Text Image with Complex Background Filtering Method Based on Harris Corner-point Detection

Fanfeng Zeng
Department of Computer Sciences, North China University of Technology, Beijing, China
Email: zengfanfeng@sina.com

Guofeng Zhang and Jin Jiang
Department of Computer Sciences, North China University of Technology, Beijing, China
Email: zhgf0618@163.com, jiangjin0127@163.com

***Abstract*—In order to solve the problem of text image containing the complex background in which the text of image cannot be read clearly and intuitively, this paper proposed a filtering method based on the Harris corner-point detection. The experimental results show that the algorithm can filter the background which contains text image commendably, making it possible for visually reading.**

***Index Terms*—corner-point detection, text feature, text location, layout reconstruction, background filter**

I. Introduction

In recent years, with the continuing development of the digital image processing technology, information transfer and expression more and more rely on images. Image can not only carry the scene information which words cannot express, but also contain the key words information which can explain it. Whereas the double functions of it challenges the computer ability of processing complex images. For instance, a beautiful book cover can perceptual attract readers attention, but it needs complicated operation to quickly obtain the title name, author name and other information.

So far, many text location algorithms[2] have been put forward for this kind of image. The typical algorithms are: the connected domain algorithm by the use of text arrangement features; the image segmentation algorithm for texture by treating the text image as a special texture; the edge algorithm by searching the edge arrangement rules of the Text and background supposing there are strong edges between them; the color localization algorithm using color clustering for image decomposition and so on. However, in view of the different text sizes and color variety of complicated background[2,7,9], a single algorithm cannot solve the problem completely. It needs to combine above algorithms to achieve an ideal effect.

Text contains abundant corner-points[4,8,16], arrange regularly and have apparent partial feature. These three characteristics are not directly related with text sizes and colors. The text image with complex background filtering method proposed in this paper is based on the Harris corner-point detection[1]. First, make corner-point detection for gray image, detecting the corner-points for words and noises in the image. Then, select the words corner-points by utilizing the features that the distribution of words corner-points is more regular while the noises corner-points are discretely distributed. Finally, locate text areas using the local region features of words and acquire the contour coordinates of text by carrying on binary processing for the valid text area, completing the complicated background filtering after generating the new image which only contains words by the use of layout reconstruction method. Experimental results show that this method can realize the background filtering for images with different color backgrounds, color words and text sizes. And the filtering accuracy is commendable.

II. Corner-Point Detection Theory

At present, since the understanding of the diagonal points is varied, there are many different descriptions for it. This paper defined the corner-point as: the point where the border curvature is high enough, from which any movements in any directions can cause great gray changes of the image[6,16,18].

Corner detection theory after decades of development, the algorithm of it has been widely used for corner detection based on the geometric characteristics. The main algorithm: Kitchen put forward the local gradient multiplied by the gradient direction which changes to extract corner; Moravec[1] proposed the use of local area as operator; Smith.M and Brday.JM posed SUSAN[4], based on a window area, using the point and the change of local area of the point to judge a corner-point; Harris algorithm[1,4,15] is proposed by C.Harris and M.Stephens, which is based on the point feature extraction operator of Moravec algorithm, using the matrix M which is autocorrelation coefficient associated. The eigenvalues of M matrix are the first-order curvature of autocorrelation function. If the two eigenvalues are comparatively large, it shows that in this point, the curvature of the autocorrelation function on two orthogonal directions is bigger, which is define as the corner-point.

Harris algorithm[1,3], which broke through the limitations of the Moravec algorithm, can get good repeatability and high detection accuracy. Although it

doi:10.4304/jsw.8.8.1827-1834

needs large amount of calculation, it has been widely used because of its better detection effect. By the judgment based on the above analysis, we have chosen the Harris algorithm for corner-point detection.

A. Harris Corner-point Detection Algorithm

The difference between Harris algorithm[1] and Moravec algorithm is that the partial autocorrelation measurement results. Harris can get gradation changes in all directions. However, the Moravec algorithm[1] is only calculated in a series of discrete moving direction. In order to facilitate the calculation, we detected only on the horizontal and vertical directions of the image. If a point in both horizontal and vertical direction has greater gray-scale variation, this point is the corner points.

Firstly, obtain the first order differential Ix and Iy using the horizontal and vertical difference operator, as the formula 1(a) (b) shows.

$$\begin{matrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{matrix} \qquad \begin{matrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{matrix}$$

(a) horizontal difference operator (b) vertical difference operator Formula 1

Then we get matrix M:

$$M = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}, \quad \text{and} \quad I_x^2 = I_x * I_x,$$

$$I_x^2 = I_x * I_x \qquad \textit{Formula 2}$$

In order to eliminate the effect of noise, introduce the Gaussian smoothing filtering. Matrix M' is obtained after Gaussian filtering by the use of GAUSS template.

$$GAUSS = \exp(-\frac{x^2 + y^2}{2\sigma^2}),$$

$$M' = GAUSS \leftarrow \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \qquad \textit{Formula 3}$$

The calculation formula of corner-point amount is defined as:

$$Cim(x, y) = \det(M') - k * trace(M') \quad \textit{Formula 4}$$

$\det(M')$ ——the determinant of the matrix M',

$trace(M')$ ——trace of the matrix M',

k ——adjustable parameter.

Use formula 4 to calculate the amount of corner-points $Cim(x, y)$ point by point, set a threshold value T, if $Cim(x, y) > T$, then mark point (x, y) as the corner-point.

The focus of this article is the Chinese text images, due to the difference between English characters and Chinese characters features, in order to achieve the purpose of the corner-point detection, the above parameters should be adjusted.

B. Harris Algorithm Flowchart

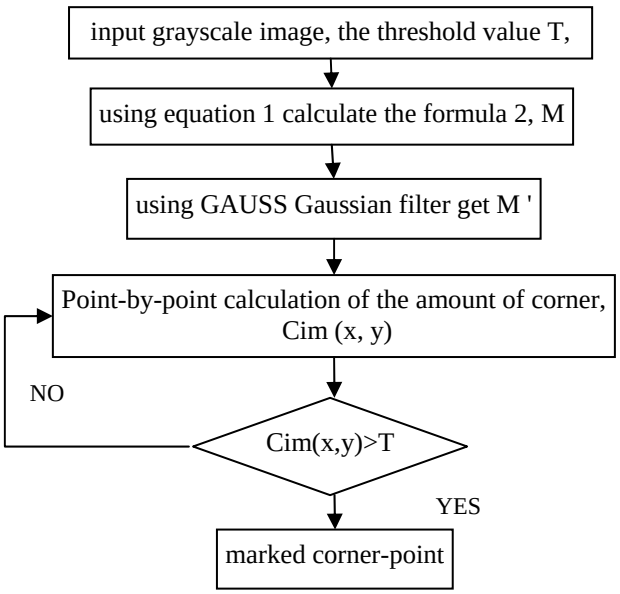

Figure 1 Harris algorithm flow chart

III. THE CHARACTERISTICS ANALYSIS OF TEXT IMAGE

Text image is different from the other image[13] for its distribution of pixels with certain regularity. For example, a line of text between lines within each line between words have obvious interval, and the height of the row is broadly consistent with the size of the text in each line, and these inherent characteristics are not because of the font size models and it will not change due to the image direction. These characteristics are reflected in the binary image as pixel distribution regularity between lines of black pixels and no black pixels between words.

The so-called complicated background image is that the characters in it cannot be segmented using the usual threshold method, where the character font, size, style, direction, alignment degree, color and background as well as texture is priori unknown. Therefore, it is necessary to analyze the character characteristics of the image if want to locate text in the complicated background. Although the words are embedded in the simple or complicated backgrounds, if only the words have certain significance to the image, it must been have some characteristics for identification.

The common characters characteristics are as follows[13]:

1. The text size should be large enough to ensure the words can be identified;

2. The amount of words which express certain meanings cannot be too few no matter they are Chinese or English. And the words should be arranged in a relatively fixed way, using vertical or horizontal distribution generally;

3. The certain contrast must be ensured for identification whether for people or the computer, so that distinguish can be made between text and the background;

4. The words of the same row or the same column should have the same characteristics, such as size, arrangement, color, etc.

The characteristics mentioned above do not depend on specific images and specific scene, which have certain commonality. This paper utilizes the above features of characters for corner-point selection and text area location, to ensure that the generality of this algorithm.

IV. Words Positioning

A. Algorithm Flow Chart

Text position[3], is to get the coordinates of the specific location of the text in the image, crucial step character recognition, in text retrieval, video retrieval and image processing, its accuracy directly determines the success or failure of the entire treatment. Therefore, there have been a lot of algorithms and has made certain achievements. But the conventional algorithms mainly deal with the simple layout of text images or fixed format text images. For irregular text images and complex background text image, the processing capabilities are to be further enhanced. In this paper, the text of the complex background image characteristics, combined with the corner detection algorithm is proposed in this algorithm.

As shown in figure 2:

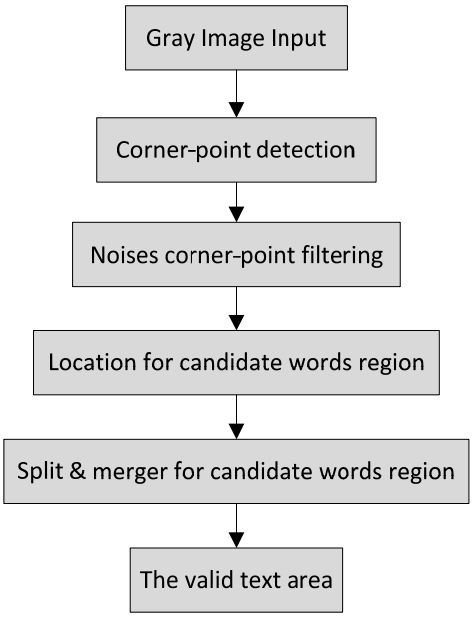

Figure 2 Algorithm flow chart

From Figure 2, we can conclude that the corner detection is the basis of the present algorithm, the accuracy affect the subsequent arithmetic processing. So for the different images, such as Chinese images and English image, the parameter setting for corner detection algorithm is particularly important. Additionally, a key step is the candidate character regions positioned for different information corner in the image, in which the characteristics are shown definitely different. So the use of the characteristics distribution of the text for the positioning of the text area is feasible. Of course, some similar text image noise on the positioning which is to be resolved by appropriate way is not excluded.

B. Filtering Noise Corner-point

By the definition above this article on the corner, if the point on the boundary and curvature high enough can be think as corner-point, object in the image, the background noise, and so can produce a corner-point. Therefore, in order to locate the text by corner-point, must be able to determine which corner-point is corner-point as text, which corner-point is the noise generated. That is, to classify them.

The corner-point obtained by corner detection includes two kinds of corner-point: words corner-points and noises corner-points. Words corner-point is obtained by edge and stroke detection of words. Noises corner-point is produced by the image background. Obviously, the corner-points produced from background are useless. The noise interference must be eliminated for locating the text information[10].

Due to the object or scene in background is uncertain, the produced corner-point distribution is of discrete and random; meanwhile, as the words are present as columns or rows arrangement, there are not obvious intervals between each line. Interval area generally will not produce a large number of corner-points. Thus, we can make the noises corner-points filtering using the above characteristics. The filtering rules formulated as follows:

Rule 1: Statistics of the number of corner-points of each row, roughly locate the coordinates of text line, and get the average width H of the Chinese text line in image.

Rule 2: As the height and width of text are broadly similar, set a moving window H*H as the filter window.

Rule 3: Set a threshold T1 for the number of window corner-point, calculate the number of the corner-points inside the window as n1, if n1<T1, argues that the corner-point inside the window is not the word corner-point; If n1> = T1, then mark corner-point inside the window as the word corner-point, and continue filtering.

Rule 4: Set a threshold T2 for the number of text line corner-point, calculate the number of corner-points from top to bottom inside H/2 height area as n2. If n2 <T2, set the corner-point within H/2 height area as noise corner-point; If n2>= T2, mark the corner-point within H/2 height area as word corner-point, and continue filtering.

Rule 5: Based on experimental data, a Chinese character generally needs more than 5 angular point, so set T1 > = 5; According to Rule 3, most isolated corner-points have been filtered, remaining some kind of words corner-points, generally set T2>=2*T1.

After the above processing, the retaining corner-points of the image are mostly the valid words corner-points, as shown in figure 3, (a) stands for the calculation of the original corner-point of the image, (b) stands for the filtered valid words corner-points.

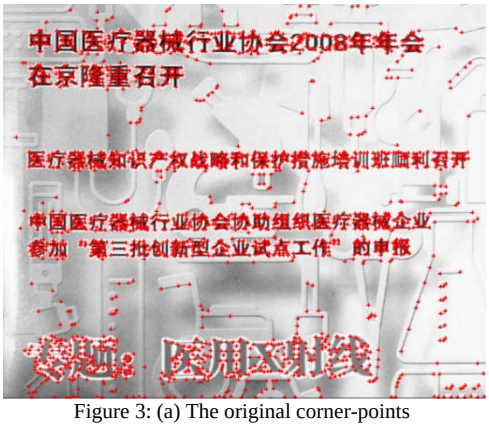

Figure 3: (a) The original corner-points

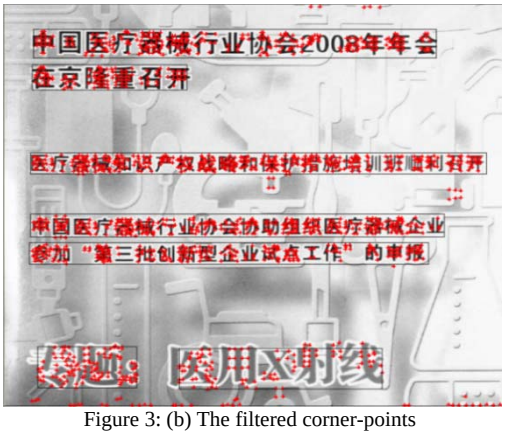

Figure 3: (b) The filtered corner-points

Chinese text relative to English characters have more corner point during the filtering of noise the focus may be appropriate to relax the limit of the number of focus, or because the conditions are too strict text corner point filter out, affecting subsequent text positioning and layout reconstruction.

C. Text Location

Text position[6,11,13] is to determine the specific coordinates of the text in the image. Text positioning is all text image processing work, because the text-only position to determine, to the subsequent extraction, identification, in order to know which part is which part of the text area background area, and then to filter background.

After detection and filtering for corner-points, the retained corner-points in image are identified as words corner-points. Using the above image characteristics of text analysis, we can locate the text.

First of all, calculate the number of corner-points of image line by line to get the y coordinates of the specific line of the text, obtaining multiple lines by this marking method. We define: the continuous text in the same line of text is called a text area. As there may be more than one text area in a line of text and the starting position of text lines are not certain, we need to calculate the number of corner-points of each text lines for marking the specific coordinates of each text areas. Define the structure of text area as follows:

```
struct sCharRect{
    // the x coordinate of upper-left corner in text area
    int iPosx;
    // the y coordinate of upper-left corner in text area
    int iPosy;
    int iWidth;// width of text area
    int iHeight;// height of text area
};
```

Text localization algorithm process is as follows:

(1) Using H acquired by section 3.3 rules 1 and T1 acquired by rule 5, set moving window width threshold as TW=H/2, corner-point amount threshold as NW=T1/2.

(2) Horizontally move x+=TW width from left to right orderly, count the number of the corner-points of the current text line within the scope of TW as N.

(3) Text area location marker iFlag = 0.

(4) Marking text area

```
IF  N<NW
  IF iFlag==0
    //not mark starting point
    Continue to move backward TW;
  ELSE
    //mark as the end of this text area
    iWidth = x + TW - iPosx;
    //Clear mark and began to mark the next text area
    iFlag = 0;
  ELSE
    IF  iFlag == 0
      iPosx = x; //text area marking begin
      iFlag = 1; //into the marking state
    ELSE
      //not mark the end point
      Continue to move backward TW;
```

(5) Repeat (2) (3) (4) to mark all the text areas of every text lines.

As shown in figure 3(b), by this method we obtain 6 text lines and 7 text areas where the sixth line contains 2 text areas.

D. Merge and Split the Text Area

After the above operation marked text area, due to some text corners and other similar text corner mixed together, will result in a text area contains noise region, or multiple small text area constitutes a large text area, which is obviously not line with our expectations, so the need for such a text area merge or split, possible zoning text relative accuracy independent, in order to improve the system correctness.

The specific[12,14] processing steps for merge and split:

(1) Calculate the average height of the text area mark *Avgh*, $Avgh = \sum_{i=0}^{i=Tot} iH \ / Tot$, Wherein, *Tot* for all markers obtained the total number of the text area, *iH* for the i-th of the height of the text area.

(2) The text area which is considered to be the normal text area of the text area is set as *AllRect*, otherwise it should be split into the composite text area set as *NeedRect.*

(3) Split each text area in the sets of *NeedRect*, based on the connected region which is considered to mark the corner-point in this area, marking them as different small connected regions from 0 to n.

(4) Mark the small regional location in this connected region and get the coordinates of the location of the small area, which are determined as the set of texts *AreaRect.*

(5) Repeat (3), (4) to get all undetermined text area *AreaRect* in sets of *NeedRect*

(6) In *AreaRect*, there may be existed some small area or isolated noise. So we need to judge the area size of each one in the *AreaRect* for determining the too small one as noise area and filtering it out directly.

(7) There may be some overlapped area in *AreaRect*, such as Figure 5 (a) shows. So it is necessary to split them according to the position coordinates of the various regions of the overlap portion, which mainly consists of the following situations:

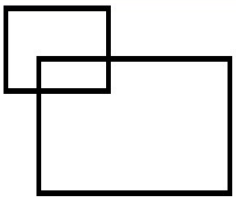

Figure 4 :(a) Corner nested

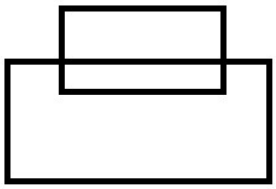

Figure 4 :(b) Edges nested

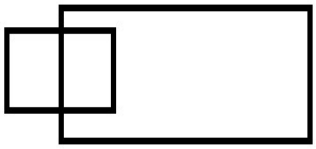

Figure 4: (c) Part nested

(8) Split a nested region. As results shown in Figure 5 (b), some independent text areas are acquired and stored in the sets of *AllRect.*

(9) Here, the sets of AllRect is all the text area of available.

Figure 5: (a) Region overlapped images

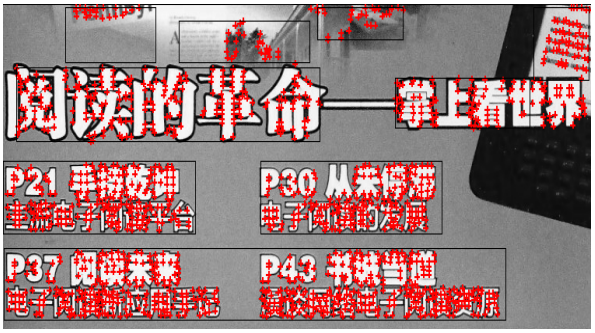

Figure 5: (b) After the split image
(The wireframe of the region is marked available text area)

V. Background Filter and Layout Reconstruction

A. Background Filter

For text image, the background image of it can the use its scenes to interpret the meaning of the image[12]. The text out of the image of the scene can still express the meaning of the image, which contains more important value is the text in the image. In order to make the text images more flexible for using in other environments, it must be able to position and extract text for clearly display. This article focuses on how to realize text positioning and filter it out of background clearly as intuitively reading the contents of the image.

After text area mark and location, it can realize the background filter for non-text area. However, the words of text area are still embedded in the background, if want to filter the background within the text area, the location for the accurate coordinates of the words outline are essential. As the text area mainly consists of words and the pixels of the words within the area are relatively obvious compared with the background pixel, we binarize the text of this area using global threshold binarization algorithm in this paper. The coordinates of the black pixel in the binary image are the specific coordinates for words, while the white parts are filter out as background.

Text area the background filtering algorithm flow chart shown in Figure 6:

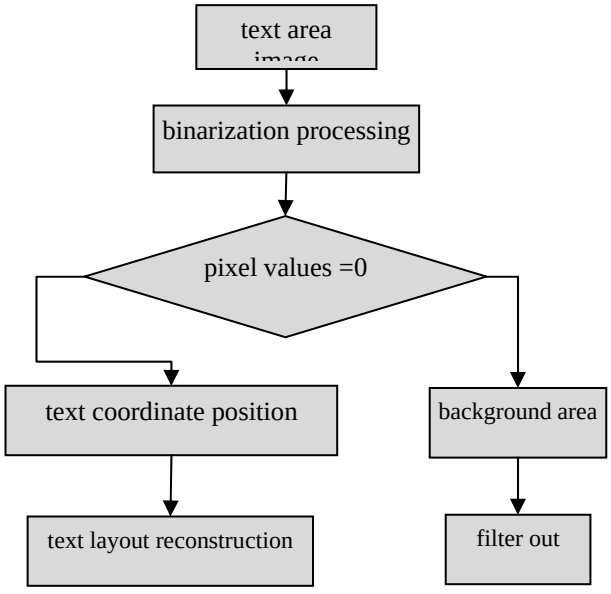

Figure 6: Algorithm flow chart

B. Text Layout Reconstruction

After the above analysis processing, first, the image is divided into two broad categories: text areas and non-text area, then the text area is divided into two sub-categories: a text coordinate position and the background area. Cut segmentation of text image this way, because our goal is filtered image in the complex background[15,17,18], so only need treatment to these four types of image area. That is, the position of the non-text area all as a background Filter out, Filter out the background area of the text area, where the rest is text area. This background images with

complex background filtering, but also retain the relative position of the text in the image does not change, easy to follow text layout reconstruction.

In order to realize the text layout reconstruction for image, we need to use the point coordinates of the black pixels in each binarized text areas to build a new image which is as the same size as the original image. Initialized the background as white and assigned the value of the black pixels in all text area of the original image directly to the corresponding points of the new image, realizing the text area reconstruction for the original image, not only restoring the original size of the image, but also keeping the format for original image. The method realizes the background filtering for the image with complex background and further retains the text information. The algorithm process is as follows:

(1) Binarize for all the marked text area, obtaining the binarized text area A1, A2... An.

(2) Newly built an image B, which is of the same size and format as the original image A.

(3) Initialize B as white background

(4) Selected image Ai, circularly search the black pixels of Ai.

(5) Assign pixels values of the points in image A to the points of image B.

(6)Repeat (4) (5) until all the text area in image are processed, completing the image reconstruction.

VI. Experimental Results

The algorithm are realized under Visual C++ 6.0 software environment，using C++ language. All the test are accomplished under Windows XP operating system, with Inter (R) Core(TM) 2 Duo CPU, 2G memory. In order to validate the validity of this algorithm, we selected magazine covers and book covers respectively for confirmatory tests in experiment and make a comparison with Hanvon OCR layout analysis module.

The experimental results are as follows table show:

TABLE 1
THE COMPARATIVE ANALYSIS BETWEEN THE ALGORITHM AND THE OCR LAYOUT ANALYSIS MODULE

Sample category	Algorithm	Accuracy for text locating	Conclusion
complicated background, hollow word, solid word	the algorithm of this paper	>95%	The locating accuracy for complex text image using the algorithm of this paper is much higher than the conventional layout analysis and text location algorithm.
	Hanvon OCR layout analysis	<50%	
complex background, different colors word	the algorithm of this paper	>95%	
	Hanvon OCR layout analysis	<50%	
complicated background, different font sizes	the algorithm of this paper	>95%	
	Hanvon OCR layout analysis	>70%	

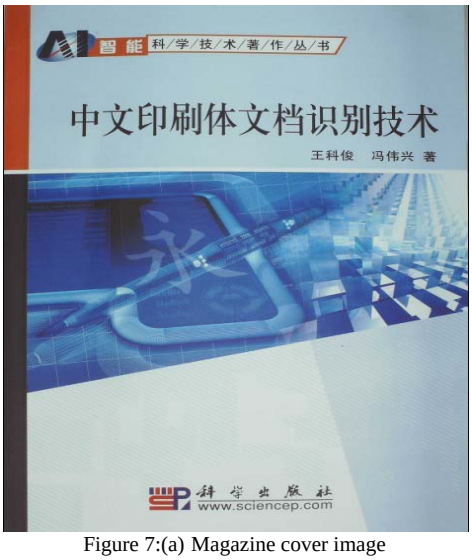

Figure 7:(a) Magazine cover image

Figure 7:(b) Filtered image using this algorithm

Figure 7:(c) Hanvon OCR layout analysis (red box for text)

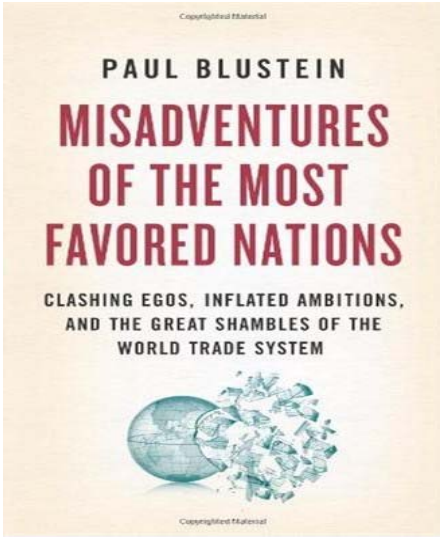

Figure 8:(a) Book cover image

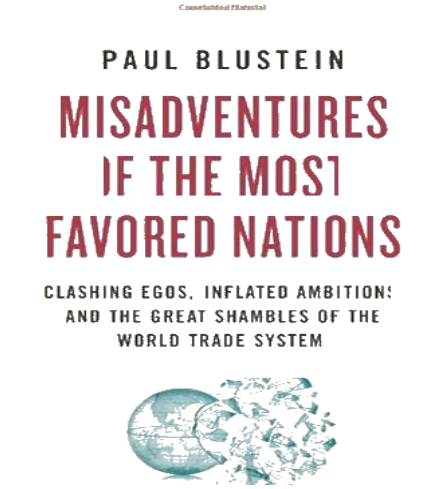

Figure 8:(b) Filtered image using this algorithm

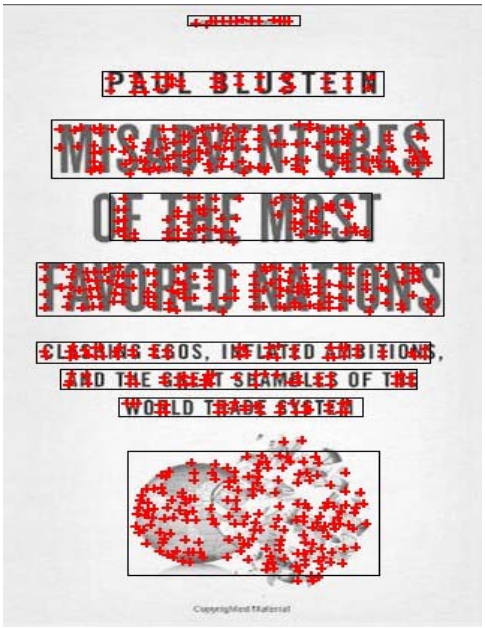

Figure 8:(c) corner-point and text area of image

From Figures 7 and 8, we can get this conclusion that by adjustment of the parameter of this method we can well filter the background of Chinese and English two types of image. For images other OCR products cannot process, while the proposed method of this article can still locate the image text, which illustrate that this algorithm can be used for not only the background of the image filter, but also for OCR processing. Firstly, by the background filtering we can get the images only contained text message, after that, by OCR processing, we can get the correct text information. It improves OCR recognition accuracy and practicality.

Based on the analysis of experimental results, it is known that the proposed complex background text image filtering method based on Harris corner-point detection has higher positioning accuracy compared with the conventional text localization algorithm. Since the conventional layout analysis and words locating is based on more regular text image, this algorithm can be used as a supplement for regular layout analysis and localization algorithm, making up for the shortage of marking text areas as the pictures. Using accurate text positioning results to filter the complicated background image, this method has a lot of practical application value.

VII. Conclusion

Within the field of corner-point detection extraction, Harris algorithm is applied more broadly. The main point of this paper is that by using the Harris corner-point algorithm for image corner-point detection, acquire the selection rules for the screening of noises corner-points and words corner-points based on analyzing the features of text structure. Realize the words marking by utilizing the characteristics of text locally accumulated. Conduct binarization for the obtained text area to extract the text contour coordinates. Finally achieve the complex background filtering by using the image reconstruction technology. The experimental results show that this algorithm overcomes the deficiency of other related algorithms by improving itself and making the utmost of the core idea of others, realizing the accurate text locating and background filtering.

The effect of this algorithm for processing the background which does not contain a large number of text produced noise is ideal. The processing ability for the image which contains a lot of word produced noise corner-points remains to be further improved. At the same time, for the English text image processing need to continue to improve and perfect. It is also the focus of the following work and the part needs improvement.

Acknowledgment

Eleventh Five-Year National Science and Technology to support the platform for key projects (2009BA171B02).

Funding Project for Academic Human Resources Development in Institutions of Higher Learning Under the Jurisdiction of Beijing Municipality PHR(IHLB) for short(PHR201007121).

REFERENCES

[1] HARRIS C,STEPHENS M, "A Combined Corner and Edge Detector," *Proceedings of Fourth Alvey Vision Conference*,UK,1988. pp. 147-152.

[2] Hua-Qing Zhou, Sheng-Hui Dai. Wavelet Descriptor for Closed Curves Detection in Complex Background. *Journal of Computers*, 2010(5):1723-1730.

[3] HOU Jian-hui, LIN Yi. "Adaptive Harris X-corner detection algorithm". *Computer Engineering and Design*, 2009, 30(20): 4741-4743.

[4] Rolf P Wurtz,Tino Lourens. "Corner detection in color images by multiscalecombination of end-stopped cortical cells". *Springer Verlag:Lecture Notes in Computer Science*,1997:901-906.

[5] Rong Zhu, Yong Wang. Application of Improved Median Filter on Image Processing. *Journal of Computers*, 2012,7(4): 838-841.

[6] Zhang X.H, Wang H.X. Robust Image Corner Detection Based on Scale Difference Evolution of Planar Curves. *Pattern Recognition Letters*, 2009(30):449-455.

[7] Youjia Fu, He Yan, Jianwei Li, Ruxi Xiang. Robust Facial Features Localization on Rotation Arbitrary Multi-View face in Complex Background. *Journal of Computers*, 2011(6), 337-342.

[8] Guo Yongfang,Yu Ming,Sun Yicai. "Study on advanced rapid corner detection method". *Computer Engineering and Applications*, 2011, 47(12):159-161.

[9] Jianping Han, Zhigeng Pan, Mingmin Zhang. Fast and Robust Moving Objects Detection based on Non-parametric Background Modeling. *Journal of Software*, 2009(4), 1084-1090.

[10] Beiji Zou, Haoyu Zhou, Hao Chen, Cao Shi. Multi-Channel Image Noise Filter based on PCNN. *Journal of Computers*, 2012,7(4):838-841

[11] Jin Jin, Ping Xi-jian, Zhang Tao, Chen Ming-gui. "Survey of Text Localization Techniques in Images". *Application research of computers*, 2007(06):8-11.

[12] Chuanxu Wang. Face Segmentation Based on Skin Color in Complicated Background and Its Sex Recognition. *Journal of Software*, 2011(6), 1209-1216.

[13] Zeng Fan-feng, Zhang Guo-feng, Chen Kan. "Chinese test image inversion fast detection algorithm". *Computer Engineering and Design*, 2012, 33(9): 3512-3516.

[14] Yun-fang Zhu. Moving Objects Detection and Segmentation Based on Background Subtraction and Image Over-Segmentation. *Journal of Software*, 2011(6), 1209-1216.

[15] LI Bo,YANG Dan,WANG Xiao Hong. "Novel image registration based on harris multiscale corner detection algorithm". *Computer Engineering and Applications*, 2006,42(35):37-40..

[16] Zheng Z Q. Wang H. Analysis of Gray Level Corner Detection. *Patter Recognition Letters*,1999, 20(2):149 -162.

[17] unfeng Jing, Guangyan Li, Pengfei Li. The Research of Automatic Registering Detection of Rotary Screen Printing Machine Based on MeanShift and Fast Block-Matching Algorithm. *Journal of Computers*, 2012,7(6):1369-1376

[18] Shen F, Wang H. Real Time Gray Level Corner Detector. *Pattern Recognition Letters*, 2002, 23(8).

Fanfeng Zeng, JiAn of Jiangxi province,China. Associate professor, master tutor, the main research direction is object-oriented technology, information security, image processing and intelligent control, system identification, etc. Proficient in many programming languages, such as C++, Java and C#.

Guofeng Zhang, TaiAn of Shandong province, China. Graduate student, the main research direction is image processing and pattern recognition.

Jin Jiang, Beijing, China. Graduate student, the main research directions are the embedded software development, positioning navigation algorithm, etc.