Retinex based low-light image enhancement using guided filtering and variational framework*

ZHANG Shi (张诗)[1], **TANG Gui-jin** (唐贵进)[1]**, **LIU Xiao-hua** (刘小花)[1], **LUO Su-huai** (罗苏淮)[2], and **WANG Da-dong** (王大东)[3]

1. *Jiangsu Key Lab on Image Processing & Image Communication, Nanjing University of Posts and Telecommunications, Nanjing 210003, China*
2. *School of Electrical Engineering and Computing, The University of Newcastle, NSW 2308, Australia*
3. *CSIRO Data61, PO Box 76, Epping NSW 1710, Australia*

(Received 11 September 2017; Revised 5 November 2017)

A new image enhancement algorithm based on Retinex theory is proposed to solve the problem of bad visual effect of an image in low-light conditions. First, an image is converted from the RGB color space to the HSV color space to get the *V* channel. Next, the illuminations are respectively estimated by the guided filtering and the variational framework on the *V* channel and combined into a new illumination by average gradient. The new reflectance is calculated using *V* channel and the new illumination. Then a new *V* channel obtained by multiplying the new illumination and reflectance is processed with contrast limited adaptive histogram equalization (CLAHE). Finally, the new image in HSV space is converted back to RGB space to obtain the enhanced image. Experimental results show that the proposed method has better subjective quality and objective quality than existing methods.

Document code: A **Article ID:** 1673-1905(2018)02-0156-5

DOI https://doi.org/10.1007/s11801-018-7208-9

Image enhancement is an important branch and a hot topic in the field of image processing for many scholars. It aims to solve the problems that images have poor visual effects, missing details and other issues under the low light, backlight, etc. In the past five years, many scholars have studied this issue and put forward many good algorithms. In 2014, Yin et al[1] proposed an algorithm of optimizing the illumination based on guided image filter theory. In 2015, Pritee[2] described the disadvantage of "mean-shift" issue of the traditional histogram equalization technique. Pooja[3] developed a method to enhance the quality of images with respect to resolution and contrast. Wan[4] proposed an algorithm that combines the image dehazing and contrast enhancement seamlessly. The major advantage of this approach is that the inaccurately estimated parameters with intentional bias tend to enhance the image contrast. Lee[5] proposed a novel contrast enhancement algorithm for low light level images, which preserves image details and color constancy based on Retinex. Chang[6] and Rajendran[7] have used different methods to enhance images. Thasni et al[8] discussed color constancy algorithms in detail. In recent years, various methods are applied to image enhancement[9-12].

In this paper, we propose an algorithm based on Retinex theory to improve the visual effects for low-light images. First, we estimate the illuminations separately by guided filtering and variational framework, and compute the average gradient of the two components, then combine the two components with gradient weight. At the same time, we get the *V* channel. Second, the new reflectance is calculated using *V* channel image and the new illumination. The illumination and the reflectance are stretched and corrected, respectively. Then two components are multiplied to obtain a new *V* channel. Finally, the *V* channel is processed by contrast limited adaptive histogram equalization (CLAHE), and then we convert the image from HSV space back to RGB space and get the enhanced image. Experimental results demonstrate that the proposed algorithm has better color fidelity and detail preserving capability compared with other algorithms.

In Retinex theory, an image *S* which can be considered as the product of illumination *L* and reflectance *R* is given by

$$S = L \times R . \quad (1)$$

Solving the unknown component *R* through the estimation of *L* is the most basic method in Retinex theory. From Eq.(1), we can see that *L* directly affects the whole enhanced results, and may cause color distortion, over enhancement and other unexpected issues due to an unreasonable estimation model of *L*. Retinex based on guided filtering not only has good edge-preserving ability, but also solves the problem of over enhancement in a certain

* This work has been supported by the China Scholarship Council, Postgraduate Research & Practice Innovation Program of Jiangsu Province (No.KYCX17_0776), and the Natural Science Foundation of NUPT (No.NY214039).

** E-mail: tanggj@njupt.edu.cn

extent. The detail-preserving and natural-preserving ability of Retinex based on variational framework can effectively mitigate the problem of color distortion. In this paper, we present a weight-based method by taking the advantages of these two methods into account. Fig.1 shows the workflow of our algorithm.

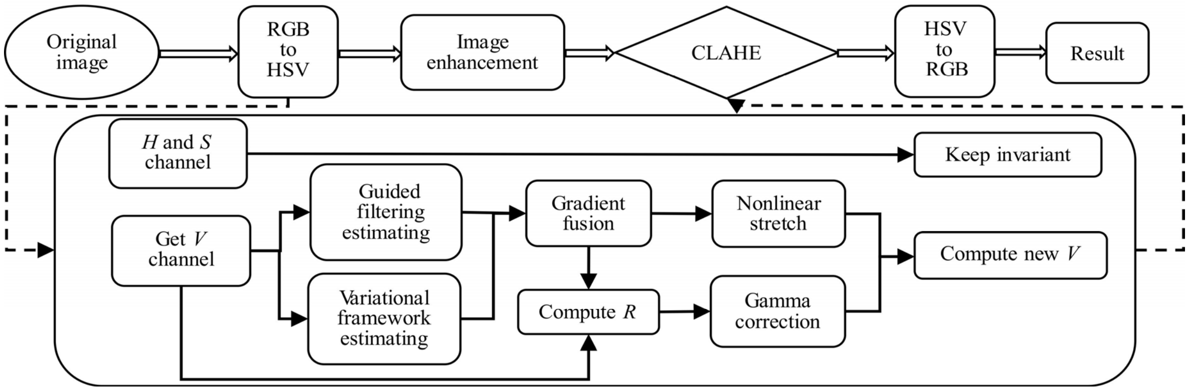

Fig.1 The workflow of the proposed algorithm

Guided filtering[13], which is derived from a local linear model computing the filtering output by considering the content of a guidance image, was proposed by Kaiming He. A guidance image can be an input image itself or another related image. Guided filtering computes the output through box filtering and integral image technique based on the least squares method to ensure that its time complexity is only O (N) and the execution speed is independent of filter window size. It has good behaviors near edges and is also a more generic concept beyond smoothing.

In guided filtering, a guidance image I, a filtering input image p and an output image q can be expressed as a weighted average:

$$q_i=\sum_j W_{ij}(I)\,p_j \quad , \tag{2}$$

where i and j are pixel indexes, and W is a filtering kernel function. The kernel function W can be obtained by

$$W_{ij}(I)=\frac{1}{|\omega|^2}\sum_{k:(i,j)\in\omega_k}\left[1+\frac{(I_i-\mu_k)(I_j-\mu_k)}{\sigma_k^2+\varepsilon}\right] \quad . \tag{3}$$

Since L is estimated by guided filtering, and it cannot guarantee the details effectively and the natural-preserving ability, we use variational framework to solve this problem.

In 2003, Kimmel et al[14] first proposed Retinex algorithm based on variational framework:

$$\mathrm{argmin}_l \int_\Omega \left(|\nabla l|^2-\alpha(l-s)^2+\beta\left|\nabla(l-s)\right|^2\right)\mathrm{d}x\mathrm{d}y,$$
$$s.t.\ l>s \text{ and } <\nabla l,n>=0 \text{ on } \partial\,\Omega \quad . \tag{4}$$

In this paper, a variational framework in spatial domain is studied[15], which not only has the detail-preserving and natural-preserving ability, but also has low complexity. This method uses a new variational model based on Retinex which has no logarithmic transformation:

$$\mathrm{argmin}_{R,L}\ \|RL-S\|_2^2+\alpha\|DL\|_2^2+\beta\|DR\|_2^2+$$
$$\gamma\|L-L_0\|_2^2,\ s.t.\ l>s \text{ and } <\nabla l,n>=0 \text{ on } \partial\Omega \quad , \tag{5}$$

where α, β and γ are weight parameters, and D is a difference operator in both horizontal and vertical directions. The first item $\|RL-S\|_2^2$ is the data-fitting term, the second item $\|DL\|_2^2$ is a constraint on L, and the third item $\|DR\|_2^2$ is a constraint on R component. The last item $\|L-L_0\|_2^2$ is an empirical term which makes L obey the Gaussian distribution with a mean of L_0. The computation in spatial domain is more complex than other variational models in frequency domain, but it can guarantee the edge and detail information of the image in the bright region.

The overall brightness of L is dark because the gray image is used as the guidance image P by guided filtering, but it has a very good hold on the edge of the image. On the contrary, the overall brightness of L is brilliant due to its piecewise smooth properties in the whole image space by variational framework, and there is no strong contrast at the edges. Fig.2 shows that the variational framework has a good effect in brightness, and guided filtering has a good edge-preserving ability. GF represents guided filtering, and VF represents variational framework.

(a) L component got by GF (b) L component got by VF

Fig.2 *L* components by (a) guided filtering and (b) variational framework

According to the above statement and visual effects of Fig.2, we design a gradient weighted fusion method based on the characteristics of the two. The gradient of the image is the change of intensity in the gray value which is the edge of the image. We can observe more edge lines in the image estimated by GF than VF in Fig.3.

(a) Edges got by GF (b) Edges got by VF

Fig.3 Comparison of two edge images

When the two L components are fused, different gradient images in Fig.3 cannot be directly used as weights. So we need to find out the average gradient through dealing with the gradient images. The average gradient is the gray change rate of a gradient image in the whole image area. The rate can be used to represent the image definition. It reflects the contrast change in small details of an image, and represents the relative clarity of the image. The average gradient is solved as follows:

$$\bar{g}=\frac{1}{(M-1)(N-1)}\times \sum_{i=1}^{M-1}\sum_{j=1}^{N-1}\sqrt{\frac{(F(i,j)-F(i+1,j))^2+(F(i,j)-F(i,j+1))^2}{2}} \quad ,(6)$$

where M and N are the width and height of an image, and $F(i, j)$ represents the gray value of the image at the pixel (i, j).

We use Eq.(6) to obtain the average gradient of the two L components which can be got from Eqs.(2), (3) and (5). Finally, the two L components are fused into a new L according to the weights:

$$G_{\mathrm{GF}}=\frac{\overline{g_{\mathrm{GF}}}}{\overline{g_{\mathrm{GF}}}+\overline{g_{\mathrm{VF}}}},$$
$$G_{\mathrm{VF}}=\frac{\overline{g_{\mathrm{VF}}}}{\overline{g_{\mathrm{GF}}}+\overline{g_{\mathrm{VF}}}}, \quad (7)$$
$$L_{\mathrm{final}}=L_{\mathrm{GF}}\times G_{\mathrm{GF}}+L_{\mathrm{VF}}\times G_{\mathrm{VF}},$$

where G_{GF} and G_{VF} are the weights. The image on the left in Fig.4 is the new image of L.

(a) New L component (b) New R component

Fig.4 The (a) new L and (b) new R components

The idea of the fusion is to offset the disadvantage of the GF by adding the VF. Meanwhile, we want to retain more edge information to enhance the visual contrast of the final enhanced image. Finally, the new R which is shown in Fig.4 can be obtained by the new L and the V channel through $R=V/L$. Then we process L by the tangent transformation as below:

$$L\_\mathrm{final}=b\times\arctan(a\times L\_\mathrm{new})/\pi \quad , \quad (8)$$

where a and b are parameters. The new R is processed by the Gamma correction shown as below:

$$R\_\mathrm{final}=(R\_\mathrm{new})^{\frac{1}{\gamma}} \quad . \quad (9)$$

A new V channel is obtained by multiplying L and R, then it is processed by CLAHE. Finally, we can get the result after converting the image from the HSV space to the RGB space.

In order to verify the effectiveness of the proposed algorithm, we use a variety of objective quality evaluation criteria to evaluate the proposed algorithm. Test images are given in Fig.5. We use MATLAB 2014b for programming, and use a computer with an eight-core 2.7 Hz CPU, 16G RAM, running on Windows 10. LIP[16] is an algorithm based on logarithmic image processing, NVE[17] is an algorithm for dark channel, VF[15] is an algorithm based on the variational framework, and SMQT[18] stands for successive mean quantization transform. Fig.6 shows the images of experimental results. The following Tabs. 1, 2, 3, 4, 5 show the objective comparison using five evaluation criteria which are mean, contrast, entropy, clarity and color colorfulness index (*CCI*), respectively.

Mean is a common method of image evaluation. Its value is proportional to the dynamic range of the image:

$$Mean=\frac{1}{3}\times\sum_{\mathrm{RGB}}\frac{1}{n\times m}\sum_{i=1}^{n}\sum_{j=1}^{m}I(i,j) \quad . \quad (10)$$

Contrast reflects the gray distribution. And it represents the strength of the contrast in the overall image. So it is stronger in the picture if the value of *Contrast* is bigger:

$$Contrast=\sum_{\delta}\delta(i,j)^2 P_{\delta}(i,j)\,, \quad (11)$$

where $\delta(i,j)=|i-j|$ represents the gray difference of adjacent pixels, and $P_{\delta}(i,j)$ is the pixel distribution probability of the gray difference which is δ between adjacent pixels.

Entropy is a quantity which is used to describe the amount of information of an image. If its value is larger, the clarity of the image will increase and the details of the image will become more abundant:

$$P_{ij}=\frac{m_{ij}}{m_i},$$
$$e_i=-\sum_{j=1}^{L}P_{ij}\log_2 P_{ij}, \quad (12)$$
$$Entropy=\sum_{i=1}^{K}\frac{m_i}{m}e_i,$$

where P_{ij} is the probability that the members in the cluster i belong to class j, m_i is the number of all members in the cluster i, m_{ij} is the number of members in the cluster i belonging to class j, e_i is the entropy of each cluster, L is the number of classes, K is the number of clusters, and m is the number of members involved in the whole clustering partition.

Clarity is also a standard for image evaluation. Its value is proportional to the expressive power of the image. Three values of clarity in the RGB channels are calculated in the experiment. Here is the single channel evaluation method and Eq.(13) is expressed in the frequency domain:

$$Clarity=\log_{10}\left[\frac{\sigma_f(\hat{I})}{\sigma_f(I)+\sigma_f(\hat{I})}\right]-\log_{10}\frac{1}{2}, \tag{13}$$

where $\sigma_f(I)$ is the variance of the original image and $\sigma_f(\hat{I})$ is the variance of the transformed image.

CCI represents the vivid and lively level of scene color:

$$\begin{cases} rg=R-G \\ yb=\frac{1}{2}\times(R+G)-B \\ \mu_{rgyb}=\sqrt{\mu_{rg}^2+\mu_{yb}^2} \\ \sigma_{rgyb}=\sqrt{\sigma_{rg}^2+\sigma_{yb}^2} \\ CCI=\mu_{rgyb}+\frac{3}{10}\times\sigma_{rgyb} \end{cases}, \tag{14}$$

where μ_{rg} and σ_{rg} are the mean and variance of *rg*, and μ_{yb} and σ_{yb} are the mean and variance of *yb*, respectively.

Fig.5 Test Images

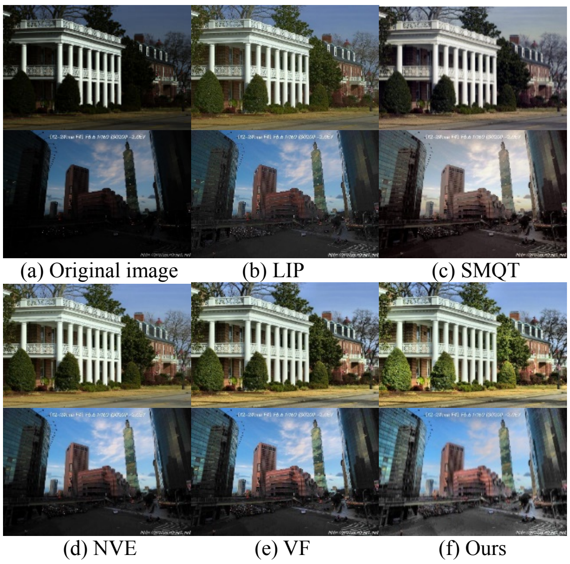

Fig.6 Comparison of various methods

It can be seen from the figures and tables that the five algorithms could enhance the contrast of the image and the amount of information in a certain extent during our comparison. The overall contrast of the LIP algorithm is not high, and the global parameter settings are difficult to make the image with uneven brightness become good. The image processed by NVE is dark and the details are not rich. SMQT makes the whole image have obvious color distortion. Some images processed by VF will have obvious ladder phenomenon. Our algorithm combines the advantages of the two methods, so that the enhanced image has a good effect in both objective and subjective quality. However, it also brings high time complexity. From Tab.6, we can see that our algorithm is several times slower than most of the algorithms in the comparison of 10 test images. In this paper, we use the advantages of the guided filtering and variational framework, and carry out nonlinear stretching and Gamma correction on *L* and *R*, respectively. These operations on the image contrast and color have introduced a big boost. The experimental results show that our method has better effects in detail information, natural-preserving and edge-preserving than other methods. The overall quality on the test images is better.

Tab.1 Comparison of *Mean*

Image/Method	LIP	SMQT	NVE	VF	Ours
House	99.48	100.42	108.43	114.42	**125.1**
Buildings	55.48	75.4	55.76	64.82	**81.83**
Crossing	74.65	91.4	75.71	85.55	**106.67**
Poster	80.79	84.48	80.98	92.84	**106.26**
Flowers	126.82	108.54	127.57	130.65	**144.53**
Horse	107.17	114.16	119.01	120.34	**128.67**
Goose	96.71	90.48	101.28	113.87	**124.25**
Room	107.37	92.78	111.46	119.04	**128.44**
Guidepost	111.23	107.04	117.39	123.43	**139.82**
Roof	127.91	108.45	133.56	138.18	**147.98**

Tab.2 Comparison of *Contrast*

Image/Method	LIP	SMQT	NVE	VF	Ours
House	95.04	106.03	115.31	118.59	**120.71**
Buildings	60.68	88.08	67.23	79.06	**97.64**
Crossing	79.83	91.85	81.93	93.23	**105.37**
Poster	82.25	97.86	95.79	100.77	**105.04**
Flowers	105.28	103.17	106.43	113.09	**122.70**
Horse	129.34	126.66	129.90	131.73	**134.61**
Goose	66.65	86.78	75.75	93.99	**112.34**
Room	68.36	84.88	83.36	85.35	**90.03**
Guidepost	104.63	106.84	106.97	115.42	**126.06**
Roof	104.13	106.76	110.97	114.87	**121.83**

Tab.3 Comparison of *Entropy*

Image/Method	LIP	SMQT	NVE	VF	Ours
House	15.92	13.27	13.80	16.56	**16.93**
Buildings	11.57	8.46	8.55	11.13	**12.74**
Crossing	11.84	9.53	9.72	12.16	**13.00**
Poster	14.30	12.31	12.34	15.31	**16.11**
Flowers	13.48	11.54	11.84	14.12	**14.86**
Horse	12.42	10.57	10.51	12.55	**13.29**
Goose	15.55	13.68	13.91	15.94	**16.12**
Room	11.77	9.765	9.90	13.53	**14.20**
Guidepost	13.66	11.31	11.34	13.90	**14.87**
Roof	14.26	12.43	12.28	14.85	**15.68**

Tab.4 Comparison of *Clarity*

Image/Method	LIP	SMQT	NVE	VF	Ours
House	0.069	0.075	0.085	0.092	**0.102**
Buildings	0.078	0.117	0.097	0.107	**0.125**
Crossing	0.078	0.112	0.093	0.102	**0.121**
Poster	0.056	0.071	0.069	0.077	**0.091**
Flowers	0.027	-0.003	0.029	0.026	**0.037**
Horse	0.051	0.057	0.072	0.069	**0.076**
Goose	0.066	0.064	0.077	0.089	**0.103**
Room	0.033	0.017	0.041	0.045	**0.054**
Guidepost	0.037	0.026	0.051	0.051	**0.064**
Roof	0.031	0.009	0.041	0.038	**0.047**

Tab.5 Comparison of *CCI*

Image/Method	LIP	SMQT	NVE	VF	Ours
House	21.23	9.58	22.93	25.12	**31.57**
Buildings	15.53	14.40	14.04	18.83	**29.69**
Crossing	26.25	9.87	30.12	31.18	**33.97**
Poster	14.53	12.18	16.39	17.55	**20.09**
Flowers	18.29	6.21	18.53	22.04	**31.34**
Horse	32.42	20.76	32.65	44.13	**61.97**
Goose	19.71	11.79	23.32	24.72	**27.37**
Room	12.98	7.40	15.78	16.21	**17.93**
Guidepost	11.76	5.34	12.14	14.99	**20.39**
Roof	22.52	8.42	24.07	23.84	**25.52**

Tab.6 Comparison of time (s)

Image/Method	LIP	SMQT	NVE	VF	Ours
House	1.455	7.226	0.472	0.822	**4.479**
Buildings	2.743	13.204	0.595	1.404	**8.229**
Crossing	1.512	6.495	0.353	0.985	**4.473**
Poster	2.238	9.212	0.384	1.274	**6.186**
Flowers	1.595	6.266	0.465	0.818	**4.192**
Horse	1.043	3.562	0.243	0.442	**2.522**
Goose	0.592	1.941	0.136	0.277	**1.312**
Room	3.179	13.124	0.602	1.302	**7.933**
Guidepost	3.969	18.081	0.891	2.396	**11.271**
Roof	12.198	48.045	2.182	4.624	**28.732**

In this paper, we propose a Retinex theory based algorithm for enhancing low-light images using guided filtering and variational framework. Low-light images suffer from some problems involving contrast decrease, detail loss and color distortion. Aiming at these, we estimate illumination through guided filtering and variational framework, and then obtain a new estimation by fusing the estimated results with gradient weights. We process the estimated component in the HSV space. Finally, the image is converted from the HSV space to the RGB space to produce the enhanced image. Experimental results show that the objective quality and subjective effect of an image can be significantly improved with the proposed method.

References

[1] Jingcao Yin, Hongbo Li, Junping Du and Pengcheng He, IEEE International Conference on Cloud Computing and Intelligence Systems, 639 (2014).

[2] Pritee Singh Rajpoot and Amit Chouksey, International Conference on Computational Intelligence and Communication Networks, 242 (2015).

[3] Pooja Bidwai and D. J. Tuptewar, IEEE International Conference on Information Processing, 511 (2015).

[4] Yi Wan and Qiqiang Chen, Visual Communications and Image Processing, 1 (2015).

[5] Hyo-Gi Lee, Seungjoon Yang and Jae-Young Sim, Signal and Information Processing Association Annual Summit and Conference, 884 (2015).

[6] Jian Chang and Jiahong Bai, 8th International Congress on Image and Signal Processing, 257 (2015).

[7] Rahul Rajendran, Shishir Paramathma Rao, Sos S Agaian and Karen Panetta, IEEE International Conference on Systems, Man, and Cybernetics, 002341 (2016).

[8] A N Thasni, V R Deepthi and Anu Bonia Francis, International Conference on Next Generation Intelligent Systems, 1 (2016).

[9] He Deng, Xianping Sun, Chaohui Ye, Xin Zhou and Maili Liu, IET Image Processing **10**, 701 (2016).

[10] Liqian Wang, Liang Xiao, Hongyi Liu and Zhihui Wei, IET Image Processing **9**, 43 (2014).

[11] Rim Walha, Fadoua Drira, Frank Lebourgeois, Adel M. Alimi and Christophe Garcia, IET Image Processing **10**, 325 (2016).

[12] Liang Zhang, Peiyi Shen, Xilu Peng, Guangming Zhu, Juan Song, Wei Wei and Houbing Song, IET Image Processing **10**, 840 (2016).

[13] Kaiming He, Jian Sun and Xiaoou Tang, IEEE Transactions on Pattern Analysis and Machine Intelligence **35**, 1397 (2013).

[14] Ron Kimmel, Michael Elad, Doron Shaked, Renato Keshet and Irwin Sobel, International Journal of Computer Vision **52**, 7 (2003).

[15] Xueyang Fu, Ye Sun, Minghui LiWang, Yue Huang, Xiao-Ping Zhang and Xinghao Ding, IEEE International Conference on Acoustics, Speech and Signal Processing, 1190 (2014).

[16] G. Deng, L.W. Cahill and G.R. Tobin, IEEE Transactions on Image Processing **4**, 506 (1995).

[17] Xuesong Jiang, Hongxun Yao, Shengping Zhang, Xiusheng Lu and Wei Zeng, IEEE International Conference on Image Processing, 553 (2013).

[18] M. Nilsson, M. Dahl and I. Claesson, IEEE International Conference on Acoustics, Speech, and Signal Processing, 429 (2005).