3D by more than 2 dB in PSNR and the visual quality improvement is clear to see.

1. Introduction

With continued advances in CMOS and CCD image sensors [26], noise in captured images is becoming smaller and smaller even in poor lighting conditions. For the noise actually captured in images, the current advanced image denoising approaches, such as BM3D [6] and EPLL-GMM [30], can attenuate it easily while maintaining the image details.

Recently, both industry and academia have proposed transmitting images and videos over wireless networks

without compression [18] or with half compression [16], which avoids the cliff effect in communication, does not require accurate channel estimation, and better supports multicast. In this way, pixels or coefficients are transmitted directly through the channel and thus received images and videos contain additive channel noise. Sometimes, the channel condition is very poor and the variance in the channel noise is even larger than that of transmitted signals, as shown in Fig. 1 (b). Fig. 1 (c) shows that current image denoising approaches, e.g. BM3D, do not work well in this case. Clearly, new approaches are required to address the denoising problem with heavy noise.

When images contain heavy noise, intuitively it is difficult for current single-image based denoising approaches to obtain good results because they only use the correlation within the noisy image, where it is hard to distinguish between the noise and image details. To better solve this problem, we propose using retrieved Web images for denoising. Before this paper, using Web images has been proposed in image colorization [19], image completion [24, 14], sketch to photo [27, 20] and computer generated graph to photo [21]. To the best of our knowledge, there have been few papers that explore using Web images for denoising.

It is not trivial to leverage *extrinsic* correlation (vs. *intrinsic* correlation within a noisy image itself) for denoising. For every noisy image block, we build two 3D data cubes, one using similar blocks in the retrieved Web images (called the *extrinsic cube*) and another using similar non-local blocks within the noisy image (called the *intrinsic cube*). In order to fully exploit both extrinsic and intrinsic correlations, we propose different denoising strategies. The denoising in the intrinsic cube, similar to BM3D, is performed in the frequency domain, whereas the denoising in the extrinsic cube is performed as median filtering in the spatial domain. We get the denoising image by combining the spatial and frequency filtered results in the frequency domain.

Although local feature descriptors extracted from noisy images can be used to retrieve correlated images, the loca-
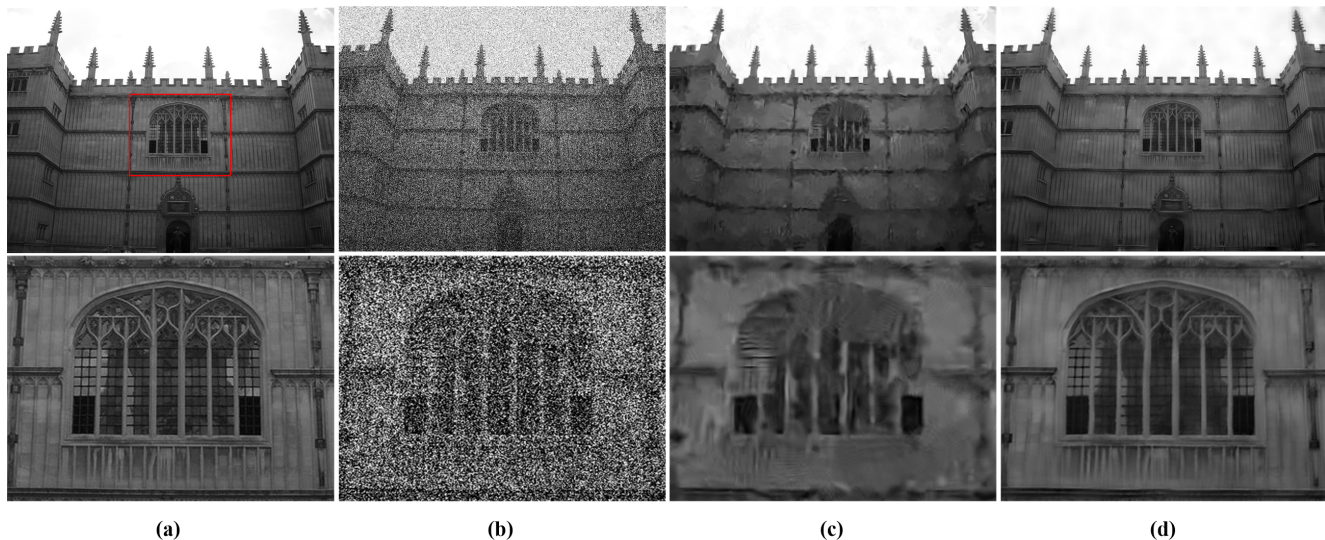

Figure 1. Image denoising results. The first row shows whole images (resized) and the second one presents regions marked by the red box correspondingly. (a) Original image (with a standard deviation of 82.3); (b) Noisy image with a noise standard deviation of 100; (c) BM3D [6] denoising result (26.23 dB); (d) Our denoising result (28.28 dB).

tions of local feature descriptors are not accurate because of heavy noise. Furthermore, the accuracy of building 3D data cubes is greatly affected by the noise. Therefore, we propose extracting local feature descriptors from the denoising image to improve the registration of the retrieved images as well as improve the 3D data building by using the denoising block for matching. At the same time, similar to BM3D, we also use the denoising results to estimate the filtering parameters in the frequency domain. With these improvements, we perform the proposed denoising approach on the noisy image again. As shown in Fig. 1 (d), our approach produces good results even when the noise variance is larger than that of the input image. It is better than BM3D by more than 2 dB.

The rest of this paper is organized as follows. Section 2 gives an overview of related work. Section 3 describes the proposed denoising scheme, including the retrieval of correlated images. Section 4 proposes combined image denoising (CID) in the spatial and frequency domains. The experimental results are presented in Section 5. Section 6 concludes this paper.

2. Related work

2.1. Image denoising

Denoising approaches can be classified into two categories: single-image-based and learning-based.

The single-image-based denoising approaches aim to remove noise by using the correlation within the noisy image. Earlier approaches removed noise pixel by pixel, through the use of bilateral filtering, anisotropy diffusion, or total-variation regularization. Recent, approaches have

performed block by block, such as non-local means filtering [1], sparse representation [8], BM3D [6], or low-rank regularization [7]. The denoising performance has been improved greatly at the block level rather than the pixel level.

Learning-based approaches use certain priors learned from noise-free images for denoising. Elad *et al*. and Mairal *et al*. proposed using sparse representation to train the priors [8, 23]. Roth *et al*. proposed training high-order Markov random field (MRF) models as priors [25]. Zoran *et al*. proposed training finite Gaussian mixture models as priors [30]. In [3], the relationship between a noisy patch and its noise-free version has been studied by a plain multi-layer perceptron (MLP) deduced from a large training image set. All these methods have demonstrated that the priors learned from noise-free images are helpful in denoising.

The above single-image-based and learning-based approaches cannot adequately deal with noisy images that have heavy noise. Thus, the multi-scale strategy is introduced in denoising when the noise level is high [2]. Since the noise is uncorrelated across neighboring pixels, down-sampling will reduce the noise level and make image structures more visible. In addition to the simple down-sampling, there is also a multi-scale sparse representation based denoising algorithm [9]. Furthermore, M. Zontak *et al*. proposed using the patch recurrence across scales to recover the signal [22]. Also, the combined internal and external image denoising methods are proposed in [10, 13].

Compared with learning-based approaches, our approach does not have a training process and tries to always retrieve the most correlated Web images for denoising. The trained models have the advantage of reduced data size. However, the training process also drops a great deal of use-

ful information contained in the training image set. Our approach is expected to have a much better performance by fully utilizing all information in retrieved images. In addition, since there is not a training process, our approach has the potential to use a large-scale image database for denoising.

2.2. Applications using Web images

As more images become available on the Internet, many problems in image processing and computer vision can be better solved by retrieving correlated images. Liu *et al*. proposed using Web images for the colorization of an image [19]. Hays *et al*. proposed using Web images to recover missed regions in an image [14]. Whyte *et al*. proposed using Web images to remove objects from an image [24]. Zhang *et al*. proposed using Web images to enhance personal photos [4]. Eitz *et al*. and Chen *et al*. proposed using Web images to generate images from their sketches [27, 20]. Johnson *et al*. proposed using Web images to generate images from computer generated graphs [21]. Yue *et al*. proposed using Web images to reconstruct an image from its small description [11].

Previously, our basic idea using Web images for denoising is proposed in [12]. That approach is similar to BM3D, which performs denoising in the frequency domain, except using similar blocks obtained from external images retrieved from Cloud. We will denote that scheme as BM3D^C in the following, with superscript C denoting Cloud. However, BM3D^C totally ignores the intrinsic correlation, resulting in poor visual quality due to the incorrect retrieval and limiting the denoising performance. Compared with BM3D^C, there are two significant contributions in this paper. First, the combination of spatial and frequency denoising is proposed to fully leverage the intrinsic and extrinsic correlations. Second, the denoising result from the first step is utilized to improve image registration and 3D data building in the second step, thus leading to a better denoising performance when the noise level is high.

3. Proposed scheme using Web images

Given a noise free image I, its noisy version $\bar{I}$ is produced as

$$\bar{I}(\cdot) = I(\cdot) + n(\cdot), \qquad (1)$$

where $n(\cdot)$ is the additive zero-mean independent identically distributed (i.i.d.) Gaussian noise with variance σ^2, $n(\cdot) \sim N(0, \sigma^2)$. We aim to recover $\tilde{I}$, the estimated noise free version of $\bar{I}$, with the help of retrieved correlated images $\{I_1, I_2, ..., I_K\}$.

The proposed denoising scheme using Web images is shown in Fig. 2. Given a noisy image $\bar{I}$, we first extract its local feature descriptors. In this paper, we use SIFT descriptors [5] but our scheme is not limited to this kind of
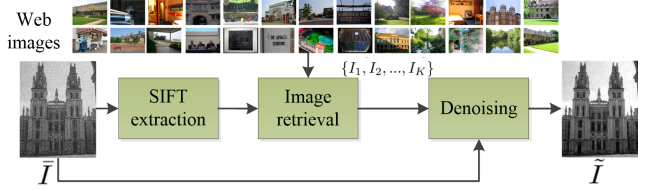

Figure 2. The proposed denoising scheme using Web images.

descriptor. Although the image $\bar{I}$ contains heavy additive Gaussian noise, we have found that the extracted SIFT descriptors are still workable for retrieval because SIFT extraction contains Gaussian filtering and down-sampling that greatly attenuate the noise. The noise mainly affects the locations of SIFT descriptors. In the next section, we will discuss how to improve the location accuracy. In addition, we also discard some SIFT descriptors that have low contrast to reduce the noise effect.

In the image retrieval module of Fig. 2, we use the approach proposed in [17], where a large scale SIFT descriptor is bundled with small scale SIFT descriptors that are completely covered by the large scale SIFT descriptor. Here we refer to them as a SIFT group. The SIFT group is more robust to noise than an individual SIFT descriptor because the relative positions of SIFT descriptors in the group are considered in retrieval. Finally, we match all SIFT groups extracted from the noisy image with those extracted from a candidate image. The images $\{I_1, I_2, ..., I_K\}$ are retrieved as correlated images with many matched SIFT groups.

Our main technical contributions in this paper are in the denoising module. With the retrieved images, we will discuss how to attenuate noise using both intrinsic and extrinsic correlations in the next section.

4. Proposed combined image denoising

Fig. 3 shows the framework of our two-step CID.

4.1. Step 1

The retrieved Web images may have different scales, viewpoints, and orientations. Using the locations of matched SIFT groups, we can estimate the geometric transform from every retrieved image to the noisy image and register it toward the noisy image. Because of heavy noise, a photometric transform cannot be estimated accurately. Thus, we must remove the mean value of every block while building the extrinsic cube, which will reduce the effect of different illuminations.

After registration, for every noisy image block $B_{\bar{I}}$, we first search for the most similar blocks in the retrieved images and assume that there are m found blocks $B_1^e, B_2^e, \cdots, B_m^e$. The similar blocks are obtained by minimizing the $\ell 2$ distance. The superscribe e indicates the ex-
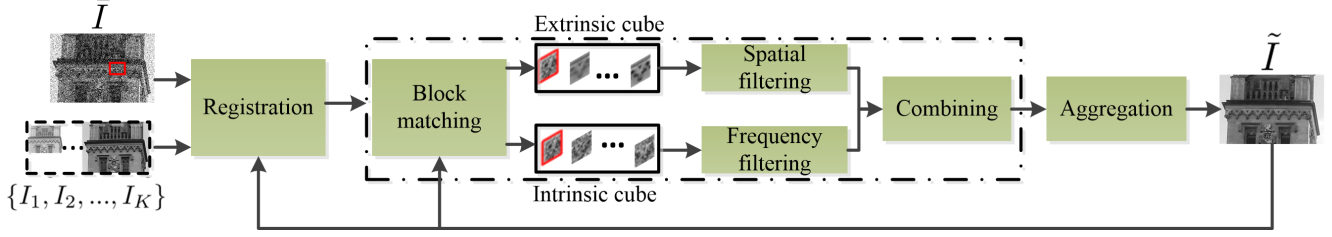
Figure 3. Flowchart of the proposed CID algorithm. The operations surrounded by dash-dot lines are repeated for each noisy block. The noisy image is $\bar{I}$, and its correlated images are $\{I_1, I_2, ..., I_K\}$. The final denoising result is $\tilde{I}$.

trinsic correlation. Similarly, there are n non-local blocks, $B_1^i, B_2^i, \cdots, B_n^i$, found in the noisy image, where the superscribe i indicates the intrinsic correlation. Therefore, two 3D data cubes are built as shown in Fig. 3, denoted as $\mathbf{B}_{3D}^e$ and $\mathbf{B}_{3D}^i$, respectively. Note that B_1^e and B_1^i are the noisy block $B_{\bar{I}}$ itself. The mean value of every block in $\mathbf{B}_{3D}^e$ has been removed to reduce illumination differences.

In the intrinsic cube $\mathbf{B}_{3D}^i$, similar to BM3D [6], we adopt a 3D transform and filter the noise by hard-thresholding in the transform domain. The denoising process is described as follows:

$$\hat{\mathbf{B}}_{3D}^i = \mathcal{T}_{3\mathcal{D}}^{-1}\big(\gamma(\mathcal{T}_{3\mathcal{D}}(\mathbf{B}_{3D}^i))\big), \quad (2)$$

where $\mathcal{T}_{3D}$ represents the 3D transform consisting of a 2D wavelet transform and a 1D Hadamard transform along the third dimension, and γ represents the hard thresholding of the transform coefficients with a given threshold $\lambda_{3D}\sigma$. After the inverse 3D transform $\mathcal{T}_{3\mathcal{D}}^{-1}$, we obtain the denoised 3D intrinsic cube $\hat{\mathbf{B}}_{3D}^i$. The first block $\hat{B}_1^i$ in $\hat{\mathbf{B}}_{3D}^i$ is the denoising result. For simplicity, we use $\hat{B}^i$ to represent $\hat{B}_1^i$.

However, the extrinsic cube $\mathbf{B}_{3D}^e$ that contains only one noisy block and $m-1$ noise free blocks is not sparse in the frequency domain. Furthermore, the hard thresholding in the frequency domain will attenuate many image details while removing noise. In the spatial domain, observing in the third dimension, the noise looks like an impulsive function in the array. Therefore, we adopt spatial filtering, namely, median filtering, to remove the noise from $\mathbf{B}_{3D}^e$. Median filtering can be described as the following $\ell 1$ minimization problem

$$\hat{B}^e(x) = \arg\min ||B^e(x) - \{B_1^e(x), \cdots, B_m^e(x)\}||_1, \quad (3)$$

where x is a 2D spatial coordinate. $\hat{B}^e$, the set of all $\hat{B}^e(x)$, is the denoising result in the extrinsic cube $\mathbf{B}_{3D}^e$.

However, since the retrieved images have different illuminations than the noisy image, as shown in Fig. 3, the denoising result $\hat{B}^e$ may have different illuminations than the noisy image, even though the mean values have been removed. In other words, the spatial filtered image may be

largely distorted in the low frequency. Fortunately, we have observed that the frequency filtered image does a good job of keeping low-frequency information because the found blocks have a similar illumination. Therefore, we propose combining the spatial and frequency filtered images $\hat{B}^e$ and $\hat{B}^i$ as follows

$$\tilde{B}^{1st} = \mathcal{T}_{2\mathcal{D}}^{-1}\big(\frac{\mathcal{T}_{2\mathcal{D}}(\hat{B}^i) \odot W^i + \mathcal{T}_{2\mathcal{D}}(\hat{B}^e) \odot W^e}{W^i + W^e}\big), \quad (4)$$

where $\mathcal{T}_{2\mathcal{D}}$ and $\mathcal{T}_{2\mathcal{D}}^{-1}$ are forward and inverse 2D DCT transform, respectively. The parameter $\odot$ is the point-wise multiplication. Matrix W^i and W^e are the weighting matrices to combine the frequency coefficients of $\hat{B}^i$ and $\hat{B}^e$. The reliability of $\mathcal{T}_{2\mathcal{D}}(\hat{B}^i)$ is decided by noise variance. The higher the noise, the less reliability the high-frequency coefficients have. Therefore, we only combine the first $\mathcal{F}$ frequency coefficients of $\hat{B}^i$ and $\hat{B}^e$. The remaining coefficients are directly copied from $\mathcal{T}_{2\mathcal{D}}(\hat{B}^e)$. The number $\mathcal{F}$ is determined by noise deviation σ as follows

$$\mathcal{F} = \frac{N}{\sqrt{\sigma} + 1}, \quad (5)$$

where N is the number of pixels in one block. We have observed that large coefficients in low frequency domains are more important for recovering the structure of a block. Therefore, the weighting matrix W^i is proportional to the value of $\big(\mathcal{T}_{2\mathcal{D}}(\hat{B}^i)\big)^2/\big(\mathcal{T}_{2\mathcal{D}}(\hat{B}^e)\big)^2$. Similarly, W^e is proportional to the value of $1/W^i$.

After obtaining the denoising image block $\tilde{B}^{1st}$ for every noisy image block $B_{\bar{I}}$ in $\bar{I}$, they are averaged together, resulting in the first step denoising image $\tilde{I}^{1st}$.

4.2. Step 2

After obtaining $\tilde{I}^{1st}$, we can use it in the second step of denoising. The noise in $\tilde{I}^{1st}$ is assumed to be greatly attenuated. Therefore, we can first use it to improve the geometric transform by extracting SIFT descriptors from $\tilde{I}^{1st}$ and recalculating the matched SIFT groups between $\tilde{I}^{1st}$ and the retrieved images. Second, since the pixel values of $\tilde{I}^{1st}$ are closer to the original image, we can estimate photometric transforms between $\tilde{I}^{1st}$ and the retrieved images. In our

Figure 4. Test images used in our experiments denoted from "a" to "g".

scheme, we use pixel values of matching SIFT keypoints to calculate a linear photometric transform.

In addition, the block matching process will be more accurate using $\tilde{I}^{1st}$. Both intrinsic and extrinsic block matching can benefit from this. We would like to point out that adding some noise to $\tilde{I}^{1st}$ will improve the accuracy of block matching because $\tilde{I}^{1st}$ may be too smooth, which will result in incorrect matching in the extrinsic cube. Therefore, we use $\tilde{I}^{basic}$, which is calculated as

$$\tilde{I}^{basic} = \tilde{I}^{1st} + \lambda(\bar{I} - \tilde{I}^{1st}), \qquad (6)$$

to search similar blocks in the registered images. λ here is a weighting factor.

Finally, $\tilde{I}^{1st}$ can be used to improve the frequency filtering by using it to calculate Wiener filtering coefficients. In the second step of denoising, we build two 3D intrinsic cubes. One is built by retrieving similar blocks in $\tilde{I}^{1st}$, denoted as $\hat{\mathbf{B}}_{3D}^{i_{2nd}}$, and the other is built from the noisy image $\bar{I}$ using the same location blocks as matched in $\tilde{I}^{1st}$, denoted as $\mathbf{B}_{3D}^{i_{2nd}}$. The frequency denoising result is obtained via Wiener filtering as follows:

$$\hat{\mathbf{B}}_{3D}^{i_{2nd}} = \mathcal{T}_{3\mathcal{D}}^{wie^{-1}}\left(\mathbf{W}^{wie} \odot (\mathcal{T}_{3D}^{wie}(\mathbf{B}_{3D}^{i_{2nd}}))\right), \qquad (7)$$

where $\mathcal{T}_{3D}^{wie}$ represents a 2D DCT transform and a 1D Hadamard transform along the third dimension. $\mathcal{T}_{3\mathcal{D}}^{wie^{-1}}$ is the inverse 3D transform. The wiener shrinkage coefficient $\mathbf{W}^{wie}$ is determined by

$$\mathbf{W}^{wie} = \frac{|\mathcal{T}_{3D}^{wie}(\tilde{\mathbf{B}}_{3D}^{i_{2nd}})|^2}{|\mathcal{T}_{3D}^{wie}(\tilde{\mathbf{B}}_{3D}^{i_{2nd}})|^2 + \sigma^2}. \qquad (8)$$

Similar to Step 1, the denoising block $\tilde{B}^{2nd}$ is obtained by combining the spatial filtering block and frequency filtering block. Then these blocks are aggregated together to generate the final denoising image $\tilde{I}^{2nd}$.

We would like to point out that when the noise level is high ($\sigma > 80$), we prefer to first down sample the noisy image by a factor of 0.8 and then hallucinate the denoised image to its original resolution.

5. Experimental results

In our test, we used an image dataset containing 505,062 images to simulate the Web images. In the dataset, 5,062 images were from the public dataset Oxford Building [15], the others were landmark and building images ($\geq 1024 \times 768$) crawled from Flickr[1]. Different levels of Gaussian noise were added to seven Oxford Building images (as shown in Fig. 4) for the denoising test.

Some empirical parameters were used to generate the following results. The number of K correlated images was four. The numbers of blocks in the extrinsic and intrinsic one were $K + 1$ and 16 ($m = 5, n = 16$), respectively. The hard thresholding parameter λ_{3D} was set to 2.7. The number of pixels in a block in Eq. (5) was 256 ($N=256$). The parameter λ in Eq. (6) was 0.1.

In the following subsections, we will first demonstrate the advantage of CID by comparing it with four state-of-the-art approaches, and then present some intermediate results of CID to show its efficiency step by step.

5.1. Comparison with other schemes

We evaluated the performance of CID in comparison to four state-of-the-art methods, including single-image based BM3D [6], learning based EPLL [30], texture enhanced based GHP [28] and correlated image based BM3D^C [12]. For a fair comparison, the same correlated images were used in both BM3D^C and CID.

Table 1 shows the objective evaluation of the five methods in terms of PSNR and SSIM [29] values. The Gaussian noise deviation σ was set at 50,70, and 90 to suit a high noise scenario. We have highlighted the best results for each noise level in bold and our CID method clearly achieved the best PSNR and SSIM results over all the test images. Compared with BM3D, GHP, and EPLL, our scheme constantly outperforms these schemes at more than 2 dB in PSNR and more than 0.1 in SSIM, on average. Compared with the second best result BM3D^C, our method achieves more than 1 dB gain, although the two methods used the same external correlated images. It further demonstrates the importance and effectiveness of the combined filtering in denoising.

Fig. 5 shows some visual evaluations of the five schemes. This figure shows the exemplified denoising results for image "a", "b", "c", and "e" when σ is 70. For each image we show two cropped regions highlighted with red boxes to visualize the details. It can be observed that our
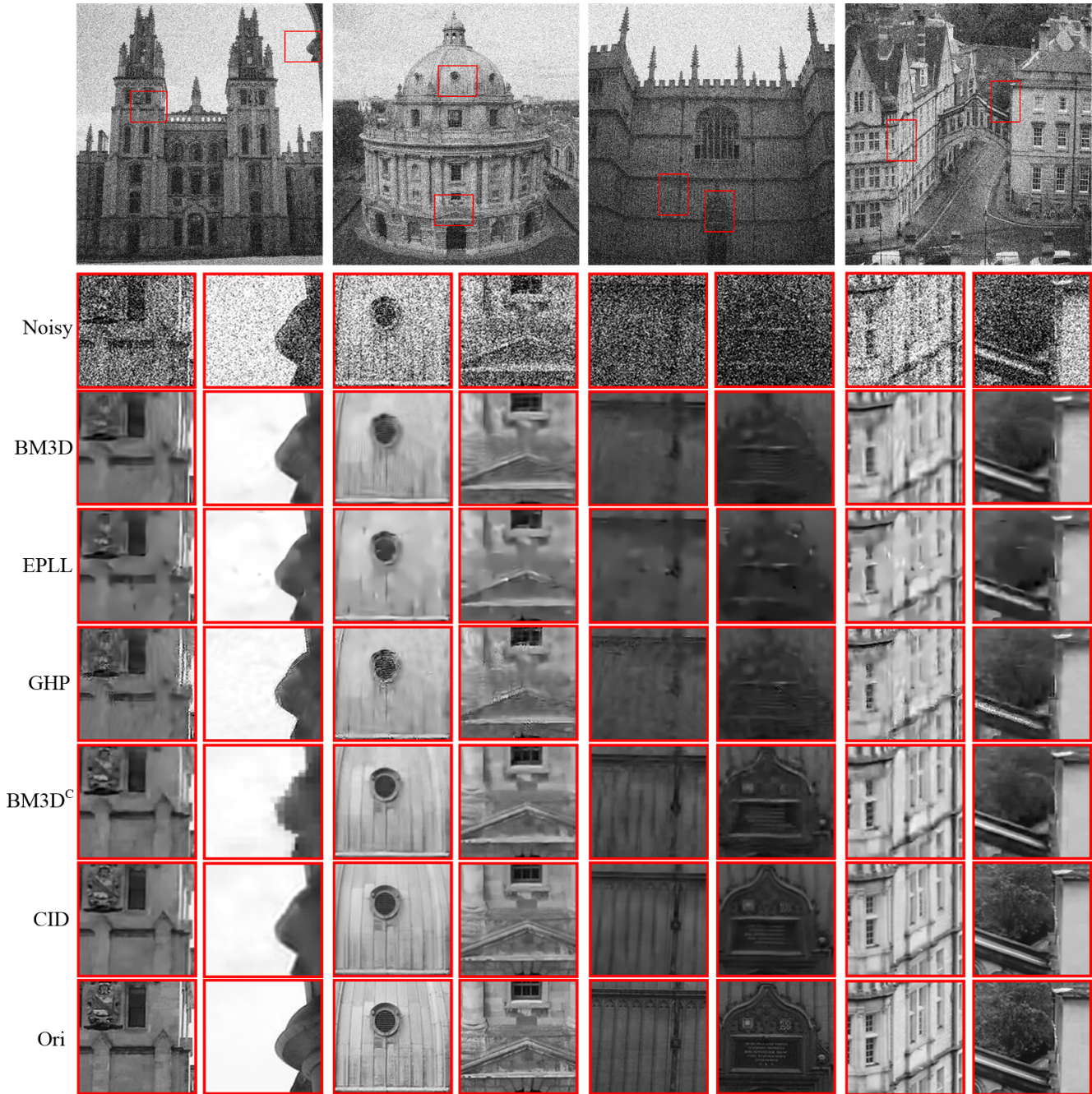
---

[1]http://www.flickr.com/

Figure 5. The denoising results for image "a", "b", "c", and "e" when the noise standard deviation was 70. For each column, the first row is the noisy image. Two crops of the regions indicated by red boxes are shown from the second row to the eighth row, obtained from the noisy image, BM3D result [6], EPLL result [30], GHP result [28], BM3D^C result [12], proposed CID result, and the original image respectively.

method reconstructs the most natural and realistic details for all the test images. For image "c", our method faithfully recovers the dense edges on the wall and the tiny characters, while all the other methods fail. We also reconstruct vivid textures, such as the decorative pattern in the wall in image "a" and the tree in image "e". BM3D generates much smoother denoising results because the frequency filtering not only alleviates the noise but also reduces the image details. EPLL and GHP produce noisy artifacts along the edges since the priors learned from general natural images or the noisy image itself (when the noise level is high) cannot clearly reveal the original information.

In brief, among all the test schemes, our CID produced the best denoising results in terms of both subjective and objective qualities.

5.2. Intermediate results

This subsection presents the intermediate results of our CID to illustrate the step-by-step enhancement of the denoising images. To clearly show the details, Fig. 6 presents the intermediate results of two regions, R1 (left) and R2 (right) in image "a", as marked by two red boxes in Fig. 5. In this figure, the top three groups of images illustrate the intermediate results of Step 1 and the bottoms ones are the results of Step 2 (as defined in Section 4). From left to right, the three columns show the results of frequency filtering only, spatial filtering only, and our combined filtering, respectively. We can observe that for region R1, the frequency filtering greatly deduces the details because the high-frequency coefficients are filtered out, whereas the spatial filtering keeps the details quite well. But for region R2, the frequency filtering provides much sharper edges whereas the spatial filtering makes a clean sky but blurred edges because of a lack of good matched blocks in the retrieved images. Our combined filtering takes advantage of both these filters and achieves the best results in both cases. We also found that improved results were achieved in Step 2 after the first step of denoising and the Step 2 CID presents the best one among all the results with clean edges and the sky.

6. Conclusion

In this paper, we have proposed denoising images that have heavy noise using combined filtering with the help of Web images. We not only built two 3D data cubes, intrinsic and extrinsic, by retrieving the similar blocks inside and outside the noise image, but also performed different filtering strategies, spatial and frequency, so as to use the data correlation adaptively. We have further proposed combining the spatial and frequency filtering and introduced the two-step denoising to greatly reduce the heavily noise while preserving the details and cleanness of the denoising image. Experimental results showed that our scheme significantly outperforms the four state-of-the-art methods in both subjective and objective evaluations.

We would like to point out that the correlated Web images play an important role in our CID. Without any correlated images, CID degenerates into one frequency filtering in the intrinsic cube and provides a similar performance as BM3D. If part of the noisy image content is too unique, we cannot find the correlated patches externally and internally. It would result in unpleasant denoising results. In addition, in our current test, no noisy images were retrieved, but this may not be true in real applications. We will work to make our scheme more feasible in the future.
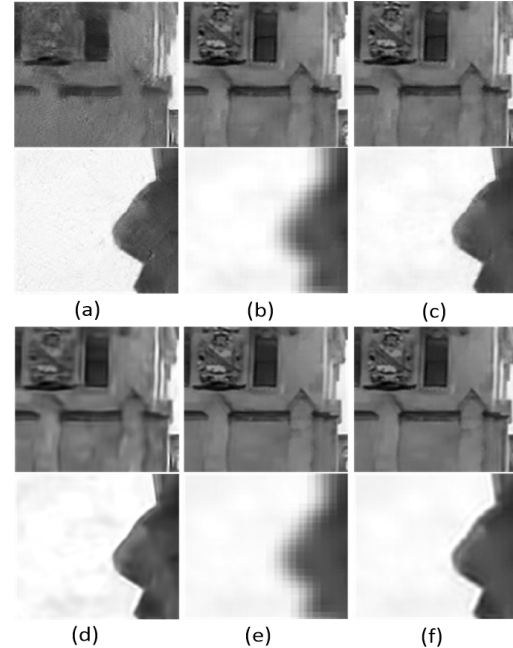

Figure 6. The intermediate results. To clearly show the details, we only show two crops of the corresponding results. (a), (b), and (c) show the frequency filtering result (25.78 dB), spatial filtering result (27.48 dB), and combining result (28.30 dB) in Step 1, respectively; (d), (e), and (f) show the the frequency filtering result (27.68 dB), spatial filtering result (27.72 dB), and combining result (29.00 dB) in Step 2, respectively. The PSNR values were computed over the whole image pixels.

References

[1] A. Buades, B. Coll, and J. Morel. A non-local algorithm for image denoising. *Proc. Computer Vision and Pattern Recognition*, pages 60–65, 2005. 2

[2] H. C. Burger and S. Harmeling. Improving denoising algorithms via a multi-scale meta-procedure. *Proceedings of the 33rd international conference on Pattern recognition (DAGM)*, pages 206–215, 2011. 2

[3] H. C. Burger, C. J. Schuler, and S. Harmeling. Image denoising: can plain neural networks compete with bm3d? *Proc. Computer Vision and Pattern Recognition*, pages 2392–2399, 2012. 2

[4] C. Zhang, J. Gao, and Oliver Wang *et al.*. Personal photo enhancement using internet photo collections. *Transactions on Visualization and Computer Graphics*, 2013. 3

[5] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004. 3

[6] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, 2007. 1, 2, 4, 5, 6, 8

[7] W. Dong, G. Shi, and X. Li. Nonlocal image restoration with bilateral variance estimation: a low-rank approach. *IEEE*

Image	BM3D[6]			EPLL [30]			GHP [28]			BM3D^C [12]			CID		
	50	70	90	50	70	90	50	70	90	50	70	90	50	70	90
a	28.56	26.99	25.86	28.52	26.98	25.90	28.14	25.82	21.79	28.99	27.85	26.72	**30.05**	**29.00**	**27.70**
	0.818	0.772	0.735	0.812	0.769	0.738	0.811	0.757	0.625	0.847	0.813	0.776	**0.882**	**0.860**	**0.845**
b	26.05	24.77	23.83	25.87	24.51	23.61	25.48	23.51	20.46	27.61	26.53	25.61	**28.56**	**27.63**	**26.83**
	0.748	0.694	0.651	0.735	0.674	0.631	0.735	0.671	0.572	0.812	0.770	0.731	**0.861**	**0.837**	**0.822**
c	28.90	27.56	26.62	28.56	27.39	26.57	28.25	26.20	22.16	29.78	28.42	27.33	**30.64**	**29.44**	**28.63**
	0.762	0.717	0.687	0.738	0.703	0.681	0.754	0.704	0.584	0.809	0.759	0.72	**0.855**	**0.816**	**0.809**
d	25.94	24.76	23.89	25.84	24.55	23.69	25.31	23.44	20.38	26.11	25.19	24.36	**26.80**	**26.03**	**25.25**
	0.704	0.648	0.604	0.698	0.635	0.593	0.693	0.628	0.526	0.721	0.675	0.630	**0.766**	**0.734**	**0.706**
e	25.63	24.14	23.02	25.38	23.86	22.76	25.24	23.12	20.01	26.40	25.31	24.48	**29.50**	**28.53**	**27.15**
	0.720	0.654	0.602	0.702	0.635	0.583	0.703	0.629	0.517	0.777	0.730	0.693	**0.909**	**0.885**	**0.873**
f	27.54	26.16	25.15	27.32	25.95	25.03	26.70	24.74	21.17	27.96	27.02	26.17	**28.50**	**27.75**	**26.91**
	0.755	0.703	0.662	0.741	0.691	0.658	0.740	0.680	0.568	0.774	0.740	0.708	**0.799**	**0.777**	**0.753**
g	28.08	26.67	25.67	27.87	26.47	25.52	27.26	25.16	21.44	28.45	27.49	26.66	**28.90**	**27.98**	**27.02**
	0.801	0.755	0.719	0.785	0.738	0.707	0.787	0.731	0.617	0.822	0.790	0.760	**0.837**	**0.819**	**0.803**
Avg	27.24	25.86	24.86	27.05	25.67	24.73	26.63	24.57	21.06	27.90	26.83	25.90	**28.99**	**28.05**	**27.07**
	0.758	0.706	0.666	0.744	0.692	0.656	0.746	0.686	0.573	0.795	0.754	0.717	**0.844**	**0.818**	**0.802**
Gain	**1.750**	**2.187**	**2.207**	**1.941**	**2.379**	**2.344**	**2.367**	**3.481**	**6.011**	**1.093**	**1.221**	**1.166**	–	–	–
	0.086	**0.112**	**0.136**	**0.100**	**0.126**	**0.146**	**0.098**	**0.133**	**0.229**	**0.050**	**0.064**	**0.085**	–	–	–

Table 1. Comparison of PSNR and SSIM results with four methods.

Transactions on Image Processing, 2013. 2

[8] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745, 2006. 2

[9] T. Gan and W. Lu. Image denoising using multi-stage sparse representations. *International Conference on Image Processing (ICIP)*, pages 1165–1168, 2010. 2

[10] H. C. Burger, C. Schuler, and S. Harmeling. Learning how to combine internal and external denoising methods. *Pattern Recognition*, pages 121–130, 2013. 2

[11] H. Yue, X. Sun, J. Yang and F. Wu. Cloud-based image coding for mobile devices-toward thousands to one compression. *IEEE Transactions on Multimedia*, 15:845–857, 2013. 3

[12] H. Yue, X. Sun, J. Yang and F. Wu. Image denoising using cloud images. *Proc. SPIE 8856, Applications of Digital Image Processing XXXVI*, 2013. 3, 5, 6, 8

[13] I. Mosseri, M. Zontak, M. Irani. Combining the power of internal and external denoising. *ICCP*, pages 1–9, 2013. 2

[14] J. Hays, and A. A. Efros. Scene completion using millions of photographs. *ACM Trans. on Graphics*, 26(3), 2007. 1, 3

[15] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Object retrieval with large vocabularies and fast spatial matching. *Proc. CVPR*, 2007. 5

[16] S. Jakubczak and D. Katabi. A cross-layer design for scalable mobile video. In *Proceedings of the 17th annual international conference on Mobile computing and networking*, MobiCom '11, pages 289–300, 2011. 1

[17] L. Dai, X. Sun, F. Wu and N. Yu. Large scale image retrieval with visual groups. *Proc. IEEE ICIP*, 2013. 3

[18] G. Lawton. Wireless hd video heats up. *Computer*, 41(12):18–20, 2008. 1

[19] X. Liu and L. W. *et al.*. Intrinsic colorization. *ACM Trans. Graph.*, 27(5):152:1–152:9, 2008. 1, 3

[20] M. Eitz, R. Richter, K. Hildebrand, T. Boubekeur, M. Alexa. Photosketcher: interactive sketch-based image synthesis. *IEEE Journal of Computer Graphics and Applications*, 31:56–66, 2011. 1, 3

[21] M. K. Johnson, K. Dale, S. Avidan, H. Pfister and W. T. Freeman. CG2Real: Improving the realism of computer generated images using a large collection of photographs. *IEEE Trans. on VCG*, 17:1273–1285, 2011. 1, 3

[22] M. Zontak, I. Mosseri, and M. Irani. Separating signal from noise using patch recurrence across scales. *Proc. Computer Vision and Pattern Recognition*, 2013. 2

[23] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. *International Conference on Computer Vision*, pages 2272–2279, 2009. 2

[24] O. Whyte, J. Sivic and A. Zisserman. Get out of my picture! Internet-based inpainting. *British Machine Vision Conference*, 2009. 1, 3

[25] S. Roth and M. J. Black. Fields of experts. *International Journal of Computer Vision*, 82(2):205–229, 2009. 2

[26] T. Suzuki. Challenges of image-sensor development. *IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pages 27–30, 2010. 1

[27] T. Chen, M. M. Cheng, P. Tan, A. Shamir, S. M. Hu. Photo-Sketch: Internet image montage. *SIGGRAPH Asia*, 2009. 1, 3

[28] W. Zuo, L. Zhang, C. Song and D. Zhang. Texture enhanced image denoising via gradient histogram preservation. *Proc. CVPR*, 2013. 5, 6, 8

[29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. on Image Processing*, 13(4):600–612, 2004. 5

[30] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. *International Conference on Computer Vision*, pages 479–486, 2011. 1, 2, 5, 6, 8