A NEW IMAGE FILTERING METHOD: NONLOCAL IMAGE GUIDED AVERAGING

Jing Zhang[⋆†], *Yang Cao*[⋆†], *Zengfu Wang*[⋆†‡]

[⋆]Department of Automation, University of Science and Technology of China
[†]National Engineering Laboratory for Speech and Language Information Processing, Hefei, China
[‡]Institute of Intelligent Machine, Chinese Academy of Sciences, Hefei, China
zjwinner@mail.ustc.edu.cn, {forrest,zfwang}@ustc.edu.cn

ABSTRACT

Image guided filtering has been widely used in many image processing applications. However, it is a local filtering method and has limited propagation ability. In this paper, we propose a new image filtering method: nonlocal image guided averaging (NLGA). Derived from a nonlocal linear model, the proposed method can utilize the nonlocal similarity of the guidance image, so that it can propagate nonlocal information reliably. Consequently, NLGA can obtain a sharper filtering results in the edge regions and more smooth results in the smooth regions. It shows superiority over image guided filtering in different applications, such as image dehazing, depth map super-resolution and image denoising.

Index Terms— nonlocal image guided averaging, image dehazing, depth super-resolution, image denoising

1. INTRODUCTION

In a wide variety of image processing applications, it is necessary to smooth an image while preserving its edges. Such edge-preserving smoothing methods have been proposed in recent literatures [1, 2, 3], including bilateral filter, guided filter.

Bilateral filter compute the weighted average of each pixels neighboring pixels as its output, where the weights are calculated according to the intensity/color similarities in the guidance image. The characteristics of intensity/color change in a local neighborhood of the guidance image should be similar to the one of the filter input. For applications such as Smoothing/HDR compression, the guidance image can be the filter input itself [1, 2]. For applications such as image matting/image dehazing, the guidance image can be another image [2, 3].

Though bilateral filter is popular for its effectiveness, it may suffer from unwanted gradient reversal artifacts near edges. To overcome this drawback, He et al. [2] propose a new edge-preserving filtering method, image guided filtering. Derived from a local linear model, the guided filter computes the linear transformation of the guidance image locally as the filter output. Sharing with the same edge-preserving property like bilateral filter, this filter is efficient and does not suffer from the gradient reversal artifacts. And its direct relation with matting Laplacian matrix enables it to be used in many image processing applications, such as image matting and image dehazing. However, the bilateral filter and guided filter are indeed local filters, which only utilize the intensity/color information in a local neighborhood of the guidance image. In addition, output of these filters may be affected by the tiny textures in the guidance image.

Image content is likely to repeat itself within some neighborhood. This nonlocal prior has been widely used in many applications. A. Buades et al. propose a popular image denoising method, nonlocal means (NLM), based on such a nonlocal prior [4, 5]. M. Protter et al. generalize the nonlocal means to super-resolution reconstruction [6]. Recently, P. Lee and Y. Wu apply the nonlocal prior to resolve image matting problem, and propose a new matting method, nonlocal matting [7]. This method can reduce the amount of user effort in the natural image matting problem.

In this paper, inspired by the work [7], a nonlocal linear model is presented for deriving a new image filtering method, nonlocal image guided averaging (NLGA). Though this new filter shares many similarities with guided filter, it is not the direct extension of guided filter to nonlocal case. It actually can be regarded as a two-step filter: i). Linear transformation, ii). Nonlocal averaging. The linear transformation step utilizes the structural similarity between patches in the nonlocal neighborhood. Therefore, it enables our filter to achieve better edge-preserving smoothing result. The following nonlocal averaging step calculates the weighted average of the linear transformation results in the nonlocal neighborhood as the final filter output. It enables our filter to get a more reliable estimation. This new filter shows superiority over guided filter in different applications, such as depth map super-resolution, image dehazing and image denoising.

This work is supported by the National Science and Technology Major Project of the Ministry of Science and Technology of China (No. 2012G-B102007) and the National Natural Science Foundation of China (No. 61303150).

2. RELATED WORK

Deriving guided filter starts from the following local linear model:

$$q_i = {a_k}^T I_i + b_k, \forall i \in \omega_k, \tag{1}$$

where I is the guidance image, q is the filter output, (a_k, b_k) are linear coefficients and ω_k is the filter window. Then, after minimizing the difference between the input image p and the filter output q, the coefficients are obtained. By adopting a neighborhood averaging strategy, the final filter output can be formulated as:

$$q_i = \frac{1}{|\omega|} \sum_{k:i\in\omega_k} \left({a_k}^T I_i + b_k\right) = \overline{a_i}^T I_i + \overline{b_i}. \tag{2}$$

Here $|\omega|$ is the number of pixels in ω_k, which can be computed as $|\omega| = (2r+1)\times(2r+1)$. r is the radius of ω_k . And $\overline{a_i} = \frac{1}{|\omega|}\sum_{k\in\omega_i} a_k$, $\overline{b_i} = \frac{1}{|\omega|}\sum_{k\in\omega_i} b_k$. From simple derivative calculation, the filter kernel weight can be expressed as:

$$\begin{gathered} W_{ij}(I) = \\ \frac{1}{|\omega|^2} \sum_{k:(i,j)\in\omega_k} \left(1 + (I_i - \mu_k)^T (\Sigma_k + \varepsilon)^{-1} (I_j - \mu_k)\right) , \end{gathered} \tag{3}$$

where μ_k and Σ_k are the mean vector and covariance matrix of I in ω_k. Such a filter kernel has a direct relation with matting Laplacian matrix. Consequently, the guided filter can be used for image matting and haze removal.

As proved in [8], the local linear model is based on the observation of color line model, which describes the locally smooth property of foreground and background of natural image. It constraints the size of local window to be small, and therefore the guided filter is a local filter. Recently, P. Lee and Y. Wu generalize the local linear model to be a nonlocal one based on the nonlocal prior [7]. And they propose a new matting method, nonlocal matting, which can reduce the amount of user effort in the natural image matting problem.

3. NONLOCAL IMAGE GUIDED AVERAGING

3.1. Nonlocal linear model

The nonlocal linear model presented in this paper is similar with the one in [7], and only differs in the weighting matrix. It can be formulated as:

$$Q_{ij} q_j = Q_{ij}\left({a_i}^T I_j + b_i\right), j \in N(i), \tag{4}$$

where $N(i)$ represents the nonlocal neighborhood of pixel i. $Q_{ij} = \sqrt{w_{ij} \big/ \sum_{j\in N(i)} w_{ij}}$, w_{ij} is the nonlocal weight. It is calculated by measuring the similarity between local patch $I(\omega_i)$ centered on pixel i and local patch $I(\omega_j)$ center on pixel j [4]:

$$w_{ij} = \exp\left(-\frac{1}{h^2}\left\|I(\omega_i) - I(\omega_j)\right\|_g^2\right), \tag{5}$$

where h is a kernel parameter. $\|\cdot\|_g$ denotes the norm which is weighted by Gaussian function.

Given an input p and a guidance image I, we can obtain the linear coefficients (a_i, b_i) by minimizing the following error energy function:

$$E(a,b) = \sum_{i\in\Lambda} E(a_i, b_i) = \sum_{i\in\Lambda} E_{data}(a_i, b_i) + E_{reg}(a_i), \tag{6}$$

where Λ is the pixel index set, and

$$\begin{aligned} E_{data}(a_i, b_i) &= \sum_{j\in N(i)} w_{ij}\|p_j - q_j\|^2 \\ &= \sum_{j\in N(i)} (Q_{ij}p_j - Q_{ij}q_j)^2 \\ &= \sum_{j\in N(i)} \left(Q_{ij}p_j - Q_{ij}\left({a_i}^T I_j + b_i\right)\right)^2 \end{aligned} , \tag{7}$$

$$E_{reg}(a_i) = \varepsilon a_i^2 = \varepsilon \sum_{j\in N(i)} (Q_{ij}a_i)^2. \tag{8}$$

Thus, obtaining the linear coefficients becomes a weighted quadratic optimization problem, which can be solved by applying the first-order condition.

$$\begin{aligned} &\tfrac{\partial E(a_i,b_i)}{\partial b_i} = 0 \\ &\Rightarrow b_i^* = \overline{p_w(N(i))} - {a_i}^T \overline{I_w(N(i))} \end{aligned} , \tag{9}$$

$$\begin{gathered} \tfrac{\partial E(a_i,b_i^*)}{\partial a_i} = 0 \Rightarrow a_i^* = (\Sigma_{i,w} + \varepsilon)^{-1} \times \\ \left(\sum_{j\in N(i)} (w_{ij}p_j \times I_j) - \overline{p_w(N(i))} \times \overline{I_w(N(i))}\right) \end{gathered} , \tag{10}$$

where $\overline{I_w(N(i))}$ and $\Sigma_{i,w}$ are the weighted mean vector and covariance matrix in $N(i)$. $\overline{I_w(N(i))} = \sum_{j\in N(i)} w_{ij} I_j$ and $\Sigma_{i,w} = \sum_{j\in N(i)} w_{ij}\left(I_j - \overline{I_w(N(i))}\right)\left(I_j - \overline{I_w(N(i))}\right)^T$. $\overline{p_w(N(i))} = \sum_{j\in N(i)} w_{ij} p_j$.

3.2. A new filter

After obtaining the linear coefficients (a_i, b_i) of each nonlocal patch $N(i)$, we can compute the corresponding linear transformation results. For each pixel p_i, since there are several nonlocal patches containing it, there are several linear transformation results. They can be classified into two kinds as follows:

$$p_i \mapsto \left\{ \begin{array}{c} {a_i}^T I_i + b_i \\ {a_j}^T I_i + b_j, j \in N(i) \end{array} \right. \tag{11}$$

Analyzing each term $E(a_i, b_i)$ in the energy function and the explicit form of (a_i, b_i) (Eq. (9) and (10)), we can conclude that those pixels which have bigger weights make larger contributions to the regression of (a_i, b_i). As we know the nonlocal weight w_{ij} reveals the structure similarity between the local patch $I(\omega_j)$ and the central local patch $I(\omega_i)$, so (a_i, b_i) has strong relation with the structure of the central local patch $I(\omega_i)$. Thus, for each pixel p_i, selecting its linear transformation result as ${a_i}^T I_i + b_i$, we obtain the linear transformation result for the whole filter input p.

Adopting a weighted averaging strategy in nonlocal neighborhood $N(i)$, we calculate the filter output as:

$$q_i = \sum\nolimits_{j\in N(i)} w_{ij}\left({a_j}^T I_j + b_j\right). \tag{12}$$

Taking the derivative w.r.t. p_j, we have the explicit form of the filter kernel weight:

$$W_{ij}(I) = \frac{\partial q_i}{\partial p_j} = \sum_{k:(i,j)\in N(k)} w_{ik} w_{kj} \times \left[1 + \left(I_k - \overline{I_w(N(k))}\right)^T (\Sigma_{k,w} + \varepsilon)^{-1}\left(I_j - \overline{I_w(N(k))}\right)\right] \tag{13}$$

In this paper, we denote this new filtering method as nonlocal image guided averaging (NLGA).

3.3. Relation with guided filter

As we have indicated before, there are two kinds of linear transformation results for each pixel. Here, we compute the weighted average of them as the filter output. Thus we have:

$$\begin{aligned} q_i &= \sum\nolimits_{j\in N(i)} w_{ij}\left({a_j}^T I_i + b_j\right) \\ &= \left(\sum\nolimits_{j\in N(i)} w_{ij} a_j\right)^T I_i + \sum\nolimits_{j\in N(i)} w_{ij} b_j \end{aligned} . \tag{14}$$

Taking the derivative w.r.t. p_j, we have the explicit form of the filter kernel weight:

$$W_{ij}(I) = \frac{\partial q_i}{\partial p_j} = \sum_{k:(i,j)\in N(k)} w_{ik} w_{kj} \times \left[1 + \left(I_i - \overline{I_w(N(k))}\right)^T (\Sigma_{k,w} + \varepsilon)^{-1}\left(I_j - \overline{I_w(N(k))}\right)\right] \tag{15}$$

Comparing it with the filter kernel weight of guided filter (Eq. (3)), we can conclude that this filter is the direct extension of guided filter to nonlocal case. In this paper, we refer this filter as nonlocal guided filter (NLGF).

Comparing Eq. (12) and Eq. (14), we find that, for smooth regions where $(a_i, b_i) \approx (a_j, b_j), j \in N(i)$, the output of NLGA at pixel i becomes:

$$q_i \approx {a_i}^T \sum\nolimits_{j\in N(i)} w_{ij} I_j + b_i = {a_i}^T \overline{I_w(N(i))} + b_i. \tag{16}$$

And the output of NLGF at pixel i becomes:

$$q_i \approx {a_i}^T I_i + b_i. \tag{17}$$

This difference can be observed obviously in image denoising applications (Fig. 3). Since the guidance image is the noisy image itself, NLGA can get a better denoising result for it involving a nonlocal averaging process.

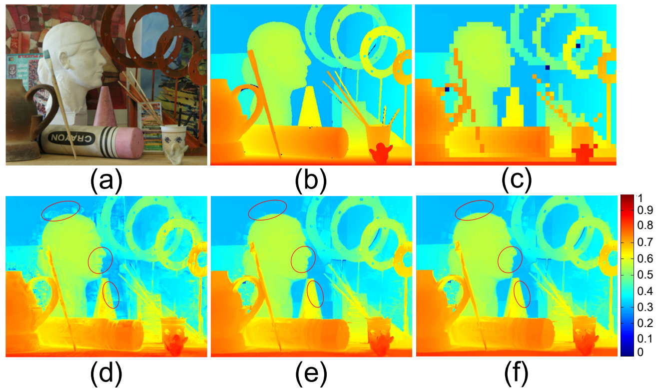

Fig. 1. Depth super-resolution results of GF, NLGF and NLGA. The low-resolution depth map is obtained using down-sampling factor of 10. (a) The guidance image “Art”. (b) The ground truth depth map. (c) Nearest neighbor interpolation result, PSNR: 22.94dB, SSIM: 0.7653. (d) Result of GF ($r = 20$, $\varepsilon = 10^{-3}$), PSNR: 25.64dB, SSIM: 0.8309. (e) Result of NLGF ($r = 20$, $\varepsilon = 10^{-3}$), PSNR: 25.92dB, SSIM: 0.8395. (f) Result of NLGA ($r = 20$, $\varepsilon = 10^{-3}$), PSNR: 26.10dB, SSIM: 0.8495. This figure is best viewed in the electronic version of this paper.

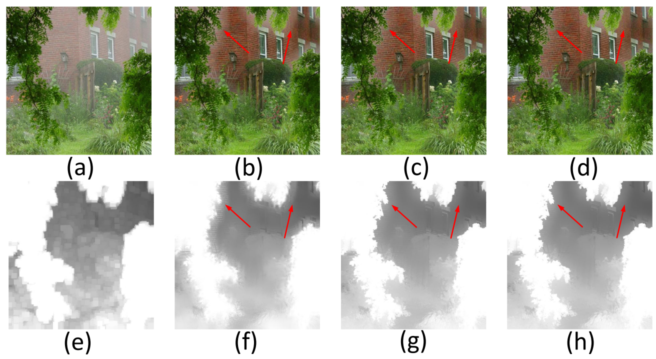

Fig. 2. Image dehazing results of GF ($r = 20$, $\varepsilon = 10^{-3}$), NLGF and NLGA ($r = 20$, $\varepsilon = 10^{-3}$). (a) The hazy image. (e) Raw transmission map [9]. (f)-(h) Filtered transmission map using GF, NLGF and NLGA respectively. (b)-(d) Recovered image using (f)-(h) respectively. This figure is best viewed in the electronic version of this paper.

Table 1. PSNR(dB) of super-resolution results obtained by different methods: Nearest Interpolation, GF, NLGF and NLGA. The down-sampling factor is 10. $r = 20, \varepsilon = 10^{-3}$.

Method	Interpolation	GF	NLGF	NLGA
Art	22.94	25.64	25.92	26.10
Books	29.91	30.52	31.31	31.23
Dolls	31.97	33.09	33.87	33.74
Laundry	27.40	29.20	29.50	29.67
Moebius	30.24	31.32	32.67	32.49
Reindeer	24.97	28.97	28.47	28.54

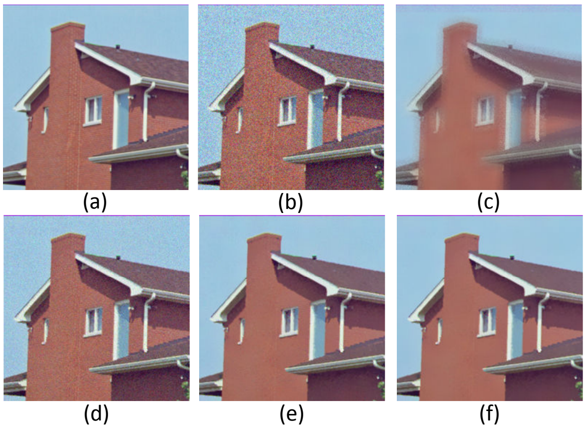

Fig. 3. Image denoising results of GF, NLGF, NLM and NLGA ($r = 10$, $\varepsilon = 0.4 \times 20^2$). (a) The original color test image: "House". (b) Noisy image with Gaussian noise: σ=20, PSNR: 22.10dB. (c) Image Guided filtering result, PSNR: 25.13dB. (d) Result of NLGF, PSNR: 27.32dB. (e) Result of NLM, PSNR: 31.59dB. (f) Result of NLGA, PSNR: 32.04dB. This figure is best viewed in the electronic version of this paper.

4. APPLICATIONS AND EXPERIMENTAL RESULTS

In this section, we conduct a series of experiments for different applications including image dehazing, depth super-resolution and image denoising to verify the effectiveness of the proposed filtering method.

Depth Super-resolution In RGBD camera system, one can get a high-resolution color image and a low-resolution depth map of a same scene [10]. Therefore, under the guidance of the color image, super-resolution of the depth map can be implemented by guided filter and the proposed filter. Figure 1 shows the super-resolution results of different filtering methods for depth map "Art" in Middlebury dataset 2005[1]. We can see that NLGF and NLGA can give sharper results in the edge regions and more smooth results in the smooth regions than the guided filter, e.g., the regions indicated by the red circle in Fig. 1. This is because that NLGF and NLGA utilize the structure similarity between patches in nonlocal neighborhood of the guidance image. The objective assessment indexes, PSNR and SSIM [11], also convince that NLGA and NLGF outperform GF significantly. Table 1 shows PSNR of super-resolution results obtained by different methods.

Image Dehazing In [2], they filtered the raw transmission map under the guidance of the hazy image using guided filter. Here, we used NLGF and NLGA to replace the guided filter to obtain the dehazed image. Figure 2 shows the results. We can see that the guided filtering result may exhibit some residual haze near the trunk and ground/braches and wall regions, e.g., the regions indicated by the red arrows. Since such regions are self-similar, the NLGF and NLGA can utilize the similarity information and obtain a clearer result with little haze residual.

Image Denoising NLGF and NLGA can achieve comparable results for above applications. To illustrate their difference, we conduct a series of experiments on image denoising. Figure 3 shows the denoising results of GF, nonlocal means (NLM), NLGF and NLGA. Guided filter obtains an over-smoothed result. NLGF gives unsatisfied result with large amounts of noise residual, especially in the smooth regions. It verifies the previous analysis of NLGF in Section 3.3. NLGA achieves the best performance, even compared with nonlocal means. Table 2 shows the denoising results for five gray test images with Gaussian noise, σ=$20, 30$. The PSNR indexes also verify that NLGA has superiority over GF, NLGF and NLM for image denoising.

Analyzing the explicit form of the proposed filter Eq. (12) and (a_i, b_i) (Eq. (9) and (10)), we can find that NLGA actually involves five nonlocal means processes. When the guidance image I is the filter input itself p, it reduces to 3. Thus, the computational cost of NLGA is approximately 3-5 times of nonlocal means filter.

[1] http://vision.middlebury.edu/stereo/data/

Table 2. PSNR(dB) indexes of denosing results obtained by GF, NLGF, NLM and NLGA ($r = 10$, $\varepsilon = 0.4 \times \sigma^2$).

σ	Method	Lena	Barbara	Boats	House	Peppers	*average*
20	Noisy	22.11	22.12	22.10	22.13	22.11	22.11
	GF	24.96	23.08	23.96	24.45	22.08	23.71
	NLGF	27.14	26.49	26.66	27.46	26.64	26.88
	NLM	30.29	28.79	28.91	30.86	29.02	29.57
	NLGA	31.01	29.22	29.31	31.78	29.53	30.17
30	Noisy	18.59	18.60	18.56	18.58	18.63	18.59
	GF	24.25	22.69	23.48	23.85	21.69	23.19
	NLGF	24.00	23.42	23.64	24.17	23.60	23.77
	NLM	27.90	26.37	26.69	28.24	26.58	27.16
	NLGA	28.94	26.90	27.33	29.47	27.22	27.97

5. CONCLUSION

In this paper, we propose a new image filtering method, nonlocal image guided filtering (NLGA), which is derived from a nonlocal linear model. Utilizing the nonlocal similarity of the guidance image, it shows superiority over the guided filter in different image processing applications, e.g., image dehazing and depth map super-resolution. In addition, we also present the direct extension of guided filter to the nonlocal case, nonlocal guided filter (NLGF) in this paper. Experimental results on image denoising verify that the proposed method NLGA outperforms the guided filter, NLGF and nonlocal means.

The nonlocal propagation capability of NLGA encourages future research for different applications. As a nonlocal filter, its computational complexity can be reduced by adopting similar techniques developed for nonlocal means.

6. REFERENCES

[1] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *International Conference on Computer Vision (ICCV)*, 1998.

[2] Kaiming He, Jian Sun, and Xiaoou Tang, "Guided image filtering," in *The European Conference on Computer Vision (ECCV)*, 2010.

[3] G. Petschnigg, M. Agrawala, H. Hoppe, R. Szeliski, M. Cohen, and K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM Trans. Graphics (SIGGRAPH)*, vol. 23, no. 3, pp. 664–672, 2004.

[4] A. Buades, B. Coll, and J.M. Morel, "A non-local algorithm for image denoising," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.

[5] A. Buades, B. Coll, and J.M. Morel, "A review of image denoising methods, with a new one," *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 490–530, 2006.

[6] M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal-means to super-resolution reconstruction," *IEEE Transactions on Image Processing*, vol. 18, no. 1, pp. 36 –51, 2009.

[7] P. Lee and Ying Wu, "Nonlocal matting," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.

[8] A. Levin, D. Lischinski, and Y. Weiss, "A closed-form solution to natural image matting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 228 –242, 2008.

[9] Kaiming He, Jian Sun, and Xiaoou Tang, "Single image haze removal using dark channel prior," *, IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341 –2353, 2011.

[10] Jaesik Park, Hyeongwoo Kim, Yu-Wing Tai, M.S. Brown, and Inso Kweon, "High quality depth map upsampling for 3d-tof cameras," in *IEEE International Conference on Computer Vision (ICCV)*, 2011.

[11] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions On Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004.