MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

A Blanket Binarization Method for Character String Extraction

Hiromi YOSHIDA
Kobe University
070w706w@stu.kobe-u.ac.jp

Naoki TANAKA
Kobe University
ntanaka@maritime.kobe-u.ac.jp

Abstract

In this paper, a binarization method based on fractal dimension for character string extraction is proposed. In character extraction from a scene image, a major problem is how to deal with much different type of characters in a complex background. The proposed method can obtain multiple threshold values which are correspond to each character regions by detecting the stable intervals of fractal dimension FD. The stable interval is a relatively low and flat valley of the FD which indicates the binarized image has the stable connected regions, and therefore fine character regions have been appeared. The character regions may contain some noise and has conflictions between the regions derived with another threshold values. We call these character region as a "Candidate Character Region Images"(CCRI), and will be processed by noise-reduction consists of two steps. After that, CCRI are integrated into one binarized image as output image through the contention resolution process. We show the performance of the proposed method by comparing Niblack's method as a local method and Otsu's method as a global method on the dataset provided at ICDAR 2003.

1 Introduction

A scene image which contains characters is still considered to be one of a challenging target for character extraction because of its wide layout variations and complexity of background. Furthermore, there exist some specific problems which make it hard to extract, i.e. lighting condition, perspective transformation and insufficient resolution. Actually, the performance of the winning method of the "Text Locating Competition" held at ICDAR 2005 remains no more than 55 percent of correct and recall rates. To improve the character extraction performance of an image, binarization proses is thought to be one of the most critical steps. A number of binarization methods have been proposed and Trier[1] evaluated 15 binarization methods as promising by the procedure called goal-directed evaluation, and showed that Niblack's method[2] has the best performance as a local adaptive method and Otsu method[3] is the best in global methods. In these 15 methods, the threshold value is selected based on local or global statistical information of the gray level such as a histogram. However the gray level statics based method can evaluate neither the shape nor the relation among the binarized connected components, the degree of the quality of a connected component as a character region can not be derived.

And also a scene image generally consist of much kind of the regions such as high or low texture, solid area, and has a contrast perturbation, the binarization for character extraction should have the ability to qualify the fineness of the binarized character regions. Yoshida[4] proposed a binarization method based on fractal dimension which is realized evaluation of the character region quality. This method utilizes originally single threshold value, and it working well for the image which contains single character string. However for the image with more than one character strings that have different gray level, we must take plural threshold values. The character images obtained by these plural threshold values may contains some noises and also there are character region conflictions with character images. In this paper, we describe the way how to get plural threshold values, reduce noises and solve the region conflictions. The proposed method has three major processes. First is an evaluation process based on FD, and it generates plural binary images named CCRI(Candidate Character Region Images). Second is a noise reduction process which consists of two steps. One is based on area ratio, the other is on contrast. Third is a confliction solving process between CCRI. Finally single integrated image is generated through this process. In following chapters, FD and binarization method based on it is described. And then, we show the algorithm of the proposed method and experimental result.

2 Fractal dimension

Fractal dimension(FD) has been proposed as a method of texture analysis named "Blanket method" by SHUMEL[5] in 1984. The range of the dimension is from 2 to 3 and is obtained from the equations 1,2,3 and 4.

$$U_\epsilon = \max\{U_{\epsilon-1}(i,j)+1, \max_{|(m,n)-(i,j)|\leq 1} U_{\epsilon-1}(m,n)\} \quad (1)$$

$$b_\epsilon = \min\{b_{\epsilon-1}(i,j)-1, \min_{|(m,n)-(i,j)|\leq 1} b_{\epsilon-1}(m,n)\} \quad (2)$$

$$A(\epsilon) = \frac{\sum_{i,j}(U_\epsilon(i,j) - b_\epsilon(i,j))}{2\epsilon} \quad (3)$$

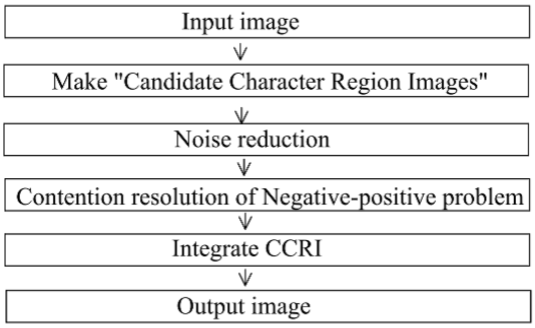
$$A(\epsilon) = F\epsilon^{2-D} \quad (4)$$

where ϵ is the number of "blanket". Noviant[6] proposed the optimal range of the FD and adapts it locally to an image and get a local fractal dimension (LFD) image. LFD image has a feature that brightness is proportional to the frequency of a region.

3 Proposed algorithm

3.1 Summary of algorithm

Summary of proposed method is shown in the following chart.



3.2 Candidate Character Region Images

Blanket method is originaly applied to a gray scale image, but [7][8][9] show that it also has the ability to evaluate the performance of the thresholding for a segmentation application by applying to a binary image. The method[8] can get a fine quality binarized images for the character extraction by evaluating the stability difference of FD against binarizing threshold change between character region and background. The algorithm is obtained by following steps.

- $\boldsymbol{Step1}$: Binarize the input image $\boldsymbol{I(x)}$ with the threshold values from $\mathbf{0}$ to $\mathbf{255}$, and obtained the 256 binarized images $\boldsymbol{I_{b_i}(x)}$ $\boldsymbol{(i = 0, 1, ..., 255)}$ respectively.

- $\boldsymbol{Step2}$: The $\boldsymbol{FD(i)}$ values can be calculated on $\boldsymbol{I_{b_i}(x)}$ images by the Blanket method.

- $\boldsymbol{Step3}$: We treat the $\boldsymbol{FD(i)}$ to be a function respect to $\boldsymbol{i}$. Then, smooth the $\boldsymbol{FD(i)}$ to remove noise of the function, and find the stable interval of $\boldsymbol{FD(i)}$ by taking the first derivation.

- $\boldsymbol{Step4}$:In the highest stable interval, detect the minimum $\boldsymbol{FD(i)}$(which is not smoothed) as the threshold value $\boldsymbol{\theta}$.

- $\boldsymbol{Step5}$: Final binarization Image $\boldsymbol{B(x)}$ can be obtained using the threshold value $\boldsymbol{\theta}$.

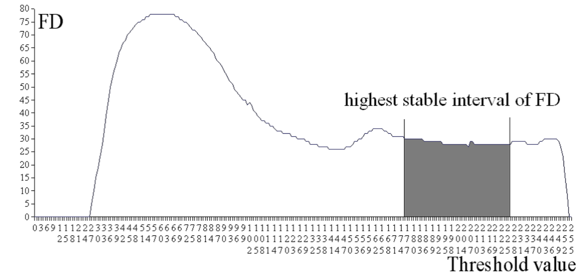
At step3, a stable interval is defined as follows: differentiate a smoothed $\boldsymbol{FD(i)}$ with respect to $\boldsymbol{i}$, and count the number of $\frac{\delta}{\delta i}\boldsymbol{FD(i)}$ which is equal $\mathbf{0}$ or nearly equal to $\mathbf{0}$ until it turns larger than $\mathbf{0}$. $\frac{\delta}{\delta i}\boldsymbol{FD(i)}$ is calculated from expression 5.

$$\frac{\delta}{\delta i}\boldsymbol{FD(i)} = \boldsymbol{FD(i+1)} - \boldsymbol{FD(i)} \qquad (5)$$

About the blanket number ϵ, Noviant showed that appropriate range of ϵ is from 34 to 53[6]. Referring to this ,we use 44 as a blanket number which is a median of the appropriate range and the FD value is set to integer of (0,100) by a quantization. The sample of the input image, binarizing result image , and graph of FD are shown at Fig.1.

This algorithm selects originaly only one threshold value with the highest stable interval of FD at step 4

(a)Input image (b)Output image



(C) FD graph

Figure 1. Graph of $\boldsymbol{FD(i)}$

and 5, however, $\boldsymbol{FD(i)}$ plural stable intervals are observes at step 3. Since some kinds of scene image have more than one character strings which may have differnt font type, luminance, size, and so on, each character strings can appear in another binarized images by different threshold values respectivly. Samples of this are shown at Fig.2.

(a)Input image (b)stable interval 1(highest) (c)stable interval 2

Figure 2. The example of multiple binarized images from a stable intervals of FD

So, our new method utilizes all of threshold values correspond to the stable intervals of FD and generate the binary images. We also generate inversed images, i.e., negative image in order to correct the reversal of character region and background each other(we call this as "Negative-positive problem"). And obtained images are named "Candidate Character Region Images"(CCRI) in this paper. The sample of input image and CCRI are shown at Fig.3.

In Fig.3, CCRI 1 and CCRI 2 is a sample of the Negative-positive problem, as well as CCRI 3 and CCRI 4. After the noise-reduction process, they will be integrated into one binary image by region confliction solving process.

3.3 Noise reduction process

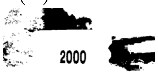
We introduce a noise reduction process to remove the non-character regions. For example, though one of

(a)Input image

(b)CCRI 1

(c)CCRI 2

(d)CCRI 3



(e)CCRI 4

Figure 3. The example of CCRI

widely known binarization method of Niblack's method [2] tends to mmake excessive extraction, it achives the top performance in a comparison experiment by a noise reduction process[1]. So we carry out two steps of the noise reduction. At first step, we use "Bounding Box"(BB) for removing the noise regions in the point of the aspect ratio of a region. At second step, we use the "Edge Contrast Feature" to remove noises whch have no strong edge in contrast on boundary.

3.3.1 Noise reduction 1 "Bounding-Box"

"Bounding-Box" (BB) is a rectangle that encloses the connected-component(CC)s on a CCRI. The noise reduction process on the height, width, aspect ratio and the size of BB are widely used to refine the image in many related researches. Same as in [10][11], we have to take some heuristics to set the parameters of the noise reduction step. And we have to pay attention that these parameters could be different depending on the nature of the script. The parameters are obtained by equation 6.

$$\begin{aligned} \frac{width(BB)}{width(CCRI)} &\leq \frac{1}{5} \\ \frac{height(BB)}{height(CCRI)} &\leq \frac{1}{5} \\ area(BB) = width(BB) \times height(BB) &\geq 8 \\ \frac{1}{8} \geq \frac{width(BB)}{height(BB)} &\geq 8 \\ \frac{area(CC)}{area(BB)} &\geq 0.15 \end{aligned} \tag{6}$$

where *width()*, *height()* and *area()* mean width, height and whole of number of pixels respectively. CCs with BB which is less than threshold range are rejected as a noise.

3.3.2 Noise reduction 2 "Edge Contrast Feature"

A character string in a scene image are designed intentionaly to be easy to read , so the included character region has high contrast against to the background. Zhu et al.[10] said that text CCs are often 'highly closed' by edge response, and they used the ratio of edge to the border to extract charcte string regions. This "Edge Contrast Feature" is presented at equation 7

$$\begin{aligned} &Feature\_Edge = \frac{Border(CC) \cap Edge(Picture)}{Border(CC)} \\ &Edge(Picture) = \\ &Canny(Picture) \cup Sobel(Picture) \end{aligned} \tag{7}$$

where *Canny(Picture)* and *Sobel(Picture)* are the normalized Canny and Sobel edge of an input image respectively. And *Border(CC)* is the border pixels of the CC on CCRI. By threshold processing of $Feature\_Edge$, the CCs with small edge closure are rejected as noises.

3.4 resolve of confliction

After the noise reduction, the "Negative-positive problem" in CCRI are solved and integrated into one binary image. As shown in Fig.4 (b), a background area which enclosed by a charcter region has a similar feature of a character region and so it hard to distinguish this kind of background from character.

(a)cCC1(character region)

(b)cCC2(noise region)

Figure 4. CCRI in a Negative-positive problem

This is the Negative-positive problem mensioned above,i.e., a relation of a charcter and its holes. However, these regions have the following property. One is "smaller than it", and surrounded by the other. So this problem can be solved by comparing the size and location of them by using a morphological operation. The process is obtained following.

$$\begin{aligned} ResultCC &= MAX(cCC1, cCC2, ..cCCn) \\ cCCl, cCCm &: CC'l \cap CC'm \neq 0 \\ CCn' &= (CCn \oplus r_3 Bsquare) \end{aligned} \tag{8}$$

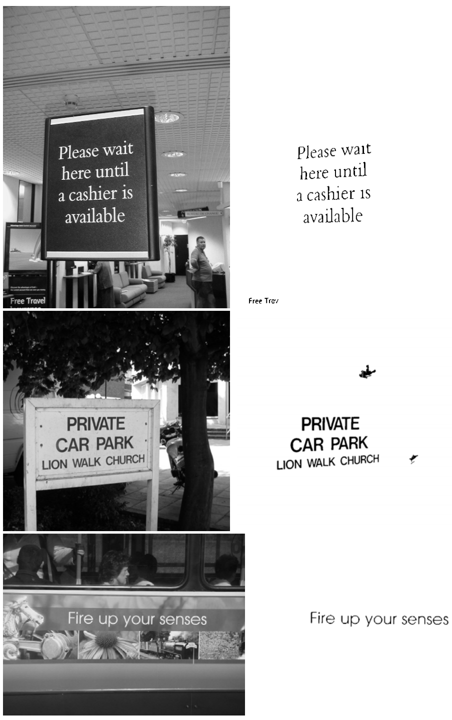
$r_3 Bsquare$ means square shape structure element with a width of 3 pixels. $\oplus$ represents a dilation operation of the mathematical morphology. This process compares CC on CCRI which are overlapped each other when they are dilated, and CCs with largest region(which is *ResultCC*) are integrated into one output image as a resulting image. Through this process, Negative-positive problems are solved.

4 Experiment result and conclusion

Binarized images by proposed algorithm are shown at Fig 5. And processing results by Otsu's method and Niblack's method are also shown for comparison in Fig 6. The former is belonging to a local method, and the latter is global one.

(a)Otsu's method (b)Niblack's method

Figure 6. The examples of comparison



(a)Input image (b)Output image

Figure 5. The examples of output image

The thresholding method we introduce here generates several binary images correspond to each character string, remaining region conflictions and a noise problem. However, by using the noise reduction process and confliction resolving process, these problems can be solved successfully. Result of experiment on the dataset of ICDAR2003 shows that the proposed method can be a promising binarization step of character detection method from a scene image. Furthermore, it shows the hight performance as a character extraction method. We are now studying to develop a character extraction method with a performance evaluation process.

Acknowledgement

This work was supported by KAKENHI(c)22500154

References

[1] O.D. Trier, A.K. Jain "Goal-Directed Evaluation of Binarization Methods," IEEE Trans. Pattern Analysis and Machine Intelligence, vol7 no.12, pp.1191-1201, July 1995

[2] W. Niblack,An Introduction to Digital Image Processing, pp.115-116. Prentice Hall,1986

[3] N.Otsu,"A threshold selection method from gray-level histograms," IEEE Trans, Systems, Man, and Cybernetics, vol. 9 no.1, pp.62-66, 1979.

[4] H.Yoshida and N. Tanaka,"A Binarization method for the Character Extraction from Scene Image by the Blanket method", IIEEJ Journal, Vol.39, No4, pp.463-472, 2010

[5] S. Peleg, J. Naor, R. Hartley, D. Avnir, "Multiple resolution texture analysis and classification", IEEE Trans. Pattern Analysis and Machine Inteligence, vol6 no.4, pp.518-523, July 1984

[6] S. Noviant,Y. Suzuki, J. Maeda, "Optimumestimation of local fractal dimension based on the blanket method", IPSJ Journal, vol.43, no.3, pp.825-828,2002

[7] H.Yoshida and N. Tanaka,"A Binarization method for Crack Detection in a Road Surface Image with the Fractal Dimension", Proceedings of MVA 2009 IAPR Conference on Machine Vision Applications, pp.70-73, 2009

[8] H.Yoshida and N. Tanaka,"A Binarization Method for a Scenery Image with the Fractal Dimension", Progress in Pattern Recognition,Image Analysis, Computer Vision, and Applications, LNCS, Springer Berlin / Heidelberg, Vol.5856, pp29-36, 2009

[9] H.Yoshida and N. Tanaka,"A New Binarization Method for a Sign Board Image with the Blanket Method", the First Joint Workshop on Pettern Recognition, IEEE,CJKPR2009,pp.845-848

[10] K.Zhu et al, M.Kimachi et al, "Using Adaboost to Detect and Segment Characters from Natural Scenes, Proc. Conference on Camera Based Document Analysis and Recognition,pp.52-59,2005

[11] Y. Wang, N.Tanaka,"Text String Extraction from Scene Image Based on Edge Feature and Morphology", Proceedings of the 8th IAPR International Workshop on Document Analysis Systems, 8th, pp.323-328, 2008

[12] H.Yoshida, N. Tanaka,"A Study on Signboard Image Identification with SIFT Features" Handout of 8th IAPR International Workshop on Document Analysis Systems, 8th, pp.10-14, 2008