Illumination Normalization Using Weighted Gradient Integral Images

Yonghuai Liu, Luigi De Dominicis, Ralph R. Martin, and Baogang Wei

Abstract—Captured images arise through interaction between objects of interest and illuminating light sources. If the latter are unevenly distributed, or are too strong or too weak, the image can have low contrast either locally or globally, impeding its interpretation and reducing its usefulness. In practice, control of illumination conditions is challenging, and not always possible. Thus, we propose a novel method to post-process captured images to reduce the effects of the illumination. We employ the Sobel operator to estimate gradients in the image, then use these gradients as weights in an averaging operation. To accelerate the operation, and remove certain unwanted features, integral images are iteratively calculated from the weighted images. This allows us to estimate the illumination component of the images, and in turn the reflectance component as their enhanced ones. A comparative study using a large number of images shows that the proposed algorithm outperforms several state of the art approaches.

Index Terms—Image enhancement, illumination removal, gradient, reflectance, Retinex.

I. INTRODUCTION

The appearances of digital images are mainly determined by the colors of objects in the scene, the colors of surrounding objects, and the colors of the illuminating lights. The interaction between the illumination and objects of interest is complicated, and difficult to model due to complexity of geometry and variation of reflectance properties from one region to another; it is further complicated by reflected lights from surrounding objects.

Consequently, captured images depend in a complex way upon the illumination conditions, and are strongly affected by them. Unless the lighting conditions are carefully controlled, the captured images may not be satisfactory for subsequent processing and interpretation, and indeed, control of lighting is not always possible. In such cases, post-processing is necessary to normalize the captured images, in the sense of removing the effect of unknown and possibly varying illumination conditions, so that the processed images depend mainly on the reflectance and geometrical properties of the object of interest rather than the light sources.

A. Previous Work

The capturing of images has a long history, and image processing and enhancement have attracted intense attention in various disciplines such as pattern recognition, computer vision, biology, and astronomy, to name just a few. Consequently, many techniques have been proposed, which can be classified as either global or local methods, or a combination of the two.

Histogram equalization (HE) is one of the most widely used and successful global methods for image enhancement. It re-assigns the colors and intensities of pixels by considering the frequencies of these colors and intensities. It has the advantage of easy implementation. However, it is most effective for images with colors and intensities concentrated in a narrow band. It is not entirely effective for images whose colors and intensities span the whole gamut of display devices, as clearly illustrated in Fig. 1. Gamma correction [1] is another widely used global method which can be effective for images which are either too dark or too bright. However, optimal choice of the parameter gamma is image dependent. In [2], [3], images are partitioned into several regions, and HE and/or gamma correction are applied to each region. A disadvantage of this approach is that transitions of colors and intensities are not smooth at the boundaries between different regions.

Fig. 1. Histogram equalisation. Columns 1 and 3: original images. Columns 2 and 4: enhanced images.

An interesting idea was proposed in [4], [5]: the intensities of pixels are proportional to the logarithm of the ratio of the intensity of the radiation falling on the pixel of interest and that of its neighbors. Such ideas have a grounding in the properties of the retina and cortex in the human visual system, leading to the name Retinex theory.

Retinex theory posits that an image $I$ can be decomposed into two components: $I = R \times L$ where $R$ and $L$ represent the reflectance and illumination respectively, and $\times$ denotes component-wise multiplication. $R$ usually reveals the reflectance and geometry of the object of interest more objectively and can thus be regarded as the enhanced image $I$. Using this elegant formulation, a number of local methods have been proposed for estimating $R$ and $L$. In [6], [7], the reflectance component is estimated as a transformation of the ratio of the original and Gaussian smoothed intensities of a pixel. To avoid a haloing effect at edges, multiple scales are
used, and the smoothed images are integrated. In [8], an iterative adaptive smoothing method is used to estimate $L$, where the weights of pixels are determined as functions of gradients and inhomogeneities. To estimate $L$, [2], [9] suggest minimizing an objective function based on the square of the first order differential and the difference of the illumination $L$ from the given image. The estimated illumination component is further enhanced in [9] using gamma correction.

**B. Our Work**

Our work is also based on Retinex theory. Again, the reflectance component is estimated as the logarithm of the ratio of intensity of the pixel of interest and that of its neighbors. Such ideas are plausible for the following reasons: on one hand, pixel intensities are affected by illumination. As long as pixel intensities change in proportion to illumination, then ratios of intensities will not change significantly. On the other hand, the logarithm function takes into account the varying response of the eye which allows it to see over a large range of intensities, perceiving similar colors under various illumination conditions. There are two key computations to be performed to determine the reflectance component based on such ideas: 1) selection of a suitable neighborhood, and 2) estimation of the intensities of neighboring pixels.

The first task is addressed by dynamically defining the size of a rectangular region centred at the pixel of interest, over which the intensities of neighboring pixels are computed. The second task is addressed by computing the weighted average of intensities of neighboring pixels with weights defined as a function of the image gradients. While gradient information is usually extracted for such tasks as corner and edge detection, it is used in this paper for illumination normalization. We extract the gradient information using the Sobel operator. To suppress the effects of large gradients, the weights of pixels are selected as a decreasing function of these gradients. To speed up the averaging operation, we use an integral image [10], exchanging memory for efficiency. An integral image holds the sum of intensities of pixels inside a window from top-left corner to any pixel of interest. Integral images of the weighted images and the weights are iteratively built and applied to estimate the illumination component $L$. Finally, the reflectance component $R$ is estimated as the logarithm of the ratio of $I$ and $L$.

To validate the proposed method for illumination normalization, a comparative study is given later using varying kinds of images: images which are too dark or too bright, and images with both dark and bright regions. We compare our proposed method with several other state of the art methods: max filtering (MAX) [11], the envelope (ENV) method [9], an adaptive smoothing (ADS) method [8] and multiple scale Retinex (MSR) [6]. The results show that our method outperforms them.

In the following, Section II details our method, experimental results are presented in Section III, and Section IV draws conclusions and indicates future research directions.

**II. THE NOVEL ALGORITHM**

In this section, we detail our proposed algorithm for illumination normalization. We argue that low contrast in an image is mainly a problem of low contrast in the achromatic intensity, rather than an issue affecting colour. We thus represent the given image of height $h$ and width $w$ in HSV color space. The enhancement is performed on the value $V$ component only without altering saturation or hue information. We first normalize the $V$ component into the interval $[0, 255]$, yielding $I$.

The Sobel operator is used to estimate $x$ and $y$ gradient information $G_x$ and $G_y$ in the normalized image $I : G_x = T \times K_x$ and $G_y = T \times K_y$, where $x$ denotes convolution, and $K_x$ and $K_y$ are Sobel kernels in $x$ and $y$ directions respectively. The gradient magnitude $G$ is then computed as: $G = \sqrt{G_x^2 + G_y^2}$. To suppress the effects of pixels with large gradients, and noisy pixels, during illumination normalization, weights $W$ are defined as: $W = 1/(1+G/G)$, where $G$ is the average of the gradients of all pixels in the image. The ratio $G/G$ in the definition of $W$ calculates the relative gradient, rather than the absolute gradient, and thus, also removes the effect of illumination on the appearance of the image.

In order to make sure that the global appearance of the image is captured and the high frequency contents are retained especially at the borders of the image, the size of the rectangular window $(i_1, j_1, i_2, j_2)$ centred at the pixel $(i, j)$ of interest is dynamically determined to be: $i_{off} = \min(i, h - i)$, $j_{off} = \min(j, w - j)$, $(i_1, j_1) = \max(i - i_{off}, 0)$, $(i_2, j_2) = \min(i + i_{off}, h)$, $j_{off} = \min(j, w - j)$, $j_{off} = \max(j - j_{off}, 0)$, $j_{off} = \min(j + j_{off}, w)$, where $(i_1, j_1)$ and $(i_2, j_2)$ are the top-left and bottom-right corners of the window and the window has a height of $2i_{off}$ and a width of $2j_{off}$. In this case, the size of the rectangular window varies from one pixel to another. The pixels at the centre of the image are associated with larger rectangular windows, while the peripheral pixels are associated with smaller rectangular windows.

The computation of the weighted average of $V$ values of the pixels inside the rectangular windows with varied sizes is time consuming. To speed up the computation, we employed the integral image [2], which is an elegant technique for efficiently carrying out an averaging operation. By pre-computing certain sums $\sum W(i, j)$ of the $V$ values of pixels, the time required for the computation of the average is
independent of the size of the rectangular window. It is constructed iteratively as:

\[
T(i, j) = \begin{cases} 
T(i, j) & \text{if } i = 0 \text{ or } j = 0 \\
T(i-1, j) + T(i, j-1) & \text{otherwise} 
\end{cases}
\]

(1)

In this paper, we employ the integral image \( T \) to accelerate the computation of weighted averages by changing \( T(i, j) \) to \( T(i, j) \times W(i, j) \). Two integral images \( T_i \) and \( T_w \) are computed from \( T \times W \) and \( W \times 1 \) using Equation (1), where 1 denotes the matrix of the same size as the original image, all of whose components are 1, and \(*\) denotes component-wise multiplication. The weighted averages \( m \) of \( V \) values of pixels are computed as follows:

\[
m(i, j) = (T_i(i, j) + T_w(i, j_1))/32
\]

The illumination component \( L \) is estimated as

\[
L = \begin{cases} 
m & \text{if } m > T \\
T & \text{otherwise} 
\end{cases}
\]

This guarantees \( T \leq L \) due to the constraint \( R \in [0, 1] \).

To further removing shadow effects, the Sobel operator is applied to \( L \) and the process repeated \( M_{iter} \) times; in general we find \( M_{iter} = 10 \) is a good choice. Finally, the reflectance component \( R \) is estimated as: \( R = \log((T + 1)/(L+1)) \). For display purposes, both \( L \) and \( R \) are then normalized to the range \([0, 1]\).

As all operations involved are pixel based, the method has a linear computational complexity of \( O(wh) \).

Computational efficiency of the different methods was evaluated using a PC with an Intel Xeon E5610 processor. Results are presented in Fig. 2 to Fig. 4 and Table I.

A. Choice of Window Size

We first consider choice of parameters \( l_{off} \) and \( f_{off} \) which define the size of the rectangular window over which the \( V \) values of pixels are averaged. These parameters can be set dynamically as proposed; we also considered a simpler approach of using constant values of \( l_{off} = f_{off} = \min(h, w)/32 \). Results are presented in Fig. 2, showing that a fixed sized window leads to a noisy face image and a stapler image in which global contrast has been changed in text and its surroundings. In contrast, dynamically sized rectangular windows produce a smoother face image and a more realistic stapler image. These results show that a trade-off needs to be made between global contrast and local detail. Optimization of this trade-off needs further investigation.

B. Choice of Number of Iterations

We next consider choice of the parameter \( M_{iter} \). Results are presented in Fig. 3, showing that when \( M_{iter} \) is small, e.g. \( M_{iter} = 2 \), a banding effect appears in the middle of the face and on the left side of the mammo1 image. As the parameter \( M_{iter} \) increases to 10, the face and mammo1 images appear much smoother, showing that estimation of the intensity of neighboring pixels plays a key role in illumination normalization of images. It must be carefully estimated by propagation of local intensity information of pixels to their surrounding regions.

C. Comparative Study

Fig. 4 shows that all methods successfully enhance the contrast of the house and tower images, with the car at the bottom-left and the color of the grass and wall easily identifiable. Our method successfully enables clear visibility of the details in all images: weighted integral images of the gradients provide a powerful tool for image enhancement. In sharp contrast, the MAX method [11] fails to enhance the contrast of the cat, widens the width of the letters in the mammo image, shrinks the eyes of maleface3 and produces a noisy train image. It fails to balance the maximum intensities in dark and bright regions, and propagates noise. The ENV method [9] is ineffective in enhancing the cat and maleface3 images due to violation of spatially piecewise smoothness. The ADS method [8] blurs the maleface3 and maleface images; the global contrast between the background and the foreground in the mammo and train images has been changed. This is because it failed to accurately estimate the weighted average of intensities of neighboring pixels. The MSR method [6] produces color shifts from cyan to blue in the sky and sea in the beach image and from red to a greenish hue on maleface; the two sides of the male face are clearly different due to shadows. The dark regions in the cat image remain quite dark. It is difficult for the method to render the original color and a number of parameters must be selected (such as the number of scales) in a data dependent way.

Table I shows that all the methods have comparable computational speed, taking several seconds to process a typical image.
IV. CONCLUSION

Land’s discovery that the reflectance component can be estimated as the logarithm of the ratio of intensity of the pixel of interest to that of its neighbors is inspiring and powerful. This paper uses it as a basis for a novel method for illumination normalization using weighted integral images of gradients. The main contributions of the paper can be summarized as follows. Firstly, an exciting finding is that gradient information can be used alone for normalizing the appearance of images captured under various lighting conditions. Secondly, weighted integral images provide an efficient approach to enhance images, and at the same time remove some unwanted features from images. Our experiments based on many more images than those shown here lead us to conclude that the proposed method is effective and robust, and outperforms selected other state-of-the-art methods. Future work will consider enhancing images captured from underwater environments.

REFERENCES


Yonghuai Liu is a senior lecturer at Aberystwyth University since 2011. He obtained his first PhD degree in 1998 from Northwestern Polytechnical University, P.R. China and second PhD degree in 2001 from The University of Hull, United Kingdom. He is currently associate editor for the international Pattern Recognition Letters journal, an editorial board member of the international Journal of ISRN Artificial Intelligence, and a guest editor for a number of international journals. He has been serving as programme committee member and referee for more than 60 international conferences and journals. He has published three books and more than 140 papers in international conference proceedings and journals. His primary research interests lie in 3D computer vision, pattern recognition, image processing, machine learning, artificial intelligence, and intelligent systems. He is a senior member of IEEE and Fellow of Higher Education Academy of United Kingdom.

Luigi De Dominicis was born in Rome, Italy, on June 22, 1966. He received the diploma in physics from the Università La Sapienza in Rome in 1992. From 1994 he is researcher at ENEA, the Italian Agency for the New Technologies, Energy and the Economic Sustainable Development. He has been director of the laboratory of non-linear laser spectroscopy from 2005 to 2008. Actually he is coordinating the activity of development of optoelectronic technologies for subsea applications. His areas of interest include 3D laser imaging, optical wireless communications and subsea positioning.

Ralph R. Martin obtained his PhD in 1983 from Cambridge University. Since then he has been at Cardiff University, as Professor since 2000, where he leads the Visual Computing research group. He is also a Guest Professor at Tsinghua and other universities in China, and Director of Scientific Programmes of the One Wales Research Institute of Visual Computing. His publications include over 200 papers and 12 books covering such topics as solid modelling, surface modelling, reverse engineering, intelligent sketch input, mesh processing, video processing, computer graphics, vision based geometric inspection, and geometric reasoning. He is a Fellow of the Learned Society of Wales, the Institute of Mathematics and its Applications, and the British Computer Society. He is on the editorial boards of several journals, including Computer Aided Design, Computer Aided Geometric Design, and Geometric Models.

Baogang Wei was born in Shenyang, China. He received the M.S. degree in computer software and the Ph.D. degree in computer application from Northwestern Polytechnical University, China, in 1993 and 1997 respectively. He worked as a post-doctor at Zhejiang University, China from October 1997 to September 1999. Since 1999, he has been a member of Chinese Association for Artificial Intelligence. He is currently a professor at College of Computer Science and Technology of Zhejiang University. So far, he has published more than 40 papers in international conference proceedings and journals. His main research interests include artificial intelligence, pattern recognition, image processing, machine learning, digital library, and information and knowledge management.