Perceptual Based Content Adaptive L_0 Smoothing

Fei Kou[1], Zhengguo Li[2], Changyun Wen[3], Weihai Chen[1],
Changchen Zhao[1], and Jianhua Wang[1]

[1] School of Automation Science and Electrical Engineering,
Beihang University, Beijing 100191, China
[2] Signal Processing Department, Institute for Infocomm Research,
Singapore 639798, Singapore
[3] School of Electrical and Electronic Engineering,
Nanyang Technological University, Singapore 138632, Singapore
koufei@yeah.net, ezgli@i2r.a-star.edu.sg, ecywen@ntu.edu.sg,
{whchenbuaa,jhwangbuaa}@126.com, zcc8843@gmail.com

Abstract. Edge preserving smoothing is a technique to decompose an image into two layers - one smoothing layer and one detail layer. It is an important image editing tool. The edges are preserved in the smoothing layer and details are decomposed into the detail layer. In this paper, we propose a content adaptive L_0 smoothing method. Unlike common smoothing schemes, we use a perceptual based content adaptive weighted fidelity term. The algorithm gives a larger weight to the region with more information, which is most likely edges, and gives a smaller weight to the region with less information, which is most likely a flat area. So the resulting smoothed image can preserve more edges and smooth the smoothing areas better. Experimental results prove that the proposed method can have better results than existing L_0 smoothing method.

Keywords: image smoothing, SSIM, L_0 sparsity, detail enhancement, perceptual, content adaptive.

1 Introduction

Edge preserving smoothing is an important image editing method, which has wide range of applications in image processing and computer vision. With an edge preserving smoothing scheme, an input image is decomposed into two layers: a base layer and a detail layer. The details are decomposed into the detail layer, while the edges are decomposed into the base layer.

Recently, a lot of edge preserving smoothing schemes has been proposed. The most widely used edge preserving smoothing scheme is bilateral filter [1] involving two filters, one domain filter and one range filter. The range filter can preserve the edges. In [2], Farbman et al. proposed a weighted least squares based edge-preserving decompositions scheme. In [9], He et al. proposed an edge preserving scheme named guided filter that filters the image by considering the content of a guidance image.

B. Huet et al. (Eds.): PCM 2013, LNCS 8294, pp. 299–307, 2013.

Xu et al. proposed an L_0 norm based smoothing scheme [3]. It was shown experimentally in [3] that an L_0 norm based scheme could get better results. Further they developed an improved L_0 norm based scheme which can smooth images with textured surfaces [4]. In [5], Shen et al. proposed an L_0 based smoothing scheme. Different from [3, 4], they used L_1 fidelity instead of L_2 fidelity which was used in [3,4]. In [6], Kou et al. proposed an L_0 smoothing based detail enhancement scheme for fusion of differently exposed images.

Structural similarity (SSIM) is a perceptual based metric for measuring the similarity between two images. Different from the most widely used measuring metric - mean squared error (MSE), SSIM can penalize errors in accordance with their visibility. In [7], Yeo et al. found the relationship between SSIM and MSE and used SSIM to replace MSE in the rate-distortion optimization in video encoding. As a result, they achieved better compression while maintaining the same perceptual quality. Inspired by their scheme, we replace MSE by SSIM in L_0 smoothing. Compared with the original L_0 smoothing optimization problem, our optimization problem gives larger weights to the regions with larger variances, which are most likely edges, while smaller weights to the regions with smaller variances, which are most likely flat areas. The new scheme can preserve more edges and smooth the smoothing areas better in the smoothed image. Experimental results prove that the proposed method can give better results than existing L_0 based smoothing method.

The remainder of this paper is organized as follows. In the next section, we introduce the optimization problem of our L_0 smoothing scheme. In Section 3, we present an approximation solver of the problem in Section 2. In Section 4 our experimental results are illustrated to verify the performance of our proposed schemes, and finally the paper is concluded in Section 5.

2 Using SSIM in L0 Smoothing

SSIM measures the similarity between two images by the following formula [8].

$$SSIM(m,n) = \Big(\frac{2\mu_m\mu_n + c_1}{{\mu_m}^2 + {\mu_n}^2 + c_1}\Big)\Big(\frac{2\sigma_{mn} + c_2}{{\sigma_m}^2 + {\sigma_n}^2 + c_2}\Big) \tag{1}$$

where m and n are two images or two image regions, μ_m and μ_n are the averages of m and n, σ_m^2 and σ_n^2 are the variances of m and n respectively, σ_{mn} is the covariance of m and n, $c_1 = (\kappa_1 L)^2$ and $c_2 = (\kappa_2 L)^2$ are two constants to stabilize the division with weak denominator, $\kappa_1 = 0.01$ and $\kappa_2 = 0.03$, L is the peak value of the two images. For a 8-bit image, L is 256.

In [7], it was concluded that

$$\begin{aligned} dSSIM &= \frac{1}{SSIM} \\ &\approx 1 + \frac{MSE}{2{\sigma_m}^2 + c_2} \end{aligned} \tag{2}$$

where dSSIM is the derivative of SSIM. It is observed from (2) that dSSIM is a weighted MSE, and the weight is $2{\sigma_m}^2 + c_2$, where σ_m is the variances of m.

The original L_0 smoothing optimization problem in [3] is

$$\min_S \left\{ \lambda \cdot C(S) + \sum_p (S_p - I_p)^2 \right\} \tag{3}$$

where λ is a smoothing parameter, which can adjust the weight of the two parts to control the smoothing. Starting from the element at position $(1, 1)$ of a matrix, we can count the elements of the matrix in an order from left to right and then top to bottom. The subscript p denotes the p-th element of the matrix from such counting. S_p is the p-th pixel in image S, I_p is the p-th pixel in image I. $C(S)$ is the L_0 norm of the gradient of S, which is expressed as

$$C(S) = \#\{p \big| \, |\partial_x S_p| + |\partial_y S_p| \neq 0\} \tag{4}$$

where # stands for the number of p which satisfies $|\partial_x S_p| + |\partial_y S_p| \neq 0$, that is the L_0 norm.

A smaller value of the first term in (3) makes the resulting image S smoother, while a smaller value of the second term makes the image S more similar to the input image I. λ is a smoothing parameter to control the weight of the two terms.

We give different weights to different pixels based on their perceptual importance. Then the optimization problem in (3) is re-written as

$$\min_S \left\{ \lambda \cdot C(S) + \sum_p W_p (S_p - I_p)^2 \right\} \tag{5}$$

where W_p is the weight of pixel p.

$$\begin{aligned} W_p &= \frac{\frac{1}{N} \sum\limits_{p'} \frac{1}{2\sigma^2(p') + c_2}}{\frac{1}{2\sigma^2(p) + c_2}} \\ &= \frac{1}{N} \sum_{p'} \frac{2\sigma^2(p) + c_2}{2\sigma^2(p') + c_2} \end{aligned} \tag{6}$$

where N is the number of pixels in the image. $\sigma^2(p)$ is the variance of the pixels in the neighborhood of the p-th pixel. p' stands for all the p in the image. After numerous experimental test, we use a 9×9 neighborhood in this paper.

3 Solver

Note that there is a discrete counting metric $C(S)$ in (5) and the optimization problem is not convex. So it is very difficult to solve it. As in [3], we change the optimization problem by minimizing the following problem to get an approximation solver

$$\min_{S,h,v}\left\{\lambda \cdot C(h,v) + \sum_p \left\{W_p(S_p - I_p)^2 + \beta\big((\partial_x S_p - h_p)^2 + (\partial_y S_p - v_p)^2\big)\right\}\right\} \tag{7}$$

where $C(h,v) = \#\{p \big| |h_p| + |v_p| \neq 0\}$, β is a parameter controlling the similarity between h, v and $\partial_x S, \partial_y S$. h, v are auxiliary matrices to solve S. h_p and v_p are the p-th element of h and v respectively. When β is large enough, the optimization problem (7) is equivalent to Eq. (5). So we can get the solution of Eq. (5) by solving (7).

Similar to [3], the problem (7) is solved through alternatively minimizing (h, v) and S. In each pass, one set of the variables is fixed with values obtained from the previous iteration. β is set as a small value β_0 at the beginning, and it is multiplied by a constant κ each time. The process ends when β is larger than β_{max}. The details are given as below.

Computing S when h and v are known: The S estimation subproblem corresponds to minimizing

$$\min_{S}\left\{\sum_p \left\{W_p(S_p - I_p)^2 + \beta\big((\partial_x S_p - h_p)^2 + (\partial_y S_p - v_p)^2\big)\right\}\right\} \tag{8}$$

Following similar analysis to [3], the global minimum of (8) is obtained. To accelerate computational speed, we diagonalize the derivative operator after Fast Fourier Transform (FFT) and this gives the solution:

$$S = \mathscr{F}^{-1}\left(\frac{W \cdot \mathscr{F}(I) + \beta(\mathscr{F}(\partial_x)^* \mathscr{F}(h) + \mathscr{F}(\partial_y)^* \mathscr{F}(v))}{W + \beta(\mathscr{F}(\partial_x)^* \mathscr{F}(\partial_x) + \mathscr{F}(\partial_y)^* \mathscr{F}(\partial_y)}\right) \tag{9}$$

where $\mathscr{F}$ is the FFT operator, $\mathscr{F}^{-1}$ is the IFFT operator and * denotes the complex conjugate.

Computing (h, v) when S is known: The objective function for (h, v) is

$$\min_{h,v}\left\{\frac{\lambda}{\beta} \cdot C(h,v) + \sum_p \left\{(\partial_x S_p - h_p)^2 + (\partial_y S_p - v_p)^2\right\}\right\} \tag{10}$$

Similar to [3], the solution is given by

$$(h_p, v_p) = \begin{cases} (0,0) & (\partial_x S_p)^2 + (\partial_y S_p)^2 \leq \dfrac{\lambda}{\beta} \\ (\partial_x S_p, \partial_y S_p) & otherwise \end{cases} \tag{11}$$

By estimating S with equation (9) and h, v with equation (11) alternatively, a smooth image S can be extracted.

4 Experimental Results

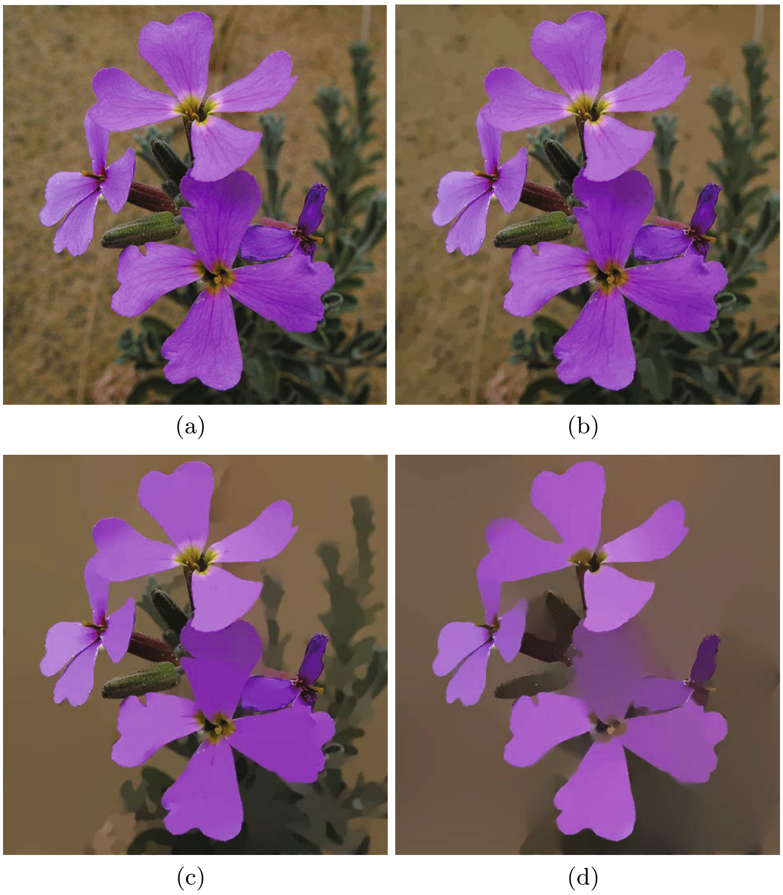


Fig. 1. Different selections of λ. Input image courtesy of Li Xu. (a) Input image. (b) Smoothing image obtained by $\lambda = 0.001$ (c) Smoothing image obtained by $\lambda = 0.01$. (d) Smoothing image obtained by $\lambda = 0.05$.

Readers are invited to read the electronic version with full-size figures in order to better appreciate the differences among images. In this section, we first evaluate the choice of λ in (7). As shown in Fig. 1, the smaller λ is, the more similar to the input image the smooth image is, the larger λ is, the less information is decomposed into the smooth layer. This is because if λ is large in (7), the first term will affect more on the result, which results in less none-zero gradients, the result image will be smoother. If λ is small, the second term will effort more on the result and this renders the result image to be more similar to the input image.

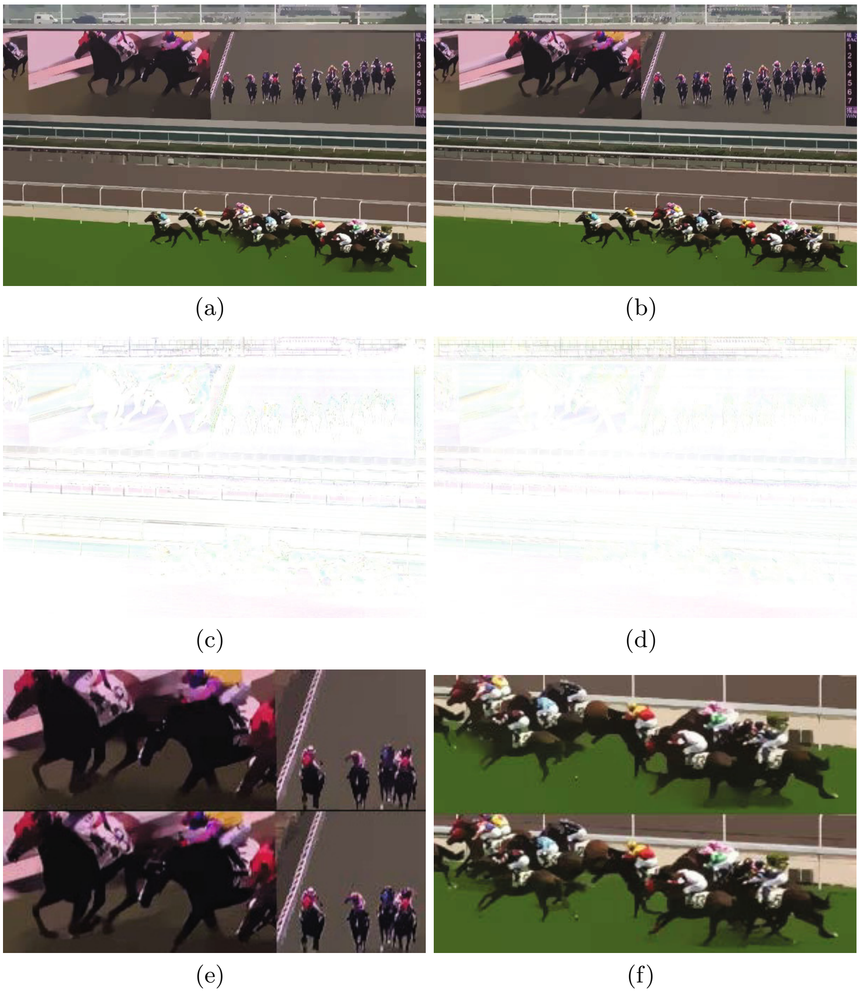
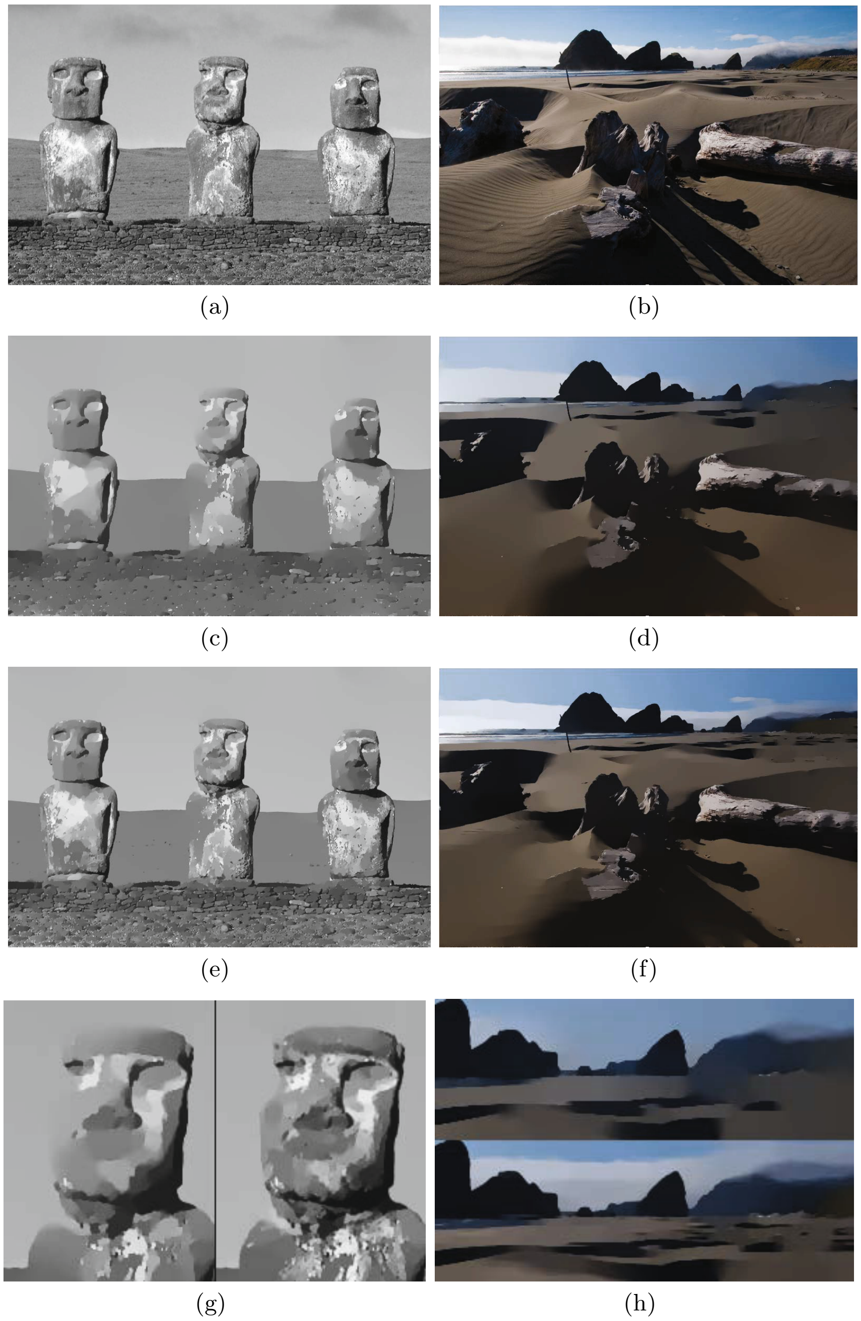


Fig. 2. Comparison of the decomposed results. To better appreciate the difference, the negative images of the detail images are given. Input image courtesy of Li Xu. (a) Smoothing layer obtained by the scheme in [3]. (b) Smoothing layer obtained by the proposed scheme. (c) Detail layer obtained by the scheme in [3]. (d) Detail layer obtained by the proposed scheme. (e) Closeups of (a) and (b). (f) Closeups of (a) and (b).

Then we compare the decomposition result by our scheme with scheme in [3] by choosing the same parameters for both schemes. From Fig. 2, it is seen that with the proposed method, the edges are more clear in the smooth layer than the result in [3] while there are less edges in the detail layer by the proposed method. For example, the horse legs are decomposed into the detail layer by the scheme in [3], while into the smooth layer with the proposed scheme. The horse legs are more clear in the smooth layer based on our scheme. All these can be



(a) (b)

(c) (d)

(e) (f)

(g) (h)

Fig. 3. Comparison of smoothing results. (a) First input image from [2]. (b) Second input image, image courtesy of Norman Koren. (c) Result image obtained by the scheme in [3]. (d) Result image obtained by the scheme in [3]. (e) Result image by the proposed scheme. (f) Result image obtained by the proposed scheme. (g) Closeups of (c) and (e). (h) Closeups of (d) and (f).

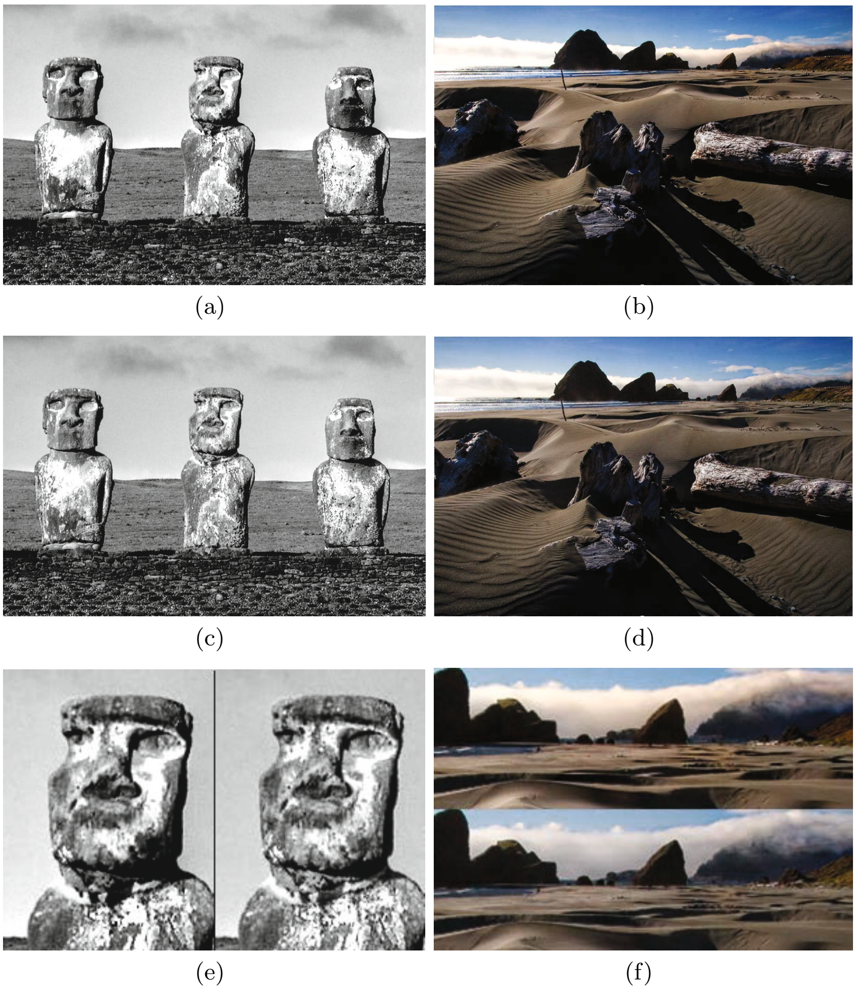


Fig. 4. Comparison of detail enhanced results. (a) Result image obtained by the scheme in [3]. (b) Result image obtained by the scheme in [3]. (c) Result image obtained by the proposed scheme. (d) Result image obtained by the proposed scheme. (e) Closeups of (a) and (c). (f) Closeups of (b) and (d).

easily observed in the closeups of Fig. 2(a) and Fig. 2(b) as shown in Fig. 2(e) and Fig. 2(f). From these results, we can conclude that our scheme can preserve edges in the smooth layer better than the existing L_0 based smooth scheme in [3]. Additional two sets of images are also tested and presented in Fig. 3.

Finally, we compare the detail enhance results by our scheme and the scheme in [3]. The detail enhanced images are generated by adding the detail layers to the source images. The original L_0 smoothing scheme in [3] may decompose edges into detail layer. As a result, the result image may be over sharpened around edges. As shown in Fig. 4(a) and 4(c) which are generated by the scheme in [3], several regions are much darker or whiter than those in Fig. 4(b) and 4(d) which are generated by the proposed scheme. A lot of details are missing in Fig. 4(a) and 4(c) in these regions. Based on all the experimental results illustrated, our scheme can give better results.

5 Conclusion

In this paper, a content adaptive L_0 norm based smoothing scheme has been proposed by using a content adaptive weighted fidelity term with larger weights given to the edges and smaller weights to the flat areas. As a result, the resulting image can preserve more edges and smooth the smoothing areas better than existing L_0 norm based smoothing scheme.

Acknowledgment. This work has been supported by National Nature Science Foundation of China under the research project 61075075 and 61175108.

References

1. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 839–846 (1998)
2. Farbman, Z., Fattal, R., Lischinshi, D., Szeliski, R.: Edge-preserving decompositions for multi-scale tone and details manipulation. ACM Transactions on Graphics 27(3), 249–256 (2008)
3. Xu, L., Lu, C., Xu, Y., Jia, J.: Image smoothing via L_0 gradient minimization. ACM Transactions on Graphics (SIGGRAPH Asia 2011) 30(6) (2011)
4. Xu, L., Yan, Q., Xia, Y., Jia, J.: Structure Extraction from Texture via Relative Total Variation. ACM Transactions on Graphics (TOG) 31(6), 139 (2012); Proc. ACM SIGGRAPH ASIA 2012 (2012)
5. Shen, C.T., Chang, F.J., Hung, Y.P., Pei, S.C.: Edge-preserving image decomposition using L1 fidelity with L0 gradient. In: SIGGRAPH Asia 2012 Technical Briefs, p. 6 (2012)
6. Kou, F., Li, Z., Wen, C., Chen, W.: L0 Smoothing Based Detail Enhancement for Fusion of Differently Exposed Images. In: in 8th IEEE Conference on Industrial Electronics and Applications (ICIEA 2013), pp. 1398–1403 (2013)
7. Yeo, C., Tan, H.L., Tan, Y.H.: On rate distortion optimization using SSIM. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2012), pp. 833–836 (2012)
8. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing 13, 600–612 (2004)
9. He, K., Sun, J., Tang, X.: Guided image filtering. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part I. LNCS, vol. 6311, pp. 1–14. Springer, Heidelberg (2010)