

Low-Light Image Enhancement Based On Retinex and Saliency Theories

Pengcheng Hao, Shuang Wang, Shupe Li, Meng Yang
Xi'an Jiaotong University
Xi'an, P.R.China
Email: mengyang@xjtu.edu.cn

Abstract—Low-light images enhancement of real scenes is a challenge task, however, existing algorithms always encounter the problems of over-enhancement, noises amplification and low subjective evaluation. One reason of these problems lies in the missing of high-level vision information. Therefore, this paper present a novel low-light image enhancement framework based on retinex and saliency theories. We first adopts a deep neural network model i.e. Saliency Attentive Model to predict the saliency map of low-light image and detect its salient regions. Then, we utilize one retinex model based method to enhance the whole low-light image. Then we fuse the generated saliency map and the enhanced image together to acquire a well-enhanced image. Experiments verify the significance of our algorithm by comparing with general low-light image enhancement method without saliency.

Keywords—low-light image enhancement, saliency map, retinex theory

I. INTRODUCTION

With fast development of technology, nowadays it's easier to capture images through diverse cameras and mobile phones. However, under low-light conditions, the amount of incoming light to the sensor is insufficient which leads captured images with low contrast and additive complex noises which lead the subsequent machine vision tasks, such as object detecting and tracking which require high quality images with sufficient details to a bad situation. What's more, people wish to capture high quality images where objects that they are interested in are clear with details. For example, when you take a picture of your friends outside at night, you will hope your friends salient enough in the captured image. As a result, low-light image enhancement algorithms are demanded and developed.

To lighten the dark regions in low-light images, the most intuitive way is to make the intensive histogram of images more uniform like histogram equalization (HE) in [1],[3] and [8] which can enhance the contrast by mapping the intensity of output image in the range $[0, 1]$. Furthermore, some methods like [4] exploit different regularization terms on the histogram to bring out better results. However, these methods that focus on the contrast enhancement will cause problems like over-enhancement in some area and amplifying noises.

Recent years, methods based on retinex theory [6] have been developed a lot to improve the contrast of the dark region. In this theory, image captured by sensor can be divided into two components, i.e. illumination and reflectance. Jobson et al. proposed single-scale retinex (SSR) [5] and multi-scale retinex (MSR) [7]. As an new approach to enhance low-light images, methods based on dehaze model also achieve promised performance like [23]. Also, with fast development of deep learning technology, methods depended deep neural network model have been proposed these years such as [18]. However, all methods mentioned above only focus on images

data themselves without considering any high-level-vision information in low-light images. As a result, their outputs have whole images strongly enhanced including irrelevant distractors and noises in background environment which lead to a bad subjective evaluation. Subjective visual perception of human tells us that people tend to focus on the most salient objects and regions in images where computer vision tasks like object detecting and tracking also focus on. Because traditional enhancement methods' results not only enhance some distractors in background but also amplify noises in environment, it's necessary to detect the high-level-vision information in low-light image and exploit it to optimize the process of low-light image enhancement. So we decide to utilize low-light image's saliency information to improve performance of traditional low-light image enhancement method.

In this paper, we proposed a framework to fuse high-level vision information into the process of general low-light image enhancement. First, we predict the saliency map of low-light image, which has not been well studied in the past. we address this problem by establishing a low-light image dataset and adopting a saliency attentive model to predict the saliency map of low-light image. Next, we utilize a retinex model based method called Frankle-McCann retinex to enhance the whole image through pixel-wise comparison which give us a fully enhanced result. Finally, we fuse saliency map and fully enhanced image generated above to ensure salient regions well-enhanced and non-salient regions properly inhibited. In experiment, we collective several real-world low-light images of variant scenes from Exclusively Dark (ExDark) Image Dataset [22] and results of our proposed method are better than results without fusing saliency map.

The paper is organized as follows. Section II presents the basic retinex model and saliency theory. Section III describes the proposed framework low-light image enhancement in detail. Experimental results and conclusion are shown in Section IV and Section V.

II. RELATED WORK

Methods based on retinex model has become the dominant way to enhance low-light image like SRIE [8] and LIME [9] mentioned above while they still only concern the data itself i.e. low-level-vision without any high-level-vision information. In order to exploit high-level-vision signals in

processing of low-light image enhancement, we decide to use saliency prediction method to extract saliency map of low-light image and make use of it for low-light image enhancement. In the rest of this section we will briefly introduce retinex theory and saliency prediction.

A. Retinex Theory

In retinex theory, the model of captured image is expressed by the formula shown below:

$$L=R \circ T, \quad (1)$$

Where we represent the captured image, reflectance map and illumination map with L , R and T respectively. As model (1) shows, captured image can be decomps into two elementwise multiplied factors which are reflectance map and illumination map. As shown in Fig. 1, it's intuitive that the imgae are captured through the illumination that the object reflects under the condition of an illuminant. However, due to the influence of uneven illumination, the reflectance map has better contast and color than the captured image. Therefore, the goal of this theory is mainly to recover the reflectance component by estimating the illumination map.

B. Saliency Prediction

The purpose of Saliency prediction is to identify the most informative and distinctive regions that grab human's attention in images and videos [10]. Existing methods can be categorized into two category: unsupervised method and supervised method. Usually unsupervised methods exploit the intrinsic cues of the captured image such as texture, contrast [11] etc. which is perceptually intuitive. Despite the contrast cue, some effective prior like the background prior [12] and center prior are fused into these methods to address some existed problems. Unlike unsupervised methods, supervised methods utilize different models to integrate different features. A pioneering work [13] learns to combine different features using conditional random field. With the trend of deep learning, massive researches based on deep neural network are proposed to extract deep features and predict saliency map like [14] and [15]. As [15] achieved a state-of-the-art result, we decided to utilize the network structure that proposed in [15] as our low-light image saliency prediction model.

III. THE PROPOSED ALGORITHM

The overall structure of our algorithm is illustrated in fig. 2. First we feed the low-light image into our saliency attentive model to derive the saliency map of the image. Then we use the low-light image enhancement method based on retinex theory to estimate the illumination map of the low-light image to produce the enhanced image. At last we fuse the saliency map and enhanced image to produce the image that enhances salient regions effectively while inhibits the distractors.

In this section, we will successively introduce our saliency attentive model and low-light image enhancement method fused with saliency map.

A. Saliency Attentive Model for Low-light Image

In [15], Marcella C et al. propose a saliency attentive model comprised of Dilated Convolutional Network unit,

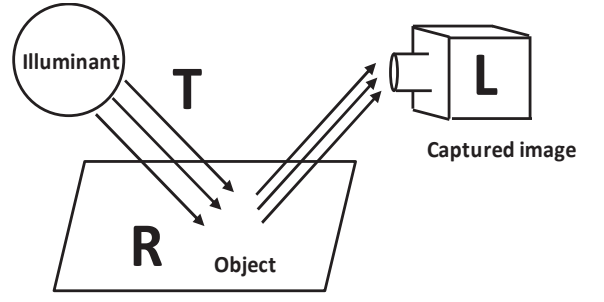


Fig. 1. Basic explanation of retinex theory.

Long Short-Term Memory (LSTM) networks unit and Convolutional Neural Network unit which is the one we adopt. First, we take a low-light image as input to Dilated Convolutional Network to extract feature maps. The reason to choose Dilated Convolutional Network rather than CNNs is to reduce the rescaling effects which can worse saliency prediction performance. We choose ResNet-50 [16] and abandon fully connected layer to construct Dilated Convolutional Network. To be specific, the stride is removed and dilated convolutions are introduced in the last two blocks. After we have feature maps X , we send it to LSTM network to process features in an iterative way which can progressively refine the saliency prediction. Though LSTM networks work on sequences of time such as video and texts, we utilize LSTM to solve problems on spatial features which is realized by substituting dot products with convolutional operations in the LSTM equations. The output of LSTM network, which is feature maps X' , is send to a CNNs unit to output the final prediction of saliency map. We simply use convolutional unit with kernel size of 1×1 and depth of 1 for the final CNNs.

To train our model, we use the dataset SALICON [17] which contains 10,000 training images, 5,000 validation images and 5,000 testing images. Since most images in SALICON are captured under natural condition with sufficient light, we decide to preprocess those images before we use them as input of Saliency Attentive Model in order to obtain a model suitable for low-light images. First, we use Gamma correction to decrease the contrast of images as shown in equation (2). We set A as 1 and γ as a random number from 2 to 5.

$$I_{\text{low-light}} = A \times (I_{\text{origin}})^{\gamma} \quad (2)$$

Then we add zero-mean-value random measure noises with Gaussian distribution to simulate noises which exist in low-light images as in [18] with standard deviation of $\sigma = \sqrt{B(25/255)^2}$, where $B \sim \text{Uniform}(0,1)$. Through this preprocess, we acquire low-light simulated dataset.

As for the implement details, we arrange the training process according to [15] without any change and the predicted saliency map is noted as Sal .

B. Retinex enhancement fused with Saliency map

In [24], Brian Funt et al. proposed a Matlab implementation of a retinex theory based low-light image enhancement method which is called Frankle-McCann retinex. So we decide to choose this method in our frame to

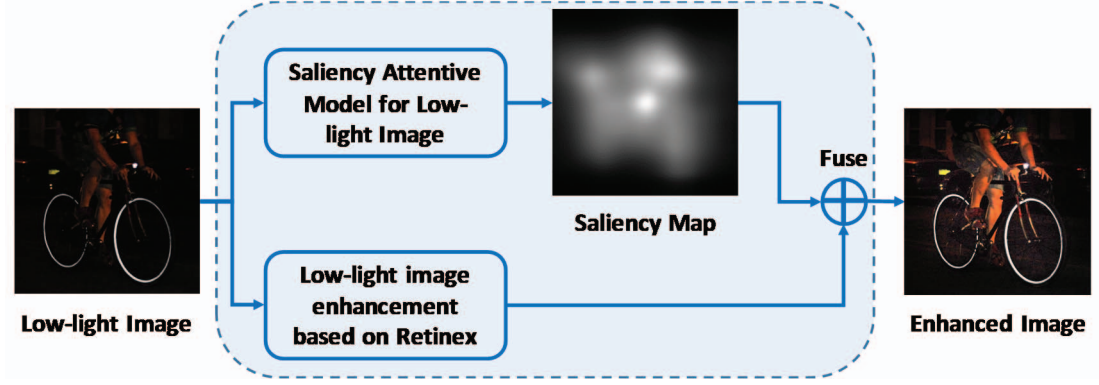


Fig. 2. Framework of the proposed algorithm.

enhance low-light image and then combine it with saliency map output from our Saliency Attentive Model.

The method in [24] is based on comparison of pixels. First we compute long-distance interactions between pixels and then the computation progressively moves to short-distance interactions. Distance between the pixels being compared decreases with each step and at each step, the comparison uses the Ratio-Product-Reset-Average operation. The process continues until the spacing decreases to 1 pixel.

First we transform the original image to the logarithmic domain in every channel shown in (3):

$$s^c(x) = \lg(S^c(x)), c \in \{R, G, B\}, \quad (3)$$

In (3), $S^c(x)$ represents each channel of original image in space domain while $s^c(x)$ refers to each channel of the original image in logarithmic domain. Then we initialize the reflectance map as $R_0(x)$ and then arrange a path around the central pixel and compare them progressively as shown in (4):

$$s_c = s_0 + \frac{s_c - s_m}{2} + \frac{s_c - s_{m-1}}{4} + \dots + \frac{s_c - s_1}{2^m}, \quad (4)$$

where s_m is the pixel on the path and the bigger the m , the closer the distance between the central pixel and s_m is. Essentially, the difference between the central pixel and the surrounding brightness is used as an estimate of the reflectance of the final central pixel. Therefore, we take the output of this process as the initial enhanced image as show in (5):

$$I_{enhanced}(x) = \exp(s_c(x)), \quad (5)$$

Then since we've acquired the enhanced image, the final step is to fuse it with its saliency map. As saliency map shows good performance on predicting salient region, we simply exploit (6) to fuse our enhanced image and saliency map where Sal refers to saliency map, $I_{low-light}$ refers to the captured image and I_{final} refers to our final output.

$$I_{final} = I_{enhanced} \times \text{Sal} + I_{low-light} \times (1 - \text{Sal}) \quad (6)$$

IV. EXPERIMENT

In this section, we will focus on the difference between performance of general enhancement method and our method. Our codes are implemented in Python and Matlab. All these experiments are conducted on a PC running Windows 10 OS with 8G RAM, NVIDIA GeForce GTX1050Ti and 2.30GHz CPU.

To compare our algorithm with the original retinex algorithm, we choose images of different scenes from Exclusively Dark (ExDark) Image Dataset [22] which is a collection of 7,363 images captured under low-light condition (i.e. 10 different conditions) with 12 object classes (similar to PASCAL VOC). The results are shown in Fig. 3, Fig. 4 and Fig. 5. From left to right successively are original low-light image, image enhanced based on retinex theory and our final result fused with saliency map.

Through the comparison, our model shows better subjective performance on enhancement of salient object, color without distortion and no noise amplification. In Fig. 3, salient objects and regions like cyclist and motorcyclist are all strongly well enhanced while the irrelevant environment and distractors like dark sky are well-inhibited unlike in original enhanced outputs whole scenes are strongly enhanced. In Fig.4, a large area of color distortion such as the color of building and sky appear in enhanced images without saliency map while results considered with saliency map show better visual perception because most of the area prone to color distortion is of no saliency. In Fig. 5, the noises in the dark sky are amplified in to some extent in directly enhanced images without saliency map, but because of low saliency of the sky the enhanced images with saliency map can effectively inhibit these noises.

In conclusion, results of our proposed method have more natural color and contrast in salient regions while the original results of produces distorted color and amplified noises.

V. CONCLUSION

In this paper, we fused saliency information into process of general retinex model based low-light image enhancement method. Experiments verified its effectiveness in improving the subjective quality of low-light image and to the best of our knowledge, it's the first attempt that fused high-level vision information into low-light image enhancement process, which may benefit subsequent computer vision tasks and satisfy subjective visual perception of people.



Fig. 3. Results of our proposed method show better performance on salient object (like cyclist and motorcyclist) enhancement and distractors inhibition



Fig. 4. Color distortion caused by over enhancement is effectively inhibited with saliency information

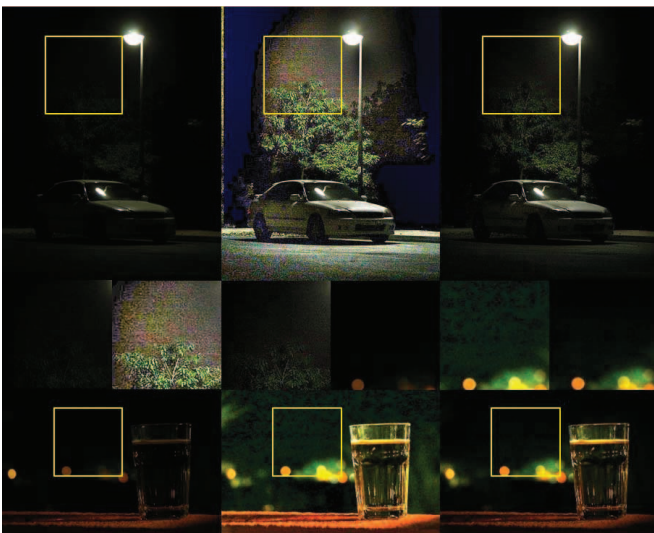


Fig. 5. Enhancement fused with saliency map has good performance on noise suppression

VI. ACKNOWLEDGEMENT

This work was supported by National Key R&D Program of China under Grant 2017YFB1302200 and the NSFC under Grant 61503294.

REFERENCES

- [1] E. Pisano et al., "Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms," *J. Digit. Imag.*, vol. 11, no. 4, pp. 193–200, 1998.
- [2] X. Fu, D. Zeng, Y. Huang, X. Zhang, and X. Ding, "A weighted variational model for simultaneous reflectance and illumination estimation," in *Proc. CVPR*, 2016, pp. 2782–2790.
- [3] M. Abdullah-Al-Wadud, M. Kabir, M. A. A. Dewan, and O. Chae, "A dynamic histogram equalization for image contrast enhancement," *IEEE Trans. Consum. Electron.*, vol. 53, no. 2, pp. 593–600, May 2007.
- [4] C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered difference representation of 2D histograms," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5372–5384, Dec. 2013.
- [5] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. Image Process.*, vol. 6, no. 3, pp. 451–462, Mar. 1997.
- [6] E. H. Land, "The Retinex theory of color vision," *Sci. Amer.*, vol. 237, no. 6, pp. 108–128, Dec. 1977.
- [7] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 965–976, Jul. 1997.
- [8] H. Cheng and X. Shi, "A simple and effective histogram equalization approach to image enhancement," *Digit. Signal Process.*, vol. 14, no. 2, pp. 158–170, 2004.
- [9] Feng Z., Hao S., "Low-Light Image Enhancement by Refining Illumination Map with Self-Guided Filtering[C]// 2017 IEEE International Conference on Big Knowledge (ICBK). IEEE Computer Society, 2017.
- [10] Borji, Ali, D. N. Sihite, and L. Itti. "Salient Object Detection: A Benchmark." *Computer Vision – ECCV 2012*. Springer Berlin Heidelberg, 2012.
- [11] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S.-M. Hu, "Global contrast based salient region detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 569–582, Mar. 2015.
- [12] Y. Wei, F. Wen, W. Zhu, and J. Sun, "Geodesic saliency using background priors," in *Proc. ECCV*, 2012, pp. 29–42.
- [13] T. Liu et al., "Learning to detect a salient object," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp. 353–367, Feb. 2011.
- [14] N. Liu and J. Han, "DHSNet: Deep hierarchical saliency network for salient object detection," in *Proc. CVPR*, Jun. 2016, pp. 678–686.
- [15] Marcella C., Lorenzo B., Giuseppe S., et al. "Predicting Human Eye Fixations via an LSTM-based Saliency Attentive Model[J]. *IEEE Transactions on Image Processing*, 2018:1-1.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
- [17] M. Jiang, S. Huang, J. Duan, and Q. Zhao, "SALICON: Saliency in context," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1072–1080.
- [18] Lore K G., Akintayo A., Sarkar S., "LLNet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement[J]. *Pattern Recognition*, 2015, 61:650-662.
- [19] Neil Bruce, John Tsotsos; "Attention based on information maximization. *Journal of Vision* 2007;7(9):950. doi: 10.1167/7.9.950.
- [20] Vanrell M., Otazu X., Parraga C A., "Saliency Estimation Using a Non-Parametric Low-Level Vision Model[C]// *Computer Vision & Pattern Recognition*. IEEE, 2011.
- [21] Tavakoli H R., Laaksonen J., "Bottom-Up Fixation Prediction Using Unsupervised Hierarchical Models[C]// *Asian Conference on Computer Vision*. Springer, Cham, 2016.
- [22] Loh Y P., Chan C S., "Getting to Know Low-light Images with The Exclusively Dark Dataset[J]. *Computer Vision and Image Understanding*, 2018.
- [23] Dong X., Pang Y., Wen J., "Fast efficient algorithm for enhancement of low lighting video[C]// *ACM SIGGRAPH 2010 Posters*. ACM, 2010.
- [24] Brian Funt, Florian Ciurea, and John McCann "Retinex in Matlab," *Proceedings of the IS&T/SID Eighth Color Imaging Conference: Color Science, Systems and Applications*, 2000, pp 112-121.