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 guided 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 halos by filtering in the gradient domain.
Extension of acceleration of multidimensional signals: The
guided image filter for color or high dimensional signals re-
quires high computational costs. The guided image filter for
high dimensional signals, such as hyperspectral image filter-
ing 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 fil-
tering 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 $t$;

$$q_i = p_i + t_i,$$

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

$$q_i' = a_k I_i + b_k, \quad \forall i \in \omega_k,$$

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

$$\arg \min_{a_k,b_k} \sum_{i \in \omega_k} ((a_k I_i + b_k - p_i)^2 + \epsilon a_k^2),$$

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

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

$$b_k = \hat{p}_k - a_k \hat{I}_k,$$

where $\hat{\cdot}$, $\text{var}$, and $\text{cov}$ 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 fre-
cquency 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.

$$q_i = \frac{1}{|\omega|} \sum_{k : i \in \omega_k} (a_k I_i + b_k)$$

$$= \hat{a}_i I_i + \hat{b}_i,$$

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

The guided image filtering utilizes a recursive representa-
tion of simple moving average [26] for mean, variance, and
covariance computation; hence, the computational time is in-
dependent 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 con-
text of signal processing societies. In this paper, the window
function supports whole image domain, but instead, we as-
sume that linear transform assumption is gradually or weight-
ily kept. For this representation, the assumption of Eq. (2) is
solved by weighted local linear regressions.

$$\arg \min_{a_k,b_k} \sum_{i \in \Omega} w_{i,k} ((a_k I_i + b_k - p_i)^2 + \epsilon a_k^2),$$

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

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

$$b_k = \hat{p}_k - a_k \hat{I}_k,$$

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

$$\hat{x}_k = \frac{\sum_{i \in \Omega} w_{k,i} x_i}{\sum_{i \in \Omega} w_{k,i}}.$$
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.

\[
q_i = \frac{\sum_{k \in \Omega} w_{i,k} (a_k I_i + b_k)}{\sum_{k \in \Omega} w_{i,k}} \quad (11)
\]

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.

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.

<table>
<thead>
<tr>
<th>noise (\sigma)</th>
<th>box</th>
<th>d-exp</th>
<th>Gauss</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>37.68</td>
<td>37.31</td>
<td>37.76</td>
</tr>
<tr>
<td>10</td>
<td>33.31</td>
<td>33.26</td>
<td>33.40</td>
</tr>
<tr>
<td>15</td>
<td>31.50</td>
<td>31.14</td>
<td>31.72</td>
</tr>
</tbody>
</table>

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 halos 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,
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 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.

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.
7. REFERENCES


