Text Location in Uneven Illumination Images Based on Homomorphic Filtering and Color Distribution

F.C. You
College of Information Engineering
Beijing Institute of Graphic Communication
Beijing, China

Q. Li
College of Information Engineering
Beijing Institute of Graphic Communication
Beijing, China

W.B. Bu
College of Information Engineering
Beijing Institute of Graphic Communication
Beijing, China

***Abstract*—Aiming at the problem that uneven illumination in natural scene image has serious interferences on accurate text location, a new method based on homomorphic filtering and color distribution is presented. The homomorphic filtering is to reduce illumination changes and enhance the edge details. Utilizing the color distribution, the text distribution is prominent and easy to get the rough location. And the prior knowledge is needed to obtain the final text location. Experiments prove the validity and the better flexibility of this method.**

Keywords-uneven illumination; text location; homomorphic filtering; color distribution

I. INTRODUCTION

As the important information sources, accurate location and recognition of text in natural scenes have directly guiding significance in image understanding. But the uneven illumination caused by the camera flash leads to the low accuracy and extraction rate of text location. So it has been widely concerned and studied. In the field of text location and recognition, the images sets and evaluation standard of ICDAR (International Conference on Document Analysis and Recognition) have been widely accepted.

At present, the main methods are based on edge, texture and connected component. The combination method and the machine learning method are also applied to text location [1, 2]. But, for the text with complicated backgrounds, the false edges seriously interfere with text location base on edge; the complicated textures have a negative influence on location based on texture; for the image with low contrast and the text with different colours, the method based on connected component cannot work well. And the uneven illumination makes it worse.

The text location in uneven illumination image is difficult because it makes parts of text strokes be with false color and submerge in background. The serious interference of the identical text structures in complicated background is also difficult to deal with [3]. A method based on homomorphic filtering and color distribution is presented to locate the text. The logarithmic transformation makes the Fourier transformation of two components additive. The homomorphic filtering achieves the image enhancement [4, 5]. And the color distribution method and the restriction of prior knowledge are utilized to complete the text location [6, 7].

II. DESIGN AND ANALYSIS OF TEXT LOCATION METHOD

A. Homomorphic Filtering

The irradiance-reflection model can be applied to the uneven illumination image f(x, y)(0<f<∞). The f(x, y) consists of the irradiance of light source i(x, y)(0<i<∞) and the amount of reflection r(x,y)(0<i<1). The relationship among them is:

$$f(x,y) = i(x,y)r(x,y) \tag{1}$$

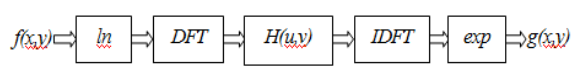


FIGURE I. HOMOMORPHIC FILTERING

The process of homomorphic filtering is shown in Fig.1. The irradiation component of images is characterized by slow varying in space domain, but the reflection is characterized by mutations especially in the linked parts. Aiming to make the two components separable, the logarithmic transformation is used, then Eqn.(1) becomes:

$$\ln f(x,y) = \ln i(x,y) + \ln r(x,y) \tag{2}$$

The Fourier transformation of Eqn.(2) is:

$$\begin{gathered}\Im[\ln f(x,y)] = \Im[\ln i(x,y)] + \Im[\ln r(x,y)] \\ \text{or: } Z(u,v) = F_i(u,v) + F_r(u,v)\end{gathered} \tag{3}$$

Z(u,v) is processed by the homomorphic filter H(u,v):

$$\begin{aligned} S(u,v) &= Z(u,v)H(u,v) \\ &= F_i(u,v)H(u,v) + F_r(u,v)H(u,v) \end{aligned} \tag{4}$$

The Fourier inversion is applied to Eqn.(4):

© 2015. The authors - Published by Atlantis Press

$$\Im^{-1}[S(u,v)] = \Im^{-1}[F_i(u,v)] + \Im^{-1}[F_r(u,v)]$$
$$s(x,y) = i'(x,y) + r'(x,y) \qquad (5)$$

Then the result of enhancement image is:

$$\begin{aligned} g(x,y) &= \exp[s(x,y)] \\ &= \exp[i'(x,y) + r'(x,y)] \\ &= i_0(x,y) r_0(x,y) \end{aligned} \qquad (6)$$

The cross-section of one homomorphic filter is shown in Fig.(2). The filter role is to influence the high and low frequency of Fourier transformation. In order to lessen the low frequency (irradiance) and enhance the high frequency(reflection), the rule are γ_H>1andγ_L<1. After that, the uneven illumination is lessened and the edge details are enhanced.

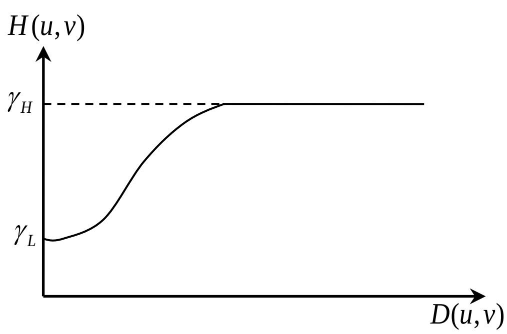
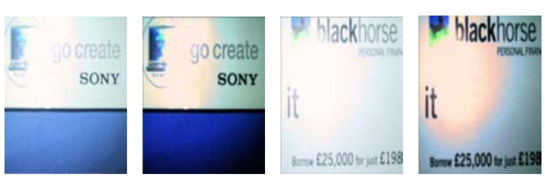


FIGURE II. CROSS-SECTION OF ONE HOMOMORPHIC FILTER

In actual processing, the color image is divided into 3 layers according to 3 channels. And each channel is processed by above steps. The homomorphic filter is:

$$H(u,v) = (\gamma_H - \gamma_L)[1 - \exp(\text{-}cD^2/\mathrm{D}_0^2)]$$
$$D(u,v) = \sqrt{(u - Height/2)^2 + (v - Width/2)^2} \qquad (7)$$

In Eqn.(7), c is used to control the slope, *Height* and *Width* are the image's height and width in processing of Fourier transformation separately. The uneven illumination images are chosen from the ICDAR scene sets. The γ_H is 2 and γ_L is 0.5. The effect of homomorphic filtering is shown in Fig.3. Fig.3-1 and Fig.3-3 are the original images, the others are enhancement images.



3-1 3-2 3-3 3-4

FIGURE III. IMAGES AFTER HOMOMORPHIC FILTERING

B. Obtaining the Color Distribution

After the homomorphic filtering, the edge details in dimmer region become clear, and the differences between text and background become larger. Generally, the text has the single color background, which has the obvious color distribution than its local backgrounds. The calculation of color distribution is as follows.

In RGB color space, the Euclidean distance of two points of image *I* is:

$$E_{(a_1,b_1),(a_2,b_2)} = \sqrt{\sum_{R,G,B} (I_{(a_1,b_1)} - I_{(a_2,b_2)})^2} \qquad (8)$$

The point (i, j) value of color distribution image S is the mean of E in its 8 neighborhoods:

$$S(i,j) = \frac{1}{8} \sum_{x=i-1}^{i+1} \sum_{y=j-1}^{j+1} E_{(x,y),(i,j)} \quad (x \neq i, y \neq j) \qquad (9)$$

It assumes that point (i, j) is the center of one window with the size w×w pixels and the density of color distribution Dens of point (i, j) is the mean of color distribution S of the widow:

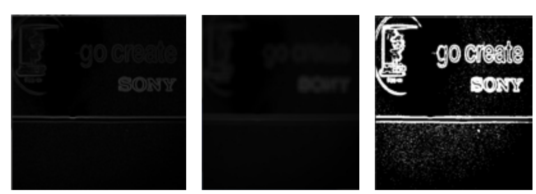
$$Dens(i,j) = \frac{1}{w^2} \sum_{x=i-w/2}^{i+w/2} \sum_{y=j-w/2}^{j+w/2} S(x,y) \qquad (10)$$

The window width w depends on the text size and the Dens reaches to its largest when w is close to the text size. However larger or smaller the window size is, the strongest response value of Dens appears in the center of single text region.

Because of the color distribution of complex textures, it is necessary to lessen the non-text color distribution. The method is utilized to obtain the reference image T:

$$T(i,j) = \begin{cases} 255 & S(i,j) \geq threshold \\ 0 & S(i,j) < threshold \end{cases} \qquad (11)$$

Because of the further restriction of prior knowledge, the threshold selection is not strict. In Eqn.(11), the threshold is the 1.5 times of the mean of image S. Here w is 20, and the S, Dens and T images of Fig.3-2 are shown in Fig.4.



a b c

FIGURE IV. COLOR DISTRIBUTION S (A), DENSITY IMAGE DENS (B) AND THE REFERENCE IMAGE T (C)

C. Rough Text Location

There is no much color distribution around the text region, which provides the foundation of rough text location. Here are the details of the method: (1) centering on the maximum value of color distribution image S, the region that has the color distribution in reference image T is surrounded by the smallest rectangle; (2) the maximum value is deleted, and a new maximum value is selected to repeat step 1; (3) if the maximum value is smaller than the threshold, the process is finished. In this process, the steps are executed in sequence, regardless of the overlapping of the rectangles. The effect is shown in Fig.5-a.

Then the surrounded rectangles are processed as follows: if the area of one surrounded rectangle is more than 60 percent of the bigger one, the bigger one is removed, or the surrounded one is removed. The effect is shown in Fig.5-b.

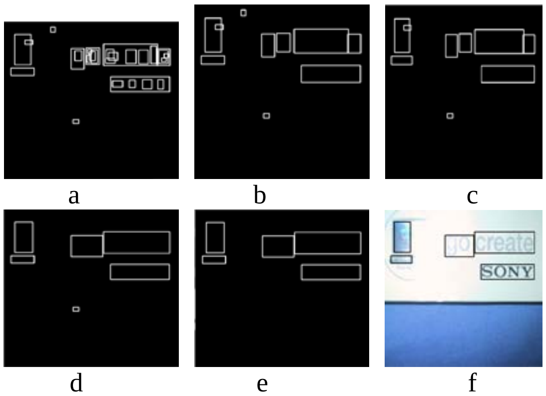


FIGURE V. PROCESS OF TEXT LOCATION

D. Restriction of Prior Knowledge

Because of greater color distribution of textures, the non-text rectangles should be removed. Here are the steps: (1) the rectangle has size requirement, which the width is more than 10 pixels and no more than 40 percent of image's width, the height is more than 10 pixels and no more than 80 percent of image's height, and the aspect ratio is restricted in the scope [0.2, 12]; (2) the rectangle area is more than 10 pixels and no more than 10 percent of the image area; (3) identical to the edge density, the density of reference image DensT is defined as:

$$Dens_T = \frac{1}{w \cdot h} \sum_{i=1}^{h} \sum_{j=1}^{w} T(i, j) \tag{12}$$

In the Eqn.(12), w and h are the rectangle's width and height, respectively. The $Dens_T$ value is more than 10, or the rectangle is non-text region. After the processes above, the effect is shown in Fig.5-c; (4) the overlapped and adjacent rectangles are merged, which is shown in Fig.5-d; (5) identical to the edge image, the x-axis projection of text region of the rectangle in reference image has the obvious wave crest and trough, while the non-text region's projection is relatively stable. The variance of the projection curve is utilized to do further sieve. The variance σ^2 is defined as:

$$\sigma^2 = \frac{1}{w \cdot h^2} \sum_{i=1}^{w} [hist(i) - mean]^2$$
$$mean = \frac{1}{w} \sum_{i=1}^{w} hist(i) \tag{13}$$

In Eqn.(13), *hist(i)* is the value of projection curve at the point i. The w and h are the rectangle's width and height, respectively. The variance of the text regions is more than 0.04. The location effect is shown in Fig.5-e and the location in original image is shown in Fig.5-f.

III. Results and Analysis

The software codes are written in C++ language in VS2010 development environment, and the images are chosen from data set of ICDAR. Here are some location results:

a b c

FIGURE VI. SOME LOCATION IMAGES

The results show that the new method can locate the text in uneven illumination images. But there are some further research points: (1) if the strokes submerge in the pollution of irradiance, the text region color becomes false color under the strong flash, thus the presented method has its scope; (2) the homomorphic filtering needs to be studied further so that the enhancement images are more suitable for this method.

IV. Conclusions

There are some uneven illuminations in natural scene images, which interfere with the text location. A new method based on homomorphic filtering and color distribution is presented to solve it. The results of location show that this method can locate the text in uneven illuminations effectively. As the threshold of color distribution is not strict, this method has better flexibility. But it has its application scope, which needs to be further studied.

Acknowledgements

This paper is supported by Scientific Research Key Project of Beijing Municipal Commission of Education (KZ201210015015) and Beijing Natural Science Foundation (4122026).

References

[1] S.L. Gu, Text Location and Extraction in Natural Scene Images, December 2008, Control theory and control engineering, Shanghai Jiao Tong University, Shanghai, China.

[2] J. Jin, X.J. Ping, T. Zhang & M.G. Ping, Survey of text localization techniques in images. *Application Research of Computers*, 24(6), pp. 8-11, 2007.

[3] S. Chen, Text location method in natural scene and its application, 3 June 2011, SIP, South China University of Technology, Guangzhou, China.

[4] Maria, P., et al, *Image Processing, The Fundamentals*, China Machine Press: Beijing, pp. 116-118, 2005.

[5] Rafael C.G., Richard E.W., *Digital Image Processing, Second Edition*, Publishing House of Electronics Industry: Beijing, pp152-154, 2007.

[6] H.C. Zhou, Q. Liu, Y.N. Wang, Text location in natural scenes based on color distribution analysis. *Computer Engineering*, 36(8), pp. 197-199+202, 2010.

[7] W.W. Ou, P.M. Zhu & C.P. Liu, Text location in natural scene. *Journal of Chinese information processing*, 18(5), pp. 42-47+63, 2004.