GUIDED IMAGE FILTERING WITH ARBITRARY WINDOW FUNCTION

Norishige Fukushima[*]

Nagoya Institute of Technology, Japan

Kenjiro Sugimoto[†] *and Sei-ichiro Kamata*

Waseda University, Japan

ABSTRACT

In this paper, we propose an extension of guided image filtering to support arbitrary window functions. The guided image filtering is a fast edge-preserving filter based on a local linearity assumption. The filter supports not only image smoothing but also edge enhancement and image interpolation. The guided image filter assumes that an input image is a local linear transformation of a guidance image, and the assumption is supported in a local finite region. For realizing the supposition, the guided image filtering consists of a stack of box filtering. The limitation of the guided image filtering is flexibilities of kernel shape setting. Therefore, we generalize the formulation of the guide image filter by using the idea of window functions in image signal processing to represent arbitrary kernel shapes. Also, we reveal the relationship between the guided image filtering and the variants of this filter.

***Index Terms*—** guided image filter, arbitrary windowed guided image filter, edge-preserving filter, linear regression, window function

1. INTRODUCTION

Guided image filtering [1] is an edge-preserving filter based on a local linearity assumption. The filtering can smooth images at a constant time with respect to kernel radii, and the response of the filtering result is sharper than bilateral filtering [2]. Also, the guided image filtering can utilize an additional image as guidance signals for defining smoothing weights similar to joint bilateral filtering [3, 4]. These properties make the guided image filtering applicable for various applications, such as high dynamic range imaging [5], texture transferring [6], haze removing [7], texture suppression [8], stereo matching, optical flow estimation [9], depth map refinement [10], free viewpoint view synthesis [11], and so on.

The guided image filter assumes that an output image is a local linear transformation of a guidance image, and also the assumption is supported in a local finite region. The guided image filter consists of a stack of box filtering to realize the supposition. The controllable parameters for the guided image filter are radii of box filtering and Lagrangian of local linear regression. The bilateral filtering, on the contrary, adjusts one more parameter, which are kernel radii, Gaussian range distribution, and spatial one. The spatial distribution of filtering is not adjustable in the guided image filter.

In this paper, we propose an extension of the guided image filtering to support flexible window functions. We named this filter arbitrary windowed guided image filtering (AWGIF). We prove the guided image filtering can be derived from the weighted local linear regression, and the weight functions can be arbitrarily defined. With this generalization, the spatial distribution becomes more flexible, and further, the filtering response becomes more controllable. Besides, we apply infinite impulse response (IIR) filtering or recursive finite impulse response (FIR) filtering for construction of the proposed guided image filter to keep constant time feature.

2. RELATED WORKS

These are three types of extensions for the guided image filtering; the first one is an extension of support regions, the second one is a parameter adaptation extension for halo, and the third is an extension of acceleration of multidimensional signals. The category of the proposed method is the first one.

Extension of support region: The guided image filter assumes that an output image is a linear transform of a guidance image, and the filtering coefficients are gathered by using square windows, i.e., box filtering. Complex texture regions, however, violate this assumption. The cross-based local multipoint filtering [12] avoids the problem by using a box filtering with distorted support windows, which is realized by cross based filtering [13]. The cross based filter changes the filtering domain of box filtering according to the difference between current and neighboring pixel intensities and then perform adjustable filters separably for 2D images. The multipoint local polynomial approximation [14] feather improves the performance. The fully connected guided image filtering [15] employs tree filtering [16] for covering the whole region without filtering across edges. These filters are adaptive filtering in the spatial domain. Instead, the proposed method is weight adaptive. If we use binary weighting, the proposed method becomes these domain adaptive methods.

Extension of parameter adaptation for halo: The guided image filter has halos at large image gaps. The weighted guided image filtering [17] spatially adopts the parameter of ridge regression for reducing halos. Also, gradient domain guided image filtering [18] suppress halo by filtering in the gradient domain.

[*]This work was supported by JSPS KAKENHI (JP17H01764).

[†]This work was supported by JSPS KAKENHI (JP16K16092, JP17H01764).

To appear in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018.*

Extension of acceleration of multidimensional signals: The guided image filter for color or high dimensional signals requires high computational costs. The guided image filter for high dimensional signals, such as hyperspectral image filtering and non-local means filtering [19], is accelerated by the principal component analysis [20, 21]. Hardware-efficient guided image filtering [22] also reduces the computation time by changing inversing matrix operations in guided image filtering into a suitable representation for hardware.

3. REVISITING GUIDED IMAGE FILTERING

We review the guided image filtering by the style that is easy to introduce the proposed extension. The guided image filter linearly transforms a patch in a guidance or filtering image and then averages the transformed patches. Let output signals q be input signals p with noise or texture signals t;

$$q_i = p_i + t_i, \tag{1}$$

where i is a pixel position in a patch. In a square patch ω_k, whose center of the pixel position is k, the whole pixel in ω_k are linearly transformed form guidance signals I. The output pixels in the patch, q', are defined as follows;

$$q'_i = a_k I_i + b_k, \ \ \forall i \in \omega_k, \tag{2}$$

where a_k and b_k are coefficients for linear transformation. We solve these coefficients by using the linear ridge regression.

$$\underset{a_k,b_k}{\arg\min} = \sum_{i\in\omega_k} ((a_k I_i + b_k - p_i)^2 + \epsilon a_k^2), \tag{3}$$

where ϵ is a parameter of Lagrangian, which represents smoothness of the guided image filtering. As a result,

$$a_k = \frac{\mathrm{cov}_k(I,p)}{\mathrm{var}_k(I) + \epsilon}, \tag{4}$$

$$b_k = \bar{p}_k - a_k \bar{I}_k, \tag{5}$$

where $\bar{\cdot}$, var, and cov are mean, variance, and covariance of a patch at the position of a center pixel k, respectively.

The coefficients a_k, b_k are solved per patch windows; thus, the resulting patches are overlapping on the output image. The patch processing is similar to patch-based frequency transformations, such as BM3D [23] and DCT denoising [24, 25]. These patch-based filters average overlapping pixels for redundant treatment. For the case of the guided image filtering, we also average the overlapping regions.

Variables in the patches are only coefficients a, b in the averaging process; thus, we can utilize a simple mean filter instead of using the patch averaging process.

$$\begin{aligned} q_i &= \frac{1}{|\omega|} \sum_{k|i\in\omega_k} (a_k I_i + b_k) \\ &= \bar{a}_i I_i + \bar{b}_i, \end{aligned} \tag{6}$$

where $\sum_{k|i\in\omega_k}$ indicates that a combination of a pixel position i in a patch k are fully averaged, and $|\omega|$ is the number of pixels in a patch.

The guided image filtering utilizes a recursive representation of simple moving average [26] for mean, variance, and covariance computation; hence, the computational time is independent of filtering kernel radii.

4. PROPOSED DEFINITION

The idea of the conventional guided image filter is based on patch-based filters. For extending the guided image filter, we consider patches as a rectangle window function in the context of signal processing societies. In this paper, the window function supports whole image domain, but instead, we assume that linear transform assumption is gradually or weightily kept. For this representation, the assumption of Eq. (2) is solved by weighted local linear regressions.

$$\underset{a_k,b_k}{\arg\min} = \sum_{i\in\Omega} w_{i,k}((a_k I_i + b_k - p_i)^2 + \epsilon a_k^2), \tag{7}$$

where $w_{i,k}$ is a weight between pixels i and k. Ω is the whole image region. Solving this equation, the coefficients a_k and b_k become as follows;

$$a_k = \frac{\hat{\mathrm{cov}}_k(I,p)}{\hat{\mathrm{var}}_k(I) + \epsilon}, \tag{8}$$

$$b_k = \hat{p_k} - a_k \hat{I_k}, \tag{9}$$

where $\hat{x}$ is a weighted average function;

$$\hat{x_k} = \frac{\sum_{i\in\Omega} w_{k,i} x_i}{\sum_{i\in\Omega} w_{k,i}}. \tag{10}$$

$\hat{\mathrm{var}}$, $\hat{\mathrm{cov}}$ are weighted variance and covariance functions by using the weighted averaging instead of simple moving averaging in the process of variance and covariance computation.

For this weighting, we can select arbitrary functions, such as Gaussian distribution $w_{i,k} = \exp(-\frac{\|i-k\|^2}{2\sigma^2})$ and Laplacian distribution $w_{i,k} = \exp(-\frac{|i-k|}{\sigma})$, which are Gaussian filter and double exponential smoothing. Note that when we use binary weighting function, such as box filtering and cross-based filtering, the representation seems domain adaptive filtering. Besides, the weight function does not limit using linear time-invariant (LTI) filters, i.e., box and Gaussian filters, but also supports spatially variant filters, which include the bilateral filter. In an extreme case, the conventional guided image filter is also the spatially variant filter; thus, we can recursively utilize the guided image filter for this weight computation.

For the selection of the weight, the aspect of the computational cost is essential, since filtering domain is the whole image. If the Eq. (10) can be computed in constant-time, the extended guided image filter also has constant time properties. In LTI filter cases, IIR or recursive representations are efficient. The first order IIR filtering expresses the double exponential smoothing, and IIR Gaussian [27] and recursive representation of FIR Gaussian approximate Gaussian filtering [28]. For edge-preserving filtering case, there are several accelerations of the bilateral filter [5, 29, 30, 31]. Bilateral

(a) box (b) Gauss (c) d-exp

Fig. 1: Visualizing kernels of AWGIF with LTI filters.

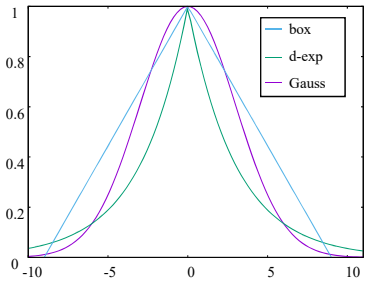

Fig. 2: 1D plots of spacial kernels of various LTI filters. $\sigma = 3$ for exp. based distribution, $r = 4.5$ for box filtering.

filtering has a similar feature of the cross-based filtering [13] and tree filtering [16], these are a component of the local multipoint filtering [12] and the fully connected guided image filtering [15], respectively. The cross-based and tree filtering are domain adaptive, i.e., binary weight adaptive filtering; thus, these filter is also represented by our extension. Furthermore, cross-based filtering is O(r), and tree filtering is more complex, hence, other fast edge-preserving filters are suitable for real-time applications.

The number of coefficient maps is double of the number of pixels. Naïve computation requires tremendous cost, but we can utilize the same approach in the conventional guided image filtering. The variable is only coefficients; thence, we can convolute coefficients instead of averaging all coefficient maps. Note that we can use the other weighted filters used in coefficient estimation in the averaging process, but we use the same weighted filter in convince.

$$\begin{aligned} q_i &= \frac{\sum_{k\in\Omega} w_{i,k}(a_k I_i + b_k)}{\sum_{k\in\Omega} w_{i,k}} \\ &= \hat{a}_i I_i + \hat{b}_i. \end{aligned} \tag{11}$$

For changing the filtering properties more, swapping this post smoothing filter is also important. This is our future work.

5. EXPERIMENTAL RESULTS

Figures 1, 2 show the kernels of the adaptive windowed guided image filtering (AWGIF) with LTI filters, which are box filtering (box), Gaussian filtering (Gauss), and dual exponential smoothing (d-exp). We focus the flat region in Fig. 1, where near the hat in Lenna image. On this region, the conventional guided image filtering becomes twice-iterated box filtering. When the kernel radius of box filtering is r, the filtering response is the tent or triangle filtering, whose radius is $2r$ (See Fig. 2.). The tent filter has Manhattan distance in 2D space; thus, the filter is isotropic. In the case of Gaussian convolution, the response becomes dual iterated Gaussian filtering; hence, the filter's distribution becomes $\sqrt{2}\sigma$. Note that Gaussian filtering is isotropic filtering. The distance in 2D space of the d-exp kernel is $L_n(n < 1)$ norm; thence, the kernel shape becomes sharper than the above filters. Besides, the filter has longer tails than the others. Focusing the edge region, where the shoulder part, each filter has edge-preserving properties. We should switch the filters in each application owing to the characteristics of these filters.

Table 1: Denoising (PSNR [dB]) with various LTI filters. Parameters of each filter are optimized to obtain the best PSNR.

noise σ	box	d-exp	Gauss
5	37.68	37.31	**37.76**
10	33.31	33.26	**33.40**
15	31.50	31.14	**31.72**

For denoising applications, Tab. 1 shows denoising results by AWGIF with various LTI filters. The results show that the guided image filtering based on Gaussian filtering has the best performance. Gaussian distribution has more power near the center pixel and is isotropic filtering. Therefore, the characteristics are suitable for denoising. The dual exponential smoothing is the farthest from the suitable property. Note that the guided image filter is not specialized for denoising. If the users need better denoising, BM3D [23] and DCT denoising [24, 25] is recommended.

For detail enhancement applications, suppression of halos is essential. We use iterative filtering of guided image filtering for a base signal generation, and detail signals are the subtraction of the base signal from an input signal. Figure 3 shows detail enhancement results. In the iterative guided image filtering, halos are inevitable, but the dual exponential filter can weaken the hales in synthetic and real images. Applying the proposed strategy to the weighted guided image filtering [17] and the gradient domain guided image filtering [18] would reduce more halos.

Figure 4 shows the dehazing results [7]. Transition regions between hazy and non-hazy regions have white halos on these results. The halo is caused by the fact that the total of the weight in the kernel of edge-preserving filtering is just one. Therefore we should distribute high contrast values to low contrast regions where should be quite larger than halo's region for reducing white halos without the changing image contrast. For this character, box filtering is suitable, because the kernel has large power at far from the center point. More suitable kernel shape is box-like kernel shape after dually iterated filtering. Note that iterated box filtering is the tent kernel, not box kernel.

Figure 5 shows the filtering result of non-local linear characteristic regions. In complexly and binary changing region,

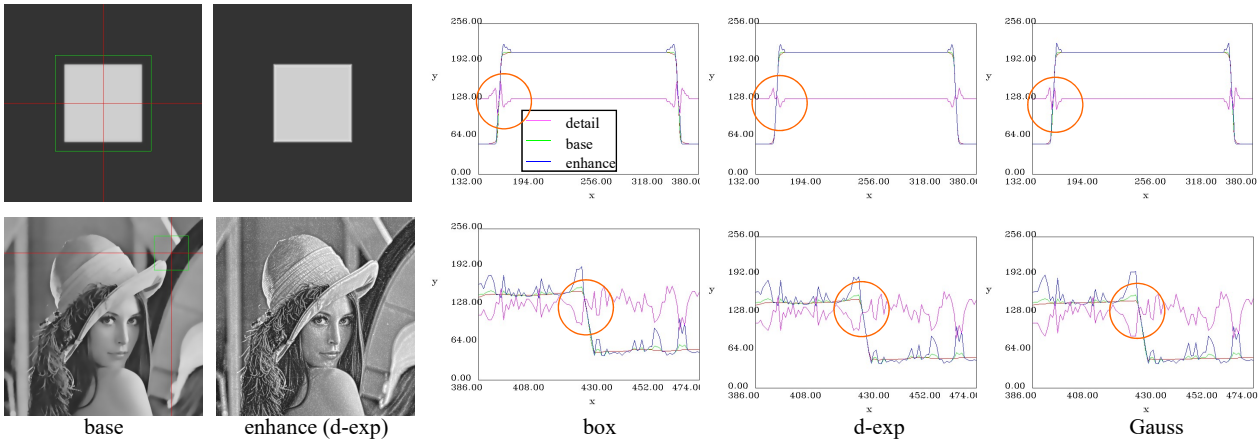

Fig. 3: Detail enhancement. The top and bottom are synthetic and real image results, respectively. The profile line is along the horizontal red line with clipping the region in the green rectangle, where indicates the base image. The orange circles in the profile signals indicate notable halos. The dual exponential kernel mostly weakens the halo, and box kernel has the largest halo.

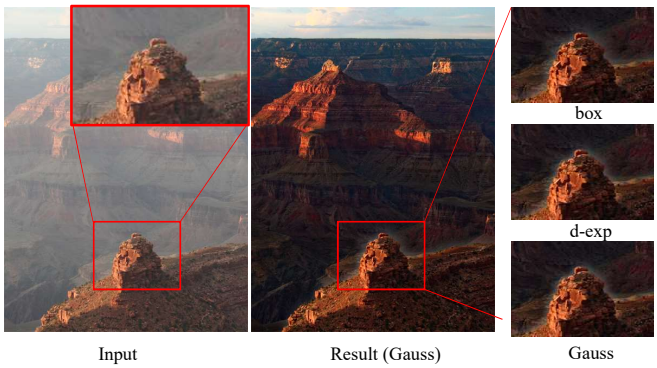

Fig. 4: Dehazing result [7]. Box based AWGIF, which is the conventional approach, has the best performance.

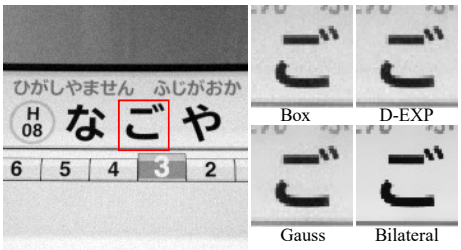

Fig. 5: Violation cases of local linearity assumption. $\epsilon = 0.3$, $r = 5$. (only bilateral case, $\sigma_r = 30$). LTI based filters are weak in this region, but the bilateral based filter works well.

such as embedded text regions, naïve guided image filtering violates the assumption of local linear transformation. Local-multipoint filtering [12] deals well such regions by using filtering domain adaptation. Our strategy of AWGIF achieves similar effects by using edge-preserving filtering in weighted averaging in Eq. (10). We used compressive bilateral filtering with joint way [31], which is O(1). Therefore, the computational time of the proposed representation with edge-preserving filtering is faster than the cross-based method. Furthermore, bilateral filtering based filtering has more continuous kernel shape than the cross-based one.

Table 2 shows the computational cost of AWGIF with various LTI filters on various size images. All filters are vectorized by AVX with single thread implementation on Intel Core i7 6700K 4 GHz and compiled by Visual Studio 2015. We used fast implementation of box [32] and Gaussian [33] filters, and also optimized codes for dual exponential smoothing. For Gaussian filtering, we used the DCT-5 based sliding $O(1)$ implementation. Box filtering is the fastest and Gaussian filtering is the second best. The d-exp smoothing is the IIR based $O(1)$ implementation, which requires much time

Table 2: Computational time [ms] of AWGIF with LTI filters.

size	box	d-exp	Gauss
512×512	4.68	6.72	6.23
1024×1024	22.58	41.74	29.14
2048×2048	88.65	198.72	118.78

of image scanning; hence, this filter is slower than the others. For more acceleration, parallelization of guided image filtering is effective. The implementation is available from our website[1].

6. CONCLUSION

In this paper, we generalize the guided image filtering to have arbitrary window shape. The proposed representation is derived by using the weighted linear regression. With this representation, we can use any weighted averaging filtering, such as Gaussian Laplacian, bilateral filtering, and even including guided image filtering itself. Each filter should be switched for each application. The better filters for applications, which are detail enhancement, denoising, haze remove, and text filtering, are reported in experimental results.

[1] http://fukushima.web.nitech.ac.jp/research/awgif/

7. REFERENCES

[1] K. He, J. Shun, and X. Tang, "Guided image filtering," in *Proc. European Conference on Computer Vision (ECCV)*, 2010.

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

[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. on Graphics*, vol. 23, no. 3, pp. 664–672, 2004.

[4] E. Eisemann and F. Durand, "Flash photography enhancement via intrinsic relighting," *ACM Trans. on Graphics*, vol. 23, no. 3, pp. 673–678, 2004.

[5] 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, 2002.

[6] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Trans. on graphics*, vol. 22, no. 3, pp. 313–318, 2003.

[7] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.

[8] Q. Zang, X. Shen, L. Xu, and J. Jia, "Rolling guidance filter," in *Proc. European Conference on Computer Vision (ECCV)*, 2014.

[9] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 504–511, 2013.

[10] T. Matsuo, N. Fukushima, and Y. Ishibashi, "Weighted joint bilateral filter with slope depth compensation filter for depth map refinement," in *Proc. International Conference on Computer Vision Theory and Applications (VISAPP)*, 2013.

[11] N. Kodera, N. Fukushima, and Y. Ishibashi, "Filter based alpha matting for depth image based rendering," in *Proc. IEEE Visual Communications and Image Processing (VCIP)*, 2013.

[12] J. Lu, K. Shi, D. Min, L. Lin, and M. N. Do, "Cross-based local multipoint filtering," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[13] K. Zhang, J. Lu, and G. Lafruit, "Cross-based local stereo matching using orthogonal integral images," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 19, no. 7, pp. 1073–1079, 2009.

[14] X.i Tan, C. Sun, and T. D Pham, "Multipoint filtering with local polynomial approximation and range guidance," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.

[15] L. Dai, M. Yuan, F. Zhang, and X. Zhang, "Fully connected guided image filtering," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2015.

[16] L. Bao, Y. Song, Q. Yang, H. Yuan, and G. Wang, "Tree filtering: Efficient structure-preserving smoothing with a minimum spanning tree," *IEEE Trans. on Image Processing*, vol. 23, no. 2, pp. 555–569, 2014.

[17] J. Zheng Z. Li, Z. Zhu, W. Yao, and S. Wu, "Weighted guided image filtering," *IEEE Trans. on Image Processing*, vol. 24, no. 1, pp. 120–129, 2015.

[18] W. Chen C. Wen F. Kou and Z. Li, "Gradient domain guided image filtering," *IEEE Trans. on Image Processing*, vol. 24, no. 11, pp. 4528–4539, 2015.

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

[20] S. Fujita and N. Fukushima, "High-dimensional guided image filtering," in *Proc. International Conference on Computer Vision Theory and Applications (VISAPP)*, 2016, pp. 27–34.

[21] S. Fujita and N. Fukushima, *Extending Guided Image Filtering for High-Dimensional Signals*, vol. 693, pp. 439–453, Springer International Publishing, 2017.

[22] L. Dai, M. Yuan, Z. Li, X. Zhang, and J. Tang, "Hardware-efficient guided image filtering for multi-label problem," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[23] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Trans. on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[24] G. Yu and G. Sapiro, "Dct image denoising: A simple and effective image denoising algorithm," *Image Processing On Line*, vol. 1, pp. 1, 2011.

[25] S. Fujita, N. Fukushima, M. Kimura, and Y. Ishibashi, "Randomized redundant dct: Efficient denoising by using random subsampling of dct patches," in *Proc. SIGGRAPH Asia, Technical Briefs*, 2015.

[26] F. C. Crow, "Summed-area tables for texture mapping," in *Proc. ACM SIGGRAPH*, 1984, pp. 207–212.

[27] R. Deriche, "Recursively implementating the gaussian and its derivatives," Research Report RR-1893, INRIA, 1993.

[28] K. Sugimoto and S. Kamata, "Fast gaussian filter with second-order shift property of dct-5," in *Proc. IEEE International Conference on Image Processing (ICIP)*, 2013.

[29] K. N. Chaudhury, "Constant-time filtering using shiftable kernels," *IEEE Signal Processing Letters*, vol. 18, no. 11, pp. 651–654, 2011.

[30] N. Fukushima, S. Fujita, and Y. Ishibashi, "Switching dual kernels for separable edge-preserving filtering," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015.

[31] K. Sugimoto and S.-I. Kamata, "Compressive bilateral filtering," *IEEE Trans. on Image Processing*, vol. 24, no. 11, pp. 3357–3369, 2015.

[32] M. Nakamura and N. Fukushima, "Fast implementation of box filtering," in *Proc. International Workshop on Advanced Image Technology (IWAIT)*, 2017.

[33] K. Sugimoto, S. Kyochi, and S. Kamata, "Universal approach for dct-based constant-time gaussian filter with moment preservation," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018.