Multimed Tools Appl (2016) 75:7909–7930
DOI 10.1007/s11042-015-2709-z

A generalized relative total variation method for image smoothing

Qiegen Liu[1] · **Biao Xiong**[2] · **Dingcheng Yang**[1] · **Minghui Zhang**[1]

Received: 30 September 2014 / Revised: 2 April 2015 / Accepted: 20 May 2015 /
Published online: 4 June 2015

Abstract Recently, two piecewise smooth models L0smoothing and relative total variation (RTV) have been proposed for feature/structure-preserving filtering. One is very efficient for tackling image with little texture patterns and the other has appearance performance on image with abundant uniform textural details. In this work, we present a general relative total variation (GRTV) method, which generalizes the advantages of both approaches. The efficiency of RTV depends on the defined windowed total variation (WTV) and windowed inherent variation (WIV), which focus on edge enhancing and texture suppressing respectively. The key innovations of the presented GRTV method are to extend the norm of WTV in RTV from 1 to [0, 1] and set the norm of WIV inversely proportional to the norm of WTV. These modifications substantially improve the structure extraction ability of RTV. The presented GRTV also improves the edge-boundary enhancing ability of L0smoothing and further enables it to deal with images containing complex textural details and noises. Furthermore, the L2-norm data fidelity term replaced by L1-norm is discussed. Experimental results demonstrate that the proposed method presents better performance as the state-of-the-art methods do.

Keywords Image smoothing · Structure preserving · non-convex regularization · Iterative Reweighed Least Square

✉ Qiegen Liu
liuqiegen@ncu.edu.cn

Biao Xiong
sailor080@126.com

Dingcheng Yang
ydcxuanyuan@msn.com

Minghui Zhang
zhangmh3529@163.com

[1] Department of Electronic Information Engineering, Nanchang University, Nanchang, People's Republic of China

[2] Faculty of Geo-Information Science and Earth Observation - ITC, University of Twente, Enschede, The Netherlands

1 Introduction

Many problems in computer vision and computer graphics involve estimating some spatially-varying image content from noisy raw data. One important requirement of such estimation is the feature-preserving filtering, which became a fundamental tool in many applications. Specifically, the feature-preserving filtering decompose a given image into structure and detail components by smoothing the image, simultaneously preserving or even enhancing image edges. Unfortunately, feature-preserving filtering is inherently challenging, because it is difficult to distinguish features from noise. There exist some different feature-preserving filtering methods, which can be roughly classified into two categories, i.e., spatial filters [6, 9, 13, 18, 25] and variational models [10, 23, 27, 28]. They differ from each other in how they define edges and how this prior information guides smoothing.

Anisotropic diffusion model [18] falls into the category of spatial filters. It employs a partial differential equation (PDE) based formulation in which pixel-wise spatially-varying diffusivities are estimated from image gradients. The gradient of the filtering image guides the diffusion process such that to avoid smoothing edges. Bilateral Filtering (BF) [9, 25] is another widely used model for removing noise from images while simultaneously performing detail flattening and edge preservation. It averages the nearby pixels by calculating weights from spatial and range domain [25], in favor of smoothing low-contrast regions while preserving high-contrast edges. Due to its simplicity and effectiveness, BF has been successfully applied to several computational photography applications [11, 26]. In [6], Buades et al. proposed a relatively simple technique to decompose an image into structure and oscillatory texture parts by using a nonlinear low pass-high pass filter pair. It is used to compute a local total variation of the image around a pixel and subsequently perform the decomposition. By considering the local linear model between the guidance image and the filtering output at a small neighbor window, a spatially guided filter (GF) was proposed in [13].

Variational models often follow the energy minimization framework comprising a data fitting term and a smooth penalty term [10, 23, 27, 28]. The data term measures the disagreement between the filtered signal s and the original signal f, while the smooth term measures the extent to which the filtered signal is not piecewise smooth. i.e.,

$$s = \arg\min_{s} \{G(f, s) + \lambda J(s)\} \tag{1}$$

where λ is a non-negative parameter controlling the weight of the smooth term. The design of the data term $G(f,s)$ is usually straightforward. For instance, the squared L2 distance between the filtered signal and the original signal $\|f-s\|_2^2$ is often used. The choice of the smooth term $J(s)$ is a critical issue. A representative work is the total variation [23], which uses L1-norm based regularization constraints to penalize large gradient magnitudes. In its original formulation, this model provides fairly good separations for structure from texture. Some studies extended the standard TV formulation with different norms for both regularization and data fidelity terms, and demonstrated that more robust norms could improve image decomposition. In [10], Farbman et al. proposed a robust method with the weighted least square (WLS) measure. The success of the WLS optimization attributes in part to the Lp norm in the Iterative Reweighed Least Square (IRLS) framework. Xu et al. [27] proposed an L0 smoothing method, using L0 term instead of L1 term, to directly measure the gradient sparsity in the context of image smoothing, and achieved some promising results. Later, they proposed a special method RTV [28] dealing with textured image. The efficiency of RTV depends on the defined

windowed total variation (WTV) and windowed inherent variation (WIV) involving spatial information. Recently, some improved algorithms for L0smmothing method were proposed, in ref. [24] Shen et al. presented a L0 smoothing-L1 method, where a L1-defility term instead of L2-norm was used to deal with large noise. In ref. [7], rather than using L0-L2 iteration algorithm to solve the model in ref. [27], Cheng et al. developed the L0smoothing_FCD (Fused coordinate descent) algorithm, which iteratively repeated coordinate descent step and fusion step to approximately solve the model.

Contribution Although most of the existing image smoothing models aim at extracting structure from noise with edge-preserving capabilities, they have their intrinsic strengths and short-comings. For instance, BF, WLS and L0smoothing work poor for tackling the image with non-uniform texture details, while RTV may smooth the edge when works on image containing number of cartoon patterns. In the present work, an effective general relative total variation (GRTV) method which absorbs the advantages of both spatial filters and variational models is proposed. The key innovations of our method are to extend the norm of WTV in RTV from 1 to [0, 1] and set the norm of WIV inversely proportional to the norm of WTV. These modifications substantially improve the structure extraction ability of RTV. The presented GRTV also improves the edge-boundary enhancing ability of L0smoothing and further enables it to deal with image containing complex textural details and noises. Furthermore, the L2-data fidelity term replaced by L1-data term is discussed in this paper. Experimental results demonstrate that the proposed method presents better performance as the state-of-the-art methods do.

The rest of this study is organized as follows. In Section 2, we briefly review the L0smoothing and RTV filter. In Section 3, a general model containing non-convex norm is proposed to alleviate the drawback of L0smoothing and RTV model. Numerical experiments will be provided in Section 4. Finally, concluding remarks and perspectives are sketched in Section 5.

2 Review of L0smoothing and RTV models

In this section, two image smoothing techniques L0smoothing and RTV, one focusing on the sparse and robust norm and the other emphasizing the spatially-varying total variation measure, will be briefly reviewed. The corresponding mathematical formulation and model will be constructed. Subsequently their strengths and drawbacks will be discussed in detail.

2.1 L0smoothing model

In ref. [27], Xu et al. proposed an image smoothing method via L0 gradient minimization. Assuming that ∇s denotes the gradients of s, the energy function of the L0smoothing model is defined as follows:

$$s = \arg\min_{s} \left\{ \frac{1}{2} \|f - s\|_2^2 + \lambda |\nabla s|_0 \right\} \tag{2}$$

where the data term is the squared L2 distance between f and s, the smooth term is the L0 norm of ∇s. The L0 norm of a vector is the number of non-zero value, which directly measures the

sparsity. Compared to the $l_p(0<p<1)$ norm regularization such as WLS filter [10], L0smoothing can remove low-amplitude structures and globally preserve and enhance salient edges, even if they are boundaries of very narrow objects. One main drawback introduced by L0smoothing is that, like other previously proposed edge-preserving smoothing approaches [10, 18, 23, 25], it mainly employs differences in the brightness values or gradient magnitudes as the main cues for edge indicator at an image pixel, and makes use of this information to guide the smoothing process. Therefore, the local contrast-based definition of edges might fail to capture high-frequency components that are related to fine image details or textures. It suggests that L0smoothing cannot fully separate textured regions from the main structures as they consider them as part of the structure to be retained during computations.

2.2 RTV model

In ref. [28], a structure extraction method from textured image was developed by employing relative total variation (RTV). For the purpose of properly removing texture, the proposed RTV model does not assume the type of textures in advance, instead it introduces a novel map of windowed inherent variation (WIV). The WIV map in a region that only contains texture is generally smaller than that in a region also including structural edges. Applying WIV map as a weight of TV model in the vertical and horizontal directions yields adaptive edge preservation and texture removal. Mathematically, the objective function can be expressed as:

$$s = \arg\min_{s} \left\{ \frac{1}{2}\|f-s\|_2^2 + \lambda \sum_i \left[\frac{D_x s(i)}{L_x s(i)+\varepsilon} + \frac{D_y s(i)}{L_y s(i)+\varepsilon} \right] \right\} \tag{3}$$

where $D_x s(i) = \sum_{j\in R(i)} g_{i,j}\left|(\partial_x s)_i\right|$. $L_x s(i) = \sum_{j\in R(i)} \left|g_{i,j}(\partial_x s)_i\right|$ are the weights and $g_{i,j}$ is a Gaussian weighting function with standard deviation σ controlling the spatial scale of the window. Both i and j index the 2D pixels, x and y are pixel coordinates. The second term in objective function (3) enforces the structure part to be sparse in the gradient domain, where ε is used to prevent numerical un-stability. RTV is modeled as variational formulation and its texture suppression weight is spatial vary dependent. However, although having good performance in dealing with highly textured images, it may overly smooth natural images by blurring major edge structures.

3 Generalized RTV (GRTV) method

As discussed in Section 2, the RTV has powerful ability to decompose image with irregular texture. However, it cannot separate the image component completely. Particularly, it may fail to enhance the edge and remove details simultaneously. On the other hand, the edge-aware filtering L0smoothing works perfect to sharpen or preserve the salient edges, although it fails to eliminate the large textures. Therefore, it is reasonable to design a model that inherits the strengths of the above approaches,

such as to achieve the purpose of jointly texture-smoothing and edge-preserving. One possible objective function we choose is expressed as follow:

$$s = \arg\min_{s} \left\{ \frac{1}{q} \|f-s\|_q^q + \lambda \sum_i \left[\frac{D_x^p s(i)}{L_x^{2-p} s(i) + \varepsilon} + \frac{D_y^p s(i)}{L_y^{2-p} s(i) + \varepsilon} \right] \right\} \tag{4}$$

Besides of the factor λ, the proposed model GRTV contains two another parameters: p ($0<p\leq 1$) promoting the sparsity of gradient and texture patterns, meanwhile q ($1\leq p\leq 2$) enforcing the degree of data consistency. Intuitively, on one hand, the operator D_y^p with $0<p<1$ has bigger edge preserving ability than that with $p=1$. On the other hand, L_x^{2-p} with $0<p<1$ will pose better texture-suppressing capability than that with $p=1$. In the following content, we will discuss the solver of this general model in the circumstance of $q=2$ and $q=1$ respectively. The relation and difference between the proposed model and the existing state-of-the-art algorithms will be revealed finally.

3.1 GRTV_L2: Solving model (4) with $q=2$

In the case of $q=2$, the objective function of Eq. (4) can be rewritten as:

$$s = \arg\min_{s} \left\{ \frac{1}{2} \|f-s\|_2^2 + \lambda \sum_i \left[\frac{D_x^p s(i)}{L_x^{2-p} s(i) + \varepsilon} + \frac{D_y^p s(i)}{L_y^{2-p} s(i) + \varepsilon} \right] \right\} \tag{5}$$

where $D_x^p s(i) = \sum_{j\in R(i)} g_{i,j} \left| (\partial_x s)_j \right|^p$ and $L_x^p s(i) = \sum_{j\in R(i)} \left| g_{i,j} (\partial_x s)_j \right|^p$ are the weights.

$$\begin{aligned} \sum_i \frac{D_x^p s(i)}{L_x^{2-p} s(i) + \varepsilon} &= \sum_i \sum_{j\in R(i)} \frac{g_{i,j}}{\sum_{j\in R(i)} \left| g_{i,j} (\partial_x s)_j \right|^{2-p} + \varepsilon} \left| (\partial_x s)_i \right|^p \\ &\approx \sum_i \sum_{j\in R(i)} \frac{g_{i,j}}{L_x^{2-p} s(i) + \varepsilon} \frac{1}{\left| (\partial_x s)_i \right|^{2-p} + \varepsilon} (\partial_x s)_i^2 \\ &= \sum_i u_{x,i}^{2-p} w_{x,i}^{2-p} (\partial_x s)_i^2 \end{aligned} \tag{6}$$

The second line in Eq. (6) is approximated by the Iteratively Reweighted Norm (IRN) approach proposed by Rodriguez et.al [21] and [20], which is closely related to the Iteratively Reweighted Least Squares (IRLS) method [12] and [19] and widely used in compressed sensing and image processing fields [15, 21]. It has been proven that IRN is one kind of Majorization-Minimization (MM) method [21], which involves good convergence property.

$$u_{x,i}^{2-p} = \sum_{j\in R(i)} \frac{g_{i,j}}{L_x^{2-p} s(i) + \varepsilon} = \left(G_\sigma * \frac{1}{\left| G_\sigma * \partial_x s \right|^{2-p} + \varepsilon} \right)_i \tag{7}$$

and

$$w_{x,i}^{2-p} = \frac{1}{\left| (\partial_x s)_i \right|^{2-p} + \varepsilon_s} \tag{8}$$

where G_σ is a Gaussian filter with standard deviation σ and ε_s is a small positive number to avoid division by zero. The division in Eq. (7) is point-wise and * is the convolution operator. Eq. (7) indicates that weight $u_{x,i}^{2-p}$ at pixel point i incorporates neighboring gradient information in an

isotropic spatial filter manner. Meanwhile, weight $w_{x,i}^{2-p}$ in Eq. (8) is only related to the pixel-wise gradient. The weights (7) and (8) pose the neighboring and the pixel constraint for the same point i. The penalty in the y-directional dimension is the same as that in x-directional dimension.

In summary, Eq. (5) can be written in a matrix form as follow:

$$v_s = \arg\min_{v_s} \left\{ \left(v_f - v_s\right)^T \left(v_f - v_s\right) + \lambda\left(v_s{}^T C_x{}^T U_x W_x C_x v_s + v_s{}^T C_y{}^T U_y W_y C_y v_s\right) \right. \quad (9)$$

where v_s and v_f are the vector representation of s and f respectively. C_x and C_y are the Toeplitz matrices from the discrete gradient operators with forward difference. U_x, W_x, U_y and W_y are diagonal matrices whose elements are defined in Eq. (7)(8). Similar to RTV, Eq. (5) is solved by a reweighted strategy, i.e., update the block variables U_x, W_x, U_y and W_y as a function of v_s calculated from the previous iteration and then update v_s by minimizing Eq. (9) with the last values of U_x, W_x, U_y and W_y. Practically speaking, using $L^k = C_x{}^T U_x^k W_x^k C_x + C_y{}^T U_y^k W_y^k C_y$, the minimization of Eq. (5) is given by

$$\left(1 + \lambda L^k\right) v_s^{k+1} = v_f \quad (10)$$

Since $(1+\lambda L^k)$ is a symmetric positive definite Laplacian matrix, efficient solvers are available for it. The whole optimization process is summarized as follows:

Algorithm GRTV_L2

1: For k=0 to K−1 do
2: update the weights according to Eqs. (7)(8)
3: update v_s^{k+1} according to Eq. (10)
4: End (For)

RTV is a special case of GRTV_L2 with p=1. The extended range with $0<p<1$ will substantially improve the geometry and texture separation capility. An intuitive demonstration is displayed in Fig. 1. As the parameter p decreases, the weights $u_{x,i}^{2-p}$ and $w_{x,i}^{2-p}$ of texture and edge differ more. Figure 1b and c correspond to the case of p=1 and p=0.5 respectively. It can be observed that both filters smooth small fluctuations and while preserving edges. Furthermore, as illustrated in Fig. 1d and e), the weights $u_{x,i}^{2-p}+u_{y,i}^{2-p}$ and $w_{x,i}^{2-p}+w_{y,i}^{2-p}$ with p=0.5 are more stiffness than those of with p=1.

We can consider the relationship of L0smoothing, RTV and GRTV_L2 in the globally and locally respects. L0smoothing uses L0 gradient minimization, which can globally control how many non-zero gradients are resulted in to approximate prominent structure in a sparsity-control manner. It does not depend on local features, but instead globally locates important edges. RTV mainly emphasizes on the texture-weights, which is locally defined. Our presented method GRTV_L2 falls in between, and inherits the advantages of both algorithms in the approximate strategy. From the viewpoint of parsimony-promoting measure, GRTV_L2 is closely related to L0smoothing method, it differs L0smoothing mainly from the additional weight $w_{x,i}^{2-p}$. One example depicted in Fig. 2 reveals the similarities and differences between their results. The test image shown in Fig. 2a contains textural grasses and slightly smooth clouds with small magnitude. The result of RTV in Fig. 2c cannot simultaneously maintain the lines on the girl's skirt and smooth the grasses. As expected, our GRTV_L2 with p=0.5 largely improves this simplification ability. For this test image, the behavior of GRTV_L2 is very near to that of L0smoothing by specifying λ=0.0004, σ=1. Seen from Fig. 2d, it can be found that the boundary of cloud obtained by our method is more rich and clear, consisting of piecewise constant intensities. Besides, by tuning parameters λ and σ, our method yields meaningful decomposition as depicted in Fig. 2e and f.

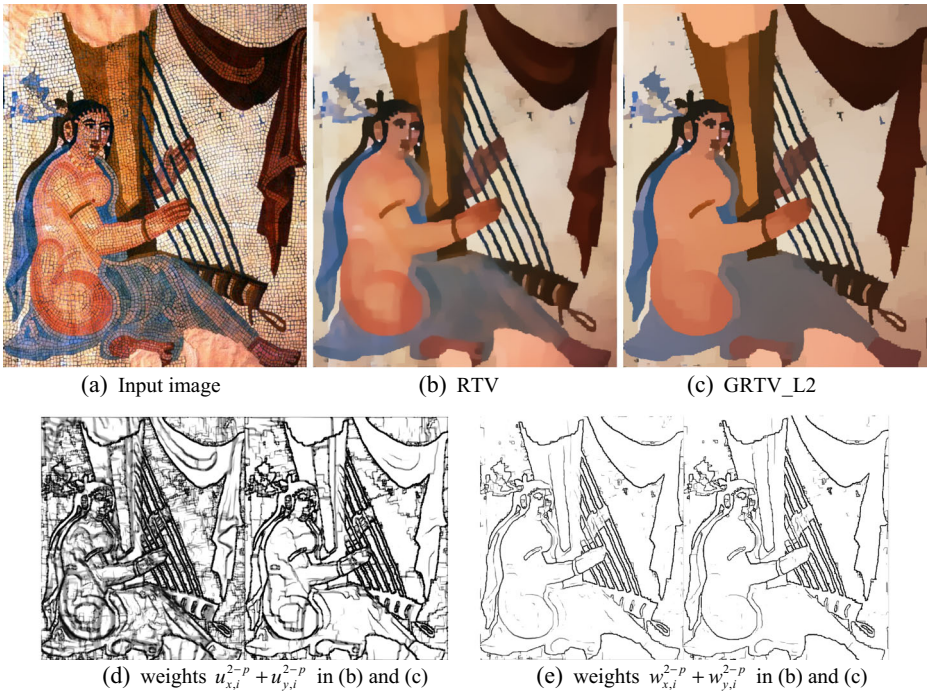


Fig. 1 Demonstration of the weights $u_{x,i}^{2-p}+u_{y,i}^{2-p}$ and $w_{x,i}^{2-p}+w_{y,i}^{2-p}$ vs p -value

3.2 GRTV_L1: Solving model (4) with q=1

In this subsection, the replacement of the L2-norm by the L1-norm as a fidelity measure w.r.t. the given image f will be considered. This modification may induce some important consequences. For example, the L1-norm prefers to better outlier in-sensivity and some

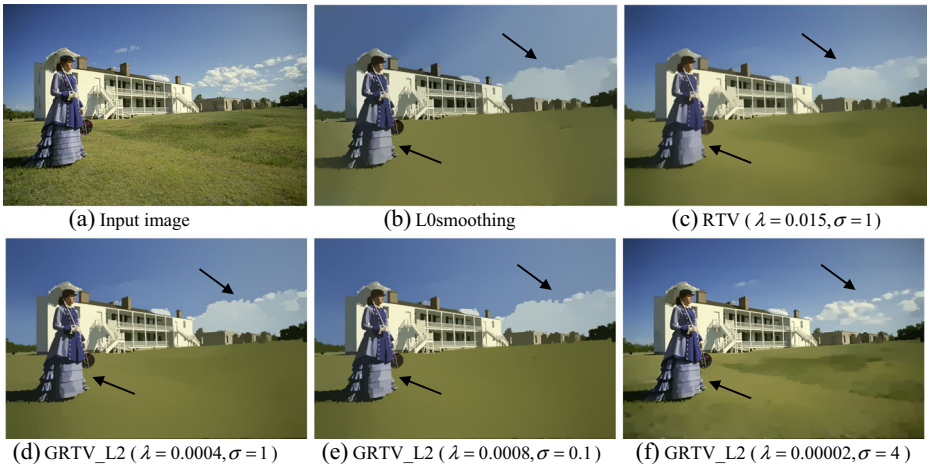


Fig. 2 Smoothing results produced by L0smoothing, RTV and GRTV

geometric properties, as stated in many references [1, 2, 16]. The objective function can be expressed as:

$$s = \arg\min_{s} \left\{ \|f - s\|_1 + \lambda \sum_i \left[\frac{D_x^p s(i)}{L_x^{2-p} s(i) + \varepsilon} + \frac{D_y^p s(i)}{L_y^{2-p} s(i) + \varepsilon} \right] \right\} \tag{11}$$

Similar to the technique used for dealing with windowed total variation (WTV), with the relation

$$\|f_i - s_i\|_1 \approx \frac{1}{\|f_i - s_i\|_1 + \varepsilon} \|f_i - s_i\|_2^2 \tag{12}$$

, Eq. (11) can be written in a matrix form as follows:

$$v_s = \arg\min_{v_s} \left\{ \left(v_f - v_s\right)^T R \left(v_f - v_s\right) + \lambda \left({v_s}^T {C_x}^T U_x W_x C_x v_s + {v_s}^T {C_y}^T U_y W_y C_y v_s \right) \right. \tag{13}$$

where R is a diagonal matrix, whose diagonal values are $R[i,i] = \frac{1}{\|f_i - s_i\|_1 + \varepsilon}$. Accordingly, the update of v_s can be achieved by

$$\left(R + \lambda L^k\right) v_s^{k+1} = v_f \tag{14}$$

Similar to the Eq. (10), $(R+\lambda L^k)$ is also a symmetric positive definite Laplacian matrix and the whole optimization process is summarized as follows:

Algorithm GRTV_L1

1: For k=0 to K−1 do
2: update the weights according to Eqs. (6)(7)
3: update v_s^{k+1} according to Eq. (12)
4: End (For)

Figure 3 shows the advantage of modifying L2 to L1 data-fidelity. As can be observed, both GRTV_L2 and GRTV_L1 with p=0.5 improve the visual perception of the input image than that of RTV. Besides, the introduction of L1 data-fidelity lowers the sensitivity of outliers. Furthermore, the L1-data fidelity term can be generalized to Lq-data term. The numerical modification is straightforward, as we have done in Eq. (6) and refs. [15, 21]. We do not discuss them due to the limit of paper length.

3.3 Iterative Reweighed Least Square (IRLS) framework and the computation cost

From the view of numerical optimization, both RTV and the developed GRTV fall into the category of IRLS framework. They are closely related to the WLS filter, which can be viewed as a single iteration of the IRLS framework. In WLS, the weight is the $\frac{1}{\left|(\partial_x \log(s))_i\right|^{2-p} + \varepsilon}$ ($0 \le p \le 0.8$), which is a modification of $\frac{1}{\left|(\partial_x s)_i\right|^{2-p} + \varepsilon}$. Although the convergence of the proposed GRTV may not be proved strictly as in ref. [21], our numerical experiments show good convergence behaviors. Figure 4 shows one example. It can be observed from Fig. 4a that the objective function values change slightly after 3 iterations. Figure 4b displays the intermediate structure images obtained at the 1-th, 2-th, and 4-th iteration, where we can find the proposed method quickly updates

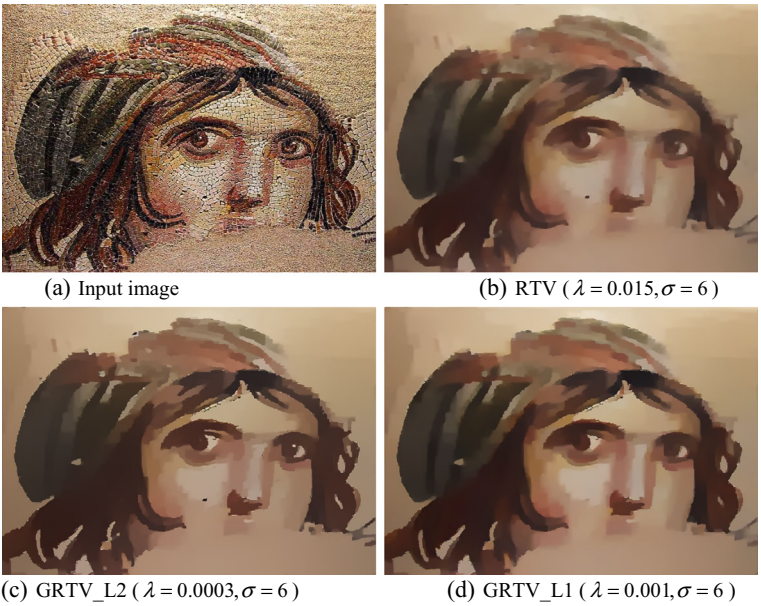


Fig. 3 Smoothing results on the Gypsy girl mosaic image

the saliency sharpened image in iterations. It indicates the effectiveness of the IRLS strategy adopted by our method.

The computation cost of GRTV is the same as that of RTV. i.e., At each iteration, the computational cost of (6) and (7) is linear with respect to problem size with $O(\sigma N)$ by taking advantage of the nice separation property of Gaussian kernel, where N is the total number of pixels in an image. As stated in ref. [28], the step of calculating v_x is to solve a linear system with a five-point spatially inhomogeneous sparse Laplacian matrix. Fast solvers such as the multi-resolution preconditioned conjugate gradient (PCG) can reach $O(N)$ complexity. In all the experiments of this article, it usually takes 3 s to process an 800x600 color image on an Intel Corei3CPU@2.93G with our matlab implementation.

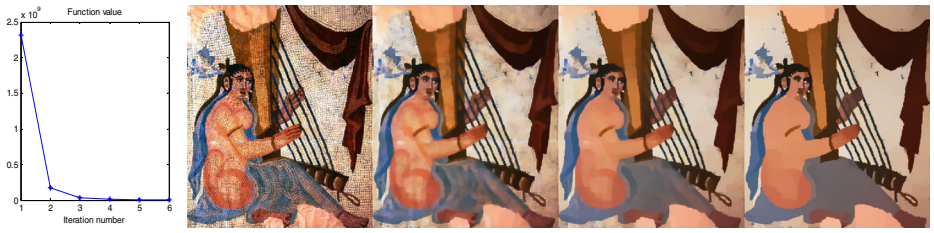


Fig. 4 The plot of function values vs iteration number, and the intermediate structure images at different iterations. **a** Objective value evolution. **b** The intimidate structure image obtained by the algorithm at the 1-th, 2-th, and 4-th iteration

4 Experimental results

In this section, at first we investigate the performance of the proposed algorithm by varying its parameter, then our method is compared with some existing edge-aware image smoothing methods, including BF [6], WLS [28] and L0smoothing [25]. The parameters of these methods are hand-tuned carefully to achieve the best performance. Finally, some potential applications are briefly exhibited.

4.1 Performance on varying values of p

Parameter p controls the degree of sparse norm. The effect of varying parameters is demonstrated in Fig. 5. Seen from Fig. 5, it can be found from the second row that GRTV_L2 changes the image more sharpen with smaller p -value. In the third row, the results of GRTV_L1 with different p -value also reflect this tendency, while they better preserve the shade and contrast information than these corresponding results obtained by GRTV_L2. By the way, although L0smoothing method shown in Fig. 5c uses the penalty with $p=0$ norm, its smoothing result may not provide impressive visual effect due to the lack of adaptive weighted fidelity term. According to the Figure, it is expected that the user can obtain the specific effect by tuning parameter p.

4.2 Comparison with state-of-the-art methods on cartoon images

Since L0smoothing targets globally preserving salient structures, even if they are small in resolution, it fails to remove the noise with large magnitude. One example is shown in Fig. 6.

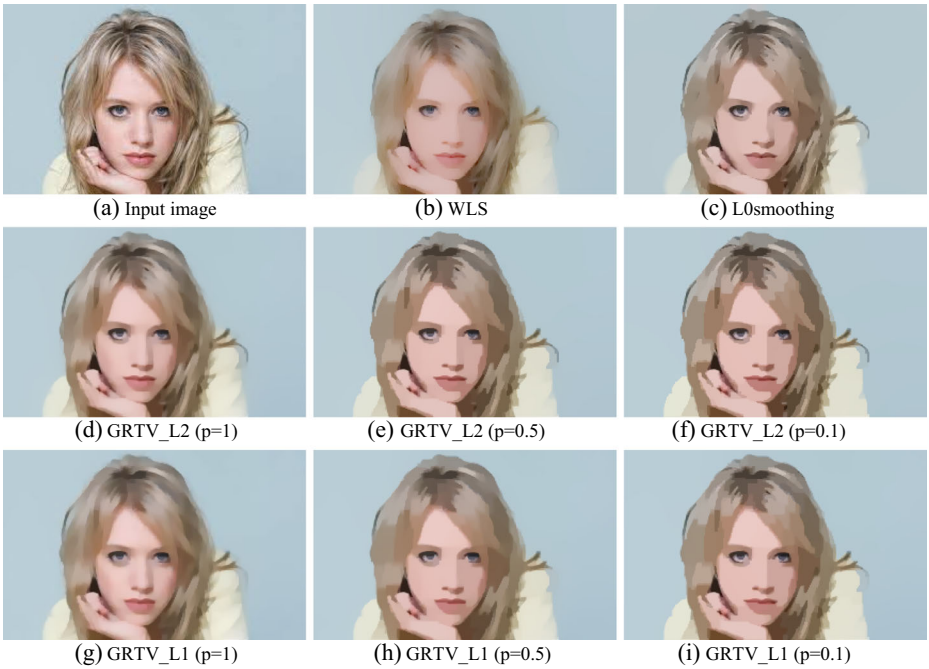


Fig. 5 Effect of varying parameters. Smaller p-values give more prominent edges sharpening results

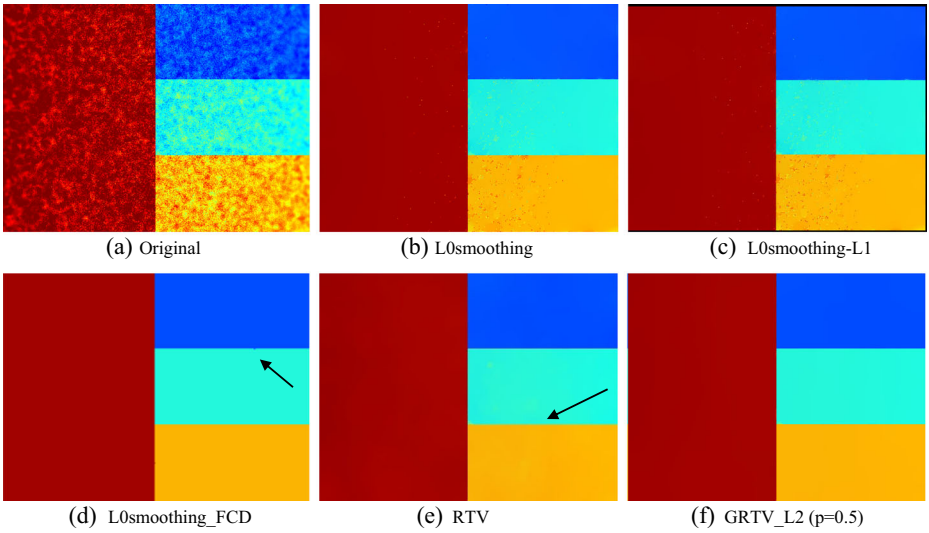


Fig. 6 Comparison with state-of-the-art methods on image created by Farbman et al

As can be observed in Fig. 6b, L0smoothing produces speckle-like artifacts. Moreover, although L0smoothing_FCD improved the noise insensitivity ability than that of L0smoothing-L1 and L0smoothing_FCD, it still retained some points near the edge, distorting the edge boundary. The result of RTV is slightly blurred. Our method performs the best to remove high frequency noise while preserving major edges.

The flower depicted in Fig. 7a contains smoothly branches and leaves and flower bud with small magnitude. The result of L0smoothing, RTV and L0smoothing_L1 cannot balance the detail smoothing and edge enhancing. Due to L0smoothing_FCD use the fuse technique such that a group of variables together rather than a single variable to be updated, it may cause over-fitting. On the other hand, our method obtains visually satisfied results; the GRTV method can better preserve the contrast around edges than the original RTV filter.

We applied four smoothing methods WLS, L0smoothing, RTV and GRTV to the image depicted in Fig. 8a. Figure 8 exhibits the difference among these methods clearly. In this example, as can be observed from Fig. 8b and c, WLS and L0smoothing algorithms cannot properly balance the conflicting goals of noise suppression and shape preservation. On the contrary, we can say that RTV and our GRTV framework with p=0.5 perform similarly. However, our method produces better result in the middle and lower right part in the cartoon image.

4.3 Comparison with state-of-the-art methods on textured images

In this subsection, some examples with ample non-uniform and anisotropic textures are shown to exhibit the flexibility of our method. It is worth noting that the TV model, BF and WLS filters were used in natural image smoothing and may not have effective terms to tackle textures. L0smoothing also has limitation in dealing with the "structure+texture" images, since it usually cannot suppress the texture component. Figure 9a shows the image "Bishapur_zan", in which the main structures are surrounded by the background formed by many tiles with

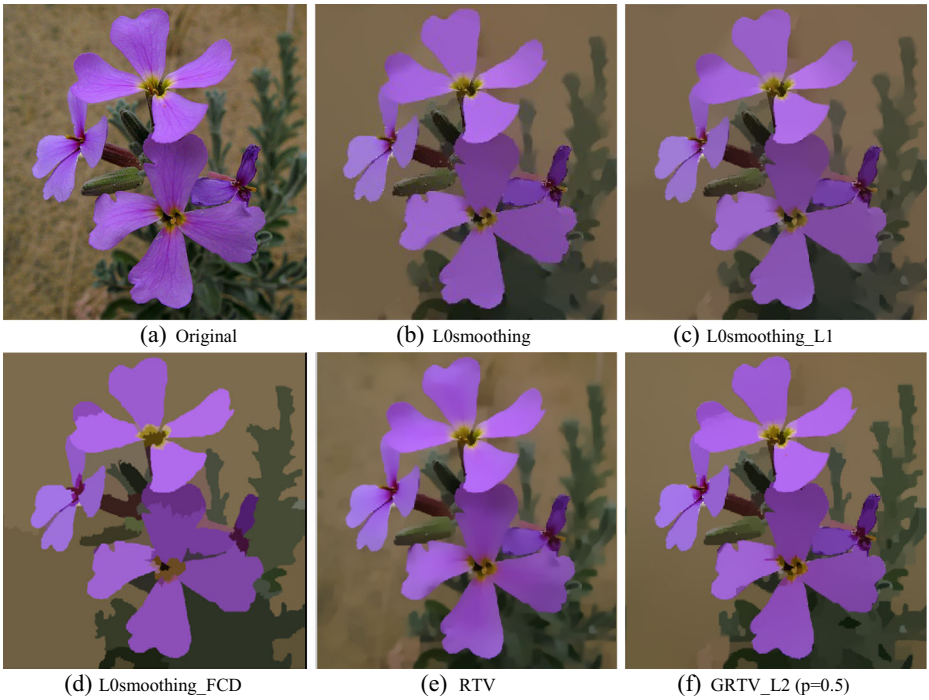


Fig. 7 Smoothing comparison with advanced method of L0smoothing, RTV and GRTV

salient but fine boundaries. From Fig. 9b and c, it can be observed that our method in the case of L1 data-fidelity can attain satisfied results. Besides, compared the results depicted in Fig. 1, it indicates that the L1-norm better preserves the contrast than the L2-norm. By the way, although tuning the regularization parameter as shown in Fig. 9d, e and f, RTV cannot well enhance the result meanwhile eliminating the outliers and irregular patterns.

The input image in Fig. 10 is an image contains persons with different shape. Since each people embedded with a different color, their texture measurements differ from each other much although they share the same background texture pattern, which makes the decomposition task very difficult. In Fig. 10b, the RTV result in the second people still has some artifacts. In Fig. 10c, GRTV_L2 with p=0.5 attains a more visual pleasure result, although it remains some outliers with strong white tone. The GRTV_L1 result with p=0.5 in Fig. 10d attains the best estimate.

Figure 11a shows an input image, where the texture is highly non-uniform and anisotropic. It can be observed in Fig. 11a, b, c and d) that these edge-difference guided approaches (e.g., BF, WLS and L0smoothing) do not have effective terms to deal with this texture-containing image. Results from RTV method is presented in Fig. 11e. Although RTV makes use of local signed gradients and the relative total variation (RTV) exhibits special properties, it fails to preserve some fine structures and blurs main edges compared to our method shown in Fig. 11f. Additionally, the decomposition of GRTV_L2 preserves the contrast between the girl and the background better. In summary, our edge-preserving texture-smoothing filtering method improves the visibility of input image than other algorithms. It can effectively achieve strong texture smoothing while keeping edges sharp. Some more comparisons with RTV on the same data from ref. [28] are shown in Fig. 12.

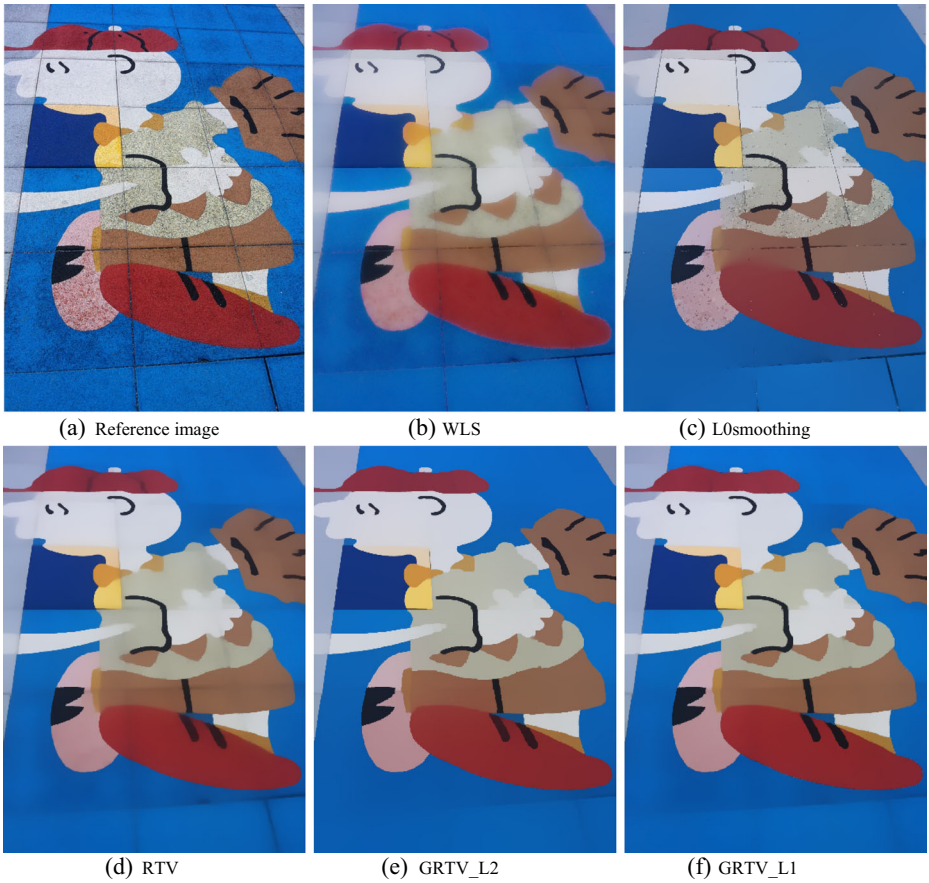


Fig. 8 Smoothing comparison with state-of-the-art methods

Our method can enhance the edge boundary by increasing steepness of transition, specifically as displayed in Figs. 2, 9 and 11. This property may in favor of the image segmentation such as the variational active contour/snake models [4, 17], there our developed weight can be incorporated to their weighted TV-norm for better segmentation.

As stated in ref. [27], although L0smoothing can not to efficiently tackle texture images, it can produce novel effects by firstly applying the local filtering BF and then the global filter L0smoothing. Figure 13 shows the example reproduced in ref. [27], there BF yields overblurrng result and L0smoothing remains fluff textures, combining them in consecutive way yields an image with sharpen prominent edges. In Fig. 13d RTV gives a slightly blurred image. Finally, we can observe that our GRTV_L2 result with p=0.5 in Fig. 13f sharpen the large-scale salient edges more than that of RTV.

4.4 Some applications

Edge extraction from images is usually the basic preprocessor to natural image editing [3] and high-level structure inference. High quality results that are continuous, accurate, and thin are generally very difficult to produce due to high susceptibility of edge

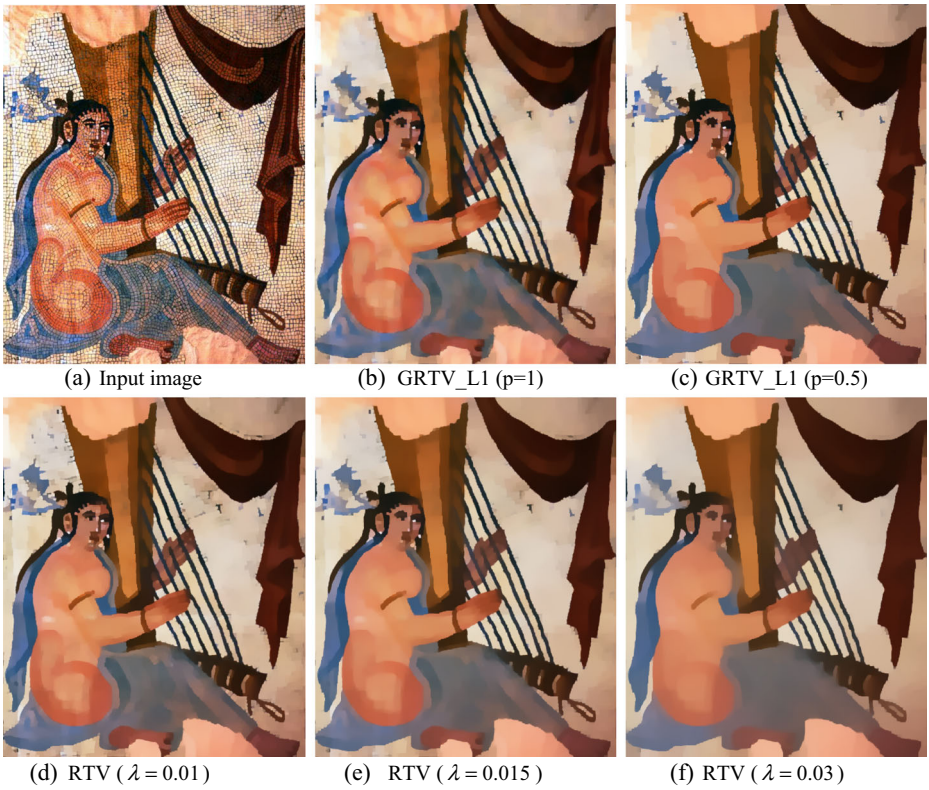


Fig. 9 Comparison with RTV on image "Bishapur_zan"

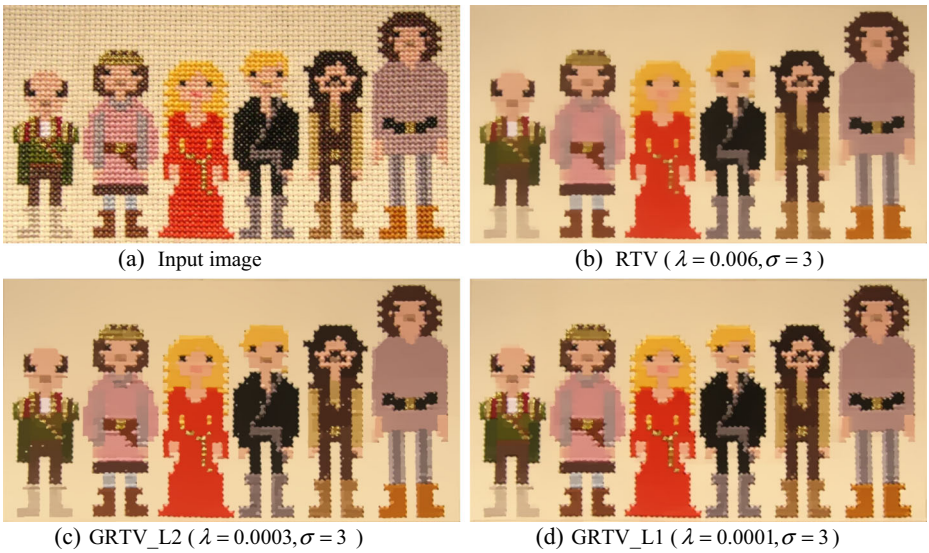


Fig. 10 Structure extraction result comparison with RTV and GRTV

Fig. 11 Structure extraction result comparison with state-of-the-art methods

detectors to complex structures and inevitable noise. Our method is able to suppress low-amplitude details, which remarkably stabilizes the extraction process. One example is shown in Fig. 14, the boundaries of the original ramp in Fig. 14a are not very sharp with overall small-magnitude gradients and it is difficult to distinguishable them from low-contrast details around. The results of L0smoothing and our method with p=0.3 are shown in Fig. 14b and c. Their corresponding gradient magnitude images are shown in (d)(e)(f) by linearly enhanced for visualization. As can be observed, our result is more faithful and the boundary is clearer.

We can see that our model, using a p-norm contrast function, gives us a better quality result because the edge function better preserves the geometry of the original features such as the corners and the largest disk. As can be seen from Fig. 15b and c), our methods GRTV_L2 with p=0.5 and p=0.1 lower the amplitudes of noise-like structures more than those of long coherent edges, meanwhile they globally sharpen prominent edges. Typically, result in Fig. 15c only contains large-scale salient edges, profiting main structure extraction and understanding. The GRTV_L1 results with p=0.5 and p=0.1 in Fig. 15e and f also display the similar estimates. The results by using our methods indicate that lower p-value has better capability on contrast- preserving.

Fig. 12 Structure extraction result comparison with state-of-the-art methods

When cartoon-like images such as the clip-art were heavily compressed it causes severe compression artifacts. In ref. [27], Xu et al. found that the general denoising

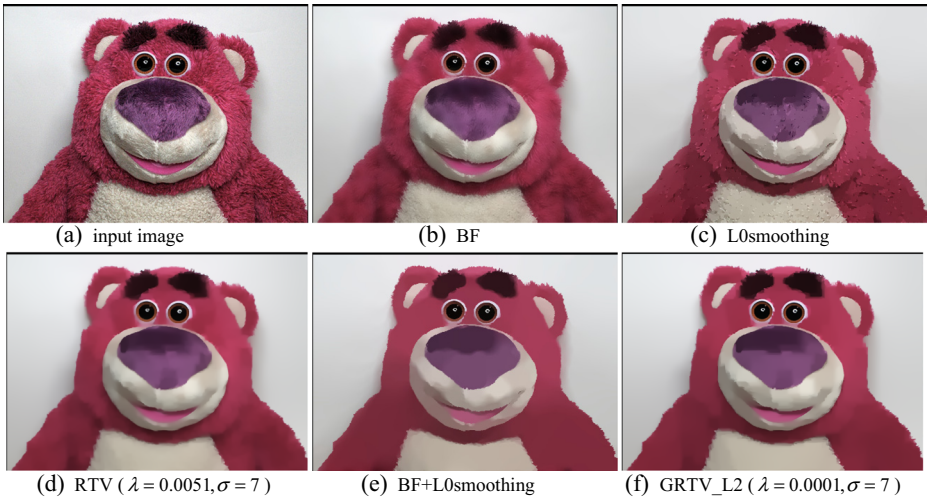


Fig. 13 Comparison with combined BF and L0smoothing

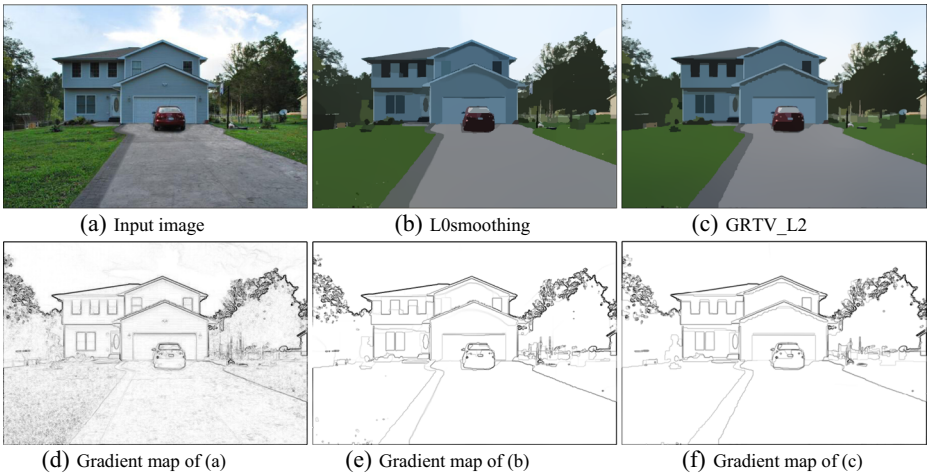


Fig. 14 Edge enhancement and extraction. Our result is more faithful and the boundary is clearer

approaches even with the state-of-the-art BM3D [8] cannot achieve satisfied output and their developed method L0smoothing gives appear performance. In fact, as can be expected, this artifact that strongly correlated with edges can be viewed as a special non-uniform texture and hence can be properly coped by our texture-smoothing and edge-preserving filters. One example is shown in Fig. 16, where the input image contains acute edges. As indicated, GRTV outperforms RTV mainly that it better lowers the amplitudes of noise-like structures on the animal's claws.

Mosaics images, paintings, and drawings sometimes cannot be directly used in image composition because the source and target textures are incompatible. Our

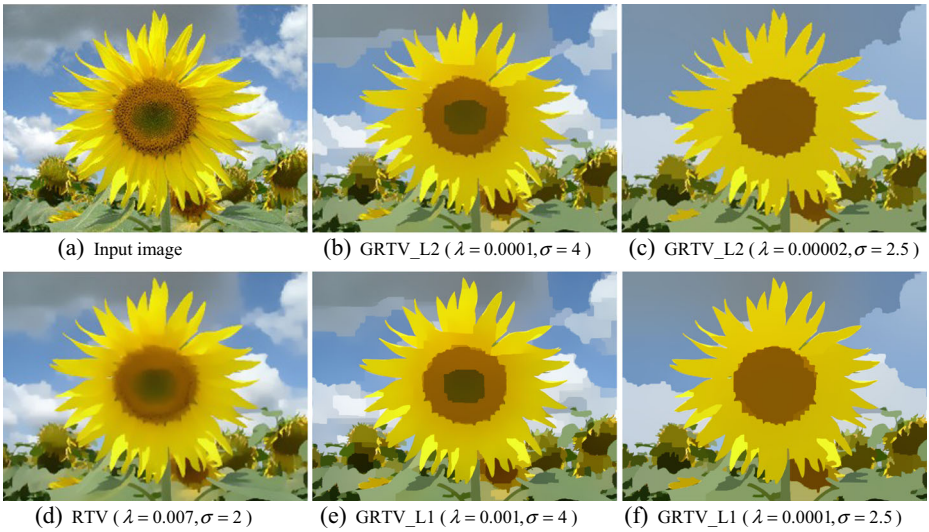


Fig. 15 Image simplification example

Fig. 16 Clip-art compression artifact removal example

proposed GRTV method can largely attenuate this shortcoming by first smoothing the original input with our structure-preserving filter and then performing composition between the resulting base layer with the another image. As shown in Fig. 17, the well smoothed image by GRTV enables the composed image in Fig. 17d to give more visual pleasure than that in Fig. 17c.

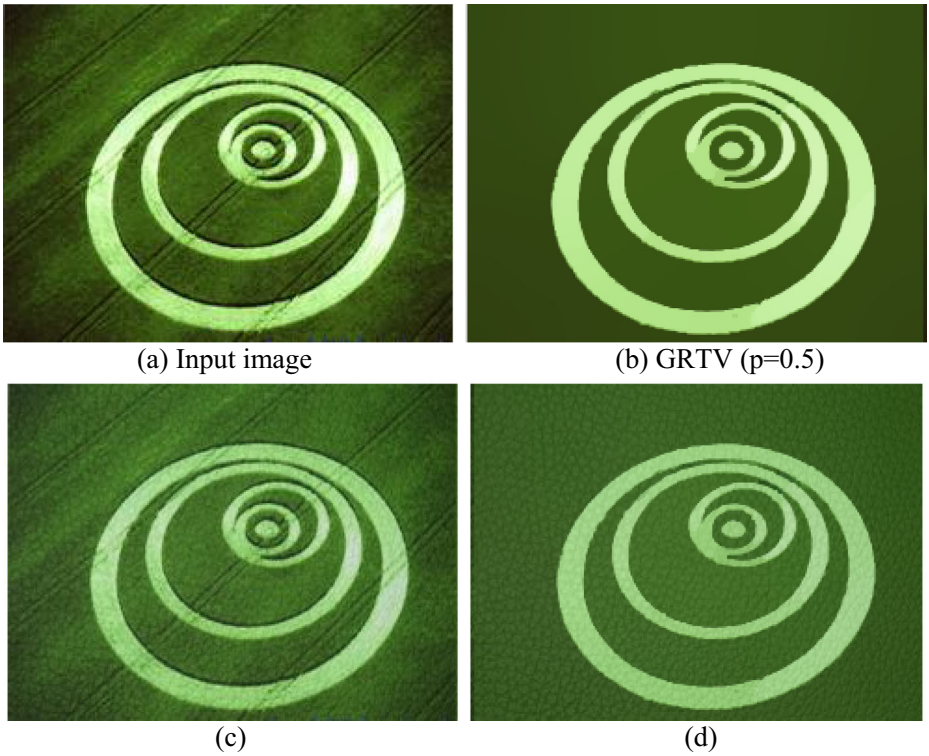


Fig. 17 Image composition example. **a** Input. **b** Structure image by GRTV_L2. **c d** Composed images obtained with the original and the structure images (**b**)

5 Conclusion and discussion

In this paper, a model can effectively eliminate texture meanwhile preserving/enhancing structure was proposed. The image smoothing results were improved by assigning with different norms for the weights, regularization and data fidelity terms. With the aid of Iteratively Reweighted Least Squares (IRLS) technique, the original non-linear problem was transformed to the alterative iterations of weights updating and the least square solver that are much easier to solve quickly. Two algorithms GRTV_L2 and GRTV_L1 were developed. Experimental results demonstrate the effectiveness and robustness of the proposed method on both cartoon-like images and texture images. Generally, empirical experiments show that GRTV_L1 better preserves the global contrast of the input image and lowering the sensitivity of outliers than GRTV_L2.

The proposed model can be applied to a large range of image patterns, and thus posses the advantages such as contrast-preserving, edge-preserving/sharpening and data-driven scale selection. Additionally, the proposed model allows a high level of randomness due to the nonlinear formulation. Compared to the L0smoothing that contains only one parameter and RTV that contains two parameters, a drawback of the proposed GRTV is that it possesses three parameters (i.e., λ, p and σ). Nevertheless, as demonstrated in subsection 4.1, GRTV makes the input image more sharpen with smaller p value. Additionally the role of λ and σ in GRTV is the same as that in RTV. Therefore, the user can obtain the specific effect by tuning these parameters.

Inspired by the fact that extending the concept of neighborhood in a non-local way to potentially include more pixels such that may in favor of smoothing [5], smoothing at non-local/semi-local regions has received a lot of attention and some schemes were proposed recently [14, 22, 29]. The nonlocal extension of total variation framework will be further considered in the forthcoming study.

Acknowledgments This work was partly supported by the National Natural Science Foundation of China under 61362001, 62162084, 61261010, 61365013, 51165033, the Science and Technology Department of Jiangxi Province of China under 20121BBE50023, 20132BAB211030, Young Scientists Training Program of Jiangxi Province under 20133ACB21007, 20142BCB23001, International Scientific Cooperation Project of Jiangxi Province under 20141BDH80001, Jiangxi Advanced Project for Post-doctoral Research Funds under 2014KY02 and Post-doctoral Research Funds under 2014 M551867.

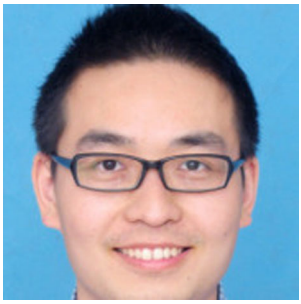
References

1. Alliney S (1992) Digital filters as absolute norm regularizers. IEEE Trans Signal Process 40(6):1548–1562
2. Alliney S (2004) A variational approach to remove outliers and impulse noise. J Math Imaging Vis 20(12):99–120
3. Bae S, Durand F (2007) Defocus magnification. Comput Graph Forum 26(3):571–579
4. Bresson X, Esedoglu S, Vandergheynst P, Thiran JP, Osher SJ (2007) Fast global minimization of the active contour/snake model. J Math Imaging Vis 28(2):151–167
5. Buades A, Coll B, Morel J-M (2005) A non-local algorithm for image denoising. CVPR 2:60–65
6. Buades A, Le TM, Morel J-M, Vese LA (2010) Fast cartoon + texture image filters. IEEE Trans Image Process 19(8):1978–1986

7. Cheng X, Zeng M, Liu X (2014) Feature-preserving filtering with L0 gradient minimization. Comput Graph 38:150–157
8. Dabov K, Foi A, Katkovnik V, Egiazarian KO (2007) Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Trans Image Process 16(8):2080–2095
9. Durand F, Dorsey J (2002) Fast bilateral filtering for the display of high-dynamic-range images. ACM Trans Graph 21(3):257–266
10. Farbman Z, Fattal R, Lischinski D, Szeliski R (2008) Edge-preserving decompositions for multi-scale tone and detail manipulation. ACM Trans Graph 27:3
11. Fattal R, Agrawala M, Rusinkiewicz S (2007) Multiscale shape and detail enhancement from multi-light image collections. ACM Trans Graph 26:3
12. Gorodnitsky B (1997) Rao, Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm. IEEE Trans Signal Process 45(3):600–616
13. He K, Sun J, Tang X (2010) Guided Image Filtering. In ECCV: 1–14
14. Karacan L, Erdem E, Erdem A (2013) Structure preserving image smoothing via region covariances. ACM Trans Graph 32(6)
15. Liu Q, Wang S, Luo J, Zhu Y, Ye M (2012) An augmented Lagrangian approach to general dictionary learning for image denoising. J Vis Commun Image Represent 23(5):753–766
16. Nikolova M (2002) Minimizers of cost-functions involving nonsmooth data-fidelity terms. SIAM J Numer Anal 40(3):965–994
17. Perazzi F et al. (2012) Saliency filters: Contrast based filtering for salient region detection. In CVPR
18. Perona P, Malik J (1990) Scale-space and edge detection using anisotropic diffusion. IEEE Trans Pattern Anal Mach Intell 12:629–639
19. Rao BD, Delgado KK (1999) An affine scaling methodology for best basis selection. IEEE Trans Signal Process 47(1):187–200
20. Rodrguez P, Wohlberg B (2008) An efficient algorithm for sparse representations with lp data fidelity term, in: Proceedings of 4th IEEE Andean Technical Conference (ANDESCON)
21. Rodrguez P, Wohlberg B (2009) Efficient minimization method for a generalized total variation functional. IEEE Trans Image Process 18(2):322–332
22. Rouselle F, Knaus C, Zwicker M (2012) Adaptive rendering with non-local means filtering, ACM Trans Graph 31(6).
23. Rudin L, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. Phys D 60(1–4):259–68
24. Shen C-T, Chang F-J, Hung Y-P, Pei S-C (2010) Edge-preserving image decomposition using L1 fidelity with L0 gradient, SIGGRAPH ASIA
25. Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In ICCV. 839–846
26. Winnem¨Oller H, Olsen SC, Gooch B (2006) Real-time video abstraction. ACM Trans Graph 25(3):1221–1226
27. Xu L, Lu C, Xu Y, Jia J (2011) Image smoothing via L0 gradient minimization. ACM Trans Graph 30(6);174:1–174:12. doi:10.1145/2024156.2024208
28. Xu L, Yan Q, Xia Y, Jia J (2012) Structure extraction from texture via relative total variation. ACM Trans Graph 31(6)
29. Ye C, Tao D, Song M, Jacobs DW (2013) Sparse norm filtering, M Wu - arXiv preprint arXiv:1305.3971, 2013 - arxiv.org.

Qiegen Liu Received the B.S. degree in Applied Mathematics from Gannan Normal College, M. Sc. degree in Computation Mathematics and Ph.D. degree in Biomedical Engineering from Shanghai Jiaotong University. Since 2012, he has been with School of Information Engineering, Nanchang University, Nanchang, P. R. China, where he is currently an Associate Professor. His current research interests include compressed sensing, image reconstruction, and pattern recognition.

Biao Xiong Received a Master of Science degree from state key laboratory of information engineering in surveying, mapping and remote sensing, Wuhan University, China in 2010. In 2014, he received a Ph.D degree from University of Twente, the Netherlands. His research interest is on lidar data processing, 3D modelling, (photogrammetric) computer vision, and computer graphics.



Dingcheng Yang Born in 1985. He received his Ph.D. degree from Wuhan University in 2012. Now, he is an associate professor in the Information Engineering School of Nanchang University. His research interests are signal processing, cooperation communications, cognitive radio techniques and wireless resource management.

Minghui Zhang Professor. He is with the Department of Electronic Information Engineering, Nanchang University, Jiangxi Province, China. He received the B.S. degree from Chongqing University, Chongqing, in 1990, majored in Biomedical engineering. Professor Zhang's research interest includes MRI reconstruction, image compression and restoration, pattern recognition, etc.