ENHANCING LOW-LIGHT IMAGES USING RETINEX THEORY: ILLUMINATION MAP ESTIMATION THROUGH LOCAL VARIANCE ANALYSIS AND REFLECTANCE ESTIMATION WITH TOTAL VARIATION WITH BILATERAL FILTER WEIGHTS

Mohammed Saidul Islam, 220232724

EECS - 6154 Project Report, Fall 23

ABSTRACT

This project aims to address the issues in images captured in dimly lit conditions that often suffer from unwanted noise and reduced visibility, which affects not just the image's visual quality, but also impairs the performance of many computer vision techniques that require high-quality inputs. This project introduces a retinex-based approach to enhance low-light images. Using retinex theory, the initial illumination map is estimated using a novel method that involves analyzing local neighborhood variance of an image. Subsequently, the initial reflectance component is estimated by dividing each pixel of the image by the initial illumination map. The reflectance component is then further refined by imposing a Total Variation Prior combined with bilateral filter weights. Finally, contrast enhancement is achieved by applying Contrast Limited Adaptive Histogram Equalization to the image. The experimental results demonstrate that the proposed approach improves the quality of low-light images.

Index Terms**—** Retinex Theory, Neighborhood variance analysis, Total Variation denoising

1. INTRODUCTION

Images captured in low-light often lack visual details and clarity and most of the time contain a great deal of noise distortions, that more likely to cause issues to computer vision algorithms. To this end, several studies [1, 2] propose Histogram Equalization based methods. Although these methods improve the overall contrast of the image, they increase the noise elements in the image as well. In recent years, the retinex theory has become a much more widely used approach to handle this problem. Retinex theory [3] states that, an image I can be represented in terms of pixel by pixel multiplication of an illumination component L, and a reflectance component R.

$$I(x, y) = L(x, y) \times R(x, y)$$

Here, R corresponds to the scene's inherent attributes, including texture details and the color information of the original scene, whereas L corresponds to the light's intensity and its spatial distribution within the scene's environment. Several methods have been proposed to decompose an input image to its illumination and reflectance component. Techniques like single-scale Retinex (SSR) [4] and multi-scale Retinex (MSR) [5] use a Gaussian filter to decompose an image to its illumination and reflectance component. To estimate the initial illumination map [6, 7] use the maximum value across all R, G, B channels of an image, while [8] follows the mean-RGB approach where the mean value across all R, G, B channel is computed, assuming the illumination is same across all channels. Another study [9] follows a different approach that converts the RGB image to its HSV counterpart and initializes the illumination component with the blurred V component. Subsequently, to further refine the illumination layer and the piece-wise smooth reflectance layer simultaneously, [10, 11, 12] propose different types of variational models. With the recent advances in deep learning architectures, several different methods [13, 7] have been proposed to tackle this problem.

To this end, this project proposes simple yet effective methods for initial illumination estimation using neighborhood variance analysis and reflectance estimation using Total variation denoising with bilateral filter weights.

The major contribution of this project is **two-folds**:

- A novel approach to estimate the initial illumination map via pixel neighborhood variance analysis which ensures a better initial estimation of illumination map.
- A novel approach to estimate the reflectance component of an image via Total variation denoising with the combination of bilateral filter weights that provides an improved estimation and allows better supervision in terms of how much denoising is performed with tunable hyper-parameters σ_s and σ_r.

2. RELATED WORKS

The natural yet effective model based methods to enhance the contrast of low-light images are the Histogram Equalization based methods like [1, 2]. While improving the overall contrast of the image, these methods increase the noise elements as well. In order to jointly denoise a low-light image and improve the contrast, retinex theory based methods have become popular recently. To remove unwanted artifacts, earliest techniques like single-scale Retinex (SSR) [4] and multi-scale Retinex (MSR) [5] employ Gaussian filters. Later, retinex theory based variational methods [11, 6, 8, 12] have been proposed for denoising and contrast enhancement of low-light images. With increasing popularity of graph based processing, [9] propose a method for contrast enhancement and denoising of low-light images with Gradient Graph Laplacian Regularizer. With the advancement of deep convolutional neural networks, methods like [13, 14] train end to end deep neural networks architectures for both denoising and contrast enhancement of an image. Furthermore, [7] employ a combination of plug and play framework and convoluted neural network to estimate the illumination component and reflectance component respectively.

3. PROPOSED METHODOLOGY

The proposed methodology is divided into the following subsections: *Illumination and Reflectance Initialization*, *Computation of Reflectance with Total Variation combined with Bilateral Filter Weights*, *Estimation of Illumination with Tikhonov Regularization*, *Contrast Enhancement using Contrast Limited Adaptive Histogram Equalization*.

3.1. Illumination and Reflectance Initialization

Illumination typically exhibits a smoother overall variation across an image, while reflectance is usually piece-wise smooth with sharp discontinuities. With this fact in mind, the proposed approach performs a neighborhood analysis in a $N \times N$ image patch and computes the variance of the pixel values within that neighborhood.

Given an image I, for each pixel at location (x, y), an initial neighborhood $N_{init}(x, y)$ is defined which is centered at (x, y). The variance $Var(N_{init}(x, y))$ within this neighborhood is calculated as:

$$Var(N_{init}(x,y)) = \frac{1}{|N_{init}|} \sum_{(i,j)\in N} (I(i,j) - \mu(N_{init}(x,y)))^2$$

where, $|N_{init}|$ is the number of pixels in the neighborhood, $I(i, j)$ is the intensity of the pixel at (i, j). $\mu(N_{init}(x, y))$ is the mean intensity of the pixels in the neighborhood, calculated as:

$$\mu(N_{init}(x,y)) = \frac{1}{|N_{init}|} \sum_{(i,j)\in N_{init}} I(i,j)$$

Next, based on the result of initial variance $Var(N_{init}(x, y))$ calculation, a new neighborhood size is selected based on the following criteria,

- If $Var(N_{init}(x, y))$ is greater than $variance\_threshold$, then a small neighborhood size (3×3) is selected

Finally, for each pixel at $(x, y))$, the $neighborhood\_size$ is compared,

- If $neighborhood\_size > small\_neighborhood\_size$, the pixel is assumed to be part of a smoothly varying region (illumination), and its value in the illumination component $L(x, y)$ is set to the maximum of the intensity value across three channels $max(I(x, y))$ (with the assumption that the illumination is at least the maximal value of three channels at a certain location [6])
- If $neighborhood\_size == small\_neighborhood\_size$, the pixel is in a region with significant variation (more likely reflectance), and its value in $L(x, y)$ is set to the average intensity of its neighborhood $\mu(N_{new}(x, y))$ [8], where $N_{new}(x, y)$ is the newly selected neighborhood size.

Next, for each pixel at (x, y), the initial reflectance component is constructed as,

$$R(x,y) = I(x,y)/L(x,y)$$

3.2. Computation of Reflectance with Total Variation combined with Bilateral Filter Weights

Total variation with bilateral filter weights. A discrete gradient operator ∇u for an image $u \in \mathbb{R}^{N\times N}$ is defined by,

$$(\nabla u)_{i,j} = \left((\nabla u)^h_{i,j}, (\nabla u)^v_{i,j}\right)$$

where,

$$(\nabla u)^h_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j}, & \text{if } i < N \\ 0, & \text{if } i = N \end{cases}$$

$$(\nabla u)^v_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j}, & \text{if } j < N \\ 0, & \text{if } j = N \end{cases}$$

The weighted version of the discrete gradient operator $\nabla_w u$ would be,

$$(\nabla_w u)_{i,j} = \left(w_{m,n} \cdot (\nabla u)^h_{i,j}, w_{o,n} \cdot (\nabla u)^v_{i,j}\right)$$

where m corresponds to coordinate $(i + 1, j)$, n corresponds to (i, j) and o corresponds to $(i, j + 1)$. Here, the weight $w_{p,q}$ is defined as,

$$w_{p,q} = \exp\left(-\frac{||p-q||^2}{2\sigma_s^2} - \frac{||u(p)-u(q)||^2}{2\sigma_r^2}\right)$$

where $||p - q||^2$ is the euclidean distance between two pixels and $||u(p) - u(q)||^2$ is the pixel intensity difference and σ_s and σ_r are smoothing parameters. The total variation is then calculated as,

$$TV(u) = \sum_{1\leq i,j\leq N} |(\nabla_w u)_{i,j}| = ||\nabla_w u||_1$$

So, the optimization equation becomes,

$$\min_u(||y-u||_2^2+\mu||\nabla_w u||_1)$$

To solve this optimization problem, the following steps are performed,

1. A new variable $z=\nabla_w u$ is defined, thus the new optimization becomes:

$$\min_{u,z}\|y-u\|_2^2+\|z\|_1 \quad \text{s.t.} \quad z=\nabla_w u$$

2. Then an unconstrained version of the problem is defined via **augmented Lagrangian**:

$$\min_{u,z}\|y-u\|_2^2+\|z\|_1+\lambda^T(\nabla_w u-z)+\frac{\delta}{2}\|\nabla_w u-z\|_F^2$$

3. First, z and λ is fixed, and solved for u,

$$\min_u\|y-u\|^2+\lambda^T(\nabla_w u-z)+\frac{\delta}{2}\|\nabla_w u-z\|^2$$

   Taking derivative w.r.t. u and setting result to 0,

$$-2(y-u)+\nabla_w^T\lambda+\delta\nabla_w^T(\nabla_w u-z)=0$$

   then,

$$(I+\delta\nabla_w^T\nabla_w)u=y+\nabla_w^T(\delta z-\lambda)$$

   and,

$$u=(I+\delta\nabla_w^T\nabla_w)^{-1}(y+\nabla_w^T(\delta z-\lambda))$$

4. Then, u, λ is fixed and solved for z,

$$\min_z\|z\|_1+\lambda^T(\nabla_w u-z)+\frac{\delta}{2}\|\nabla_w u-z\|^2$$

   Proximal operator for L_1-norm (soft thresholding operator):

$$\text{prox}_{\alpha\|\cdot\|_1}(v)=\text{sign}(v)\max(|v|-\alpha,0)$$

5. Finally, λ is updated,

$$\lambda^{k+1}=\lambda^k+\delta(\nabla_w u-z)$$

Finally, the denoised version of the reflectance component is achieved.

3.3. Estimation of Illumination with Tikhonov Regularization

If the initial illumination component is too noisy, then it is further refined using Tikhonov Regularization.

3.4. Contrast Enhancement using Contrast Limited Adaptive Histogram Equalization

After the combining the final version of the illumination and the reflectance part to construct the image, the Contrast Limited Adaptive Histogram Equalization [1] is performed to boost the overall contrast of the image.

4. EXPERIMENTAL RESULTS

For the experimental setup, a python environment is used to perform the image processing operations on the input image. The hyper-parameters are empirically set to the following values: $neighborhood\_size$, and $small\_neighborhood\_size$ is set to (5×5) and (3×3) respectively, and $variance\_threshold$ is set to 300. The optimal $variance\_threshold$ is found after several rounds of experiment. The σ_s and σ_r values are empirically set to 0.2 and 1 respectively.

4.1. Initial Illumination Estimation Results

Experimental results of initial illumination in Figure 1 shows that the initial illumination layer attained by pixel neighborhood analysis is more accurate and smoother than other existing methods.

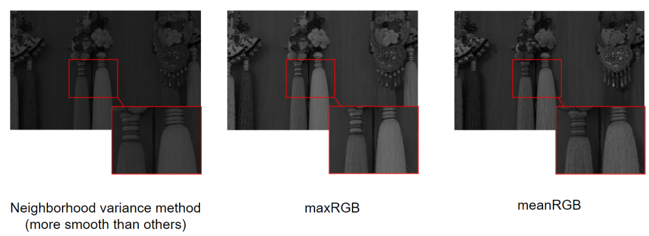

Fig. 1. Comparison of initial illumination map estimation between neighborhood variance analysis and other methods

4.2. Reflectance Estimation Results

With the introduction of bilateral filter weights in total variation, experiments shows promising results in estimating the reflectance component. With weighted Total Variation denoising, the horizontal and vertical gradients of an image is more likely to retain the edge information with higher accuracy than without using weighted total variation and edges are also more visible. Figure 2, 3 show the comparison between horizontal and vertical gradients when the regular and weighted version of TV is used. Figure 4 shows comparison between the refined version of the reflectance component using weighted TV and using regular TV denoising. Finally, Figure 5 represents the final results of the enhanced low-light image after going through the whole image processing pipeline. From the figures it is evident that the weighted TV provides more control of smoothing, thus providing a

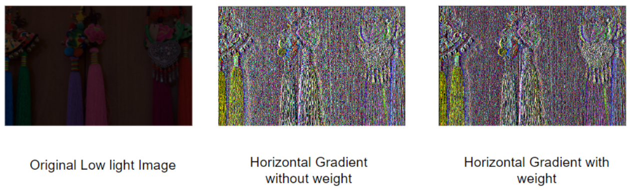

Fig. 2. Comparison of horizontal gradient between with/without using weighted total variation

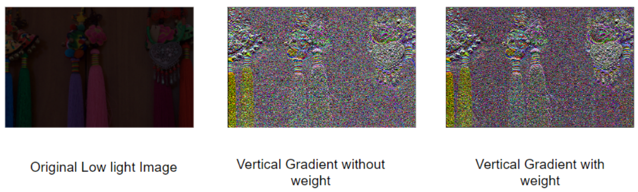

Fig. 3. Comparison of vertical gradient between with/without using weighted total variation

means to retain sharp edges of an image by having supervision over the hyper-parameters σ_s and σ_r.

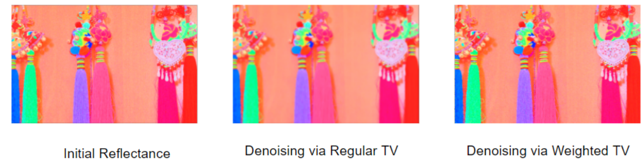

Fig. 4. Comparison of estimated reflectance component achieved using regular and weighted TV

Fig. 5. Comparison of final contrast enhanced image achieved using regular and weighted TV

5. REFERENCES

[1] Etta D Pisano, Shuquan Zong, Bradley M Hemminger, Marla DeLuca, R Eugene Johnston, Keith Muller, M Patricia Braeuning, and Stephen M Pizer, "Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms," *Journal of Digital imaging*, vol. 11, pp. 193–200, 1998.

[2] H.D. Cheng and X.J. Shi, "A simple and effective histogram equalization approach to image enhancement," *Digital Signal Processing*, vol. 14, no. 2, pp. 158–170, 2004.

[3] Edwin H. Land and John J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Am.*, vol. 61, no. 1, pp. 1–11, Jan 1971.

[4] D.J. Jobson, Z. Rahman, and G.A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Transactions on Image Processing*, vol. 6, no. 3, pp. 451–462, 1997.

[5] 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 Transactions on Image Processing*, vol. 6, no. 7, pp. 965–976, 1997.

[6] Xiaojie Guo, "Lime: A method for low-light image enhancement," 2016.

[7] Tingting Wu, Wenna Wu, Ying Yang, Feng-Lei Fan, and Tieyong Zeng, "Retinex image enhancement based on sequential decomposition with a plug-and-play framework," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–14, 2023.

[8] Xutong Ren, Wenhan Yang, Wen-Huang Cheng, and Jiaying Liu, "Lr3m: Robust low-light enhancement via low-rank regularized retinex model," *IEEE Transactions on Image Processing*, vol. 29, pp. 5862–5876, 2020.

[9] Yeganeh Gharedaghi, Gene Cheung, and Xianming Liu, "Retinex-based image denoising / contrast enhancement using gradient graph laplacian regularizer," 2023.

[10] Ron Kimmel, Michael Elad, Doron Shaked, Renato Keshet, and Irwin Sobel, "A variational framework for retinex," *International Journal of computer vision*, vol. 52, pp. 7–23, 2003.

[11] Michael K. Ng and Wei Wang, "A total variation model for retinex," *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 345–365, 2011.

[12] Wenye Ma, Jean-Michel Morel, Stanley Osher, and Aichi Chien, "An l1-based variational model for retinex theory and its application to medical images," in *CVPR 2011*, 2011, pp. 153–160.

[13] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu, "Deep retinex decomposition for low-light enhancement," 2018.

[14] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun, "Learning to see in the dark," 2018.