



## A semantic edge-aware parameter efficient image filtering technique

Anonymous Author(s)

---

### ARTICLE INFO

#### *Article history:*

Received August 24, 2024

---

Jensen Shannon divergence, semantic-aware filtering, skewness, structures, textures.

---

### ABSTRACT

The success of a structure preserving filtering technique has relied on its capability to recognize structures and textures present in the input image. In this paper a novel structure preserving filtering technique is presented that first, generates an edge-map of the input image by exploiting semantic information. Then, an edge-aware adaptive recursive median filter is utilized to produce the filter image. The technique provides satisfactory results for a wide variety of images with minimal fine-tuning of its parameters. Moreover, along with the various computer graphics applications the proposed technique also shows its robustness to incorporate spatial information for spectral-spatial classification of hyperspectral images. A MATLAB implementation of the proposed technique is available at - <https://github.com/K-Pradhan/A-semantic-edge-aware-parameter-efficient-image-filtering-technique>

© 2024 Elsevier B.V. All rights reserved.

---

### 1. Introduction

The key feature of structure-preserving filtering techniques is to smooth out textures regardless of their spectral variations while maintaining the structures of the significant objects. The filtered images generated by such techniques are used for image editing [1], tone mapping [2], detail enhancement [3], image up-sampling [4], mesh denoising [5], artistic rendering [6], and image abstraction [7] which play important role in numerous applications of computational photography and image processing.

In an image, textures are some repetitive, erratic visual patterns. Significant objects or structures are made up of a group of similar textures separated from one another by structural edges. Two fundamental approaches for accomplishing structure preserving texture filtering are: i) finding textural patterns [8, 9, 10], and ii) detecting structural edges [11, 12, 13, 14, 15, 16, 17]. To differentiate textural patterns, the former approach extracts textural features like total variation (TV), relative total variation (RTV), region covariance, etc. All such features are extracted by using a fixed-size window. As a result, for irregular and scale-varying textural patterns these features are often failed to differentiate textures and structures. To mitigate this

problem, the later approach detects structural edges for filtering. Despite the benefits of such edge-aware filtering methods, the efficiency of these methods is dependent on their ability to correctly identify the structural edges of the input image. Since in many images, structural and textural edges confuse each other by providing similar tonal and sharpness information. It is very difficult to separate them by considering only their gradient information. To this end, the recent studies focus on the inclusion of semantic information, which is a challenging task and regarded as an open research topic [13, 14, 18].

In this paper, we propose a novel method that exploits semantic information of the image for discriminating structural and textural edges. Our technique generates an edge-map of the input image by exploiting semantic information in two phases. In the first phase two novel features, i) the semantic gradient: by analyzing the local distribution of the input image, and ii) the semantic skewness: by analyzing the local distribution of morphological gradient image, are proposed to generate the semantic gradient image (SGI). In the second phase, with the help of the generated SGI, the size of the window associated with each pixel of the input image is automatically defined to determine whether it is a structural edge or non-edge pixel by exploiting

Jensen Shannon (JS) divergence. Thus, the proposed technique generates an edge-map of the input image by considering semantic information. Once the semantic edge-map is generated, an edge-aware adaptive recursive median filter is utilized to produce the filter image. Note that the proposed technique requires multiple windows and the size of most of these windows are defined automatically or kept fixed irrespective of the considered input image. Moreover, it provides satisfactory results for a variety of input images with minimal fine-tuning of its parameters, which is a useful attribute of the proposed technique in comparison to the current state-of-the-art techniques. Along with computer graphics applications such as image denoising, detail enhancement, and tone mapping, the proposed technique also shows its robustness in incorporating spatial information for image classification.

The rest part of the paper is structured as follows. The background literature and the evaluation of the filtering approaches are discussed in Section 2. The details of the proposed filtering technique are presented in Section 3. Experimental results and analysis are given in Section 4. Section 5 presents different applications of the proposed technique. Finally, Section 6 concludes this work.

## 2. Background and literature review

For edge-preserving image smoothing, anisotropic diffusion [19, 20] and bilateral filtering [21, 22] are the two most fundamental and widely used nonlinear methods. These methods have gained popularity for various applications due to their simplicity, effectiveness, and versatility. Later more sophisticated edge-preserving filtering techniques have emerged, such as weighted least squares (WLS) [50, 24], edge-avoiding wavelets [3], local histogram filtering [25], local Laplacian filtering [26], domain transform [27], and L0 gradient minimization [28, 29]. Furthermore, different forms of guided filtering [51, 52] and recursive filtering [32] have also been proposed. Although these techniques were not specifically developed for handling texture, they have shown moderate success in achieving the goal of structure-preserving filtering, which involves smoothing texture (fine-scale details) without adversely affecting image structures. In [33], the oscillations between local extrema of the input image are considered to distinguish textures from structural edges. For handling textures several patch-based filtering techniques are proposed in the literature. In [8, 9], textural features such as spatial variations, frequency, and symmetry are used for smoothing regular or nearly regular textures. Total variation (TV) regularization is one such feature that has proven its effectiveness for filtering arbitrary regular textures [34]. The original TV regularization has been further modified in [35, 12, 36]. Relative total variation (RTV), a spatially-varying TV measure has been proposed in [35]. For texture smoothing, patch-based filtering techniques either used local directional gradient patterns described by TV, RTV *etc*, or region covariance descriptors. As a result, they are able to characterize the textural features of the pixels more accurately than the pixel-based techniques. Despite their robustness for texture characterization, patch-based techniques have certain limitations in dealing with

irregular textural patterns of varied scales. Moreover, the overlapping patches near the edges provided similar statistics which leads to the potential over-blurring of the structural edges.

Recent research focuses on the development of edge-aware texture filtering techniques [9, 13, 14, 18, 16, 17, 15, 37, 38, 24, 39, 40]. To incorporate structural edge features these techniques either leverage optimization [9, 18, 37], supervised learning [38, 24, 39], unsupervised learning [40], or semantic edge-aware structure texture decomposition [13, 14, 18, 16, 17, 15]. For instance, in [9] modified RTV and edge-aware window shifting are applied to enhance the identification of textural patterns. In [13, 14] edge detection techniques are employed to develop learning-based edge-aware filtering. In [37, 39] artificial neural networks (ANNs) are trained for edge-aware smoothing. A convolutional neural network (CNN) based framework with deep variational prior is utilized in [38]. An edge-aware filtering method that combines bilateral filter with weighted least squares is introduced in [24]. An optimization framework for filtering by combining edge detection and  $L_0$  gradient minimization is presented in [18]. A generalized framework that used the truncated Huber penalty function for filtering is proposed in [41]. A soft clustering based technique by exploiting a constrained Gaussian mixture model is presented in [40]. Although these edge-aware filtering techniques have yielded moderate results even for irregular textures of different scales, they mostly had multiple parameters and the filtering outcome heavily relied on the wise manual setting of these parameters.

## 3. Proposed technique

To remove the impulse noises and reduce the textural oscillations present in the input image  $I$ , the proposed technique pre-process it by applying morphological opening ( $\delta_{SE}(\varepsilon_{SE}(I))$ ) and closing ( $\varepsilon_{SE}(\delta_{SE}(I))$ ) operations as follows:

$$J = (I + \delta_{SE}(\varepsilon_{SE}(I)) + \varepsilon_{SE}(\delta_{SE}(I))) / 3 \quad (1)$$

For further details of the usefulness of this pre-processing step reader may refer to [17]. Once the pre-process image  $J$  is obtained, our technique follows two main steps to generate the filter image. In the first step, it proposes a novel method that generates a semantic-aware binary edge-map of the input image. Then, in the second step, with the help of the generated edge-map an edge-aware adaptive recursive median filter is used to produce the filter image. The details of the proposed technique are presented in the following subsections.

### 3.1. Generation of semantic-aware edge-map

We proposed a novel method that generates a semantic-aware edge-map of the input image. Our method considers semantic information in two phases. In the first phase, for each pixel  $p$ , it computes the mean gradient and average local skewness to generate the semantic gradient image (SGI). To this end, a set  $W = \{w_1, w_2, \dots, w_n\}$  containing increasing size windows are considered. For each window  $w_i \in W$ , four lines  $H, V, D1$ , and  $D2$  passing through the center pixel  $p$  in four directions (as shown in Fig. 1) are taken. To compute the mean gradient  $MD_W^p$  of the pixel  $p$  in the pre-processed image  $J$ , each line

$L \in \{H, V, D1, D2\}$  inside the windows is partitioned into two equal halves  $L_1$  and  $L_2$  at  $p$  and then calculate their average intensity difference as follows:

$$\begin{cases} MD_W^p = \frac{1}{|L| \times |W|} \times \sum_W \sum_L |mean(J_{L_1}) - mean(J_{L_2})| \\ L = L_1 \cup p \cup L_2 \end{cases} \quad (2)$$

where  $mean(J_{L_1})$  and  $mean(J_{L_2})$  represent the average intensity of the pixels on the line segments  $L_1$  and  $L_2$ , respectively. If  $p$  is a structural edge pixel, then for at least one  $L$  (i.e., for  $L = H, V, D1$ , or  $D2$ ) the average intensity difference between the neighbor pixels in  $L_1$  and  $L_2$  will be larger. As a result, the value of  $MD_W^p$  will be higher. If  $p$  is in the textural or smoother region, then for all  $L$  (i.e., for  $L = H, V, D1$ , and  $D2$ ) the average intensity difference between the neighbor pixels in  $L_1$  and  $L_2$  will be smaller. As a result, the value of  $MD_W^p$  will be lower. Thus, the higher value of  $MD_W^p$  indicates  $p$  is more probable to be a structural edge pixel (as some of its neighbor pixels are belonged to the different object) and the lower value of  $MD_W^p$  indicates that  $p$  is more probable to be a non-edge pixel (as all the neighbor pixels are in the same object). In the proposed work such semantic information is incorporated with the help of Eq. 2.

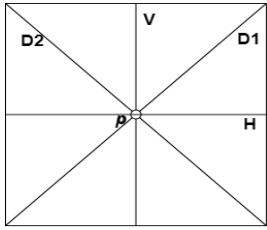


Figure 1. A fixed-size square window and the four lines  $H$ ,  $V$ ,  $D_1$ , and  $D_2$  passing through the center pixel  $p$

To compute the average local skewness of the pixel  $p$ , our technique generates the morphological gradient image  $J_{mg}$  as follows [17]:

$$J_{mg} = (\delta_{SE}(J) - \varepsilon_{SE}(J)) \quad (3)$$

Considering the pixel  $p \in J_{mg}$  as the center pixel of each window  $w_i \in W$ , a local histogram  $h_{w_i}$  is constructed by taking into account only the pixels inside the window  $w_i$  lie on the four lines  $H$ ,  $V$ ,  $D_1$ , and  $D_2$ . Then the average local skewness  $Sk_W^p$  of the pixel  $p$  is computed as:

$$Sk_W^p = \frac{1}{|W|} \sum_{w_i \in W} \frac{\max(h_{w_i}) - \text{median}(h_{w_i})}{\max(h_{w_i}) - \min(h_{w_i})}. \quad (4)$$

This local skewness basically provides the distribution of the gradient values of the neighbor pixels. If  $p$  is a structural edge pixel, then there will be a few higher gradient neighbor pixels of  $p$  for the structural edge and the rest will have lower gradients. This will lead to a right-skewed local histogram having a long gradient tail, which provides a higher skewness value. If  $p$  is a pixel in a textural or smoother region, then the gradients of its neighbor pixels have either higher (for textural region) or

lower (for smoother region) values. This will lead to either a left-skewed or a non-wide local histogram, which have lower skewness values. So, the higher value of  $Sk_W^p$  indicates  $p$  is more probable to be a structural edge pixel and the lower values of  $Sk_W^p$  indicates  $p$  is more probable to be a non-edge pixel. In the proposed work such semantic information is incorporated with the help of Eq. 4. After computing the average gradient  $MD_W^p$  and the average local skewness  $Sk_W^p$ , the proposed semantic gradient image (SGI) associated to the input image  $I$  is generated by combining these two features as follows:

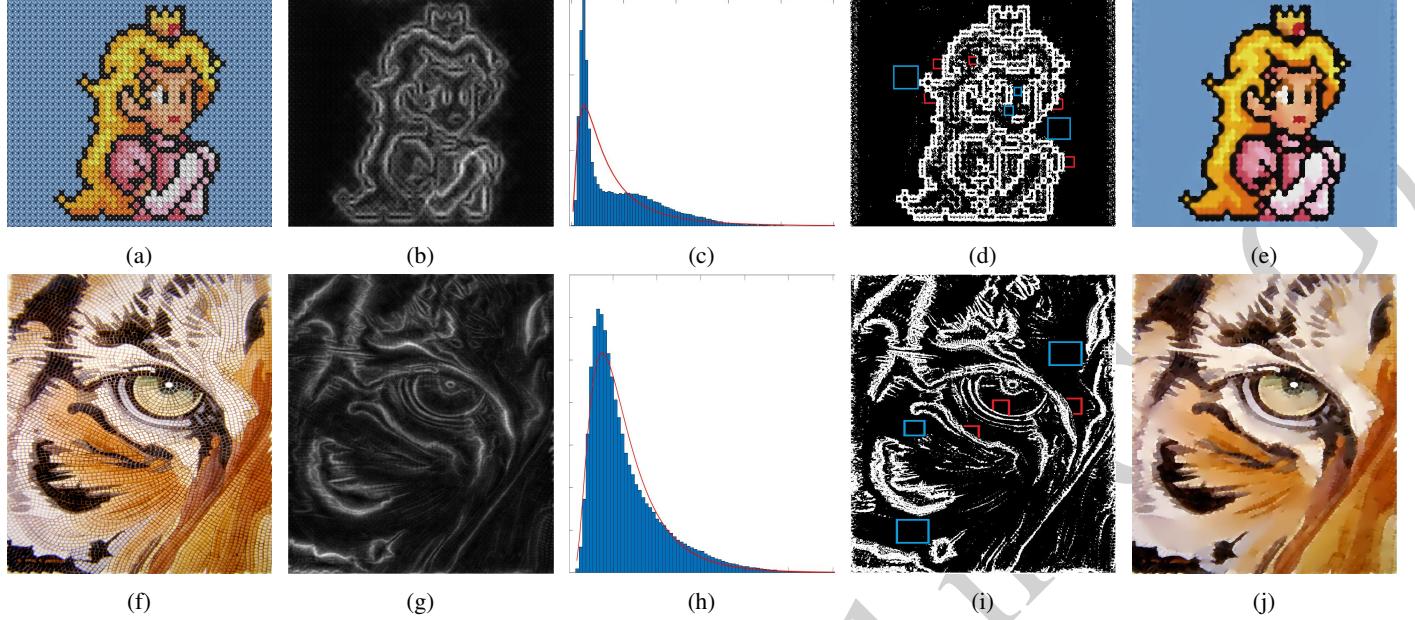
$$SGI(p) = MD_W^p \times Sk_W^p \quad (5)$$

$MD_W^p$  and  $Sk_W^p$  consider semantic information of a pixel  $p$  by analyzing the local distribution of pre-processed image  $J$  and morphological gradient image  $J_{mg}$ , respectively. Both provided higher values for structural edge pixels and lower for textural or non-edge pixels. Thus, Eq. 5 combines two different semantic information of the pixels for generating the semantic gradient image (SGI). Hence, the proposed SGI image is capable of better differentiating the structural and textural edge pixels of the input image. Figs. 2 (b) and (g) show the semantic gradient images generated by the proposed technique from the input images shown in Figs. 2 (a) and (f). From these figures, one can see that the higher intensity pixels of SGI are associated with the structural edge and the lower intensity pixels are associated with either non-edge or textural edge. Note that for the input image of irregular and varying scale textures, there is a possibility of having some pixels in SGI with high intensity values associated with textural edge pixels. In such a situation, even SGI provides little information to discriminate structural and textural edge pixels.

In order to discriminate the structural edge and non-edge pixels of the input image more accurately, we proposed a novel method that exploits JS divergence to consider semantic information of the pixels by defining appropriate windows. Since the number of structural edge pixels in an image is much smaller than the number of non-edge pixels. The histogram of the SGI is likely to follow a positive-skewed distribution as shown in Figs. 2(c) and (h). Thus, the histogram of SGI can be approximated with lognormal distribution by defining its two parameters  $\mu$  and  $\sigma$  as follows:

$$\begin{cases} f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{-(\ln(x)-\mu)^2}{2\sigma^2}} \\ \mu = \ln\left(\frac{v}{\sqrt{1+\frac{v^2}{\gamma^2}}}\right) \\ \sigma = \sqrt{\ln\left(1 + \frac{\gamma^2}{v^2}\right)} \end{cases} \quad \text{where } x > 0, \sigma > 0 \quad (6)$$

The red curves in Figs. 2(c) and (h) are the lognormal distributions that approximated the histograms of SGIs by estimating  $\mu$  and  $\sigma$  using the maximum likelihood estimator. The pixels in SGI associated with the left tail and the right tail of the approximated histogram are more certain to be non-edge pixels and structural edge pixels, respectively. The certainty level of the pixels between the left and right tails is less. In our work to consider semantic information of a pixel, the size of the window is defined based on its certainty level. Pixels that are certain to be non-edge and structural edge are used in large and



**Figure 2.** (a)(f) Original images and their corresponding generated (b)(g) semantic gradient images (SGIs), (c)(h) histograms of SGIs, (d)(i) edge-maps, and (e)(j) filtered images.

small windows, respectively and the moderate size windows are used for uncertain pixels. The certainty level of the pixels are defined with help of the approximated histogram. After estimating  $\mu$  and  $\sigma$ , the histogram of SGI is divided into four intervals  $(0, e^{\mu-\sigma})$ ,  $[e^{\mu-\sigma}, e^\mu]$ ,  $[e^\mu, e^{\mu+\sigma})$  and  $[e^{\mu+\sigma}, +\infty)$ . The intensity values of the pixels in SGI within the first interval  $(0, e^{\mu-\sigma})$  and the fourth interval  $[e^{\mu+\sigma}, +\infty)$  are more certain to be non-edge pixels and structural edge pixels, respectively. So, the window  $w_{01}, w_{12}, w_{23}$ , and  $w_{34}$  (such that the size of  $w_{01} > w_{12} > w_{23} > w_{34}$ ) is used to consider semantic information of the pixels associated to the intervals  $(0, e^{\mu-\sigma})$ ,  $[e^{\mu-\sigma}, e^\mu]$ ,  $[e^\mu, e^{\mu+\sigma})$  and  $[e^{\mu+\sigma}, +\infty)$ , respectively. Once the window for each pixel of the input image is defined, the proposed technique exploits JS divergence to consider semantic information for determining whether it belongs to a structural edge or not [16].

In probability theory, JS divergence is a useful tool for measuring the similarity between two probability distributions. The JS divergence  $JSD(P, Q)$  between the two discrete random variables  $P$  and  $Q$  gives a symmetrical score in the range  $[0,1]$ . Lower the value of  $JSD(P, Q)$  indicates  $P$  and  $Q$  have similar distributions. In this research, for each pixel  $p$  of the pre-processed image  $J$ , its corresponding window  $w \in \{w_{01}, w_{12}, w_{23}, w_{34}\}$  is used to incorporate semantic information. To this end, considering  $p$  as the center pixel of  $w$ , four lines, namely horizontal  $H$ , vertical  $V$ , two diagonal  $D1$  and  $D2$  pass through  $p$  as shown in Fig. 1 are taken. The probability density function of each line  $H, V, D1$ , and  $D2$  is defined as follows:

$$\left\{ P_L(q) = \frac{|J_q - \min(\text{mean}(J_{L_1}), \text{mean}(J_{L_2}))|}{\sum |J_q - \min(\text{mean}(J_{L_1}), \text{mean}(J_{L_2}))|}, \quad q \in L_1 \cup L_2, \quad L \in \{H, V, D1, D2\} \right. \quad (7)$$

where  $J_q$  represents the intensity value of the pixel  $q \in L_1 \cup L_2$ . The probability density function of at least one of the four lines (i.e.,  $P_H, P_V, P_{D1}$  or  $P_{D2}$ ) will be similar to a discrete step distribution if the pixel  $p$  is in the structural edge. Otherwise, the

probability density functions of all four lines will be more akin to discrete uniform distributions. To incorporate such semantic information two reference functions, a uniform function  $U$  and a step function  $T$  are defined as follows:

$$T(q) = \begin{cases} 1, & q \in L_1 \\ 0, & q \in L_2 \end{cases} \quad (8)$$

$$T(q) = \begin{cases} 0, & q \in L_1 \\ 1, & q \in L_2 \end{cases} \quad (9)$$

Let  $Q_L$  and  $R_L$  be the probability density function of the step function  $T$  and the uniform function  $U$ , respectively. If the pixel  $p$  is a structural edge pixel then for at least one line  $L \in \{H, V, D1, D2\}$ , the distribution of  $P_L$  will be similar to the step distribution  $Q_L$  i.e., the value of  $JSD(P_L, Q_L)$  will be low. If the pixel  $p$  is not a structural edge pixel then for all four lines the distribution of  $P_L$  will be similar to the uniform distribution  $R_L$  i.e., the value of  $JSD(P_L, R_L)$  will be low. In our work such semantic information is exploited to discriminate structural edge and non-edge pixels of the input image and generate the binary edge-map  $E_b$  as follows:

$$E_b(p) = \begin{cases} 1, & \text{if } \min_L \{JSD(P_L, Q_L)\} \\ & \leq \frac{1}{|L|} \sum_L \{JSD(P_L, R_L)\} \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

Figs. 2 (d) and (i) show the semantic-aware edge-maps generated by the proposed technique.

### 3.2. Edge-aware adaptive recursive median filter

Once the semantic edge-map of the input image is obtained, an edge-aware recursive median filter [42, 43] is used to generate the filtered image. The adaptive window of this filter is defined with the help of the edge-map [17]. A brief description of how to define such an adaptive dynamic window is given below.

The filtering technique follows two different approaches to define the dynamic window. For filtering non-edge pixels, initially, it considers a small square window. Then gradually increases the size of the window  $L_w$  until it meets a structural edge pixel or the size of  $L_w$  becomes sufficiently large. The blue squares shown in Figs. 2(d) and (i) are the adaptive dynamic windows defined for filtering some non-edge pixels.

For filtering structural edge pixels, the dynamic window is defined by identifying the object in which it belongs. Let  $p$  be a structural edge pixel and  $I_p$  be its intensity value. Initially, a fixed-size square window  $W_s$  is considered. Then, within the window  $W_s$ , the sub-image  $I_{w_s}$  and  $E_{w_s}^e$  is taken from the image  $I$  and the edge-map  $E_b$ , respectively. The average intensity value  $R_{mid}$  of all the edge pixels within  $W_s$  (represented as  $I_{w_s}^e$ ) is computed as  $R_{mid} = (\max(I_{w_s}^e) + \min(I_{w_s}^e))/2$ .  $I_p$  is greater than  $R_{mid}$  implies that the pixel  $p$  belongs to the brighter object. Otherwise, it belongs to the darker object. Thus, the dynamic window  $W_s^a$  is defined as follows:

$$\begin{cases} R_{mid} = (\max(I_{w_s}^e) + \min(I_{w_s}^e))/2 \\ W_s^l = \text{pixels in } W_s \text{ for which } (I_{w_s} \leq R_{mid}) \\ W_s^u = \text{pixels in } W_s \text{ for which } (I_{w_s} > R_{mid}) \\ \text{if } (I_p \leq R_{mid}), \text{ then } W_s^a = W_s^l \\ \text{if } (I_p > R_{mid}), \text{ then } W_s^a = W_s^u \end{cases} \quad (11)$$

The red boxes shown in Figs. 2(d) and (i) are some adaptive dynamic windows obtained by this approach for filtering edge pixels. Figs. 2(e) and (j) show the filtered images generated by the proposed technique.

### 3.3. Parameters of the proposed technique

Although the proposed technique considers several windows but most of them are of fixed-size. In pre-processing step, irrespective of the input image, a fixed-size  $3 \times 3$  square window is used as SE to generate the pre-processed image  $J$ . To generate SGI, the proposed technique computes the average gradient and the average local skewness of each pixel by considering semantic information. For the image having varying scale irregular textures, it is not possible to capture sufficient semantic information by using a single window. To incorporate appropriate semantic information, a set of windows  $W = \{w_1, w_2, \dots, w_n\}$  of increasing sizes are considered. The present work suggested to consider only four windows  $w_1, w_2, w_3$  and  $w_4$  to capture semantic information of the variety of textures starting from smaller to moderate to larger, and the size of these windows are defined by taking 3<sup>rd</sup>, 5<sup>th</sup>, 7<sup>th</sup>, and 9<sup>th</sup> numbers from the Fibonacci series [1, 1, 2, 3, 5, 8, 13, 21, 34, ...]. In our work, the size of a window is determined by using the formula  $2x + 1$ ,

where  $x$  is a number taken from the Fibonacci series. So, irrespective of the considered input image, four windows of size  $5 \times 5$ ,  $11 \times 11$ ,  $27 \times 27$ , and  $69 \times 69$  are used to generate the SGI.

After generating the SGI, to incorporate semantic information of each pixel of the input image for determining whether it belongs to a structural edge or not, our technique considers one of the four windows  $w_{01}, w_{12}, w_{23}$ , or  $w_{34}$ . In this research irrespective of the considered input image the size of the window  $w_{01}$  and  $w_{34}$  is fixed as the size of  $w_1$  (i.e.,  $69 \times 69$ ), and  $w_4$  (i.e.,  $5 \times 5$ ), respectively. Moreover, depending upon the textures present in the input image, the size of the other two windows  $w_{12}$  and  $w_{23}$  are manually fixed from four discrete options  $F_1, F_2, F_3$  and  $F_4$ . Where,  $F_1 = [w_{12} = 43 \times 43, w_{23} = 27 \times 27]$ ,  $F_2 = [w_{12} = 27 \times 27, w_{23} = 17 \times 17]$ ,  $F_3 = [w_{12} = 17 \times 17, w_{23} = 11 \times 11]$ , and  $F_4 = [w_{12} = 11 \times 11, w_{23} = 7 \times 7]$  is defined by considering four consecutive pairs of Fibonacci numbers [21, 13], [13, 8], [8, 5] and [5, 3], respectively. Note that the size of windows  $w_{12}$  and  $w_{23}$  are the only parameters of the proposed technique that need to be fixed manually by choosing one of the four options  $F_1, F_2, F_3$  or  $F_4$ . Irrespective of the input images, the other parameters are either fixed or defined automatically. On the other hand, most of the existing state-of-the-art methods have multiple parameters, and their results are highly sensitive to their parameter values which need to be fine tuned manually within a wide interval. Fig. 3 shows the filtered images produced by the proposed technique considering the options  $F_1, F_2, F_3$ , and  $F_4$  for a simple cartoon image. From this figure, one can see that the options  $F_3$  and  $F_4$  are suitable for preserving smaller details, whereas the options  $F_1$  and  $F_2$  are suitable for smoothing.

### 3.4. Computational complexity

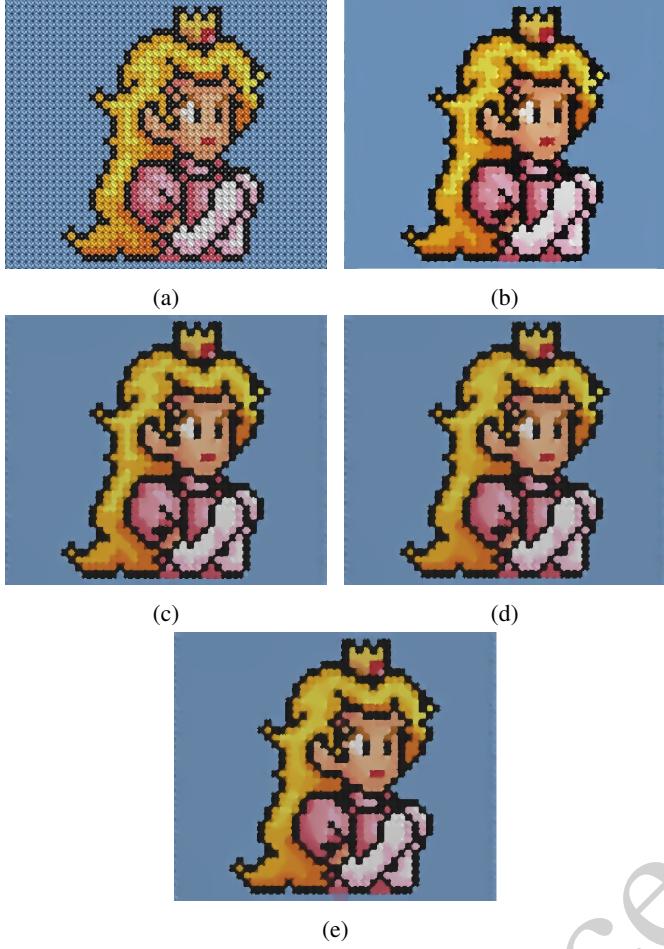
The computational complexity of the proposed technique is dependent on the algorithm used to generate the semantic-aware edge-map and the recursive median filter. For an input image of size  $M \times N$ , the computational complexity of the algorithm that generates the semantic-aware edge-map is  $\Theta(M \times N \times S_1 \log S_1)$ , where  $S_1$  is the size of the largest window in  $W$  used to compute the local skewness of the pixels. The computational complexity of the adaptive median filter is  $\Theta(M \times N \times S_2 \log S_2)$ , where  $S_2$  is the size of the largest window among the adaptive windows  $L_w$  and  $w_s^a$  automatically defines for filtering non-edge and edge pixels, respectively. Therefore, the computational complexity of the proposed technique is  $\Theta(M \times N \times (S_1 \log S_1 + S_2 \log S_2))$ .

## 4. Experimental results and analysis

In order to assess the robustness of the proposed technique, in the experiment both qualitative and quantitative comparisons with several state-of-the-art techniques are shown.

### 4.1. Qualitative comparison

In this experiment the quality of the filtered images produced by the proposed technique are compared with the filtered images generated by the six popular state-of-the-art structure-preserving filtering techniques: Region Covariance (Reg-Cov)



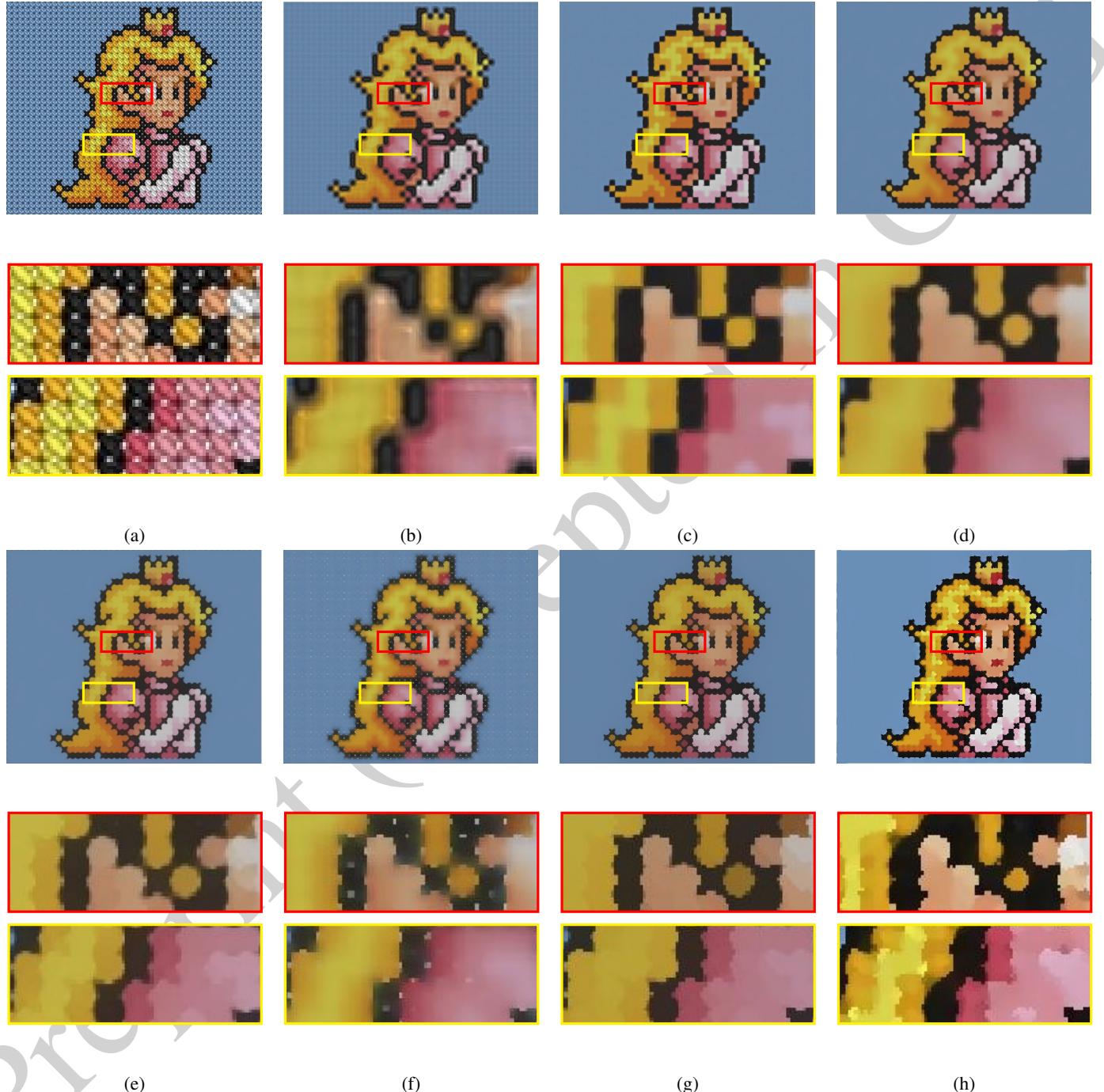
**Figure 3.** (a) Original image and the filtered images generated by the proposed technique by using the options (b)  $F_4$ , (c)  $F_3$ , (d)  $F_2$ , and (e)  $F_1$ .

[8], Bilateral Texture Filteringing (BTF) [9], Scale-Aware Texture Filtering (SATF) [44], Structure Adaptive Total Variation (SATV) [10], Real-Time Iterative Least Square (RILS) [45], and Generalized Image Smoothing Framework (GISF) [41]. The first two i.e., Reg-Cov and BTF are textural features based techniques. The next two, i.e., SATF and SATV are structural edge-aware techniques. The last two, i.e., RILS and GISF are optimization-based techniques. In qualitative comparison, the filtered images produced by the different techniques are visually analyzed to find out how well the structures and textures present on the input images are discriminated against for removing minute details of textures while retaining the key structures.

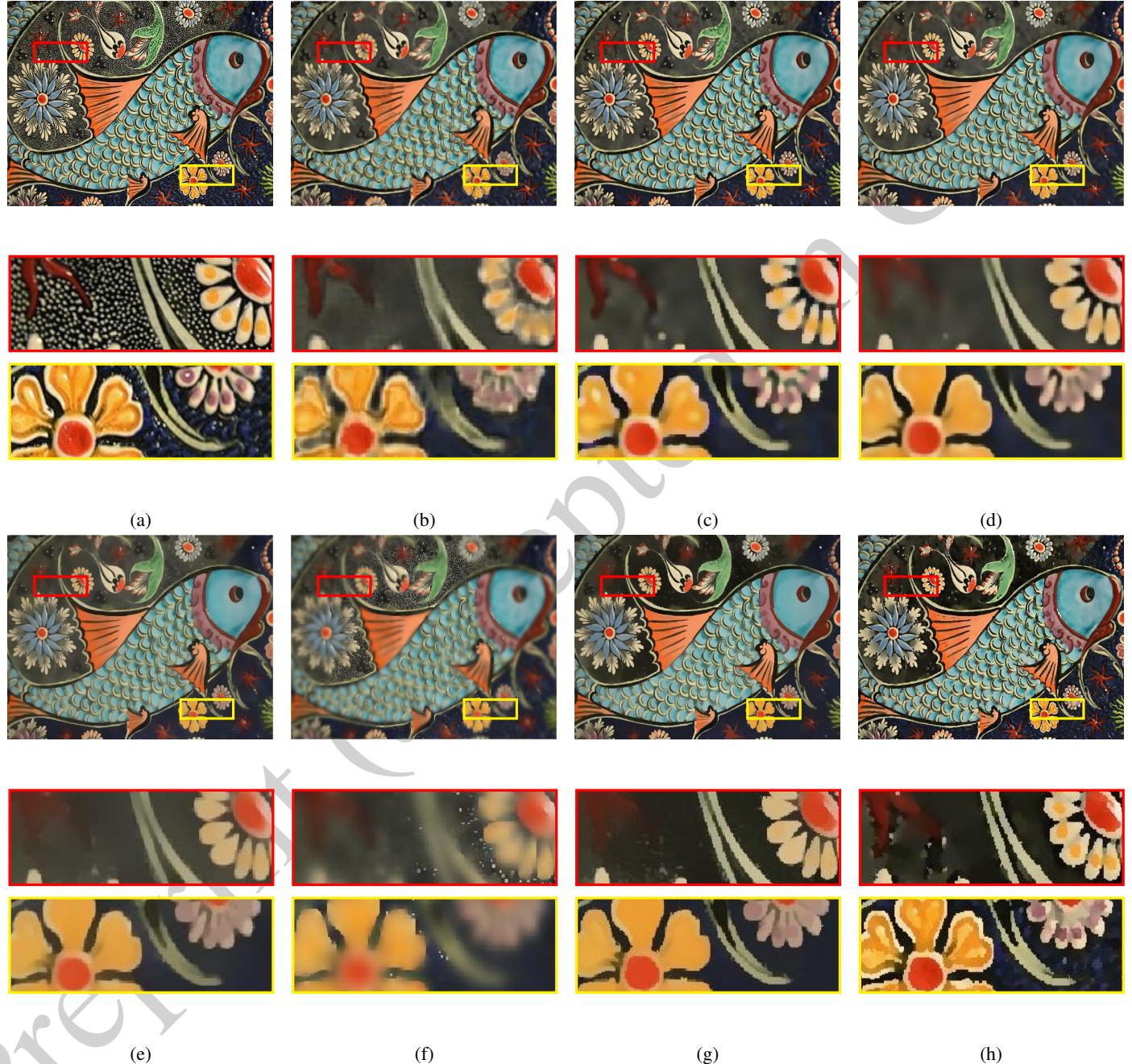
For wide varieties of input images, the filtered images produced by the different techniques are shown in Figs. 4, 5, 6, 7 and 8. By looking at these figures one can see that the proposed technique is outperformed in terms of both, texture smoothing and structure preserving. For example, the input image shown in Fig. 4(a) is a simple cartoon image with regular textures and smaller structural details of different colors. The filtered images produced by the different techniques and some zoomed portions of these filtered images are shown in Fig. 4. From the figure, one can see that the proposed technique better preserves the smaller color details while smoothing out the textu-

ral details. Fig. 5(a) shows another input image with random textures and complex structural details. The corresponding filtered images and their zoomed portions are shown in Fig. 5. Again, from these images one can see that compared to the existing state-of-the-art methods, the proposed technique better preserves the smaller structures while smoothing out the textural details. Figs. 6 and 7 show the filtering results for another two input images with varying scale textures. By visually analyzing these results one can see that the proposed technique is able to preserve the structures and smooth out the textures much better than most of the state-of-the-art techniques. Note that all the state-of-the-art techniques considered here have multiple parameters and their filtering results are heavily dependent on these parameters. The values of these parameters are varied from image to image and need to be fixed manually within a wide range of continuous intervals. For a fair comparison, in the present experiment, if possible, the values of these parameters are taken from the respective papers for which the best results are reported. Otherwise, the best possible values of these parameters are set manually by using the trial and error method. On the other hand, as described in Subsection 3.3 the size of the windows  $w_{12}$  and  $w_{23}$  are the only parameters of the proposed technique that needs to be fixed manually by choosing one of the four options of  $F_1, F_2, F_3$  or  $F_4$ . This seems to be a great advantage of the proposed technique as compared to the state-of-the-art techniques. Our proposed technique is tested on a large number of images with a variety of regular and irregular textural patterns and always provides satisfactory results. Fig. 8 shows a few more filtering results provided by the proposed technique for different types of input images. From these results, one can observe that our technique accomplishes numerous contradictory aims, such as removing textures, maintaining structural boundaries, safeguarding subtle features like corners, and avoiding over-sharpening and/or over-blurring artifacts.

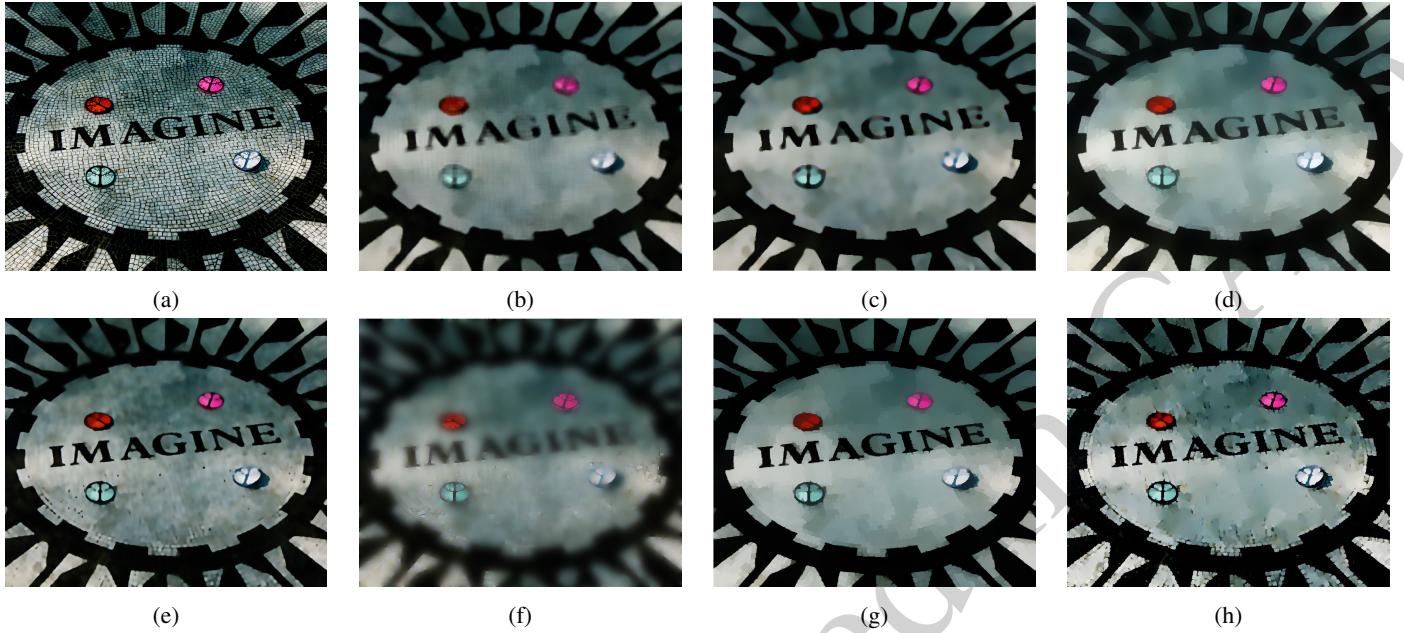
In order to further analyze the effectiveness of the proposed technique, it compared with the two semantic-aware filtering techniques recently proposed in [16] (called SSTF) and [17] (called SMMF). Note that the basic difference between these two techniques and the proposed one is the way they exploit semantic information for generating the edge-map of the input image. To incorporate semantic information both, the SSTF [16] and the SMMF [17] used a single window which is defined manually by using trial-and-error method. Whereas, the proposed technique uses multiple windows and the size of these windows are defined from four discrete options. Fig. 9 shows the filtered images generated by the SSTF, the SMMF, and the proposed techniques for two different type of input images, one having regular texture and the other one of having varying scale irregular texture. From this figure, one can see that for the image with regular texture all the three techniques provided satisfactory results. But for the image with varying scale irregular texture, the proposed technique provides better filtering results. Since for the images with varying scale irregular textures a single window will not be capable of capturing sufficient semantic information. Both the SSTF and the SMMF failed to provide good results. Whereas, the proposed technique uses multiple windows to capture better semantic information.



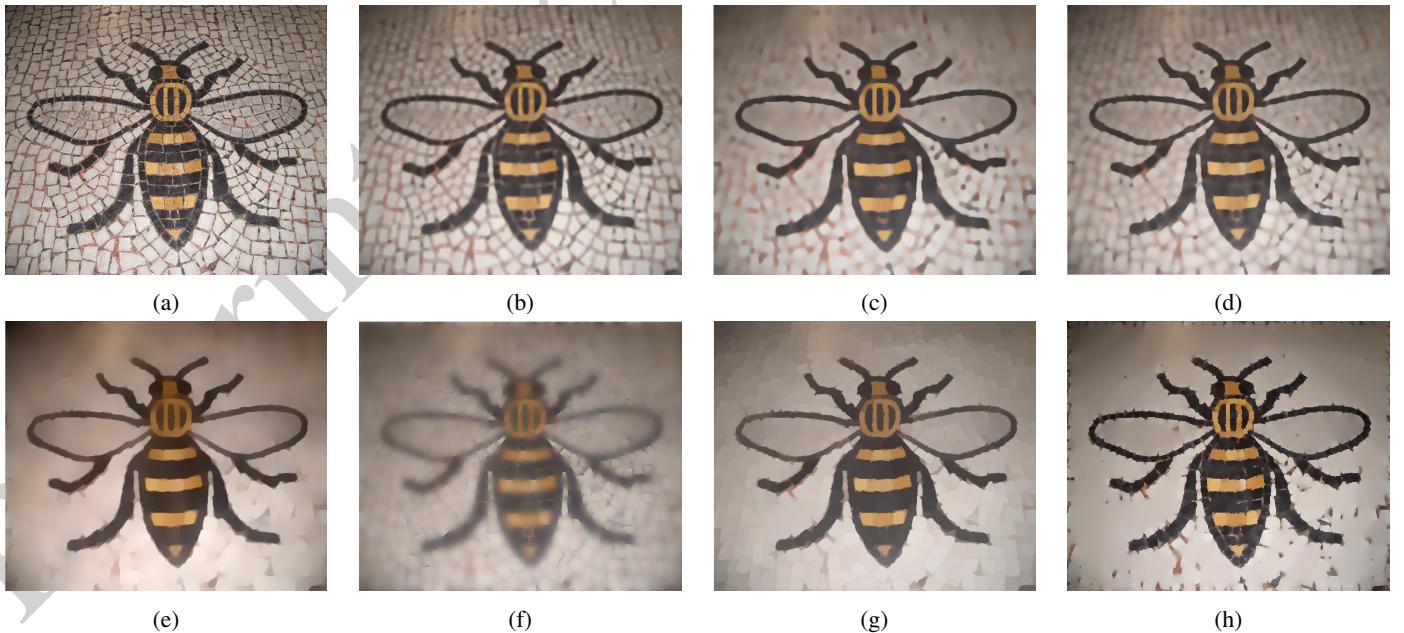
**Figure 4.** Filtering results of a cross-stitch cartoon image having regular texture: (a) the original image and the filtered images obtained by (b) Reg-Cov ( $k = 15, ps = 6, \sigma = 0.2$ ) [8] (c) BTF,  $k = 9, n_{itr} = 7$  [9], (d) SATF  $ss = 3, sr = 0.1, st = 0.1, n_{itr} = 7, div = 30$  [44], (e) SATV,  $\lambda = 2.5$  [10], (f) RILS  $\rho_{smooth} = 3, \rho_{sharp} = 5$  [45], (g) GISF,  $\lambda = 50, \gamma = 20/255, n_{itr} = 15$  [41], and (h) Proposed ( $F_4$ ) techniques.



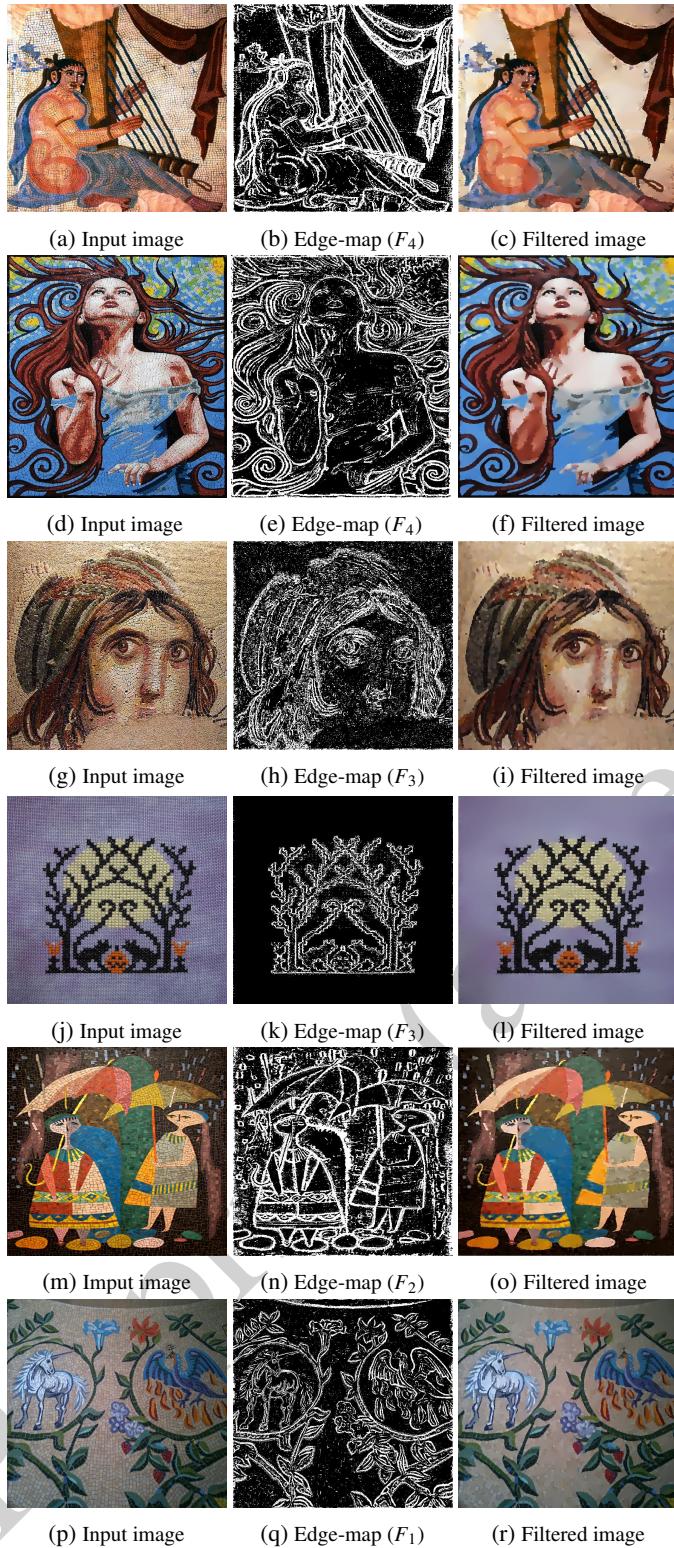
**Figure 5. Filtering results of a mosaic image having complex textures and structures:** (a) the original image and the filtered images obtained by (b) Reg-Cov ( $k = 15, ps = 6, \sigma = 0.2$ ) [8] (c) BTF,  $k = 9, n_{itr} = 7$  [9], (d) SATF  $ss = 3, sr = 0.1, st = 0.1, n_{itr} = 7, div = 30$  [44], (e) SATV,  $\lambda = 2.5$  [10], (f) RILS  $\rho_{smooth} = 3, \rho_{sharp} = 5$  [45], (g) GISF,  $\lambda = 50, \gamma = 20/255, n_{itr} = 15$  [41], and (h) Proposed ( $F_3$ ) techniques.



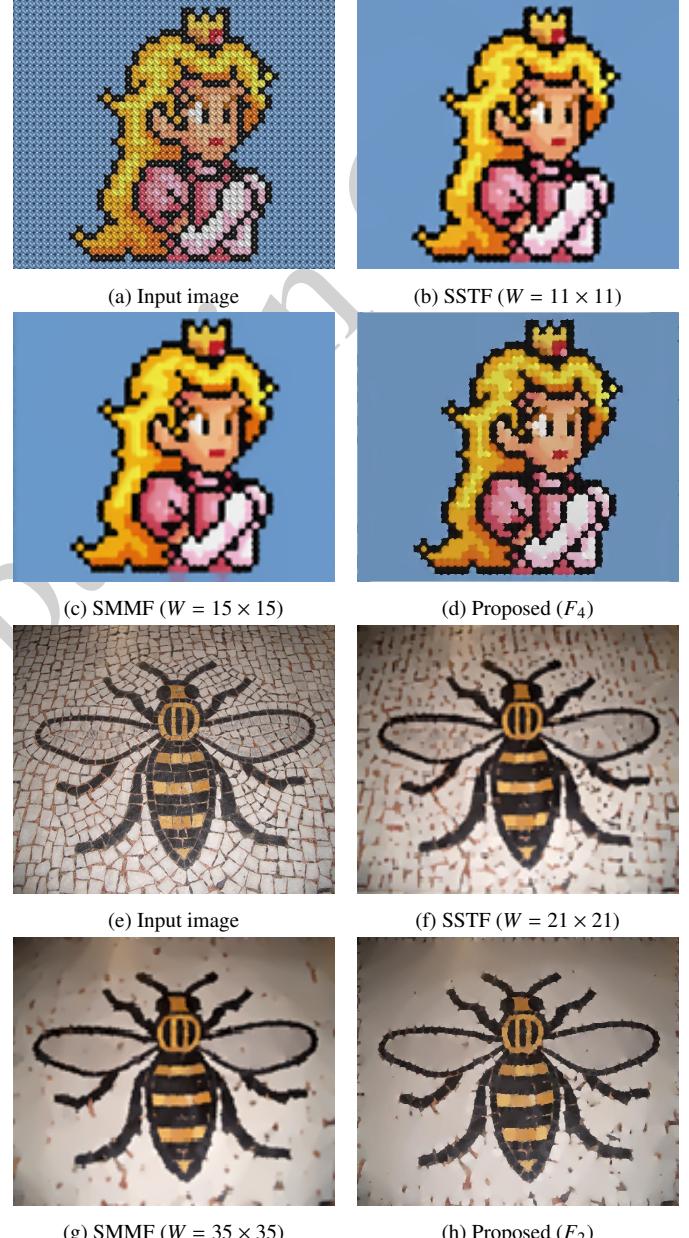
**Figure 6.** Filtering results of a floor image having irregular texture: (a) the original image and the filtered images obtained by (b) Reg-Cov ( $k = 15, ps = 6, \sigma = 0.2$ ) [8] (c) BTF, ( $k = 9, n_{itr} = 7$ ) [9], (d) SATF ( $ss = 3, sr = 0.1, st = 0.1, n_{itr} = 7, div = 30$ ) [44], (e) SATV, ( $\lambda = 2.5$ ) [10], (f) RILS ( $\rho_{smooth} = 3, \rho_{sharp} = 5$ ) [45], (g) GISF, ( $\lambda = 50, \gamma = 20/255, n_{itr} = 15$ ) [41], and (h) Proposed ( $F_3$ ) techniques.



**Figure 7.** Filtering results of a bee image having irregular varying scale textures: (a) the original image and the filtered images obtained by (b) Reg-Cov ( $k = 15, ps = 6, \sigma = 0.2$ ) [8] (c) BTF, ( $k = 9, n_{itr} = 7$ ) [9], (d) SATF ( $ss = 3, sr = 0.1, st = 0.1, n_{itr} = 7, div = 30$ ) [44], (e) SATV, ( $\lambda = 2.5$ ) [10], (f) RILS ( $\rho_{smooth} = 3, \rho_{sharp} = 5$ ) [45], (g) GISF, ( $\lambda = 50, \gamma = 20/255, n_{itr} = 15$ ) [41], and (h) Proposed ( $F_2$ ) techniques.



**Figure 8.** (a)(d)(g)(j)(m)(p) Original images with various textures and structures and their corresponding (b)(e)(h)(k)(n)(q) edge-maps and (c)(f)(i)(l)(o)(r) filtered images generated by the proposed technique.



**Figure 9.** Original (a) Cartoon image with regular texture and (e) Bee image with varying scale irregular texture. Corresponding filtered images generated by (b)(f) SSTF [16] (c)(g) SMMF [17], and (d)(h) Proposed techniques.

#### 4.2. Quantitative comparison

In this experiment, a quantitative comparison between the filtered images produced by the proposed and the state-of-the-art techniques is carried out. As the conventional Image Quality Assessment (IQA) measures such as Peak Signal to Noise Ratio (PSNR) and Signal to Noise Ratio (SNR) does not help to assess the quality of the filtered images produced by the structure-preserving filtering techniques [46]. Here two subjective reference IQA metrics Structural Similarity Index (SSIM) [47] and Multi-scale SSIM (MSSIM) and the subjective no-reference metric Perception Image Quality Evaluator (PIQE) [49] are used for quantitative comparisons. Table 1 reports the IQA values computed from the filtered images generated by different techniques. From this results, one can see that the proposed technique always provides better IQA values than the existing state-of-the-art techniques.

### 5. Applications

Structure preserving image filtering techniques are extensively applied for different computer graphics applications such as image denoising, details enhancement, tone mapping, etc. They are seldom exploited for image classification. Here we demonstrate the effectiveness of our proposed technique not only for computer graphics applications but also for incorporating spatial information in image classification.

#### 5.1. Computer graphics applications

The proposed filtering technique can be used for different computer graphics applications such as image denoising, details enhancement, and tone mapping. The details are given in the following subsections:

##### 5.1.1. Image denoising

Image denoising involves the removal of unwanted or random variations in pixel values, known as noise, from digital images. The noises can be introduced during image acquisition, transmission, or processing, leading to degradation in image quality. The objective of denoising is to restore the original image by minimizing the impact of noise. Fig. 10 shows some noisy images and their corresponding denoised (filtered) images produced by the proposed technique. By looking into the filtered images one can see the ability of the proposed technique to preserve structures while removing noises.

Table 1. The SSIM, MSSIM, and PIQE for the filtered images produced by different techniques. Bold typeface indicates the best results.

Images	Metrics	RegRov	BTF	SATF	SATV	RILS	GIFS	Proposed
Cartoon- Fig. 4 PIQE=44.5055	SSIM	0.5227	0.5772	0.6006	0.6091	0.5017	0.5424	<b>0.6236</b>
	MSSIM	0.5161	0.5401	0.5613	0.5704	0.4276	0.4348	<b>0.5939</b>
	PIQE	66.5025	87.1371	87.1248	82.4621	100	85.5485	<b>53.8407</b>
Fish- Fig. 5 PIQE=31.6049	SSIM	0.4968	0.5527	0.6353	0.6759	0.4945	0.4544	<b>0.6924</b>
	MSSIM	0.6909	0.7431	0.8171	0.8284	0.6063	0.5238	<b>0.8522</b>
	PIQE	69.88529	87.2677	84.6146	73.6599	83.8448	80.7166	<b>64.4851</b>
Imagine- Fig. 6 PIQE=37.2748	SSIM	0.36	0.38	0.43	0.31	0.26	0.44	<b>0.4583</b>
	MSSIM	0.58	0.61	<b>0.68</b>	0.62	0.54	0.62	0.6551
	PIQE	86.78	88.85	87.05	84.74	88.19	86.26	<b>82.6128</b>
Bee- Fig. 7 PIQE=16.2815	SSIM	0.49	0.44	0.57	0.51	0.47	0.53	<b>0.5588</b>
	MSSIM	0.58	0.74	0.70	0.51	0.50	0.53	<b>0.76</b>
	PIQE	86.74	88.86	89.96	86.47	100	87.91	<b>72.2437</b>



**Figure 10.** (a)(d)(g) Original images, (b)(e)(h) images with Gaussian noise ( $\sigma = 0.03$ ) and their corresponding (c)(f)(i) denoised images obtained by applying proposed filtering technique.

### 5.1.2. Detail enhancement

Image enhancement encompasses the modification of an image to enhance its visual quality, rendering it more suitable for human perception or specific computer vision tasks. Image enhancement techniques aim to accentuate particular features, amplify contrast, reduce noise, and elevate the image quality. For image enhancement, the input image is decomposed into two components: a smooth image and an image containing textural details. Then the enhanced image is obtained by adding the textural details back to the original image. Fig. 11 shows some enhanced images produced by the proposed technique.



**Figure 11.** (a)(c)(e) Original images and their corresponding (b)(d)(f) enhanced images obtained by the proposed technique.

### 5.1.3. Tone mapping

Tone mapping is a technique employed in computer graphics to transform high-dynamic-range (HDR) images into low-dynamic-range (LDR) images, a format suitable for display on devices such as computer monitors, TVs, or printed media. HDR images capture a broader range of luminance values compared to standard images and tone mapping assists in rendering these images on regular display while retaining visual information as much as possible. Figs. 12(a) and (c) show two LDR images obtained from HDR images by applying the tone mapping method presented in [50]. Figs. 12(b) and (d) show the LDR images generated by the tone mapping method presented in [2], where in place of bilateral filtering our proposed filtering is embedded. From these figures, one can see the usefulness of the proposed filtering technique for tone mapping.



**Figure 12.** (a)(c) Tone mapped images by [50] and (b)(d) tone mapped images generated by [2] that embedded proposed filtering technique.

## 5.2. Incorporating spatial information for image classification

The aim of the structure-preserving filtering techniques is to smooth out textures regardless of their spectral variations while maintaining the structures of the significant objects. So the filtered images generated by such techniques can provide rich spatial information. It has already been proven that the inclusion of appropriate spatial information of the pixel along with its spectral features significantly improves image classification results. So, the filtering technique can become an effective tool for image classification which has diverse applications in the fields of remote sensing, medical imaging, text analysis, etc. Although the effectiveness of the existing filtering techniques is assessed for different computer graphics applications, their ability to provide spatial information for image classification is seldom tested. In this work, the robustness of the proposed filtering technique to incorporate spatial information for image classification is tested on hyperspectral datasets.

### 5.2.1. Spectral-spatial HSI classification

The proposed filtering technique filtered out insignificant details and noises while minimizing distortion of the underlying structures in the image. So, the filtered images produced by this technique provide useful spatial information. To consider sufficient spatial information, for an input image multiple filtered images are generated by varying the size of the windows  $w_{12}$  and  $w_{23}$ . However, when dealing with HSI with hundreds of spectral bands, the generation of multiple filtered images for each band is cumbersome. To circumvent this problem, a reduced set of principal components (PCs) extracted from the HSI, which preserves sufficient information is considered. For each of the considered PCs, four filtered images are generated by varying the size of  $w_{12}$  and  $w_{23}$  using the options  $F_1, F_2, F_3$ , and  $F_4$ . Finally, for spectral-spatial classification of the HSI, all the filtered images derived from the considered PCs are concatenated together to construct an extended semantic filtering profile (ESFP).

Let  $I$  be the hyperspectral image with dimension  $M \times N \times P$ , where  $M$  and  $N$  denote the spatial dimension, and  $P$  represents the number of spectral bands. To effectively reduce the dimension of HSI, principal component analysis (PCA) is applied to  $I$ . Consequently, the first  $m$  PCs denoted as  $I^{pc_i}, i = 1, 2, \dots, m$  are chosen to represent the HSI. The proposed filtering technique is applied on each PC by varying the size of the windows  $w_{12}$

and  $w_{23}$ . For a PC  $I^{pc_i}$ , the four filtered images  $I_{F_1}^{pc_i}, I_{F_2}^{pc_i}, I_{F_3}^{pc_i}$ , and  $I_{F_4}^{pc_i}$  are generated by choosing the options  $F_1, F_2, F_3$ , and  $F_4$ . Now by concatenating all the generated filtered images the ESFP of HSI is formed as follows:

$$I_{ESFP} = [\{I_{F_1}^{pc_1}, I_{F_2}^{pc_1}, \dots, I_{F_4}^{pc_1}\}, \{I_{F_1}^{pc_2}, I_{F_2}^{pc_2}, \dots, I_{F_4}^{pc_2}\}, \dots, \{I_{F_1}^{pc_m}, I_{F_2}^{pc_m}, \dots, I_{F_4}^{pc_m}\}] \quad (12)$$

The constructed *ESFP* derived from the HSI contains rich spatial information. Fig. 13 shows the first PC and its four filtered images produced by the proposed technique for the Pavia University dataset. Once  $I_{ESFP}$  is obtained, each pixel of the HSI is represented with the corresponding spectral features in  $I_{ESFP}$ . In this work Random Forest (RF) classifier is used for classification. To assess the usefulness of the proposed filtering technique for considering spatial information, the classification results provided by it are compared with several state-of-the-art techniques such as EMEP[53], MFASRC[54], Gabor-CNN [56], TEAP [57], and EPF [58] by considering same training and test sets as used in [55]. Table 2 reports the average class-wise accuracy (AA), overall accuracy (OA), and kappa coefficient ( $\kappa$ ) provided by the different techniques for the three real HSI datasets, viz. Indian Pines, Pavia University, and University of Houston. From this table, one can see that for all three datasets, our proposed ESFP provides either the best classification result or a similar result to that of the best state-of-the-art method. The results confirmed the usefulness of the proposed filtering technique in incorporating spatial information for the classification of HSI. Note that this approach is also applicable to the spectral spatial classification of medical images, text images, etc.

## 6. Conclusion

The most challenging task in structure preserving filtering is to discriminate important structures from textures, especially when irregular textural patterns with different scales are there. As in many cases, spectral and spatial variations of the input image are not sufficient to discriminate structures from textures. In such cases, semantic information may provide additional useful insight. In this regard, the existing methods are very much parameter sensitive. Here, we proposed a novel method to exploit



**Figure 13.** First PC of the Pavia University data set and the four filtered images generated from it by varying the size of the windows  $w_{12}$  and  $w_{23}$  using the options  $F_1$ ,  $F_2$ ,  $F_3$ , and  $F_4$ .

**Table 2.** Average class-wise accuracy (AA), overall accuracy (OA), and kappa coefficient ( $\kappa$ ) provided by the EMEP, MFASRC, Gabor-CNN, TEAP, EPF, and the proposed ESFP techniques for Indian Pines, Pavia University, and University of Houston datasets.

Accuracy Measures	EMEP[53]	MFASRC[54]	Gabor-CNN[56]	TEAP[57]	EPF[58]	ESFP
Indian pines						
AA	96.00	97.25	95.24	93.86	93.84	<b>97.29</b>
OA	93.70	95.22	92.84	91.04	90.99	<b>95.91</b>
$\kappa$	0.9279	0.9452	0.9161	0.8975	0.8969	<b>0.9531</b>
Pavia University						
AA	<b>96.57</b>	88.26	87.83	92.47	88.59	94.68
OA	<b>95.46</b>	74.39	91.62	89.17	85.30	93.33
$\kappa$	<b>0.9407</b>	0.6828	0.8914	0.8562	0.8072	0.9300
University of Houston						
AA	83.64	84.36	82.94	85.33	75.32	<b>85.90</b>
OA	80.83	82.09	<b>84.12</b>	82.82	70.21	83.16
$\kappa$	0.7920	0.8058	<b>0.8251</b>	0.8138	0.6773	0.8173

the semantic information of the image for better discriminating the structural and textural edges. Our technique first generates a semantic-aware edge-map of the input image by exploiting semantic information. Then an edge-aware adaptive recursive median filter is utilized to produce the filter image. Although the proposed filtering technique requires multiple windows, the size of most of these windows is either fixed or defined automatically irrespective of the considered input image. There exists only a pair of windows that need to be fixed manually by choosing one of the four discrete options. The proposed technique provides satisfactory results for a broad range of input images with minimal fine-tuning of its parameters, which is an important benefit in comparison to the current state-of-the-art techniques. Moreover, in contrast to the majority of existing techniques, our technique accomplishes numerous contradictory aims, such as locating and removing texture, maintaining structural boundaries, safeguarding subtle features like corners, and avoiding over-sharpening and/or over-blurring artifacts. Furthermore, along with different computer graphics applications, the proposed technique also shows its robustness to incorporate spatial information for image classification.

## Data availability

All the data (images) utilized in this study are available in the public domain like <http://www.cse.cuhk.edu.hk/~leojia/projects/texturesep/index.html> and <http://cg.postech.ac.kr/publications>.

## Conflict of Interest:

The authors declare that they have no conflict of interest.

## References

- 1 Jiawen Chen, Sylvain Paris, and Frédéric Durand. Real-time edge-aware image processing with the bilateral grid. *ACM Transactions on Graphics (TOG)*, 26(3):103–es, 2007.
- 2 Durand Frédéric and Dorsey Julie. Fast bilateral filtering for the display of high-dynamic-range images. In *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH ’02*, page 257–266. Association for Computing Machinery, 2002. ISBN 1581135211.
- 3 Raanan Fattal, Maneesh Agrawala, and Szymon Rusinkiewicz. Multiscale shape and detail enhancement from multi-light image collections. *ACM Transactions on Graphics (TOG)*, 26(3):51, 2007.
- 4 Johannes Kopf, Michael F Cohen, Dani Lischinski, and Matt Uyttendaele. Joint bilateral upsampling. *ACM Transactions on Graphics (TOG)*, 26(3):96–es, 2007.
- 5 Shachar Fleishman, Iddo Drori, and Daniel Cohen-Or. Bilateral mesh denoising. In *ACM SIGGRAPH 2003 Papers*, pages 950–953. 2003.
- 6 Henry Kang, Seungyong Lee, and Charles K Chui. Flow-based image abstraction. *IEEE Transactions on Visualization and Computer Graphics*, 15(1):62–76, 2008.
- 7 Zoya Shahcheraghi, John See, and Alfian Abdul Halin. Image abstraction using anisotropic diffusion symmetric nearest neighbor filter. In *Pacific Rim Conference on Multimedia*, pages 343–352. Springer, 2014.
- 8 Levent Karacan, Erkut Erdem, and Aykut Erdem. Structure-preserving image smoothing via region covariances. *ACM Transactions on Graphics (TOG)*, 32(6):1–11, 2013.
- 9 Hojin Cho, Hyunjoon Lee, Henry Kang, and Seungyong Lee. Bilateral texture filtering. *ACM Transactions on Graphics (TOG)*, 33(4):1–8, 2014.
- 10 Jinjoo Song, Heeryon Cho, Jungho Yoon, and Sang Min Yoon. Structure adaptive total variation minimization-based image decomposition. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(9):2164–2176, 2017.
- 11 Li Xu, Jimmy Ren, Qiong Yan, Renjie Liao, and Jiaya Jia. Deep edge-aware filters. In *International Conference on Machine Learning*, pages 1669–1678. PMLR, 2015.
- 12 Guoyun Zhang, Jinping Wang, Xiaofei Zhang, Hongyan Fei, and Bing Tu. Adaptive total variation-based spectral-spatial feature extraction of hyperspectral image. *Journal of Visual Communication and Image Representation*, 56:150–159, 2018.
- 13 Qingxiang Yang. Semantic filtering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4517–4526, 06 2016.
- 14 Wei Ye and Kai Kuang Ma. Semantic image content filtering via edge-preserving scale-aware filter. In *IEEE International Conference on Image Processing (ICIP)*, pages 2443–2447, 2017.
- 15 Lei Zhu, Xiaowei Hu, Chi Wing Fu, Jing Qin, and Pheng Ann Heng. Saliency-aware texture smoothing. *IEEE Transactions on Visualization and Computer Graphics*, 26(7):2471–2484, 2018.
- 16 Kunal Pradhan and Swarnajyoti Patra. Structure preserving semantic texture filtering. In *2022 IEEE Calcutta Conference (CALCON)*, pages 283–287. IEEE, 2022.
- 17 Kunal Pradhan and Swarnajyoti Patra. Semantic-aware structure-preserving median morpho-filtering. *The Visual Computer*, 40(2):505–521, 2024.
- 18 Linggang Chan and Gang Fu. Structure-preserving image smoothing with semantic cues. *The Visual Computer*, 36:2017–2027, 2022.
- 19 Pietro Perona and Jitendra Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on pattern analysis and machine intelligence*, 12(7):629–639, 1990.
- 20 Jon P Ewins, Marcus D Waller, Martin White, and Paul F Lister. Implementing an anisotropic texture filter. *Computers & Graphics*, 24(2):253–267, 2000.
- 21 Carlo Tomasi and Roberto Manduchi. Bilateral filtering for gray and color images. In *6th International Conference on Computer Vision*, pages 839–846. IEEE, 1998.
- 22 Sanjay Ghosh, Ruturaj Gavaskar, Debasisha Panda, and Kunal N Chaudhury. Fast scale-adaptive bilateral texture smoothing. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(7):2015–2026, 2019.
- 23 Zeev Farbman, Raanan Fattal, Dani Lischinski, and Richard Szeliski. Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Transactions on Graphics (TOG)*, 27(3):1–10, 2008.
- 24 Wei Liu, Pingping Zhang, Xiaogang Chen, Chunhua Shen, Xiaolin Huang, and Jie Yang. Embedding bilateral filter in least squares for efficient edge-preserving image smoothing. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(1):23–35, 2018.
- 25 Michael Kass and Justin Solomon. Smoothed local histogram filters. *ACM Transactions on Graphics (TOG)*, 29, 07 2010.
- 26 Sylvain Paris, Pierre Kornprobst, Jack Tumblin, and Frédéric Durand. *Bilateral filtering: Theory and applications*. Now Publishers Inc, 2009.
- 27 Eduardo SL Gastal and Manuel M Oliveira. Domain transform for edge-aware image and video processing. *ACM Transactions on Graphics (TOG)*, 30(4):1–12, 2011.
- 28 Li Xu, Cewu Lu, Yi Xu, and Jiaya Jia. Image smoothing via  $L_0$  gradient minimization. *ACM Transactions on Graphics (TOG)*, 30(6):1–12, 2011.
- 29 Xuan Cheng, Ming Zeng, and Xinguo Liu. Feature-preserving filtering with  $l_0$  gradient minimization. *Computers & Graphics*, 38:150–157, 2014.
- 30 Mukhalad Al-nasrawi and Guang Deng. Modified iterative guided texture filtering algorithm. *Computers & Graphics*, 79:81–100, 2019.
- 31 Hongbin Jia, Qingbo Yin, and Mingyu Lu. Weighted guided image filtering with entropy evaluation weighting. *Computers & Graphics*, 117:114–123, 2023.
- 32 Wei-Chih Tu and Shao-Yi Chien. Two-way recursive filtering. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(11):4255–4268, 2021.
- 33 Kartic Subr, Cyril Soler, and Frédéric Durand. Edge-preserving multiscale image decomposition based on local extrema. *ACM Transactions on Graphics (TOG)*, 28(5):1–9, December 2009. ISSN 0730-0301.
- 34 Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1–4):259–268, 1992.

- 35 Li Xu, Qiong Yan, Yang Xia, and Jiaya Jia. Structure extraction from texture via relative total variation. *ACM Transactions on Graphics (TOG)*, 31(6):1–10, 2012.
- 36 Lei He, Yongfang Xie, Shiwen Xie, and Zhipeng Chen. Structure-preserving texture smoothing via scale-aware bilateral total variation. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(4):1493–1506, 2022.
- 37 Fei Zhou, Qun Chen, Bozhi Liu, and Guoping Qiu. Structure and texture-aware image decomposition via training a neural network. *IEEE Transactions on Image Processing*, 29:3458–3473, 2019.
- 38 Youngjung Kim, Bumsuk Ham, Minh N Do, and Kwanghoon Sohn. Structure-texture image decomposition using deep variational priors. *IEEE Transactions on Image Processing*, 28(6):2692–2704, 2018.
- 39 Kaiyue Lu, Shaodi You, and Nick Barnes. Deep texture and structure aware filtering network for image smoothing. In *Proceedings of the European conference on computer vision (ECCV)*, pages 217–233, 2018.
- 40 Yang Yang, Hongjun Hui, Lanling Zeng, Yan Zhao, Yongzhao Zhan, and Tao Yan. Edge-preserving image filtering based on soft clustering. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(7):4150–4162, 2021.
- 41 Wei Liu, Pingping Zhang, Yinjie Lei, Xiaolin Huang, Jie Yang, and Michael Ng. A generalized framework for edge-preserving and structure-preserving image smoothing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6631–6648, 2021.
- 42 Gopinath Qiu. An improved recursive median filtering scheme for image processing. *IEEE Transactions on Image Processing*, 5(4):646–648, 1996.
- 43 Hanli Zhao, Dandan Gao, Ming Wang, and Zhigeng Pan. Real-time edge-aware weighted median filtering on the gpu. *Computers & Graphics*, 61:11–18, 2016.
- 44 Junho Jeon, Hyunjoon Lee, Henry Kang, and Seungyong Lee. Scale-aware structure-preserving texture filtering. *Computer Graphics Forum*, 35(7):77–86, October 2016. ISSN 0167-7055.
- 45 Wei Liu, Pingping Zhang, Xiaolin Huang, Jie Yang, Chunhua Shen, and Ian Reid. Real-time image smoothing via iterative least squares. *ACM Transactions on Graphics (TOG)*, 39(3):1–24, 2020.
- 46 Panpan Xu and Wencheng Wang. Improved bilateral texture filtering with edge-aware measurement. *IEEE Transactions on Image Processing*, 27(7):3621–3630, July 2018.
- 47 Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.
- 48 Du-Yih Tsai, Eri Matsuyama, and Yongbum Lee. A mutual information-based image quality metric for medical imaging systems. In *Medical Imaging*. IntechOpen, 2011.
- 49 N Venkatanath, D Praneeth, Maruthi Chandrasekhar Bh, Sumohana S Channappayya, and Swarup S Medasani. Blind image quality evaluation using perception based features. In *2015 Twenty First National Conference on Communications (NCC)*, pages 1–6. IEEE, 2015.
- 50 Zeev Farbman, Raanan Fattal, Dani Lischinski, and Richard Szeliski. Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Transactions on Graphics (TOG)*, 27(3):1–10, 2008.
- 51 Mukhalad Al-nasrawi and Guang Deng. Modified iterative guided texture filtering algorithm. *Computers & Graphics*, 79:81–100, 2019.
- 52 Hongbin Jia, Qingbo Yin, and Mingyu Lu. Weighted guided image filtering with entropy evaluation weighting. *Computers & Graphics*, 117:114–123, 2023.
- 53 Pedram Ghamisi, Roberto Souza, Jon Atli Benediktsson, Leticia Rittner, Roberto Lotufo, and Xiao Xiang Zhu. Hyperspectral data classification using extended extinction profiles. *IEEE Geoscience and Remote Sensing Letters*, 13(11):1641–1645, 2016.
- 54 Leyuan Fang, Cheng Wang, Shutao Li, and Jón Atli Benediktsson. Hyperspectral image classification via multiple-feature-based adaptive sparse representation. *IEEE Transactions on Instrumentation and Measurement*, 66(7):1646–1657, 2017.
- 55 Pedram Ghamisi, Emmanuel Maggioli, Shutao Li, Roberto Souza, Yuliya Tarablaka, Gabriele Moser, Andrea De Giorgi, Leyuan Fang, Yushi Chen, Mingmin Chi, et al. New frontiers in spectral-spatial hyperspectral image classification: The latest advances based on mathematical morphology, markov random fields, segmentation, sparse representation, and deep learning. *IEEE geoscience and remote sensing magazine*, 6(3):10–43, 2018.
- 56 Yushi Chen, Lin Zhu, Pedram Ghamisi, Xiuping Jia, Guoyu Li, and Liang Tang. Hyperspectral images classification with gabor filtering and convolutional neural network. *IEEE Geoscience and Remote Sensing Letters*, 14(12):2355–2359, 2017.
- 57 Kaushal Bhardwaj, Swarnajyoti Patra, and Lorenzo Bruzzone. Threshold-free attribute profile for classification of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 57(10):7731–7742, 2019.
- 58 Xudong Kang, Xuanlin Xiang, Shutao Li, and Jon Atli Benediktsson. Pca-based edge-preserving features for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12):7140–7151, 2017.