Presented at the DLSU Research Congress 2019
De La Salle University, Manila, Philippines
June 19 to 21, 2019

Structure-Preserving Guided Filtering

Carlo Noel Ochotorena[1,2,*] and Yukihiko Yamashita[2]
[1] *De La Salle University, Manila, Philippines*
[2] *Tokyo Institute of Technology, Tokyo, Japan*
**Corresponding Author: carlo.ochotorena@dlsu.edu.ph*

Abstract: Image filtering is a vital aspect in many imaging and computer vision technologies. The guided filter is one of the most widely used of the modern filters, in part due to its low complexity construction. However, the original guided filter also has several limitations, particularly the production of halos, detail halos, and mishandling of inconsistent structures. A newer filter, the anisotropic guided filter, addresses most of these problems but is, under certain conditions, sensitive to the density of details in the image. This work proposes a multiscale adaptation to address these new limitations. Experiments on edge-preserving smoothing and texture removal demonstrate the improvements in filtering quality of the proposed filter on real images.

Key Words: image filtering, adaptive filtering, guided filtering

1. INTRODUCTION

Computer vision and image processing technologies are omnipresent in the modern world and have integrated themselves into daily living. Some of these technologies empower fully-automated systems that are making their way from being used, solely, on an industrial level to more tangible systems that can be felt by the average person. A good example of this would be the development of self-driving vehicles in the past decade (Paden, Cáp, Yong, Yershov, & Frazzoli, 2016).

On a less grand scale, these technologies can also be used in assistive tasks. In the context of vehicles, for instance, driver-assistance technologies allow for added convenience and safety factors without removing humans from the equation (Sun, Bebis, & Miller, 2006; Geronimo, Lopez, Sappa, & Graf, 2010). Likewise, simple tasks such as object identification can now, readily, be performed using accessible computer vision technologies on mobile platforms (Tiefenbacher, Gillich, Schott, & Rigoll, 2016).

The term "computer vision", itself, encompasses a wide range of complex technologies dealing with the processing of visual information. Despite this wide variation in design, a common component in many of these systems is a filtering stage. Such a filter can be used to remove noise from the image or, perhaps, isolate salient features of the image for further processing.

The importance of filtering has led to much work in the development of filtering techniques on visual data. On one hand of these developments, we have global filters that focus on delivering high-quality output at the cost of computation. Some important filters on this front include the weighted least-squares (WLS) filter (Farbman, Fattal, Lischinski, & Szeliski, 2008), the fast weighted l1 filter (FWL1) (Kim, Min, Ham, & Sohn, 2017), and the relative total variation (RTV) filter (Xu, Yan, Xia, & Jia, 2012).

On the other side, there are filters that focus on computational efficiency, while trading off the filtering quality. These include the bilateral filter (Tomasi & Manduchi, 1998) and the guided filter (He,

Sun, & Tang, 2013). The latter is known to operate very efficiently but is still capable of producing good output. This filter, however, has its own share of limitations including the production of haloing artefacts, addressed by the newer weighted guided image filter (WGIF) (Li, Zheng, Zhu, Yao, & Wu, 2015), as well as the production of detail halos and mishandling of inconsistent structures, tackled by our previous work, the anisotropic guided filter (AnisGF) (Ochotorena & Yamashita, Anisotropic Guided Filtering, 2019).

Even with the significant improvements introduced by the AnisGF design, some notable problems remain due to the scale-dependent formulation of the filter. For instance, dense details appearing in an image, sometimes remain unfiltered due to the scale-dependent nature of the filter.

This work addresses such limitations by combining information across multiple filtering scales, as was implemented by another one of our filter designs, the structure-preserving image filter (SPIF) (Ochotorena & Yamashita, Multi-scale structure-preserving image filtering, 2017). We discuss, in the succeeding section, the mathematical development of the new filter, referred to as the structure-preserving guided filter (SPGF). In addition, we highlight the performance of the filter, particularly for edge-preserving smoothing and texture removal tasks.

2. FILTER DESIGN

2.1 Original Guided Filter

The original guided filter introduced by He, et al. (2013) describes a filtering operation given a small patch $\boldsymbol{x}_i$ in an input image and a corresponding patch $\boldsymbol{g}_i$ from a guide image, both taken from the same position in the images described by the index i.

The input, in this case, is the noisy image to be filtered while the guide is either the input image itself or a separate image that is structurally related to the input. The availability of guidance data, while seemingly rare, actually appears in many computer vision tasks.

The guided filter operates based on a linear transform of the guide patch to replace the corresponding input patch:

$$a_i = \frac{\mathrm{cov}(\boldsymbol{x}_i - \bar{x}_i, \boldsymbol{g}_i - \bar{g}_i)}{\mathrm{var}(\boldsymbol{g}_i - \bar{g}_i) + \gamma} \qquad \text{(Eq. 1)}$$

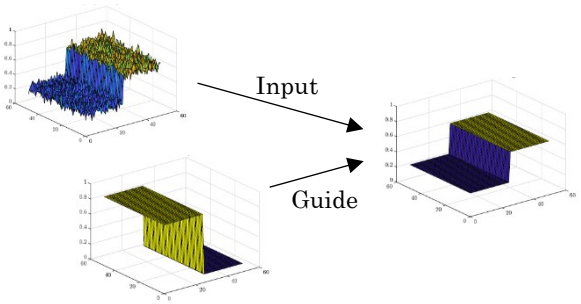

Fig. 1. The guided filtering process performs a linear transformation on the guide signal such that it closely resembles the input image.

$$b_i = \bar{x}_i - a_i \bar{g}_i \qquad \text{(Eq. 2)}$$

where $\bar{\cdot}$ denotes the vector mean operator and n represents the number of pixels in each patch. A more visual representation of this operation can be seen in Fig. 1.

Because the guided filter operates on a patch basis, every pixel in the full image is considered for multiple overlapping patches. In order to resolve this, the average transformation is considered to produce the final filter output:

$$\bar{a}_i = \frac{1}{n} \sum_{j \in \mathcal{N}(i)} a_j \qquad \text{(Eq. 3)}$$

$$\bar{b}_i = \frac{1}{n} \sum_{j \in \mathcal{N}(i)} b_j \qquad \text{(Eq. 4)}$$

$$\hat{x}_i = \bar{a}_i g_i + \bar{b}_i \qquad \text{(Eq. 5)}$$

2.2 Anisotropic Guided Filter

The original guided filter presents a computationally efficient yet highly effective filter in many computer vision tasks. Nonetheless, researchers, including the original authors, have identified certain limitations, particularly the manifestation of halos (Li, Zheng, Zhu, Yao, & Wu, 2015). In addition to the said artefacts, we noted in our previous work that the guided filter is not capable of stronger filtering as it produces detail halos under such conditions (Ochotorena & Yamashita, Anisotropic Guided Filtering, 2019). The guided filter also behaves erratically when the structures in input and the guide patches are not similar to each other.

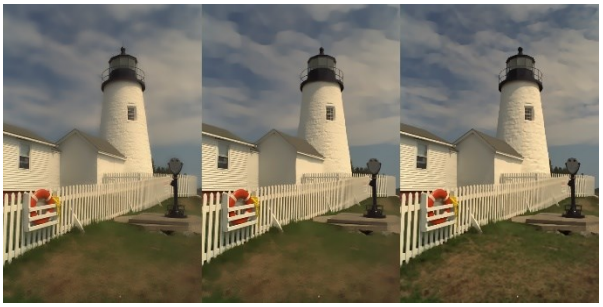

Fig. 2. Comparison of the edge-preserving smoothing using the original guided filter (left), the AnisGF (centre), and the SPGF (right).

Fig. 3. Comparison of the texture removal using the original guided filter (left), the AnisGF (centre), and the SPGF (right).

Our analysis of these limitations showed that the problem was centred on the weakly anisotropic (i.e. only minimally adaptive to the spatial direction) filtering behaviour enforced by the unweighted averaging in Eq. 3 and 4. To resolve this, we introduced a strongly anisotropic weighted averaging term:

$$\tilde{a}_i = \sum_{j \in \mathcal{N}(i)} w_j \, a_j \qquad \text{(Eq. 6)}$$

$$\tilde{b}_i = \sum_{j \in \mathcal{N}(i)} w_j \, b_j \qquad \text{(Eq. 7)}$$

$$\hat{x}_i = \tilde{a}_i g_i + \tilde{b}_i \qquad \text{(Eq. 8)}$$

where the weights w_j are obtained using a local smoothness criterion:

$$\underset{w_j}{\operatorname{argmin}} \sum_j \left[w_j E_s(j) \right]^2 + \epsilon \sum_j \left(e^{-\frac{(\boldsymbol{p}_i - \boldsymbol{p}_j)^2}{2\sigma^2}} - w_j \right)^2 \qquad \text{(Eq. 9)}$$

In this formulation, the smoothness cost $E_s(j)$ is used to decide the weight assigned to that neighbourhood and, in practice, can be derived from the variance in Eq. 1 to save on computation. A spatial regularisation term, controlled by ϵ, encourages a Gaussian-like weighting scheme with respect to Euclidean pixel distance $(\boldsymbol{p}_i - \boldsymbol{p}_j)^2$. Apart from these two terms, the resulting weights are normalised to ensure that they will always have a sum of 1. The resulting the weighting scheme is described below:

$$w_j = \frac{\epsilon e^{-\frac{(\boldsymbol{p}_i - \boldsymbol{p}_j)^2}{2\sigma^2}} \Big/ \left(\operatorname{var}(\boldsymbol{g}_j + \bar{g}_j)^{\alpha} + \epsilon \right)}{\sum_j \epsilon e^{-\frac{(\boldsymbol{p}_i - \boldsymbol{p}_j)^2}{2\sigma^2}} \Big/ \left(\operatorname{var}(\boldsymbol{g}_j + \bar{g}_j)^{\alpha} + \epsilon \right)} \qquad \text{(Eq. 10)}$$

where α is used to control the degree of anisotropy of the filter.

2.3 Structure-Preserving Guided Filter

An inherent problem of the AnisGF is that it operates, solely, within one scale (i.e. the filter size is fixed). For most applications, this has little consequence, however, some dense details may remain unfiltered under certain conditions.

For this reason, it is useful to describe a multiscale variation of the AnisGF. We adapt the dyadic scale filtering from another of our previous works, the structure-preserving image filter (SPIF) to efficiently incorporate multiple scales into the filter (Ochotorena & Yamashita, Multi-scale structure-preserving image filtering, 2017). The new filter, referred to as the structure-preserving guided filter (SPGF) modifies the weighting function in Eq. 10 as follows:

$$w_j = \frac{\dfrac{1}{\left[n_s \cdot \operatorname{var}(g_{s,j} - \bar{g}_{s,j}) \right]^{\alpha} + \epsilon \beta^{\lambda - s}}}{\sum_{s=0}^{\lambda} \sum_j \dfrac{1}{\left[n_s \cdot \operatorname{var}(g_{s,j} - \bar{g}_{s,j}) \right]^{\alpha} + \epsilon \beta^{\lambda - s}}} \qquad \text{(Eq. 11)}$$

where s and λ are integers that describe the current and maximum scale, respectively, n_s is the number of pixels in a patch at the current scale, and $\beta \geq 1$ is a constant that controls the multiscale behaviour of the filter. The scales, themselves are connected to the window size $(k \times k)$ of the filter as follows:

$$\mathcal{N}(s) = k \times k, \qquad k = 2^s \qquad \text{(Eq. 12)}$$

3. RESULTS AND DISCUSSION

3.1 Experiments

In order to validate the performance of the new filter, we conduct experiments comparing the original guided filter, the AnisGF, and the proposed SPGF. The first of these experiments is with edge-preserving smoothing where the goal is to obtain a smoothened version of the image without destroying the major structures. Looking at the results in Fig. 2, one can easily recognise that the guided filter fails at smoothing the image and produces an image with less contrast than the original. On the other hand, both AnisGF and SPGF perform smoothing well. A subtle distinction between the two can be observed in the adobe area above the orange buoy. In AnisGF, this region is unfiltered as it contains dense details. In contrast, with SPGF, the region has been smoothed.

A second experiment, dealing with the removal of textures from an image can be seen in Fig. 3. Again, the original guided filter is incapable of operating at these higher filtering strengths and tends to blur important details. With the AnisGF, some mosaiced areas are too dense which affects the filtering behaviour. In these areas, the tiling of the original image make their way to the filtered output. In comparison, the SPGF is scale-adaptive and can more readily handle filtering in these tight regions. These results demonstrate the improvements of the SPGF design over its predecessor, particularly when dealing with densely clustered details.

3.2 Complexity

An important aspect to any filter design is the computational complexity associated with it. As previously mentioned, high-quality filters are often associated with larger complexities. On the other hand, fast filter designs are usually limited in quality. In our previous work, we demonstrated how the AnisGF, despite being a linear-time $\mathcal{O}(n)$ filter, delivers high-quality filtering. The SPGF, being a multiscale extension on the latter, is a bit more complex operating at $\mathcal{O}(\lambda n)$ complexity. For typical filtering scenarios, however, the value of λ remains low. Considering, for instance, that a $\lambda = 5$ will result in a maximum window size of 33×33, it can be said that this complexity remains relatively low compared to optimisation-based filtering techniques. At the same time, it should be noted that the calculations involved at each scale is independent of the other scales, allowing for the parallel calculation of the results prior to the final normalisation process. This property makes it easier to distribute computation in hardware.

4. CONCLUSIONS

This work introduced a novel filter, the structure-preserving guided filter (SPGF) that builds upon the framework of its predecessor, the anisotropic guided filter (AnisGF). The new filter features a multiscale formulation and was shown to handle filtering even when important structures appear at various densities. This makes the proposed filter more versatile compared to the previous versions, while maintaining a relatively low complexity of $\mathcal{O}(\lambda n)$.

While this new filter performs well, it operates exclusively with square neighbourhoods. Future work may be done in incorporating shape-adaptive neighbourhoods in the formulation to further improve the adaptivity of the filter.

5. REFERENCES

Farbman, Z., Fattal, R., Lischinski, D., & Szeliski, R. (2008). Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Transactions on Graphics (TOG)*, *27*, p. 67.

Geronimo, D., Lopez, A., Sappa, A., & Graf, T. (2010). Survey of Pedestrian Detection for Advanced Driver Assistance Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*(7), 1239-1258.

He, K., Sun, J., & Tang, X. (2013). Guided Image Filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35*(6), 1397-1409.

Kim, Y., Min, D., Ham, B., & Sohn, K. (2017). Fast Domain Decomposition for Global Image Smoothing. *IEEE Transactions on Image Processing, 26*(8), 4079-4091.

Li, Z., Zheng, J., Zhu, Z., Yao, W., & Wu, S. (2015). Weighted Guided Image Filtering. *IEEE Transactions on Image Processing, 24*(1), 120-129.

Ochotorena, C., & Yamashita, Y. (2017). Multi-scale structure-preserving image filtering. *2017 IEEE*

19th International Workshop on Multimedia Signal Processing (MMSP), (pp. 1-6).

Ochotorena, C., & Yamashita, Y. (2019). Anisotropic Guided Filtering. *Manuscript submitted for publication*.

Paden, B., Cáp, M., Yong, S., Yershov, D., & Frazzoli, E. (2016, March). A Survey of Motion Planning and Control Techniques for Self-Driving Urban Vehicles. *IEEE Transactions on Intelligent Vehicles, 1*(1), 33-55.

Sun, Z., Bebis, G., & Miller, R. (2006). On-road vehicle detection: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28*(5), 694-711.

Tiefenbacher, P., Gillich, J., Schott, P., & Rigoll, G. (2016). Comparison of mobile touch interfaces for object identification and troubleshooting tasks in augmented reality. *2016 IEEE Virtual Reality (VR)*, (pp. 297-298).

Tomasi, C., & Manduchi, R. (1998). Bilateral filtering for gray and color images. *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, (pp. 839-846).

Xu, L., Yan, Q., Xia, Y., & Jia, J. (2012). Structure extraction from texture via relative total variation. *international conference on computer graphics and interactive techniques, 31*(6), 139.