IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 15, 2020, accepted February 26, 2020, date of publication March 2, 2020, date of current version March 12, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2977408

Structure-Aware Texture Filtering Based on Local Histogram Operator

YANG LIU, GUANGDA LIU, HONGLIANG LIU, AND CHANGYING LIU
College of Instrumentation and Electrical Engineering, Jilin University, Changchun 130021, China

Corresponding author: Changying Liu (liuchangy@jlu.edu.cn)

This work was supported by the Jilin Scientific and Technological Development Program under Grant 20150204053GX and Grant 20160101286JC.

ABSTRACT Smoothing the multi-scale textures with strong gradient while maintaining the weak structures is still a challenging work for the existing texture filtering methods. In this paper, we propose a variant of the bilateral filter to address this issue. Unlike the classical bilateral filter, which usually uses a fixed range kernel for texture filtering, the variant uses an adaptive range kernel, that is, if a pixel is located near a structure, the range kernel assigned to the pixel is small, otherwise, the range kernel is large. This mechanism ensures the preservation of sharp edges. In order to perform structure measurement, we propose a local histogram operator, which can identify structures from textures by comparing the difference in color distribution between the left and right half neighboring windows centered on the pixel of interest. In addition, after once bilateral filtering, instead of multiple iterations, a novel anisotropic filter is designed to remove residual textures, which can avoid edge blurring to some extent. Experimental results demonstrate the effectiveness of our method by comparing it with the state-of-the-art methods in preserving weak structures as well as suppressing textures with strong gradient or varying scales.

INDEX TERMS Texture filtering, bilateral filter, structure measurement, local histogram operator, adaptive range kernel, anisotropic filter.

I. INTRODUCTION

Texture filtering is a fundamental and important research direction in computational photography and computer vision. The goal of texture filtering is to suppress textures as much as possible while preserve the main structures, which is of immediate use for a variety of applications in the fields of detail enhancement [1], image segmentation [2], optical flow estimation [3], super-resolution reconstruction [4], and high dynamic range compression [5], etc. However, it is still a challenge due to the variety and complexity of image textures (as shown in Fig. 2).

In recent years, many methods for texture filtering have been proposed, of which bilateral filtering algorithm [6] is widely used because of its good performance. The bilateral filter algorithm is based on Gaussian filtering, which uses the spatial kernel along with the range kernel. That is, when calculating the filtering weight, the algorithm not only considers the influence of spatial distance on the filtering weight, but also uses the intensity difference between the pixel of interest and its neighbors. The neighboring pixel farther from the pixel of interest is given a small weight, otherwise, given a large weight. Similarly, the neighboring pixel having a large intensity difference (e.g., the pixels are from different sides of an edge) with the pixel of interest is given a small weight, otherwise, given a large weight. This mechanism, which avoids the mixing of pixels with large intensity differences, ensures the edge pixels with continuous intensity gradation can be preserved to some extent.

In [7], the bilateral filter was extended by substituting a texture description image in the range kernel, which is called joint bilateral filter. Since the standard bilateral filter usually adopted a fixed range kernel, it cannot perform sharpening, as shown in Fig. 3. Then in [8], [9], an adaptation of range kernel was used to handle images of high complexity, where features of multiple scales coexist. The basic idea is to discriminate structures from textures in advance, then give small smoothing weights to pixels at texture edges to ensure the preservation of sharp edges and give large smoothing weights to pixels inside textures to avoid texture residue.

The associate editor coordinating the review of this manuscript and approving it for publication was Qichun Zhang.

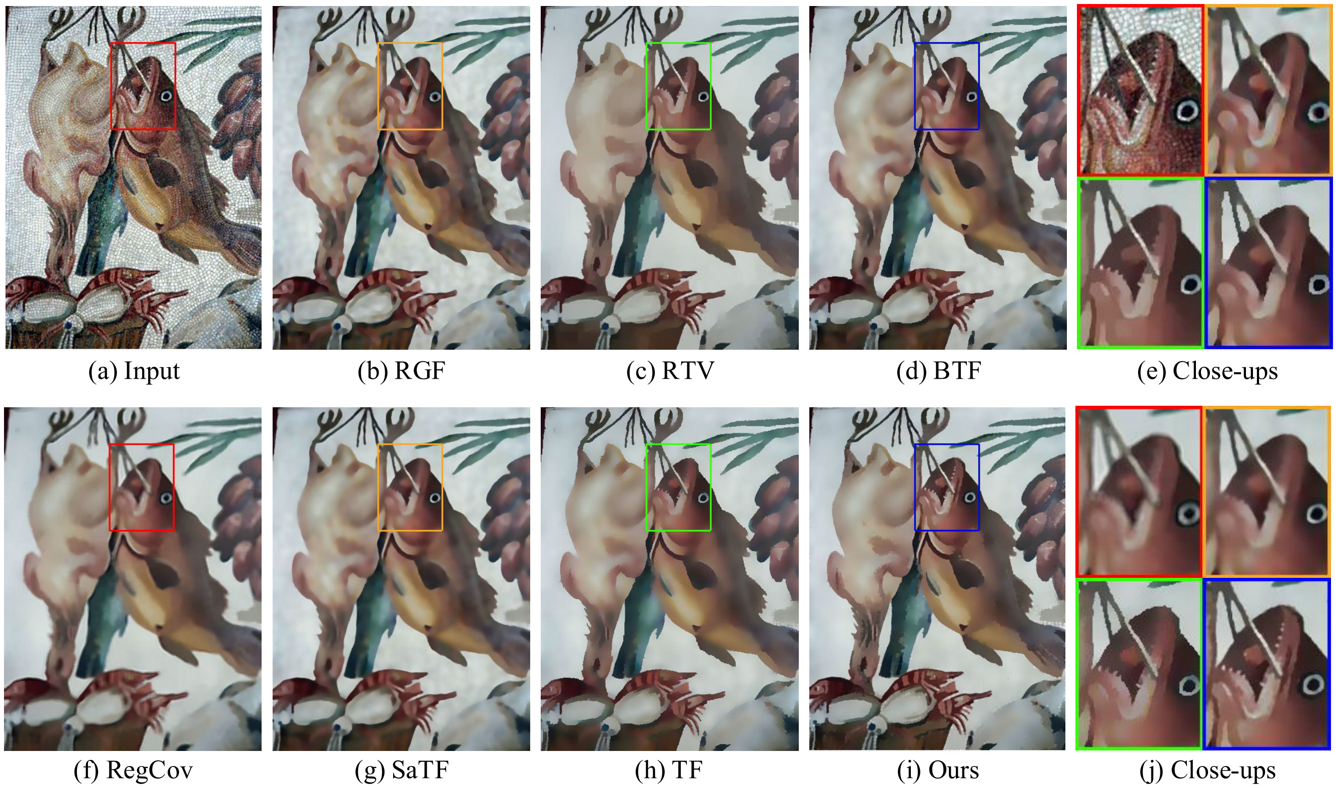


FIGURE 1. Results and comparison on the mosaic art 'fish'. (a) Input image, (b) RGF [$\sigma_s = 4$, $\sigma_r = 0.05$, $n_{iter} = 5$], (c) RTV [$\lambda = 0.015$, $\sigma = 2$, $\varepsilon_s = 0.02$, $n_{iter} = 4$], (d) BTF [$k = 2$, $n_{iter} = 5$], (f) RegCov [$k = 15$, $\sigma = 0.2$, *Model* 1], (g) SaTF [$\sigma = 3$, $\sigma_r = 0.1$, $n_{iter} = 3$], (h) TF [$\sigma = 0.01$, $\sigma_s = 3$], and (i) our method. In this example, the input image contains textures and small structures, such as teeth of fish and the details of shrimp. Our method outperforms the methods of RGF, RegCov, and SaTf in terms of both image structure preservation and texture smoothing. In the method of RTV, some shading is lost in their results where the back of the fish looks flat, which is effectively restored by our method. The method of TF over-blurs some small structures, such as the palate teeth of the fish, which are preserved better in our method.

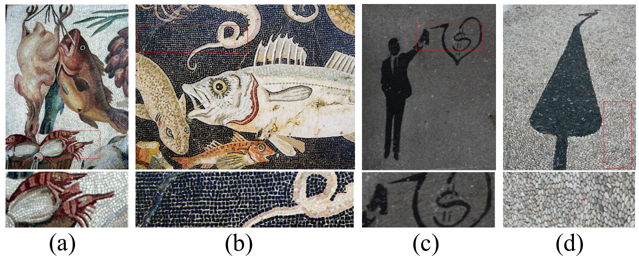


FIGURE 2. Examples of textures with different patterns. Images with (a) small details, (b) high-contrast textures, (c) irregular textures, and (d) multi-scale textures.

These methods heavily rely on the effect of structure measurement, which is still a challenge. In addition, in order to smooth textures with high-contrast and large-scale texture patterns, multiple iterations of bilateral filtering are necessary, which inevitably leads to edge blurring.

Local histograms of images contain a great deal of useful information, which can be used in texture filtering. Since a variety of popular image filters can be expressed as functions of the local histogram, they can be classified as histogram-based filters, such as median filter, bilateral filter, and mean-shift. For these histogram-based filters, the computation is still a challenge work when dealing with large neighborhoods. To solve this problem, [10] present an efficient and practical computing method, which is based on look-up table and well suited to modern GPU hardware. Similarly, [11] propose a new joint-histogram representation and other schemes to accelerate the weighted median solver or filter, which can reduce the computation complexity from $O\left(r^2\right)$ to $O\left(r\right)$, where r is the kernel size.

In this paper, instead of accelerating the calculation of filters, we use the local histogram to design a novel structure measurement operator, which distinguishes structures from textures by calculating the difference in color distribution on the left and right half neighboring windows centered on each pixel and in multiple directions. Then, the structure measurement result of the local histogram operator is used to design the smoothing weights of the joint bilateral filter. In addition, to avoid edge blurring by multiple iterations of bilateral filtering, a variant of anisotropic filter is proposed to optimize filtering results.

The experimental results show the effectiveness of our improvements compared with the state-of-the-art methods, particularly with regard to smoothing out textures with high contrast while preserving small structures, prominent edges and shading well. The main contributions of this paper are follows: (1) we introduce a structure measurement operator

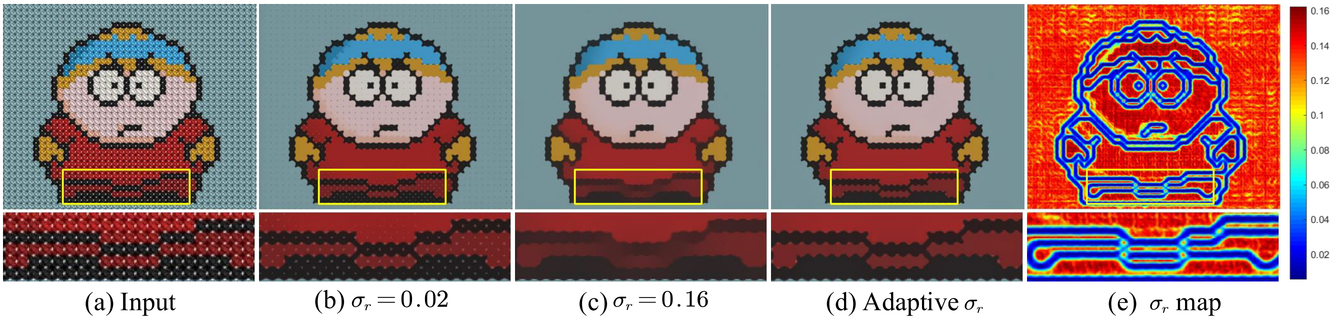


FIGURE 3. Texture filtering using the classical bilateral filter and ours. If we use the former, small enough range kernel σ_r is needed to preserve the strong edges; however, this causes coarse textures to not be smoothed out, as shown in (b). On the other hand, as shown in the enlarged box in (c), a large σ_r causes the edges to get blurred. As shown in (d), our method with adaptive σ_r achieves both objectives, and the distribution of σ_r is shown in (e).

based on local histogram, which can separate structures from textures effectively; (2) we improved the classical joint bilateral filter by setting the range kernel and guidance image adaptively, which enables the filter to smooth multi-scale textures; and (3) we propose a variant of anisotropic filter to remove the residual textures with strong gradient by changing the patch size and simply designing the diffusion coefficient.

The rest of this paper is organized as follows. In Section II, we review some other related works. In Section III, we briefly explain the technical background associated with our work. In Section IV, we introduce our proposed method; the scheme of structure measurement, the adaptive bilateral filtering, the anisotropic filter, and other details regarding our filtering technique are provided. Section V gives our experimental results and comparisons with the state-of-the-art and demonstrates the applicability of our method on many tasks, such as edge detection, detail enhancement, image segmentation, and stereo matching. Finally, a summary is provided in Section VI.

II. RELATED WORK

The existing texture filtering methods can be roughly divided into local texture filtering and global texture filtering.

Bilateral filtering is a classical local texture filtering. In order to achieve edge-preserving while denoising, it introduces a range kernel based on Gaussian filtering and takes spatial information and color similarity into account, but the method cannot filter out textures with strong gradient. Based on the framework of bilateral filtering, various efficient texture filtering methods are proposed. The method proposed by Zhang *et al.* [12] continuously updates the guidance information through iterative bilateral filtering to restore the main structure of the image. But with the increasing of iterations, the structure will appear fuzzy and color cast, and even passivation phenomenon. Ham *et al.* [13] proposed the idea of patch offset on the framework of bilateral filtering, and generated a smooth image based on patch offset as the guidance image, but it could not suppress textures with strong gradient. Karacan *et al.* [14] proposed a block similarity texture filtering method based on regional covariance. Gastal and Oliveira [15] proposed the idea of domain transformation to achieve a certain improvement in smoothing results. Li *et al.* [16] used the Gaussian pyramid mixing and smoothing to improve the effect of structure-preserving. Hua *et al.* [17] design a filtering framework for local diffusion in the gradient domain to preserving structures. Jeon *et al.* [8] proposed a scale-aware structure-preserving texture filtering (SaTF), which uses directional relative total variation (dRTV) to identify texture from structure and finds an optimal per-pixel smoothing scale for the guidance image. In order to better denoise the piece-wise smoothness image, [18] proposed an iterative range-domain weighted filter method combined with Gaussian filtering, which performs better performance on structural information preservation in image smoothing and can be applied into many applications.

The global-based filtering method is based on the idea of global optimization. It usually defines an objective function, including data items and smoothing items. The data item requires that the difference between the filtering result and the original image is as small as possible, and the smoothing item requires the texture region to be smoothed. The total variation method proposed by Rudin *et al.* [19] is a classical model of global filtering. It uses image gradient as constraints to smooth, but it only has a certain smoothing effect on textures and noises with smaller gradients. Farbman *et al.* [3] proposed a weighted least squares (WLS) method to process images with multi-scale textures, but this method cannot suppress the texture of strong gradients, and color rendering problems may occur. Zang *et al.* [20] proposed a directional adaptive image smoothing method based on anisotropic structure measurement. Ham *et al.* [13] designed an iterative method combining dynamic and static. Although good results can be achieved, the termination conditions are difficult to set. Xu *et al.* [21] proposed a method for L_0 gradient minimization, which obtains global optimization filtering results by controlling the number of non-zero gradients; Xu *et al.* [22] proposed the RTV method by improving the total variation model to further improve the filtering quality, but the parameters are difficult to set, and the texture with strong gradient cannot be filtered out. Inspired by the RTV method, [23] proposed a local activity-driven RTV method (LAD-RTV) for image smoothing and scale representation, which adopts the way of gradient product and the local activity measurement. Magnier et al. [24] proposed a smoothed rotation filter that can discriminate textons and combines anisotropic diffusion to obtain the texture filtering results of the preserved structure, but this method is not suitable for images with strong gradient textures. Based on iterative global optimization, [25] proposed a scale-aware filtering technique, which performs the scale-aware measure by finding the local extrema of the RTV within a neighboring window. The method has the advantages of avoiding halos and preventing from over-sharpening of edge.

III. TECHNICAL BACKGROUND

A. JOINT BILATERAL FILTER

Following the state-of-the-art methods for texture filtering, we adopt joint bilateral filter as our underlying framework to obtain the initial texture filtered result, written as:

$$J_p = \frac{1}{k} \sum_{q \in \Omega_p} g_{\sigma_s}(p, q)\, g_{\sigma_r}\left(G_p, G_q\right) I_q \tag{1}$$

where, k is the normalizing term, Ω_p is the spatial neighborhood with pixel p at its center, G is the guidance image, $g_{\sigma_s}(\cdot)$ and $g_{\sigma_r}(\cdot)$ give spatial and range weights with corresponding Gaussian kernel sizes σ_s and σ_r, respectively.

The guidance image G is generated by employing varying scale isotropic Gaussian smoothing on the input image I, written as:

$$G_p = \frac{1}{k} \sum_{q \in \Omega_p} \exp\left(-\frac{\|p-q\|^2}{2K_p^2}\right) \tag{2}$$

where $\|\cdot\|$ is the Euclidean distance. The output G_p at pixel p is a weighted average of the input image in the spatial neighborhood Ω_p.

B. ANISOTROPIC FILTER

Anisotropic filtering has been widely used in image processing since Perona and Malik [26] applied the anisotropic diffusion model to image filtering. The basic idea of anisotropic filtering is to smooth the interior region and preserve the edge region by estimating the gradient of the image. To implement it on a computer, this model can be discretized using the P-M equation [27] as follows:

$$\begin{cases} S_p^{t+1} = S_p^t + \dfrac{\lambda}{|N_p|} \sum_{q \subset N_p} g\left(\nabla S_{p,q}\right) \nabla S_{p,q} \\ S^0 = J \end{cases} \tag{3}$$

where, J is the initial input image, p is the pixel coordinates, S_p is the value at pixel p, $\nabla S_{p,q} = S_q^t - S_p^t$, t is the number of iterations, λ is the constant that determines the diffusion rate, N_p describes the neighborhood of pixel p, and $|N_p|$ is the number of pixels in the neighborhood of pixel p (except for the edge of the image, $|N_p|$ is usually set to 4 or 8).

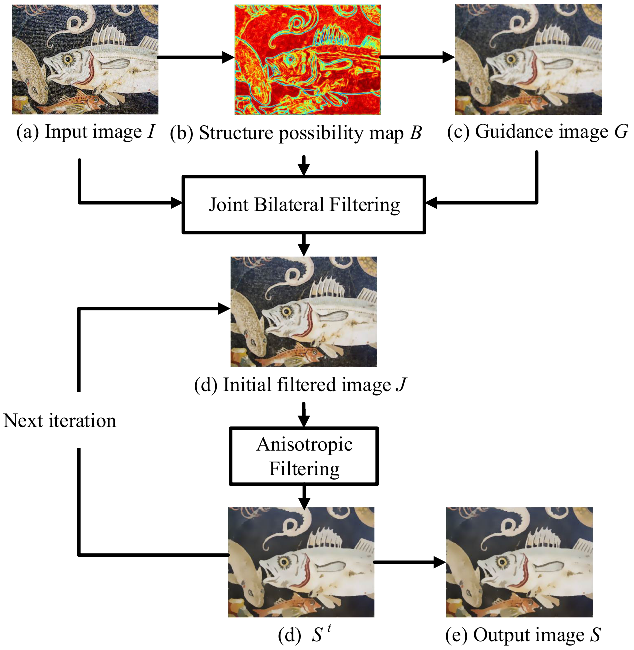


FIGURE 4. Overall process and intermediate images of our method.

The essence of the P-M model is to solve the nonlinear diffusion equation with the initial image as the initial value and adjust the diffusion intensity according to the gradient. $g(x)$ plays an essential role in controlling the filter and it is a bounded nonnegative decreasing function: if $x \to \infty$, then $g(x) \to 0$, this ensures that the diffusion stops at the edges of the image to maintain edge features; and if $x \to 0$, then $g(x) \to 1$, this is equivalent to a smoothing filter.

The ideal diffusion coefficient should make the equation produce forward diffusion in the flat region to facilitate the elimination of noises/textures, and produce backward diffusion in the edge region to facilitate edge sharpening. Although anisotropic diffusion can smooth the image while maintaining the edge characteristics, the nonlinear partial differential equation of the P-M equation, in addition to being difficult to solve, may not be able to obtain stable solutions in some cases.

IV. PROPOSED METHOD

In this section, the proposed method is introduced in detail. The method includes the following three parts: (1) structure measurement, (2) adaptive bilateral filtering, and (3) anisotropic filtering.

A. STRUCTURE MEASUREMENT

1) COLOR QUANTIZATION

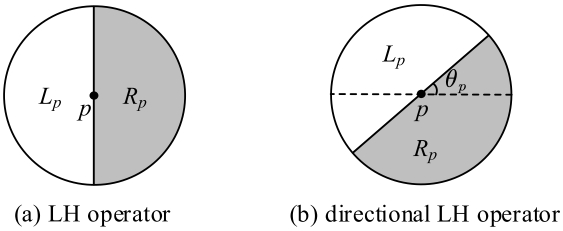
Color quantization is the basic process of the subsequent structure measurement. The process reduces the color complexity of the original image yet keeps a sufficient number of representative colors. It not only reduces the computational cost but also makes our proposed method more robust for noise. Color quantization algorithms mainly contain two parts, i.e., palette design and pixel mapping. In this paper, the mean-shift (MS) algorithm [28] is used in the palette design and pixel mapping. That is, the clustering centers are the palette, and the clustering process is the pixel mapping, which is performed to find the closest color from the palette to represent the pixel in the original image with minimum distortion.

2) LOCAL HISTOGRAM OPERATOR

In order to find the optimal filter scale for each pixel, it is necessary to distinguish structures from textures firstly. In this section, a local histogram operator is designed to perform the structure measurement. Suppose the input image is I, the color quantized image is $\bar{I}$ with N colors, the local histogram (LH) operator is defined as:

$$\mathscr{L}\left(h_{L_p},\ h_{R_p}\right) = \frac{\sum_{i=1}^{N} h_{L_p}(i)\cdot h_{R_p}(i)}{\sqrt{\sum_{i=1}^{N} h_{L_p}(i)\cdot h_{L_p}(i)}\cdot\sqrt{\sum_{i=1}^{N} h_{R_p}(i)\cdot h_{R_p}(i)}} \tag{4}$$

where, L_p and R_p are the left and right halves of the circular neighborhood centered on the pixel p of $\bar{I}$, which are shown in the white and gray sections in Fig. 5 (a); h_{L_p} and h_{R_p} denote the local histograms of the left and right halves of the circular neighborhood, respectively; $h_{L_p}(i)$ denotes the i-th element of h_{L_p}.



(a) LH operator (b) directional LH operator

FIGURE 5. Illustration of the LH operator.

The LH operator only along one direction is not enough to obtain all the structure information of images, which may be in any direction. Considering this problem, we define the directional local histogram operator as:

$$\mathscr{L}\left(h_{\varphi(L_p)},\ h_{\varphi(R_p)}\right) \tag{5}$$

where, $\varphi(\cdot)$ is the angle rotation operator, $\varphi(L_p)$ and $\varphi(R_p)$ represent semi-circular windows on both sides after a certain rotation angle, which is shown in Fig. 5 (b).

For each pixel, we find θ_p, which we call *structure direction*, in the direction the LH operator has the smallest value, which we define as *structure possibility* and is obtained as:

$$B_p = \underset{\theta_p}{\operatorname{argmin}}\left(\mathscr{L}\left(h_{\varphi(L_p)},\ h_{\varphi(R_p)}\right)\right) \tag{6}$$

for $0 \leq \theta_p < 2\pi$. In our implementation, we compared 8 different directions to efficiently find structure direction θ_p.

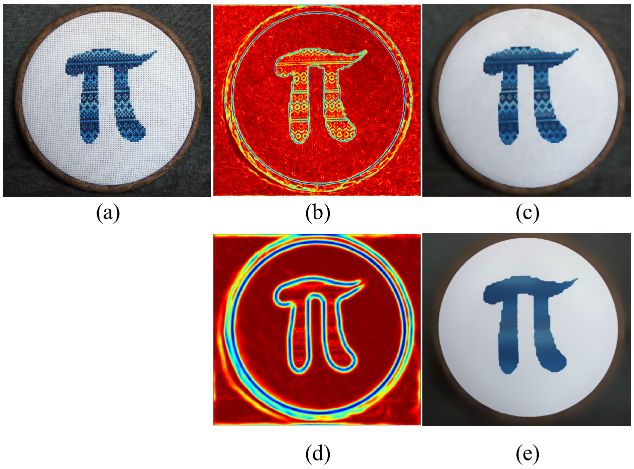


FIGURE 6. Results of LH operator with varying radius r. (a) Input image; (b) structure impossibility map and (c) filtered image using LH operator when $r = 1$; (d) structure impossibility map and (e) filtered image using LH operator when $r = 2$.

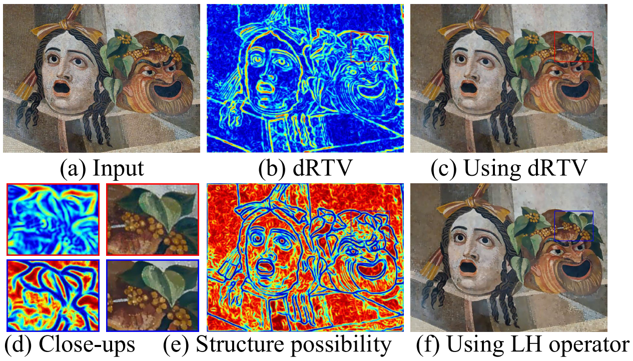

(a) Input (b) dRTV (c) Using dRTV
(d) Close-ups (e) Structure possibility (f) Using LH operator

FIGURE 7. Comparison of SaTF and ours. Red color indicates large value while blue color indicates small value. For dRTV, the pixels with higher values are likely to be in texture regions, while the pixels with lower values are on structure edges. For the LH operator, the situation is reversed.

The LH operator can quantify the difference in color distribution between the left and right half neighboring windows centered on the pixel of interest. Intuitively, if the pixel is around texture edges, the color distributions of the two half neighborhood windows are different and the structure possibility is high. Otherwise, the color distributions are similar and the structure possibility is low. The LH operator has only one parameter r, which controls the maximum scale of the texture to be removed. As shown in Fig. 6, when r is larger, more small details can be smoothed out. Therefore, our LH operator can be used to obtain multi-scale image structures. Fig. 7 (b, e) compare the structure measurement results generated by the method of SaTF and ours, which shows that our method delivers better texture-structure separation.

B. ADAPTIVE BILATERAL FILTERING

In section III, we have briefly reviewed the joint bilateral filtering. In this section, we propose an adaptive bilateral

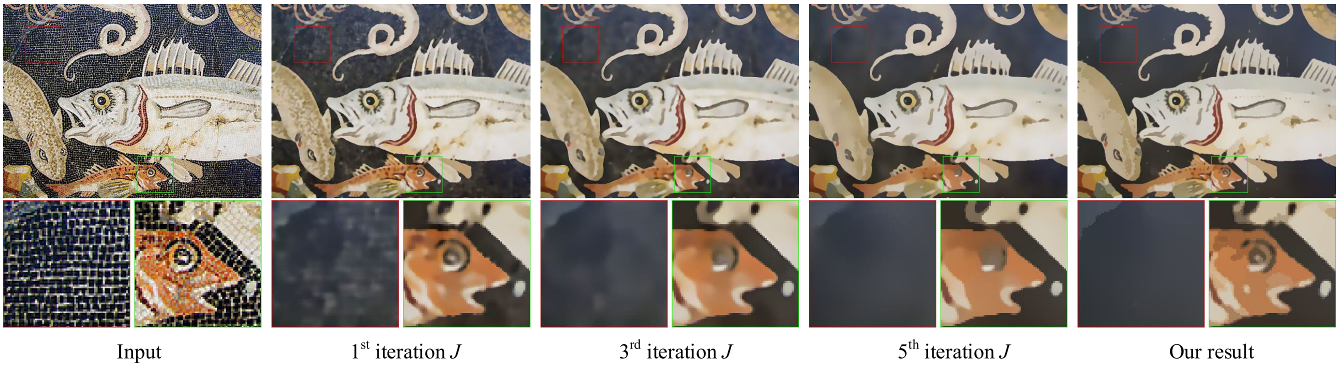


FIGURE 8. Results of the joint bilateral filtering with various iterations and our method. Through multiple iterations, the joint bilateral filtering can smooth the textures with strong gradient, while causing small structures blurring simultaneously.

filtering by setting the parameters of joint bilateral filtering adaptively. It can be seen from equation (1), the effect of joint bilateral filtering is decided by the three parameters, namely, the spatial/range kernel sizes σ_s and σ_r, and the guidance image G. The spatial neighborhood size σ_s determines the smoothness of the filter output and is set to $\sigma_s = r$, where r is the radius of the LH operator. As for σ_r, unlike the classical joint bilateral filtering, we propose the adaptive bilateral filter by adjusting σ_r per-pixel, that is, $\sigma_r(p) = 2B_p$, where B_p is the structure possibility and comes from the LH operator. As shown in Fig. 7, B_p is high for the pixel within flat or texture regions, and low for the pixel around texture edges. Therefore, $\sigma_r(p)$ is high whenever B_p is high and vice versa. This ensures that aggressive smoothing is performed in texture regions, but without excessively blurring the texture edges.

The effect of the joint bilateral filtering also heavily depends on the guidance image G. We use equation (2) to generate the guidance image. An ideal guidance image must satisfy the following two conditions, namely suppressing textures as much as possible and guarantying structures not to be blurred. Therefore, the filtering weight assigned to the pixels within flat or texture regions should be large, and small to the pixels around texture edges. In this regard, we use the structure possibility B to design the kernel scale K of the guidance image G, which is defined as:

$$K_p = 0.05B_p + 1 \tag{7}$$

Fig. 7 (c, f) compares the guidance images generated by the method of SaTF and ours, which shows that ours performs better in texture suppression and structure preservation.

C. ANISOTROPIC FILTERING

As shown in Fig. 8, most textures can be removed through once bilateral filtering, while there are still some residues for textures with strong gradient. The existing methods usually adopts multiple iterations to improve the filtering results. However, as shown in Fig. 8, this inevitably leads to the small structures blurring due to the isotropic nature of the bilateral filter. In order to solve this problem, we propose a novel anisotropic filter based on the basic mechanism of anisotropic filter. In the follows, the proposed filter is introduced in details.

The RGB color space is usually used to perform the physical encoding of color images but it does not fit with human vision; therefore, the LAB color space is used to represent the image in our anisotropic filter, where the structural component of pixel p in the given image J is computed by:

$$\begin{cases} S_p^{t+1} = \dfrac{\sum_{q\in N_p} f_q^t S_q^t}{\sum_{q\in N_p} f_q^t} \\ S^0 = J \end{cases} \tag{8}$$

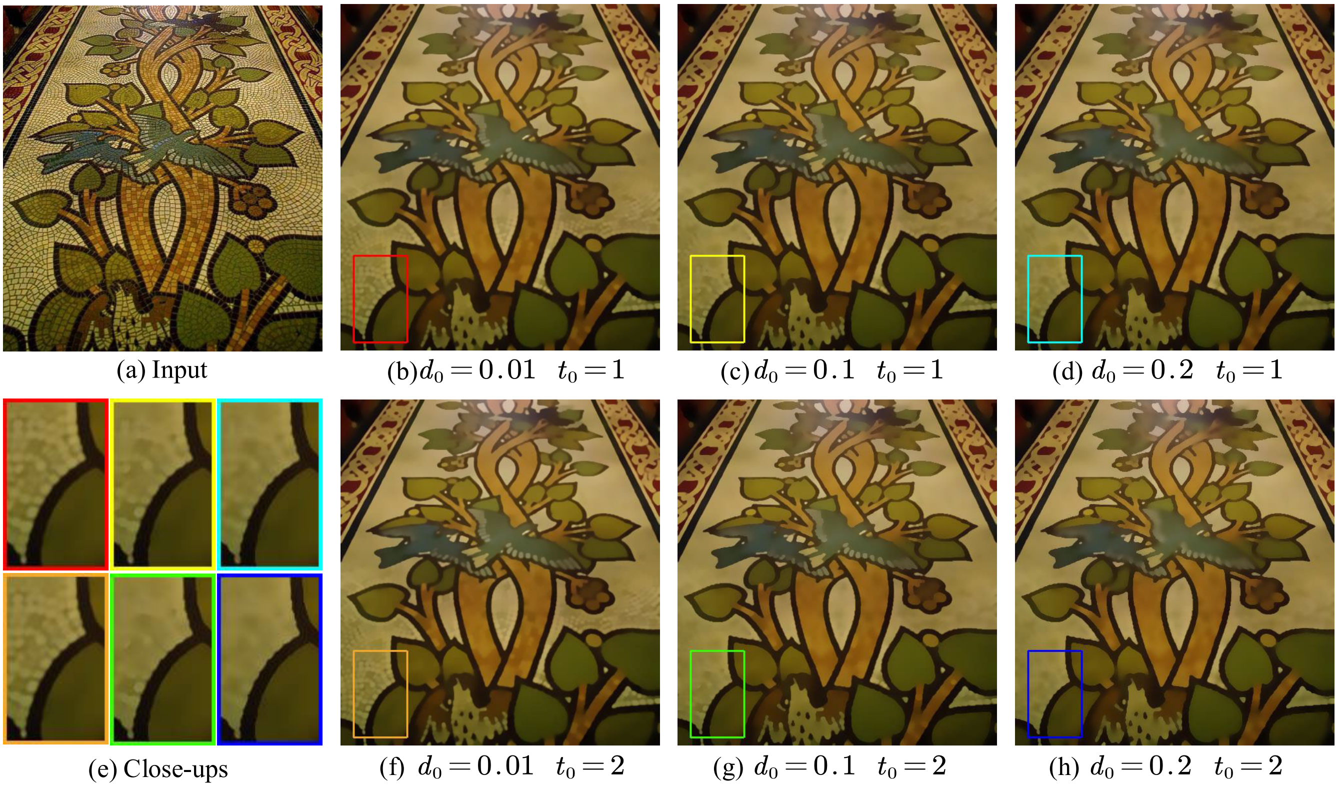
where, the range coordinates $S_p = \left(L_p, A_p, B_p\right)$, S^t denotes the filtering results with t iterations, N_p denotes a squared neighborhood centered at p and of patch size $w \times w$ pixels, and f_q^t denotes diffusion coefficient function, which is defined as:

$$f_q^t = \begin{cases} 1, & \|S_q^t - S_p^t\| < d_0 \\ 0, & \|S_q^t - S_p^t\| \geq d_0 \end{cases} \tag{9}$$

where, $\|\cdot\|$ is the Euclidean distance in the LAB color space, and d_0 denotes the edge threshold and is used to determine whether pixel p is in the edge or texture/flat regions.

It can be seen from equation (8) and equation (9): when $\|S_q^t - S_p^t\| \geq d_0$, pixel p is in the edge region, a small diffusion coefficient is assigned to the filter to achieve edge-preserving; when $\|S_q^t - S_p^t\| < d_0$, pixel p is in the texture/flat region, a large diffusion coefficient is assigned to the filter to achieve textures/noise removal. This mechanism can overcome the defect of Gaussian filter. In addition, equation (8) can also be written as:

$$\begin{aligned} S_p^{t+1} &= \frac{\sum_{q\in N_p} f_q^t S_q^t}{\sum_{q\in N_p} f_q^t} \\ &= S_p^t + \frac{\sum_{q\in N_p, q\neq p} f_q^t \left(S_q^t - S_p^t\right)}{1 + \sum_{q\in N_p, q\neq p} f_q^t} \\ &= S_p^t + \frac{\sum_{q\in N_p, q\neq p} f_q^t \nabla S_{p,q}}{\sum_{q\in N_p} f_q^t} \end{aligned} \tag{10}$$


(a) Input

(b) $d_0 = 0.01 \quad t_0 = 1$

(c) $d_0 = 0.1 \quad t_0 = 1$

(d) $d_0 = 0.2 \quad t_0 = 1$

(e) Close-ups

(f) $d_0 = 0.01 \quad t_0 = 2$

(g) $d_0 = 0.1 \quad t_0 = 2$

(h) $d_0 = 0.2 \quad t_0 = 2$

FIGURE 9. Results of anisotropic filtering with varying parameters: (from left to right) bigger scale textures are removed with the increasing of d_0; (from top to bottom) smoother results are obtained with the increasing of t_0.

By comparing equation (3) and equation (10), it can be seen that our proposed filter is a variant of anisotropic filter and the differences are two-fold: one is the patch size and the other is the diffusion coefficient. In the variant, the patch size is flexible, and is no longer limited to 4 neighborhoods or 8 neighborhoods. In addition, the simplified design of the diffusion coefficient makes the variant easier to be solved.

The anisotropic filtering comes with its original parameters; patch size w, edge threshold d_0 and iterations t_0. Patch size w determines the size of the region to perform filtering. Larger patch size can help to remove larger textures, but more small detailed structures may be damaged and jaggy artifact will appear around the edges. Thus, patch size w is fixed to 3, which is enough to yield good results. The anisotropic filtering can work with d_0 and t that control the degree of smoothness; larger value of d_0 can help to remove textures with strong gradient and the increasing t_0 yields more smoothing effects. We show the results with various values of d_0 and t in Fig. 9. In this paper, $d_0 \in \{0.6,\ 0.8,\ 1.0\}$ and $t_0 \in \{1,\ 2,\ 3\}$ are used for most examples. Table 1 summarizes the entire process of our method, and its intermediate images are shown in Fig. 4.

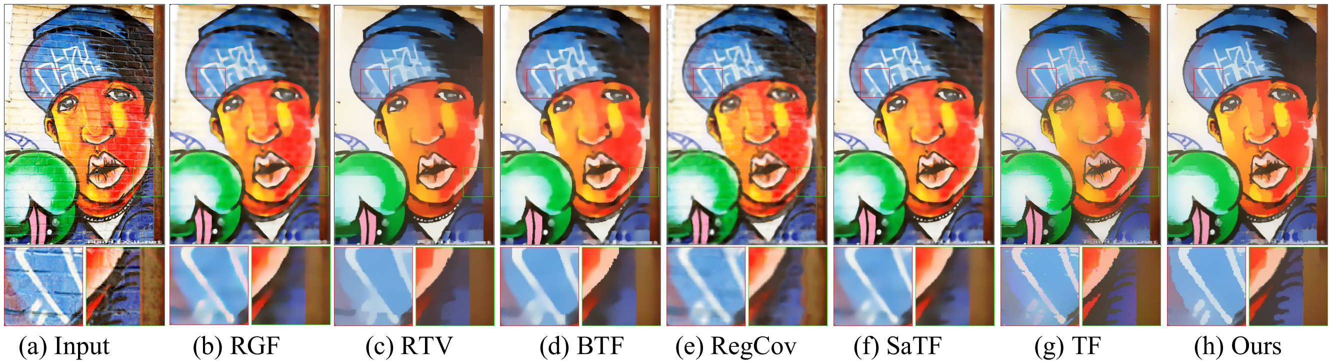
TABLE 1. Structure-aware texture filtering.

```
Input: image $I$
Output: filtering result $S$
$\overline{I} \leftarrow$ Color quantization of $I$
$B \leftarrow$ Minimum value of $B$ using $\overline{I}$          △ Eq. (6)
$K \leftarrow$ Kernel scales using $B$                              △ Eq. (7)
$G \leftarrow$ Guidance image from $\overline{I}$ using $K$         △ Eq. (2)
$J \leftarrow$ Joint bilateral filtering of $I$ using $G$           △ Eq. (1)
$S^0 \leftarrow J$
for $t = 1 \cdots t_0$ do
    $S^t \leftarrow$ Anisotropic filtering using $S^{t-1}$          △ Eq. (8)
end for
$S \leftarrow S^{t_0}$
```

V. EXPERIMENTS

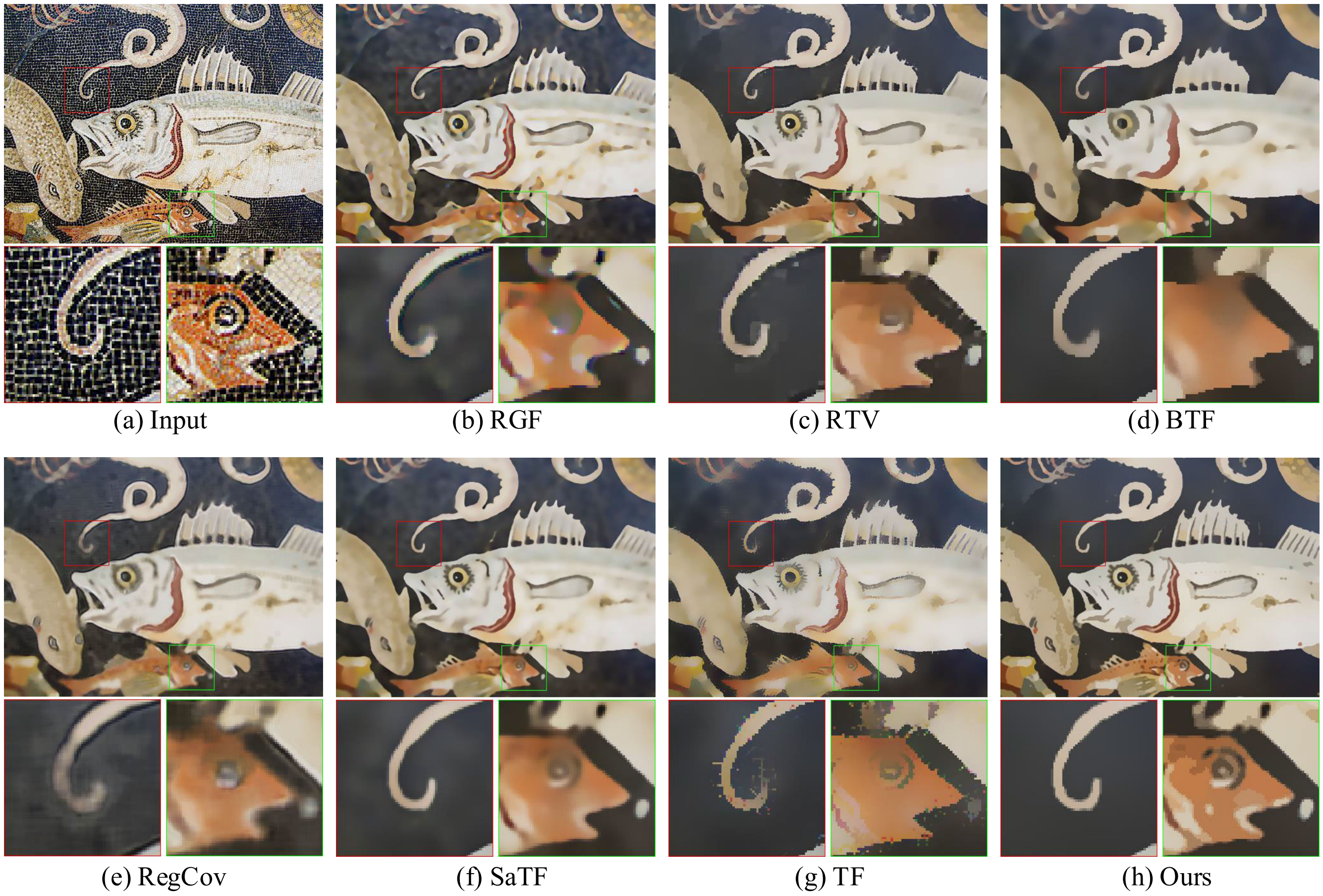
A. COMPARISON WITH THE STATE-OF-THE-ART

In this section, we compare our method with the state-of-the-art texture removal methods, including rolling guidance filter (RGF) [12], RTV texture smoothing [22], bilateral texture Filtering (BTF) [7], region covariance filter (RegCov) [14], scale-aware texture filtering (SaTF) [8], and tree filtering (TF) [29]. They are representative as both local- and global- schemes are included and the effect of smoothing, sharpening, and texture removal can be produced. For fair comparison, we used the implementations provided online by the authors and carefully tuned the parameters manually to generate results from these techniques. Since there is no reasonable objective evaluation index in the field of texture filtering, the subjective evaluation is used to compare the effects of different texture filtering methods in this paper.



(a) Input (b) RGF (c) RTV (d) BTF (e) RegCov (f) SaTF (g) TF (h) Ours

FIGURE 10. **Filtered results for comparison on graffiti on brick wall. (a) Input image, (b) RGF [$\sigma_s = 4$, $\sigma_r = 0.1$, $n_{iter} = 5$], (c) RTV [$\lambda = 0.02$, $\sigma = 3$, $\varepsilon_s = 0.02$, $n_{iter} = 4$], (d) BTF [$k = 3$, $n_{iter} = 4$], (e) RegCov [$k = 13$, $\sigma = 0.4$, *Model* 2], (f) SaTF [$\sigma = 3$, $\sigma_r = 0.1$, $n_{iter} = 3$], (g) TF [$\sigma = 0.025$, $\sigma_s = 6$], and (h) our method.**



(a) Input (b) RGF (c) RTV (d) BTF

(e) RegCov (f) SaTF (g) TF (h) Ours

FIGURE 11. **Filtered results for comparison on graffiti on brick wall. (a) Input image, (b) RGF [$\sigma_s = 4$, $\sigma_r = 0.05$, $n_{iter} = 5$], (c) RTV [$\lambda = 0.015$, $\sigma = 4$, $\varepsilon_s = 0.02$, $n_{iter} = 4$], (d) BTF [$k = 2$, $n_{iter} = 5$], (e) RegCov [$k = 19$, $\sigma = 0.2$, *Model* 1], (f) SaTF [$\sigma = 3$, $\sigma_r = 0.1$, $n_{iter} = 3$], (g) TF [$\sigma = 0.02$, $\sigma_s = 3$], and (h) our method.**

Fig. 10 shows the filtered results on a graffiti image, where the graffiti is the structures and the bricks are the background textures. It can be seen that all the texture filtering methods can extract the prominent structure of the graffiti rightly. However, the methods of RGF, RegCov and SaTF cannot completely smooth the brick textures as shown in Fig. 10 (b, e, f). The methods of RegCov and TF may over-blur structural detailed edges such as eyes shown in Fig. 10 (e, g). The methods of RTV and BTF can efficiently smooth the brick textures while maintain the main structures of the graffiti. However, they also undergo unsatisfactory edge-preserving performance, causing the slender structures

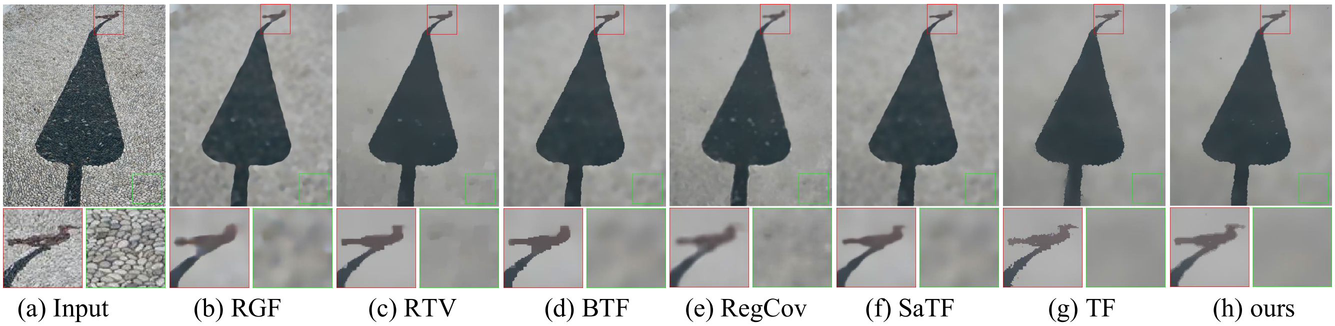

(a) Input (b) RGF (c) RTV (d) BTF (e) RegCov (f) SaTF (g) TF (h) ours

FIGURE 12. Filtered results for comparison on graffiti on brick wall. (a) Input image, (b) RGF [$\sigma_s = 5$, $\sigma_r = 0.05$, $n_{iter} = 4$], (c) RTV [$\lambda = 0.015$, $\sigma = 4$, $\varepsilon_s = 0.02$, $n_{iter} = 4$], (d) BTF [$k = 3$, $n_{iter} = 5$], (e) RegCov [$k = 19$, $\sigma = 0.2$, *Model* 1], (f) SaTF [$\sigma = 4$, $\sigma_r = 0.1$, $n_{iter} = 3$], (g) TF [$\sigma = 0.08$, $\sigma_s = 3$], and (h) our method.

to be blurred as shown in the enlarged boxes in Fig. 10 (c, d). In contrast, as shown in Fig. 10 (h), our method can properly remove the brick textures while preserve the graffiti without damaging it.

To further demonstrate the effectivity of our method in preserving small details and removing textures, Fig. 11 shows the texture removal results of an ancient mosaic image from Pompeii, which contains textures with strong gradient and complex structures with small detailed structures, such as eyes and eyebrows of fishes. Although all the methods can extract the prominent image structure, there are some differences. As shown in Fig. 11 (b, e, f), the methods of RGF, RegCov, and SaTF cannot effectively filter out high-contrast mosaic textures in the background. The methods of RGF, BTF, RegCov, and TF have problem of preserving fine and complex structures and may over-blur structural detailed edges to a certain extent, such as the enlarged boxes in Fig. 11 (b, d, e, g). In addition, in Fig. 11 (b, e), the methods of RGF and RegCov may cause the ringing phenomena. In Fig. 11 (c, d), jaggy artifact appears around the edges, since they are over-sharpened by the method of RTV. Fig. 11 (d) displays the eye of the small fish disappeared by the method of BTF, because it is smaller than the patch size. The pixel-based TF method may misidentify some pixels as being in regions that do not belong to them, thus producing discrete pixels around the edge, as shown in the enlarged boxes in Fig. 11 (g). Compared to the state-of-the-art texture filtering methods, ours generally performs better in terms of preserving small detailed structures while removing textures with strong gradient.

Fig. 12 (a) shows an image whose cobbled ground are textures and the scales of the textures are different because of the perspective. Since the existing texture filter methods usually adopt fixed-scale kernels for structure-texture separation, it is a challenge for them to preserve salient but small-scale structures when remove large-scale textures shown in the enlarged boxes in Fig. 12 (b, c, d, e, f, g). The performances on suppressing the large-scale textures are unsatisfactory and the bird-like structure is destroyed to a varying extent. In contrast, as shown in Fig. 12 (d), since our method automatically identifies the smoothing scale per pixel, small but salient structures (bird-like structures) are well preserved even when large textures are smoothed out. This experiment demonstrates that our method outperforms other methods in filtering out multi-scale textures.

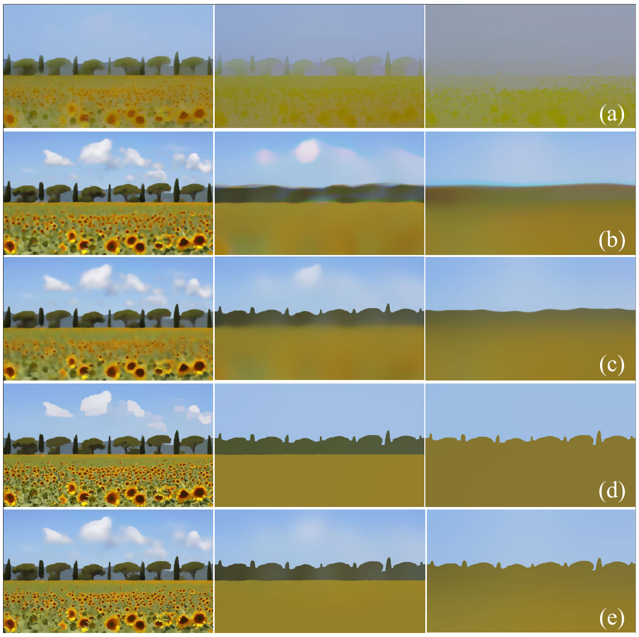


FIGURE 13. Examples of the scale-space representation obtained by. (a) WLS[(from left to right) $\lambda = \alpha = 2, 4, 8$], (b) RGF [(from left to right) $\sigma_s = 4, 30, 70$, $\sigma_r = 0.05$, $n_{iter} = 4$], (c)SaTF [(from left to right) $\sigma = 3, 20, 40$, $\sigma_r = 0.1$, $n_{iter} = 3$], (d) LAD-RTV [(from left to right) $\lambda = 0.003, 0.08, 0.3$, $n_{iter} = 4$], (e) ours [(from left to right) $r = \sigma_s = 4, 10, 30$].

In order to further validate the efficiency of our method in scale-space filtering, Fig. 13 shows examples of the scale-space representation obtained by WLS [3], RGF, SaTF, LAD-RTV [23], and our method. As shown in Fig. 13 (a), the method of WLS has a certain ability to remove details, but it cannot preserve structures at coarse scales. The method of RGF does not consider the structural information of the input image, and uses the isotropic Gaussian kernel to control the filtering scale, which results in poor performance in preserving structures and boundary localization at coarse

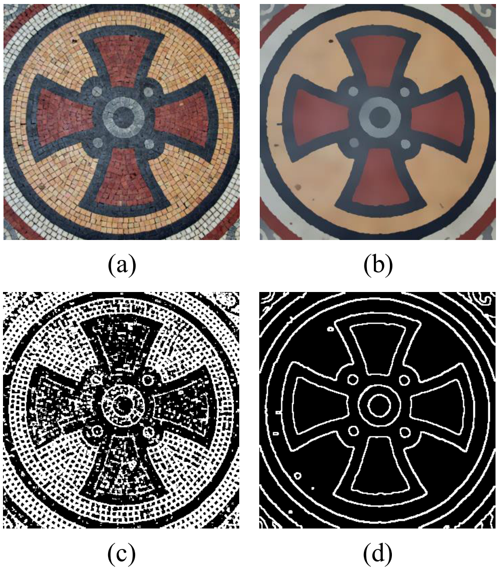


FIGURE 14. Detail enhancement. (a) Input image. (b) Our filtered image. (c) Edge detection of the input image. (d) Edge detection of the filtered image.

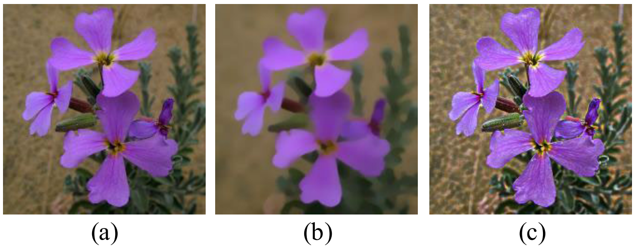


FIGURE 15. Detail enhancement. (a) Input image. (b) Our filtered image. (c) The enhancement result.

scales (Fig. 13 (b)). Although the method of SaTF still uses the isotropic Gaussian kernel to control the filtering scale, the scheme that using the structural information to obtain the guidance image enables it to coarsely represent an image in different scales and preserve the structures, but poor boundary localization is still a problem for the method (Fig. 13 (c)). As shown in Fig. 13 (d, e), the method of LAD-RTV and ours performs well in in scale-space filtering, because object boundaries are sharp and coincide well with the meaningful boundaries at each scale.

B. APPLICATIONS

Texture filtering has been widely used in different fields. In the following, we briefly show several applications, which will highlight the effectiveness of our proposed filter. Because of the structure measurement property, the most intuitive application of our method would be edge detection.

Edge detection is an application often used in image processing. However, in many cases it is difficult to detect pure structural edges due to the interference of textures in the image.As shown in Fig. 14, our filter is first used for texture/noise removal, which makes the result of edge detection very clear and reasonable.

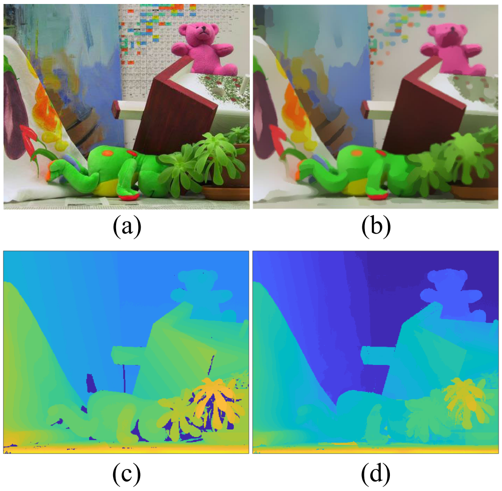


FIGURE 16. Stereo matching. (a) Input left images. (b) Our filtered image. (c) Ground truth. (d) Our result.

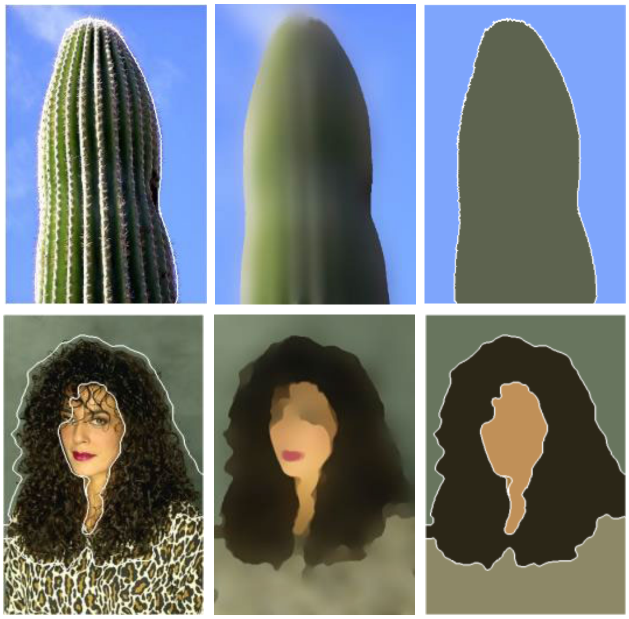


FIGURE 17. Texture image segmentation. (left) Input images. (middle) filtered results. (right) Segmentation results.

Fig. 15 demonstrates the application of our method in detail enhancement, which could generate much clearer details of images and obtain distinct visual effects. The texture information of the image is obtained by subtracting the filtered image from the original image. And then through superimposing the texture information on the original image, we can obtain the result of enhanced detail.

Fig. 16 shows an application of stereo matching. Cost aggregation is an important step in stereo matching. In this application, we use the image filtered by our method to guide cost aggregation, which can ensure better matching results. In the obtained disparity map, the surface of the same object is smoother and the edges between different objects are clear.

Fig. 17 shows some texture image segmentation results based on our proposed filter. The texture interference is removed by our method, which makes the main structure of the image more obvious. The super-pixel segmentation is performed on the filtered image first, and then the region fusion is performed to obtain the final segmentation result.

VI. CONCLUSION

In order to smooth the multi-scale textures with strong gradient while maintaining the weak structures, we propose a structure-aware bilateral filter, which uses an adaptive range kernel. Since texture filtering depends on high-quality texture measurement, the texture edge operator based on the local histogram is designed to separate structures from textures. In addition, to avoid edge blurring because of multiple iterations of bilateral filtering, a novel anisotropic filter is designed to remove residual textures after once bilateral filtering. A lot of experimental results demonstrate the superiority of our method in maintaining the weak structure while removing the texture with strong gradient and varying scales.

REFERENCES

[1] F. Kou, Z. Wei, W. Chen, X. Wu, C. Wen, and Z. Li, "Intelligent detail enhancement for exposure fusion," *IEEE Trans. Multimedia*, vol. 20, no. 2, pp. 484–495, Feb. 2018.

[2] O. Sener, K. Ugur, and A. A. Alatan, "Efficient MRF energy propagation for video segmentation via bilateral filters," *IEEE Trans. Multimedia*, vol. 16, no. 5, pp. 1292–1302, Aug. 2014.

[3] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–10, Aug. 2008.

[4] L. Zhao, H. Bai, J. Liang, B. Zeng, A. Wang, and Y. Zhao, "Simultaneous color-depth super-resolution with conditional generative adversarial networks," *Pattern Recognit.*, vol. 88, pp. 356–369, Apr. 2019.

[5] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 257–266, Jul. 2002.

[6] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. ICCV*, 1998, vol. 98, no. 1, pp. 839–846.

[7] H. Cho, H. Lee, H. Kang, and S. Lee, "Bilateral texture filtering," *ACM Trans. Graph.*, vol. 33, no. 4, pp. 1–8, Jul. 2014.

[8] J. Jeon, H. Lee, H. Kang, and S. Lee, "Scale-aware structure-preserving texture filtering," *Comput. Graph. Forum*, vol. 35, no. 7, pp. 77–86, Oct. 2016.

[9] R. G. Gavaskar and K. N. Chaudhury, "Fast adaptive bilateral filtering," *IEEE Trans. Image Process.*, vol. 28, no. 2, pp. 779–790, Feb. 2019.

[10] M. Kass and J. Solomon, "Smoothed local histogram filters," *ACM Trans. Graph.*, vol. 29, no. 4, p. 1, Jul. 2010.

[11] Q. Zhang, L. Xu, and J. Jia, "100+ times faster weighted median filter (WMF)," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2830–2837.

[12] Q. Zhang, X. Shen, L. Xu, and J. Jia, "Rolling guidance filter," in *Proc. Eur. Conf. Comput. Vis.* Springer, 2014, pp. 815–830.

[13] B. Ham, M. Cho, and J. Ponce, "Robust image filtering using joint static and dynamic guidance," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 4823–4831.

[14] L. Karacan, E. Erdem, and A. Erdem, "Structure-preserving image smoothing via region covariances," *ACM Trans. Graph.*, vol. 32, no. 6, pp. 1–11, Nov. 2013.

[15] E. S. L. Gastal and M. M. Oliveira, "Domain transform for edge-aware image and video processing," *ACM Trans. Graph.*, vol. 30, no. 4, p. 69, Jul. 2011.

[16] X.-Y. Li, Y. Gu, S.-M. Hu, and R. R. Martin, "Mixed-domain edge-aware image manipulation," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1915–1925, May 2013.

[17] M. Hua, X. Bie, M. Zhang, and W. Wang, "Edge-aware gradient domain optimization framework for image filtering by local propagation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2838–2845.

[18] L. Zhao, H. Bai, A. Wang, and Y. Zhao, "Iterative range-domain weighted filter for structural preserving image smoothing and de-noising," *Multimedia Tools Appl.*, vol. 78, no. 1, pp. 47–74, 2019.

[19] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.

[20] Y. Zang, H. Huang, and L. Zhang, "Guided adaptive image smoothing via directional anisotropic structure measurement," *IEEE Trans. Vis. Comput. Graphics*, vol. 21, no. 9, pp. 1015–1027, Sep. 2015.

[21] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via L_0 gradient minimization," *ACM Trans. Graph.*, vol. 30, no. 6, pp. 174:1–174:12, Dec. 2011.

[22] L. Xu, Q. Yan, Y. Xia, and J. Jia, "Structure extraction from texture via relative total variation," *ACM Trans. Graph.*, vol. 31, no. 6, p. 139, Nov. 2012.

[23] L. Zhao, H. Bai, J. Liang, A. Wang, B. Zeng, and Y. Zhao, "Local activity-driven structural-preserving filtering for noise removal and image smoothing," *Signal Process.*, vol. 157, pp. 62–72, Apr. 2019.

[24] B. Magnier, P. Montesinos, and D. Diep, "Texture removal preserving edges by diffusion," in *Proc. Scandin. Conf. Image Anal.* Springer, 2015, pp. 3–15.

[25] Z. Zhou, B. Wang, and J. Ma, "Scale-aware edge-preserving image filtering via iterative global optimization," *IEEE Trans. Multimedia*, vol. 20, no. 6, pp. 1392–1405, Jun. 2018.

[26] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.

[27] Y. Yu and S. T. Acton, "Speckle reducing anisotropic diffusion," *IEEE Trans. Image Process.*, vol. 11, no. 11, pp. 1260–1270, Nov. 2002.

[28] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.

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

YANG LIU received the B.S. degree (Hons.) from the China University of Geosciences, in 2013. He is currently pursuing the Ph.D. degree with Jilin University.

His research interests include stereo vision, 3D reconstruction, and image segmentation.

GUANGDA LIU received the Ph.D. degree in optical engineering from the Changchun Institute of Optics, Fine Mechanics, and Physics, Chinese Academy of Sciences, Changchun, China, in 2000.

He joined Jilin University, China, where he is currently a Professor carrying out research and working on biomedical signal processing and medical instrument development. His current research interests include human blood flow parameters detection and hepatic functional reserves assessment.

HONGLIANG LIU received the B.S. degree from the Changchun University of Science and Technology, in 2016. He is currently pursuing the master's degree with Jilin University. His research interests include image processing and point clouds matching.

CHANGYING LIU received the Ph.D. degree from the Harbin Institute of Technology, Harbin, China, in 2006.

He joined Jilin University, China, where he is currently a Professor carrying out research and working on visual measurement of railway vehicles. His current research interests include large scale 3D precise visual measurement and 3D Reconstruction.