

Spatially guided local Laplacian filter for nature image detail enhancement

Shijie Hao · Meng Wang · Richang Hong · Jianguo Jiang

Received: 27 November 2013 / Revised: 20 March 2014 / Accepted: 27 April 2014 /

Published online: 7 May 2014

© Springer Science+Business Media New York 2014

Abstract Nature images make up a significant proportion of the ever growing volume of social media. In this context, automatic and rapid image enhancement is always among the favorable techniques for photographers. Among the image representation models, the Gaussian and Laplacian image pyramids based on isotropic Gaussian kernels were once considered to be inappropriate for image enhancement tasks. The recently proposed Local Laplacian Filter (LLF) updates this view by designing a point-wise intensity remapping process. However, this model filters an image with a consistent strength instead of a dynamical way which takes image contents into account. In this paper, we propose a spatially guided LLF by extending the single-value key parameter into a multi-value matrix that dynamically assigns filtering strengths according to image contents. Since it is still very challenging to recognize arbitrary image contents with machine learning methods, we propose a simple but effective technique, which only approximates the richness of image details instead of specific contents. This trade-off between concrete semantics and algorithm efficiency enables filtering strengths to be spatially guided in the LLF process with little extra computational cost. Experimental results validate our method in terms of visual effects and a conditionally faster LLF implementation.

Keywords Image enhancement · Image content · Local Laplacian filter · Distance transform

S. Hao · M. Wang · R. Hong · J. Jiang

School of Computer and Information, Hefei University of Technology, Hefei, China

M. Wang

e-mail: eric.mengwang@gmail.com

R. Hong

e-mail: hongrc@hfut.edu.cn

J. Jiang

e-mail: jgjiang@hfut.edu.cn

S. Hao (✉)

Computer Science and Technology Postdoctoral Research Station,

Hefei University of Technology, Hefei, China

e-mail: hfut.hsj@gmail.com

1 Introduction

With the rapid development of Web 2.0 technologies and imaging devices, there are more and more personal images available on the Internet. There would be hundreds or even thousands of new photographs uploaded to social network websites just after a single personal travel. Since most people are amateur photographers, it would be attractive for them to develop techniques which are able to effectively enhance the quality of their works. For example, before taking landscape pictures, a smart imaging system automatically suggests the optimal scope of the visual field [14]. After taking pictures, some post-processing technique enhances image details or generates special effects [15]. In the process of sharing pictures onto social networks, an intelligent system helps people to select best photos from all their works and provide intuitive knowledge on photograph shooting [12]. In this paper, we primarily concern image post-processing techniques such as detail enhancement.

As for detail enhancement, various image processing algorithms [1, 2, 9, 16, 17] have been designed to achieve this goal such as anisotropic filtering, bilateral filtering and their extensions. Generally, these methods consider the anisotropic nature of image edges and bring about edge-preserving results. However, they often have technical problems such as parameter tuning, edge halos, and optimization. Different from the above-mentioned methods, Paris et al. [15] propose the Local Laplacian Filter (LLF) and obtain promising results in editing image details or tones. The LLF method distinguishes itself from other methods in the following aspects. First, the backbones of LLF, i.e. Gaussian and Laplacian image pyramids, are based on isotropic Gaussian kernels, which are opposed to the anisotropic nature of image edges and were thought to be inappropriate for constructing edge-preserving methods. In this sense, the LLF method makes contributions in enlarging the role of Laplacian image pyramid in image processing. Second, the LLF can be simply implemented with basic manipulations, e.g. convolution, resizing and linear remapping, in a non-iterative style.

With in-depth exploration of LLF, we can further observe the following issues about the method. First, the original LLF process is pixel-wise and each pixel manipulation involves a whole-image remapping sub-process. Obviously, the computational cost would be very large (In the following sections, detailed analysis on the complexity would be given). Second, the remapping process is solely determined by pixel intensities, which totally ignores the spatial information of pixels and thereof misses the spatial information of image contents. As for the first issue, Paris et al. have proposed a fast version of LLF in their seminal research [15], where each remapping is only conducted in a neighboring region. As for the second issue, it is potentially helpful to borrow ideas from the fields such as image understanding or content analysis.

In terms of image understanding, an image can be usually viewed as a spatial combination of several semantic regions. For example, a scenery photograph often contains buildings, sky, mountains, to name but a few. Ideally, our goal is to design a spatially guided remapping function for the LLF model according to these semantic regions. Then we can dynamically assign parameters to different image regions to achieve user-preferred effects. However, this “classification and filtering” roadmap seems impossible due to some technical difficulties. As the semantic gap always exists between low-level features and image contents [20], it is still a very challenging task to have an image semantically partitioned. To bridge this gap, various learning methods have been developed and achieved state-of-the-art results in image/video

tagging [19, 22], index [6, 21], and 2D/3D object detection and searching [4, 5, 7, 11]. However, these methods are not suitable for our goal due to the following reasons. On one hand, in terms of model training, as negative samples come from an infinite semantic space, the trained model may be still unreliable to an ad hoc test image. On the other hand, from the practical view, integrating these models may bring too much extra computational cost, especially for the cases where images are expected to be post-processed in real time on imaging devices.

One possible solution to the afore-mentioned problem could be a trade-off between the efficiency of improved LLF and the concrete semantics of image contents. This strategy is supported by the common observation that the image contents can be divided into only two kinds, i.e. regions with rich image details and regions with few details. Taking landscape images for example, they usually contain flat regions such as the blue sky and texture-rich regions such as buildings or mountains. According to the richness of image details, we can design a dynamical remapping function and therefore improve the LLF model. In this roadmap, by proposing a simple technique for approximating richness of image details, we construct a spatially guided remapping function and embed it into the LLF model. Base on the low-level image features directly from the Laplacian image pyramid, we use Chamfer distance transform to realize this approximation. The contributions of our improved LLF method are two-fold. On one hand, the filtering process can be spatially guided by the new remapping function. In this way, we can assign different parameters to different regions. For detail-absent regions, e.g. flat regions, it is reasonable to put small or, practically, no filtering strength on them, which would structurally enhance the efficiency of photo processing. On the other hand, the extra computational cost brought from new remapping function is very low. Even when there is no flat region throughout an image, the improved LLF simply degenerates to the original version without introducing much computation cost.

1.1 Related research

As for research on image detail enhancement, Fattal et al. [3] propose a multi-scale method based on the bilateral filter. As mentioned before, the bilateral filter can be essentially viewed as a trade-off between edge-preserving and edge-smoothing. So tuning the parameters to achieve an optimized result has to be taken into consideration. Based on a psychophysical law, Gu et al. [8] propose a locally nonlinear model to tune image tone mapping. Ling et al. [13] propose a detail enhancement framework, including an energy optimization process, to preserve both tone and fidelity. Different from these methods, our approach considers the image detail richness of image contents and constructs a simple guiding map to spatially determine the enhancing strength. In our method, estimating image detail richness is a key ingredient. Similarly, Hong et al. [10] utilize low-level image feature to approximate detail-absent image regions, according to which movie captions are dynamically placed. Our method is partially inspired by this idea but differentiates in that we directly utilize image features from Laplacian image pyramid and thus introduce no extra feature extracting process.

The rest of the paper is organized as follows. In Section 2, we demonstrate the principle of original LLF and analyze its complexity. Then we present our spatially guided LLF version in Section 3. Experiments for validating the effectiveness of our method are demonstrated in Section 4. Finally, Section 5 concludes the paper and discusses possible future work.

2 Local Laplacian Filter

In this section, we first introduce the principle of basic local Laplacian filter (LLF) and its faster version. Then we analyze their complexities as well as the feasibility on further reducing the complexity.

2.1 LLF principle

Here we assume that the size of an image I is $N \times N$ and its intensity is defined as a scalar function $I(x,y) \in [0, 1]$ after normalization. As for color images, they can be simply viewed as three independent channels (RGB) in the LLF model. With the target image, we can then formulate the well-known Gaussian image pyramid and Laplacian image pyramid respectively, i.e. $\{G_l\} = C_G(I)$ and $\{L_l\} = C_L(\{G_l\})$, where $l \in 0 \dots n$, $n = \log_2 N$. Equation 1 demonstrates the relationship between these two pyramids:

$$\begin{cases} L_l = G_l - \text{upsample}(G_{l+1}) & l = 1 \dots n-1 \\ L_n = G_n & l = n \end{cases} \quad (1)$$

In this formulation, we observe that a Laplacian image pyramid can be derived from a Gaussian image pyramid and they share a common spatial structure and a full point-wise correspondence. Computationally, apart from the identical top layers in $\{G_l\}$ and $\{L_l\}$, the rest layers of $\{L_l\}$ can be obtained by subtracting the upsampled layer G_{l+1} from G_l for $l=0 \dots n-1$. With $\{G_l\}$ and $\{L_l\}$ at hand, an image in its original resolution can be reconstructed through a recursive process $I = A(\{L_l\})$, i.e. $G_l = L_l + \text{upsample}(G_{l+1})$ from the l^{th} layer to the 0^{th} layer.

LLF model defines a remapping function $r(i,g)$ on the normalized image intensity:

$$r(i,g) = \begin{cases} g + \text{sign}(i-g) \left(|i-g|/\delta_r \right)^\alpha & \text{if } |i-g| \leq \delta_r \\ g + \text{sign}(i-g) \left(\beta \left(|i-g| - \delta_r \right) + \delta_r \right) & \text{if } |i-g| > \delta_r \end{cases}, \quad (2)$$

$r(i,g)$ is a scalar function and its input and output range are both between $[0, 1]$. The aim of this function is, given a center g , to remap i at the position of (x_i, y_i, l) in $\{G_l\}$ into $r(i,g)$, where g and i are both scalar image intensities. Here α and β are two key parameters that control the LLF effects. The former parameter α generates effects of detail enhancing/smoothing while the later parameter β generates effects of image tone mapping. In this paper, we primarily concern the detail manipulation and set $\beta=1$. As for the parameter α , we can observe that the remapping functions represent an ‘‘S’’ shape in Fig. 1, of which the geometry reveals the principles of LLF detail manipulation. Specifically, when the parameter α is bigger than 1, the intensity range near the center g are compressed into a smaller one. Since the image edges can be perceived with bilateral contrast of imaging intensities, this inverted S-shape mapping decreases this kind of contrast and thus eliminates many weak details. In the contrary, when parameter α is smaller than 1, the remapping function produces a detail-enhancing effect, i.e. the S-shape mapping enlarges subtle intensity differences and thereof strengthens the image details. As is validated by the results in Fig. 2, we can observe that remapping functions with smaller α produce more intensified detail enhancing effects.

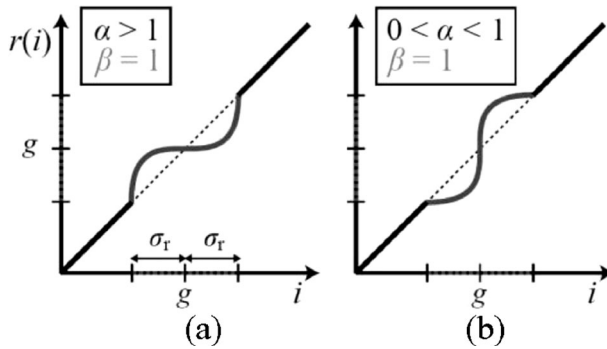


Fig. 1 Two kinds of remapping function in detail smoothing (a) and enhancing (b) in LLF model

By embedding this remapping function into the image reconstruction process $I=A(\{L_l\})$, the LLF can be summarized into the following process, which we call as naïve LLF:

- Inputs:** Original image I , remapping function $r(g,i)$, parameter δ_r .
- Output:** Reconstructed image I'

- Step1: $\{G_l\}=C_G(I), \{L_{0l}\}=C_L(\{G_l\})$
- Step2: for all $g=G_l(x_l,y_l)$
- Step3: $I_{remap}=r(I,g)$
- Step4: $\{G_l\}=C_G(I_{remap}), \{L_l\}=C_L(\{G_l\})$
- Step5: $L_{0l}(x_l,y_l)=L_l(x_l,y_l)$

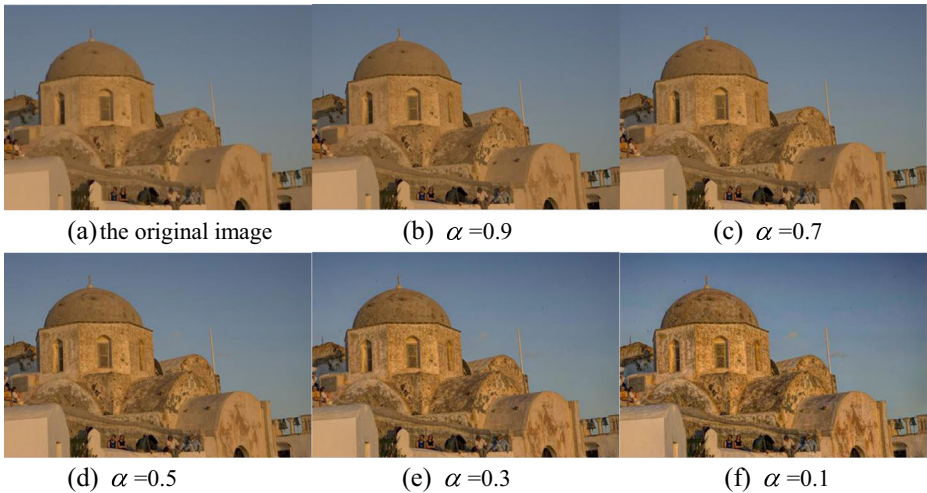


Fig. 2 Effects of LLF detail enhancement under different α a the original image b $\alpha=0.9$ c $\alpha=0.7$ d $\alpha=0.5$ e $\alpha=0.3$ f $\alpha=0.1$

Step6: end for
 Step7: $I' = A(\{L_{0l}\})$

2.2 Complexity analysis and our idea for improving LLF

From the above algorithm, we can conclude that the LLF method includes the following process: image pyramids construction, point-wise remapping, and image reconstruction. It is obvious that the main computation load comes from the remapping process. For a single point in $\{G_l\}$, in the naïve LLF, the complexity of remapping function is CN^2 . And the whole

complexity of LLF can be approximated as $CN^2 \sum_{j=0}^{n-1} (\frac{1}{4})^j N^2 \approx C_1 N^4$, which is so expensive that

it is not practical in most image processing applications. To tackle this problem, Paris et al. [15] propose an improved sub-region LLF. Instead of the whole image region, the spatial range for each point-wise remapping can be reduced into a $K \times K$ sub-region ($K=3(2^l+1)$), where the image pyramids are built. Therefore, the complexity of sub-region LLF is reduced to

$C_1 \sum_{j=0}^{n-1} (\frac{1}{4})^j N^2 K^2$, which can be further derived as $C_2 N^2 \log_2 N^2$.

The above LLF methods adopt a universal parameter α for all pixels in an image, which is irrelevant to spatial location or, to some extent, image contents. In our roadmap, we propose to improve LLF by spatially guiding the parameter value of α , i.e. distributing different values of α to different image regions, according to the richness of image details. In this way, we can extend the former universal filtering process to a dynamical one. Also, the improved model under our model can be faster than the sub-region LLF. For example, we can empirically assign no strength ($\alpha=1$) to the flat or detail-absent regions so that we can omit the $C_1 K^2$ remapping process for these pixels. This would have the potential to structurally reduce the whole complexity of the LLF process.

We then quasi-quantitatively analyze the feasibility for reducing the complexity. Suppose the complexity of the sub-region LLF and the detail approximation process are O_1 and O_2 , the ratio of non-LLF region is η ($\eta < 1$). To reduce the computation cost, it should satisfy the following in equation $(1-\eta)O_1 + O_2 < O_1$, i.e. $O_2/O_1 < \eta$. This means that, on one hand, given a fixed method on approximating α s for each pixel, the lower bound for η , at which the computation cost can be saved, is O_2/O_1 . On the other hand, given an object image (η is fixed), it is demanded that the complexity of introduced detail approximation method be as small as possible.

3 Content based Local Laplacian Filter

As the Laplacian operator is essentially the second order differential, it eliminates the 0th and first order intensity variation, i.e. producing zero response in flat regions and gradually varying regions. Instead of supervised learning methods aiming at classifying the image contents, we directly utilize the bottom layer of $\{L_l\}$ to estimate the richness of image details. This procedure produces an auxiliary matrix of the same size as image I , which can be incorporated into the remapping function.

We normalized the L_0 layer as a scalar function $f(x,y)=\gamma \in [0,1]$ defined on the grids. The locations of the non-zero γ s are the places where image details respond to Laplacian filtering. However, it is inappropriate to only manipulate these discrete non-zero dots or edges in the remapping process. Instead, we propose to estimate the richness of image details by overlapping the impacts of these dots or edges.

To this end, we define a function $h(f^{-1}(\gamma))$ that simulates the impact of γ -valued points in L_0 , where they are regarded as electric charges in the image domain. So the $h(\cdot)$ function simulates a truncated electric potential produced by these charges. Then the auxiliary image can be represented as an integral of these $h(\cdot)$ functions:

$$map(\mathbf{I}) = \int_{\tau} h(f^{-1}(\gamma))d\gamma, \tag{3}$$

where $\tau \in (0,1]$.

To achieve the discretization of the integral operator in Eq. 3, we divide the integrating range into k intervals $\{I_i\} = \{((i-1)\Delta, i\Delta]\}, i=1 \dots k$, where Δ is the length of an interval. We use the Chamfer distance transform to approximate the electric potentials:

$$map(\mathbf{I}) = \sum_{i=1}^k Cham(f^{-1}(I_i)). \tag{4}$$

In practice, the Laplacian response image is decomposed and translated into a set of binary images, i.e. $f^{-1}(I_i)$. The positions of zero-value pixels in a binary image are those whose intensities in L_0 fall into the interval I_i . The complexity of the conducting these Chamfer distance transforms is $kC_0N^2=C_3N^2$. The distance transform, whose complexity is only proportional to N^2 , i.e. the number of pixels of an image, comprises the main body for building the auxiliary image.

Empirically, after accumulating the histogram of an image’s Laplacian response, e.g. Fig. 3c, we observe that the non-zero responses in L_0 layer distribute in a small range after the normalization. Therefore, by reducing the number of intervals, i.e. k , we can further lower the computation cost of Eq. 4. We choose a smaller k' so that the $(0, k' \Delta]$ range includes 95 % of the non-zero responses. As the value of k'/k is usually much smaller than 1, we can thus reduce the number of distance transforms.

With the constructed auxiliary images, e.g. the second column of Fig. 5, we can dynamically assign non-single α s to different image regions. As the parameter α is inversely proportional to the enhancing strength, we linearly map the values in an auxiliary image onto the interval $[\alpha_{min}, \alpha_{max}]$, of which the range can be determined by users. Then we can obtain a parameter matrix α_{map} and incorporated into the sub-region LLF framework to construct the

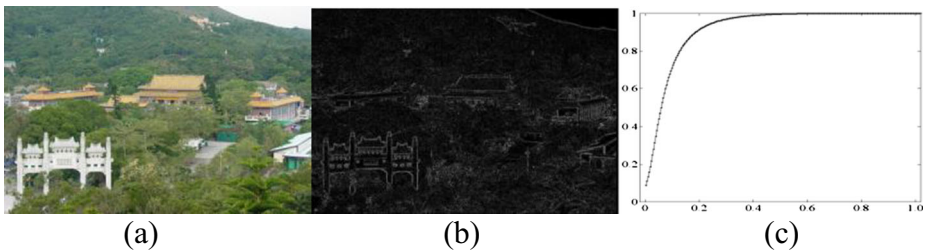


Fig. 3 **a** The original image, **b** the L_0 layer, and **c** the accumulation of **(b)**'s histogram

spatially guided LLF. Usually, we can choose $\alpha_{\max}=1$ so that the detail-absent regions can be omitted by the LLF process. Recall that the spatially guided LLF is able to save running time on condition that the complexities satisfy $O_2/O_1 < \eta < 1$. Specifically, compared with sub-region LLF, our LLF model achieves a faster implementation when the ratio between areas of detail-absent region and the whole image region is larger than $O_2/O_1 = C_3/(C_2 \log_2 N^2)$. Although we cannot give a precise value of this ratio, this condition could be easily satisfied for most images due to a qualitative observation that the bound can be very small in most cases. The following experiments also empirically validate this observation.

4 Experiments

In this section, we first analyze parameter selection of our algorithm. Then we validate our proposed spatially guided LLF in the following two aspects, i.e. the qualitative visual quality and the quantitative temporal costs. As for the experimental data, we choose three images used in [15] with a resolution of 800*500 and three images collected from the Internet with a resolution of 1800*1200. The algorithms were implemented with un-optimized codes and no parallel techniques were used. Since the naïve LLF is obviously much more time-consuming, we only compare our method with the sub-region LLF.

4.1 Parameter selection

In the sub-region LLF and our improved LLF, the main parameters are α , β , δ_r and k and k' . As for the sub-region LLF, the parameter α is set to 0.4 in the experiments. As for our methods, the dynamical range for α is set to [0.2,1]. The parameters β and δ_r for both methods are respectively set to 1 and 0.2.

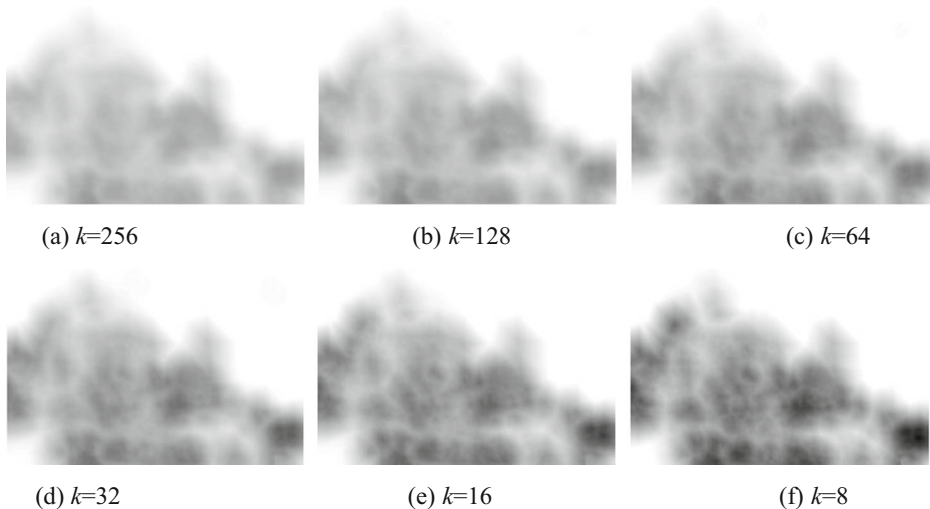


Fig. 4 Auxiliary images for guiding α values with different k s **a** $k=256$ **b** $k=128$ **c** $k=64$ **d** $k=32$ **e** $k=16$ **f** $k=8$

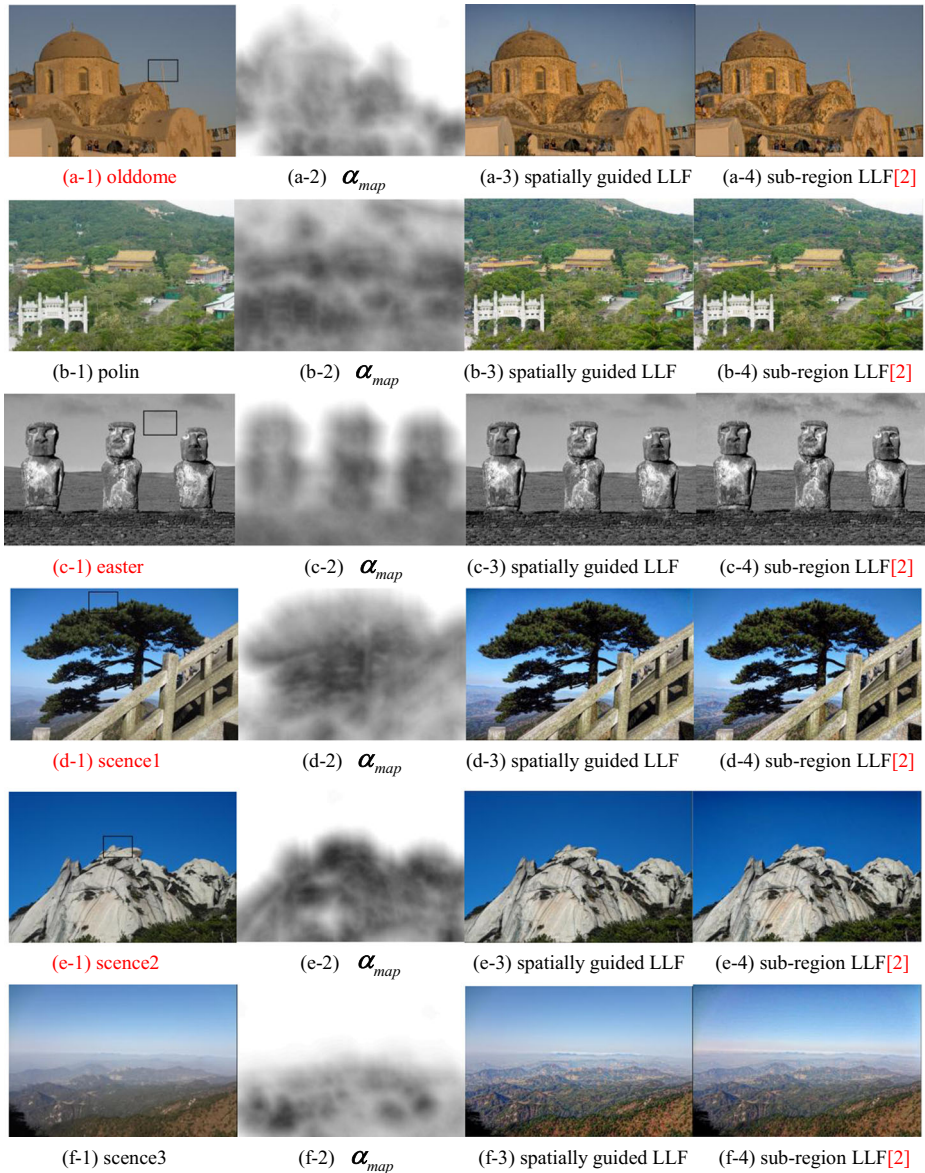


Fig. 5 Visual comparisons between sub-region LLF and our method (The *black frames* in (a-1), (c-1), (d-1) and (e-1) are for enlarged comparisons in Fig. 6) (a-1) olddome (a-2) α_{map} (a-3) spatially guided LLF (a-4) sub-region LLF[2] (b-1) polin (b-2) α_{map} (b-3) spatially guided LLF (b-4) sub-region LLF[2] (c-1) easter (c-2) α_{map} (c-3) spatially guided LLF (c-4) sub-region LLF[2] (d-1) scene1 (d-2) α_{map} (d-3) spatially guided LLF (d-4) sub-region LLF[2] (e-1) scene2 (e-2) α_{map} (e-3) spatially guided LLF (e-4) sub-region LLF[2] (f-1) scene3 (f-2) α_{map} (f-3) spatially guided LLF (f-4) sub-region LLF[2]

In the following, we discuss the selection of parameter k . As we have chosen 8-bit digital images as the experimental data, the pixel values are discretized from 0 to 255. In this context,

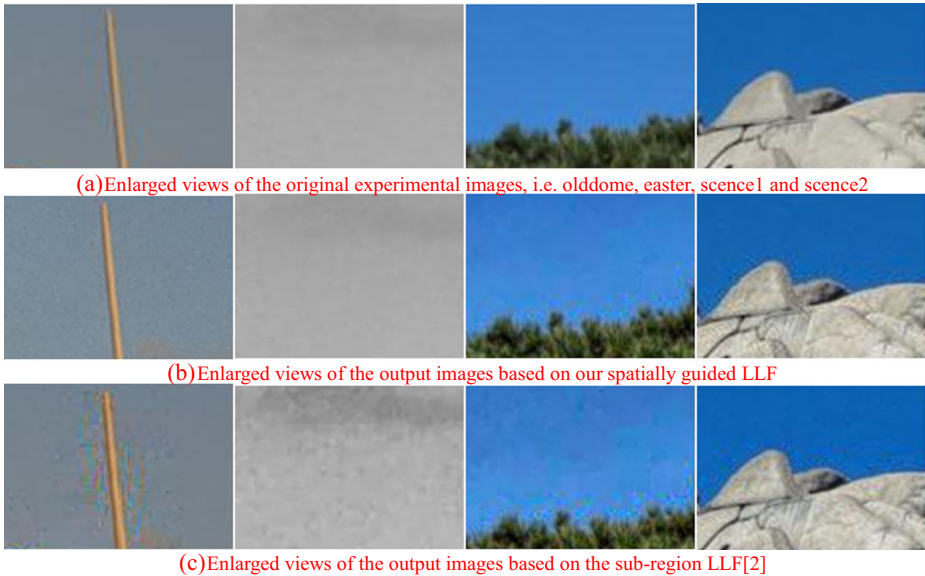


Fig. 6 Enlarged sub-regions of several experimental images for visual comparisons between sub-region LLF and our method **a** Enlarged views of the original experimental images, i.e. olddome, easter, scene1 and scene2 **b** Enlarged views of the output images based on our spatially guided LLF **c** Enlarged views of the output images based on the sub-region LLF[2]

the parameter k can be chosen as an integer which equally divides the range of image intensities. The α_{map} based on the discretized Eq. 3 can be best approximated with $k=256$. Practically, we can further decrease the number of intervals, i.e. k , to save the computational cost. From Fig. 4, we observe that larger k s produce smoother maps, and small ones generate block effects since less number of fields, which is generated by Chamfer distance transform, are accumulated. In our method, we empirically choose $k=64$ as it generally keeps the smoothly varied α s meanwhile further decreases the computational cost. With the determined k in our experiments, the values of parameter k' fall around 16 in our experiments.

4.2 Comparison between sub-region LLF and spatially guided LLF

Figure 5 demonstrates the visual effects of the sub-region LLF and spatially guided LLF on the six experimental images. Generally, these two methods achieve similar effects of detail enhancement on these images (both for grayscale images and color image), e.g. the enhanced distant mountains in

Table 1 Comparisons of running time between sub-region LLF and our method

	<i>olddome</i>	<i>polin</i>	<i>easter</i>	<i>scene1</i>	<i>scene2</i>	<i>scene3</i>
sub-region LLF (s)	3.51	3.46	1.39	22.19	22.35	22.46
spacially-guided LLF (s)	1.94	3.43	0.92	18.13	12.58	12.87
α_{map} (ms)	1.08	1.32	1.15	5.55	5.61	5.04

“scence1” and “scence3”. Differently, our method does not bring too much turbulence for the flat regions since it uses an auxiliary image to detect these regions and put no enhancing strength on them. In contrast, the sub-region LLF puts a universal strength (in our experiments the α is set to 0.4 as mentioned above) on all pixels, which would mistakenly exaggerate the microscopic variations in these regions. For the validation at a finer scale, we select some representative sub-regions from the experimental images and enlarge the results of both methods for further comparison. From Fig. 6, we can see that our method generates much less unnecessary turbulence than its counterpart 1) in flat regions and 2) around salient boundaries. As for the former case, it is mainly due to the zero filtering strength determined by our model; as for the later case, the distance transform modeling contributes to the descending filtering strength around salient boundaries. However, from the second column of Figure 5, we have to note that our technique only roughly detect detail-absent regions so that an α_{map} is far from a precise image content partition. As is discussed in Section 1, this fast unsupervised estimation is a trade-off between accuracy and efficiency. So we choose to distribute gradually descending strengths near the true boundaries between detail-rich and detail-absent contents.

In terms of temporal costs, we can observe that our method can be “conditionally” faster than the sub-region LLF from Table 1. On one hand, besides the second image “polin”, the running time of our methods on all other five images shows impressive improvements over its counterpart. Moreover, the time for computing α_{map} s only account for 0.1 % of the whole algorithm, which satisfies the demand for detail richness approximation. On the other hand, although this temporal advantage no longer exists for images with little or no flat regions, our method is still able to dynamically assign the filtering strength with certain user preference, without introducing much computational cost.

5 Conclusions

In this paper, we propose an improved local Laplacian filter that spatially guides the filtering strength by approximating the richness of image details. This method endows the LLF with the ability to dynamically assign appropriate parameter values to different image contents. We use a simple distance transform technique to achieve the detail richness estimation, which introduces little computation expense compared to the whole framework. Experimental results on some representative images validate the effectiveness of our method. Our method is also extendable to other image editing tasks such as LLF tone mapping. As for the future work, we can conduct parallel implementation to further speed up the algorithm. Also, as our method is totally unsupervised, we can consider a less supervised way to achieve better results by taking advantages of user-provided interactions [18].

Acknowledgments The authors sincerely appreciate the useful comments and suggestions from the anonymous reviewers. This work was supported by National Natural Science Fund of China (Grant No. 61301222), China Postdoctoral Science Foundation (Grant No. 2013M541821), Fundamental Research Funds for the Central Universities (Grant No. 2013HGQC0018, 2013HGBH0027, 2013HGBZ0166)

References

1. Bhat P, Zitnick CL, Cohen M, Curless B (2010) Gradientshop: a gradient-domain optimization framework for image and video filtering. *ACM Trans Graph* 29(2), Article 10

2. Buades A, Coll B, Morel JM (2006) The stair-casing effect in neighborhood filters and its solution. *IEEE Trans Image Process* 15(6):1499–1505
3. Fattal R, Agrawala M, Rusinkiewicz S (2007) Multiscale shape and detail enhancement from multi-light image collections. *ACM Trans Graph* 26(3), Article 51
4. Gao Y, Wang M, Ji RR, Wu XD, Dai QH (2014) 3-D object retrieval with hausdorff distance learning. *IEEE Trans Ind Electron* 61:2088–2098
5. Gao Y, Wang M, Tao DC, Ji RR, Dai QH (2012) 3-D object retrieval and recognition with hypergraph analysis. *IEEE Trans Image Process* 21:4290–4303
6. Gao Y, Wang M, Zha ZJ, Shen JL, Li XL, Wu XD (2013) Visual-textual joint relevance learning for tag-based social image search. *IEEE Trans Image Process* 22:363–376
7. Gao Y, Wang M, Zha ZJ, Tian Q, Dai QH, Zhang NY (2011) Less is more: efficient 3-D object retrieval with query view selection. *IEEE Trans Multimed* 13:1007–1018
8. Gu HX, Wang Y, Xiang SM, Meng GF, Pan CH (2012) Image guided tone mapping with locally nonlinear model. In: *Proceedings of European Conference on Computer Vision*
9. He K, Sun J, Tang XO (2010) Guided image filtering. In: *Proceedings of European Conference on Computer Vision*
10. Hong RC, Wang M, Yuan XT, Xu MD, Jiang JG, Yan SC, Chua TS (2011) Video accessibility enhancement for hearing-impaired users. *ACM Trans Multimed Comput, Commun, Appl* 7: Article 24
11. Li HJ, Tang JH, Wu S, Zhang YD, Lin SX (2010) Automatic detection and analysis of player action in moving background sports video sequences. *IEEE Trans Circ Syst Video Technol* 20:351–364
12. Li HJ, Yi L, Tang JH, Wang XH (2011) Capturing a great photo via learning from community-contributed photo collections. In: *Proceedings of ACM Multimedia*
13. Ling Y, Yan CP, Liu CX, Wang X (2012) Adaptive tone-preserved image detail enhancement. *Vis Comput* 28:733–742
14. Ni BB, Xu MD, Cheng B, Wang M, Yan SC, Tian Q (2013) Learning to photograph: a compositional perspective. *IEEE Trans Multimed* 15(5):1138–1151
15. Paris S, Hasinoff SW, Kautz J (2011) Local Laplacian Filters: edge-aware image processing with a Laplacian pyramid. In: *Proceedings of ACM SIGGRAPH*
16. Paris S, Kornprobst P, Tumblin J, Durand F (2009) Bilateral filtering: theory and applications. *Found Trends Comput Graph Vis*
17. Perona P, Malik J (1990) Scale-space and edge detection using anisotropic diffusion. *IEEE Trans Pattern Anal Mach Intell* 12(7):629–639
18. Wang M, Hua XS (2011) Active learning in multimedia annotation and retrieval: a survey. *ACM Trans Intell Syst Technol* 2:10–31
19. Wang M, Hua XS, Hong RC, Tang JH, Qi GJ, Song Y (2009) Unified video annotation via multigraph learning. *IEEE Trans Circ Syst Video Technol* 19:733–746
20. Wang M, Ni BB, Hua XS, Chua TS (2012) Assistive tagging: a survey of multimedia tagging with human-computer joint exploration. *ACM Comput Surv* 44(4): Article 25
21. Wang M, Yang KY, Hua XS, Zhang HJ (2010) Towards a relevant and diverse search of social images. *IEEE Trans Multimed* 12:829–842
22. Yang KY, Hua XS, Wang M, Zhang HJ (2011) Tag tagging: towards more descriptive keywords of image content. *IEEE Trans Multimed* 13(4):662–673



Shijie Hao is a lecturer in School of Computer and Information, Hefei University of Technology (HFUT). He received PhD degree from Hefei University of Technology, majoring in signal and information processing, in 2012. His current main research interests include digital image processing and analysis, multimedia content analysis and shape analysis.



Meng Wang is a professor in School of Computer and Information, Hefei University of Technology (HFUT). He received PhD degree from University of Science and Technology of China (USTC), majoring signal and information processing, in 2008. His research interests include multimedia content analysis, information retrieval and machine learning.



Richang Hong is a professor in School of Computer and Information, Hefei University of Technology (HFUT). He received PhD degree from University of Science and Technology of China (USTC), majoring signal and information processing, in 2008. His research interests include multimedia content analysis, image and video retrieval and multimedia question answering.



Jianguo Jiang is a professor in School of Computer and Information, Hefei University of Technology (HFUT). He was born in Oct., 1955. His research interests include artificial intelligence, image processing, and video coding.