Cross-Based Adaptive Guided Filtering

Ming YAN[a, b], Yueli HU[a, b, 1], Kai LI[a,b] and Jianeng ZHAO[a,b]

[a] *Shanghai University Microelectronic Research and Development Center, Shanghai 200072, China*

[b] *School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, China*

Abstract. Edge-preserving and structure-preserving smoothing filtering has attracted much interest in the last decades. A conventional linear filter effectively smoothens noise in homogeneous regions but blurs the edges of an image. This study aimed to present an adaptive guided filter using a cross-based framework. The proposed method outperformed many other algorithms in terms of sharpness enhancement and noise reduction. Moreover, the cross-based adaptive guided filter had a fast and nonapproximate linear-time algorithm as the guided filter.

Keywords: Adaptive denoising, cross-based framework, edge-preserving smoothing, guided filter

1. Introduction

Image enhancement and image noise reduction are two opposing technologies. The former aims to improve the high-frequency components of the image and enhance the edge and texture features, while the latter aims to smooth the signal and remove high-frequency noise. These two technologies have been the research hotspots in the field of image processing. Nowadays, many filtering techniques, such as bilateral filtering [1], anisotropic diffusion [2-5], guided filtering [6], and so on are available, which can remove high-frequency details while removing noise in smooth areas.

Anisotropic diffusion can be used to preserve and strengthen large edges, but its oversmoothness filters out noise and small textures at the same time. Bilateral filtering has been widely used [7], which achieves image noise reduction according to spatial correlation and gray correlation. However, it has certain limitations due to the complexity of the adaptive value of the coefficients. At the same time, it has a staircase effect during smooth operations.

The effect of guided filtering is similar to that of bilateral filtering. However, its algorithm complexity is much lower than that of bilateral filtering, and it has unique advantages in terms of detail enhancement and high dynamic range compression [8]. Different from bilateral filtering, guided filtering is essentially a local multi-point estimation [9]; that is, the calculation of multiple observation points in the neighborhood is needed to estimate a certain point. Bilateral and guided filtering can be combined to make an overall combined estimation because the points in these neighborhoods are also

---

[1] Corresponding Author: Yueli Hu, [a] *Shanghai University Microelectronic Research and Development Center, Shanghai 200072, China;* [b] *School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, China;* E-mail: huyueli@shu.edu.cn.

used for other points. The estimates are worthy of calculation. This type of combined estimation of multiple points has a great improvement over previous operations for domain window estimation [10-12].

Edge enhancement leads to clearer subjective perception of the image and more comfortable visual effect for people. However, the a forementioned algorithm is only for the noise reduction and smoothing effect in the neighborhood. Therefore, the selection of the domain window is crucial in edge enhancement. Compared with global stereo matching, local stereo matching has more advantages in terms of calculation amount, parameter setting, and implementation efficiency [13]. For video noise reduction, the real-time requirement cannot be ignored. A previous study focused on real-time matching and introduced speed comparison in stereo matching [14]. The principle of bilateral filtering and stereo matching was described in previous studies for the effectiveness of dynamic stereo processing architecture [15-16].

In response to the need for edge enhancement, this study proposed a cross-based adaptive guided filter. The cross-based neighborhood was introduced into a guided filter, and a locally adaptive variable was added [17]. The adaptive change in the neighborhood might guide the smoothing effect of filtering.

2. Guided filtering

A previous study proposed guided filtering to make a linear transformation of the guided image, which had edge protection characteristics like those of bilateral filtering. Compared with bilateral filtering, guided filtering avoided the gradient reversal effect [6].

2.1. Definition

The main idea of guided filtering is to have a linear relationship between output image Q and guided image I, that is:

$$Q_i = a_k I_i + b_i, \forall i \in w_k \tag{1}$$

w_k is a square filter window with radius r, and (a_k, b_k) is a linear coefficient. Within a certain range w_k, their values are fixed. Eq. (1) shows that if image I has edges, image Q also retains the edges, and $\nabla Q = a \nabla I$.

The value of (a_k, b_k) is determined by minimizing the difference between input image P and filtered output image Q as shown in Eq. (2):

$$E(a_k, b_i) = \sum_{i \in w_k} ((a_k I_i + b_i - P_i)^2 + \varepsilon a_k^2) \tag{2}$$

ε is a normalization coefficient used to adjust a_k, preventing its value from being too large. (a_k, b_k) is calculated using Eq. (3) and Eq. (4):

$$a_k = \frac{\frac{1}{|w|} \sum_{i \in w_k} I_i P_i - \mu_k \bar{P}_k}{\sigma_k^2 + \varepsilon} \tag{3}$$

$$b_k = \bar{P}_k - a_k \mu_k \tag{4}$$

μ_k and σ_k^2 are the mean value and variance value of guided image I in the filter window, respectively. $|w|$ is the number of pixels in w_k, and $\bar{P}_k = \frac{1}{|w|} \sum_{i \in w_k} P_i$ is the mean value of input image P in the filter window w_k.

The proposed linear model was used for all partial windows of the entire image. The occurrence of pixel i in different windows Q_i leads to its uncertain values in different windows. Therefore, the average value of Q_i is calculated using Eq. (5):

$$Q_i = \frac{1}{|w|} \sum_{k:i \in w_k} (a_k I_i + b_k) = \bar{a}_i I_i + \bar{b}_i \tag{5}$$

In Eq. (5), $\bar{a}_i = \frac{1}{|w|} \sum_{k \in w_i} a_k$, $\bar{b}_i = \frac{1}{|w|} \sum_{k \in w_i} b_k$.

After such adjustments, the linear coefficient $(\bar{a}_i, \bar{b}_i)$ refers to the spatial mean value, and a direct linear relationship no longer exists between ∇Q and ∇I. $(\bar{a}_i, \bar{b}_i)$ is the output of an average filter, and its gradient is much smaller near the strong edge than I. However, $\nabla Q \approx \bar{a} \nabla I$ is still used here, indicating that the gradient changes in I can still be largely retained in Q.

A function that contains the guided image I, input image P, and output image Q is obtained as follows:

$$Q_i = \sum_j W_{ij}(I) P_j \tag{6}$$

W_{ij} is a function that guides the image and is independent of input image P.

a_k can be written as a weighted sum of $P : a_k = \sum_j A_{kj}(I) P_j$, which is relevant to only the weights of. If Eq. (5) is rewritten as Eq. (6), then:

$$W_{ij}(I) = \frac{1}{|w|^2} \sum_{k:(i,j) \in w_k} (1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \varepsilon}) \tag{7}$$

2.2. Edge Protection Filtering

If $\varepsilon = 0$, Eq. (2) has an optimal solution when $a_k = 1, b_k = 0$. If $\varepsilon > 0$, two situations occur:

Flat area: If the pixel value of image I is constant in the window, then the solution of Eq. (2) should be $a_k = 0, b_k = \overline{P}_k$.

Fluctuation area: If the image pixel value of image I changes greatly within window w_k , that is, its variance value is large, then the value of a_k tends to 1 and the value of b_k tends to 0.

$\overline{a}_i$ and $\overline{b}_i$ are the average values of a_k and b_k , respectively, which can be obtained using Eq. (5): If a pixel is a point in the undulating region, its value does not change. If it is in a flat region, its value is replaced by the average value of the surrounding pixels. The difference between the flat area and the undulating area is defined by ε . For image blocks with large variance values, that is, when σ^2 is larger than ε , the pixel values in the image blocks are retained, and conversely, the pixel values are replaced by the mean value. The effect of ε is similar to the gray variance in bilateral filtering.

2.3. Summary

The edge protection smoothing characteristics of guided filtering can be illustrated by a 1D signal as an example. As shown in Figure 1, if I_i and I_j are on the same side of the edge, $I_i - \mu_k$ and $I_j - \mu_k$ have the same +/– sign. However, if they are on

$$1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \varepsilon}$$

different sides, they have opposite signs. Then, the value of
in this case is much smaller than the value in the case of the same side, which tends to 0. This means that the pixel values on both sides of the edge are not averaged; when

$\sigma_k^2 \ll \varepsilon$ (flat area), $\quad W_{ij}(I) = \frac{1}{|w|^2} \sum_{k:(i,j)\in w_k} 1$, it is equivalent to a low-pass filter.
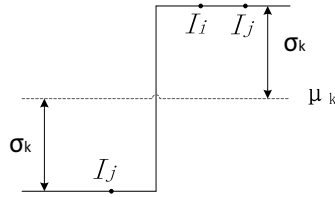

Figure 1. A 1D example of an ideal step edge.

3. Improved cross-based adaptive guided filtering

3.1. Offset ξ

This study introduced an adaptive offset ξ in the guided filtering to ehance the ability of bilateral filtering to strengthen edges [18]. The improved adaptive guided filtering was found to have better filtering effects at the edges and better detail retention.

In the filter window of guided image I, the maximum value MAX_k, minimum value MIN_k, and average value $MEAN_k$ can be used to replace the pixel values in the image. Supposing that $\Delta_{ki} = I_i - MEAN_k$, Δ_{kj} is the difference between each pixel in the image and the average value. ξ is defined as Eq. (8):

$$\xi_i = \begin{cases} MAX_k - I_i, if \Delta_{ki} > 0 \\ MIN_k - I_i, if \Delta_{ki} < 0 \\ 0, if \Delta_{ki} = 0 \end{cases} \tag{8}$$

$(I_i + \xi_i - \mu_k)$ is used instead of the $(I_i - \mu_k)$ in the original guided filter. If the pixel value I_i to be obtained is greater than the average value, its value is replaced with MAX_k; if the pixel value is less than the average value, it is replaced with MIN_k; and it is not replaced if it is equal to the average value. Substituting the original value with the lowest value helps strengthen the edges, thereby making the border texture clearer.

3.2. Cross-Based Adaptive Guided Filtering Algorithm

Scharstein and Szeliski proposed a stereo matching algorithm [19]. Compared with global stereo matching, local stereo matching is more efficient and easier to implement [20-22]. A cross-based local stereo matching algorithm was proposed to work out the size and shape of the adaptive window of the pixel [17]. A cross-based adaptive guided filtering algorithm is more practical through orthogonal integral image technology. The present study introduced this method into guided filtering and adjusted the size and shape of the neighborhood window.

The local cross-based adaptive guided algorithm is divided into two steps: First, a cross-based with a base radius is established with the pixel to be measured as the center, and the central pixel can be called a base pixel p. As shown in Figure 2, the color similarity is calculated for the pixel values in its four directions to adjust the arm length $\{h_p^-, h_p^+, v_p^-, v_p^+\}$. Using h_p^-, for example, to calculate the difference between consecutive pixels in this direction and p, and to find the maximum arm length r^* similar to the base pixel value, r^* is calculated as shown in Eq. (9):

$$r^* = \max_{r\in[1,L]}(r\prod_{i\in[1,r]}\delta(p,p_i)) \tag{9}$$

$p_i = (x_p - i, y_p)$, L is the preset maximum arm length value, $\delta(p_1,p_2)$ is the color similarity indicator function between pixels p_1 and p_2, and Eq. (10) provides its definition:

$$\delta(p_1,p_2) = \begin{cases} 1, \max_{c\in\{R,G,B\}}(|I_c(p_1)-I_c(p_2)|)\leq\tau \\ 0, otherwise \end{cases} \tag{10}$$

τ is the pixel color difference threshold between p_1 and p_2. When the maximum arm length r^* is obtained, it equals the value of h_p^-.

Second, the determination of the cross-based shows that each pixel only needs to store four arm length values, and the field determination of the base pixel is the superposition of the horizontal basis on its vertical basis, that is:

$$U(p) = \bigcup_{q\in V(p)} H(q) \tag{11}$$

q is the pixel on the vertical basis of p, as shown in Figure 2, $H(q)$ is the horizontal basis of q, and $U(p)$ is the adaptive neighborhood of pixel p.
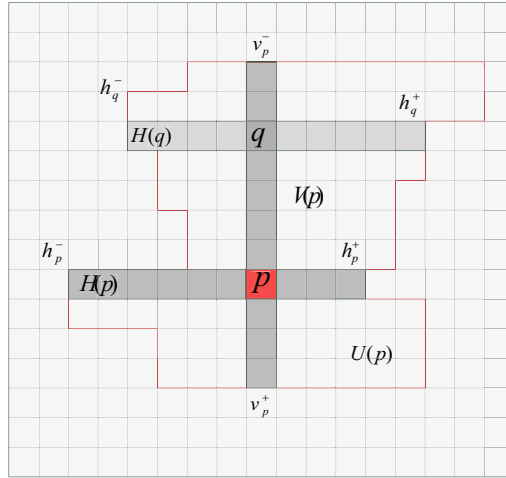

Figure 2. Configuration of a cross-based frame.

4. Experimental results

The value of ε in the ζ was determined. The three images in Figure 3 are the processing results of waterfall images under different ε values. Through the combined results of peak signal to noise ratio and subjective vision, $\varepsilon = 0.8\tau$ was finally selected, where τ is the threshold value of the difference in the cross-based pixel.
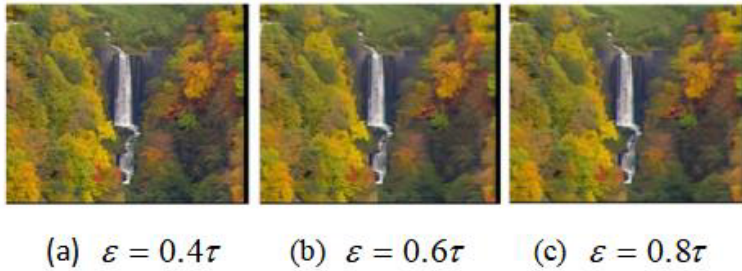
(a) $\varepsilon = 0.4\tau$ (b) $\varepsilon = 0.6\tau$ (c) $\varepsilon = 0.8\tau$

Figure 3. Effect of different coefficients of guided filtering.

Standard sequences such as Akiyo and Foreman were selected for experimental simulation to compare the advantages and disadvantages of the proposed algorithm and the original guided filtering algorithm.

Figure 4 is the 20th frame of the Akiyo sequence. Random noise with a magnitude of 0.05 was added. Bilateral filtering, guided filtering, and cross-based adaptive guided filtering proposed in this study were used to reduce the noise. Figure 4a and 4d are original images of noise reduction, and Figure 4e to 4h are corresponding enlarged views. Figures 4a and 4e are images with random noise added, Figure 4b and 4f are images processed with bilateral filtering on the noise map, and Figure 4c and 4g are directed noise filtering with guided filtering, Figure 4d and 4h are the effect diagrams of noise reduction after cross-based adaptive guided filtering. The cross-based adaptive guided filtering had better filtering effects at the edges and had obvious advantages in terms of detail processing.
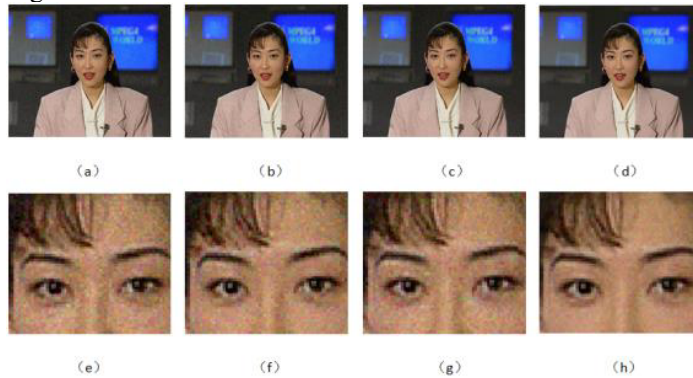

Figure 4. Denoising comparison of Akiyo.

Figure 5 is an effect diagram of noise reduction processing on a night-time real shot image. From left to right are an original image, a bilateral filter image, and a cross-based improvement guided filter image. At night, the noise of the image is large due to the insufficient exposure, which causes problems such as large color spots and blurred borders. The figure shows that the bilateral filtering and the improved cross-based guided filtering in this study both had a better smoothing effect and a better effect on removing speckles, but the over-smoothing of the bilateral filtering caused the distortion of the picture. The detailed enlarged view of Figure 5(d‑f) reveals that the smoothing effect of bilateral filtering blurred the boundary at the texture and edges, while the cross-based improved guided filtering still had clear boundaries.
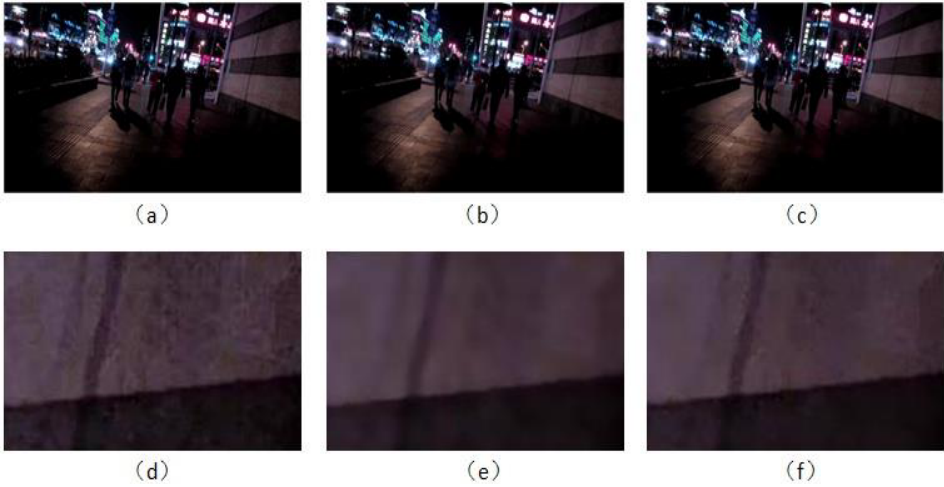
Figure 5. Denoising process on the picture of real objects.

Table 1 shows the results of the peak signal-to-noise ratio, and the structural similarity of standard sequence noise reduction based on previous studies [1, 6, 18] and the cross-based adaptive guided filtering algorithm.

Table 1. PSNR and SSIM comparison of four methods

	GF [6]	BF [1]	ABF [7]	Proposed method
PSNR comparison (dB)				
Akvio	27.0561	27.105	28.4098	29.3644
Foreman	26.7908	26.8976	27.549	28.7079
Waterfall	24.9809	25.4488	25.9767	27.6231
Bus	28.4903	29.099	29.9811	30.9946
Mobile	28.3293	28.0561	29.3002	30.724
Bottle	25.3579	25.6792	26.0593	27.4684
Flower	26.1684	26.5467	27.4918	28.3674
Luggage	27.6492	27.7843	28.8139	28.4627
Tempete	31.0575	31.1183	32.0243	33.0527
Bridge-close	26.8309	26.8976	27.9832	28.2753
SSIM comparison				
Akvio	0.827	0.835	0.842	0.860
Foreman	0.834	0.841	0.856	0.873
Qaterfall	0.825	0.834	0.844	0.865
Bus	0.838	0.846	0.862	0.884
Mobile	0.819	0.836	0.854	0.868
Bottle	0.805	0.826	0.848	0.861
Flower	0.821	0.837	0.853	0.875
Luggage	0.839	0.854	0.876	0.902
Tempete	0.828	0.842	0.865	0.887
Bridge-close	0.827	0.840	0.849	0.868

Compared with GF, BF, and ABF, the proposed method achieved better performance in terms of PSNR and SSIM when denoising the image. At the same time, the proposed algorithm still retained the advantages of guided filtering in computing time and implementation complexity.

5. Conclusions

This study proposed a cross-based adaptive guided filter. In this method, the adaptive selection of the filtering neighborhood was added to the framework of the guided filter. Filter blocks were generated whose size could be adjusted automatically and their shape could be adjusted adaptively based on gray similarity and stereo matching, which was beneficial to the smoothing effect of the filter. The improvement of the offset was added to the processing of the guided image, which enhanced the slope of filtering at the edges, reduced the blur, and made the image sharper. The selected value of the offset changed proportionally with the gray threshold of the filter window, which had a better effect than the hard threshold. The improved guided filtering still retained the original characteristics of easy hardware implementation with low algorithm complexity, and had good prospects in engineering.

References

[1]   Tomasi C, Manduchi R. Bilateral filtering for gray and color images. In: Proceedings of the IEEE International Conference on Computer Vision. 1997 23(1): 45-78.
[2]   Buades A, Coll B and Morel J M. A review of image denoising algorithms, with a new one. Multiscale Modeling Simul. 2005 4(2):490-530.
[3]   Buades A, Coll B, Morel J M. A non-local algorithm for image denoising. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005 2: 60-5.
[4]   Coup'e P, Yger P, Barillot C. Fast non-local means denoising for 3D MR images. MICCAr06, Copenhagen. 2006.
[5]   Buades A, Coll B, Morel J. Nonlocal image and movie denoising. Int. J. Comp. Vision. 2008:123-139.
[6]   He K, Sun J, Tang X. Guided image filtering. International Proceeding of ECCV, 2010. 1-6, 8.
[7]   Paris S, Kornprobst P, Tumblin J and Durand F. Bilateral filtering: theory and applications. Foundations Trends Comp. Graphics Vis. 2008 4(1):1–73.
[8]   Rhemann C, Hosni A, Bleyer M, Rother C and Gelautz M. Fast cost-volume filtering for visual correspondence and beyond. In *Proc. CVPR* 2011, 1, 2, 6, 7.
[9]   Katkovnik V, Foi A, Egiazarian K, Astola J. From local kernel to nonlocal multiple-model image denoising. Int J Comp Vis, 2010 86(1):1–32.
[10]   Hosni A, Bleyer M, Rhemann C, Gelautz M and Rother C. Realtime local stereo matching using guided image filtering. IEEE International Conference on Multimedia and Expo .2011 pp. 1- 6.
[11]   De-Maeztu L, Mattoccia S, Villanueva A and Cabeza R. Linear stereo matching. International Conference on Computer Vision, 2011, pp. 1708-1715.
[12]   Rhemann C, Hosni A, Bleyer M, Rother C and Gelautz M. Fast cost-volume filtering for visual correspondence and beyond. *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 3017-3024.
[13]   Lu J, Lafruit G and Catthoor F. Anisotropic local high-confidence voting for accurate stereo correspondence in *Proc. SPIE-IS&T Electron. Imaging*, Jan. 2008, 6812: 605822-1-605822-10.
[14]   Gong M, Yang R, Wang L and Gong M. A performance study on different cost aggregation approaches used in real-time stereo matching, Int. J. Comp. Vis., 2007 75(2):283-296.
[15]   Mattoccia S, Giardino S, Gambini A. Accurate and efficient cost aggregation strategy for stereo correspondence based on approximated joint bilateral filtering. *Computer Vision–ACCV 2009*, ed: Springer, 2010, pp. 371-380.
[16]   Wang L, Liao M, Gong M, Yang R and Nister D. High-quality realtime stereo using adaptive cost aggregation and dynamic programming. *Third International Symposium on 3D Data Processing, Visualization, and Transmission,* 2006 pp. 798-805.
[17]   Zhang K, Lu J and Lafruit G. Cross-based local stereo matching using orthogonal integral images. IEEE Trans. CSVT, July 2009. 19(7):1073–1079.
[18]   Zhang B, Allebach J P. Adaptive bilateral filter for sharpness enhancement and noise removal. IEEE Transaction on Image Process 2008 17(5), 664-678.
[19]   Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms Int. J. Comput. Vision, May 2002 47(1): 7–42.f

[20] Tombari F, Mattoccia S, Di Stefano L and Addimanda E. Classification and evaluation of cost aggregation methods for stereo correspondence. 2018 IEEE Conference on Computer Vision and Pattern Recognition, 2018 pp. 1 –8.

[21] Gong M, Yang R, Wang L, Gong M. A performance study on different cost aggregation approaches used in real-time stereo matching. *Int. J. Comp. Vision* 2007, 75(2):283-296.

[22] Hirschmuller H and Scharstein D. Evaluation of cost functions for stereo matching. 2007 IEEE Conference on CVPR, 2007 pp. 1 –8.