The Research on Effective Video Scene Character Extraction Algorithm in Natural Scene Images

Jinghua Hu[1] and Huafeng Kong[2]
[1] *Wuhan University of Technology, Wuhan 430070*
[2] *Key Laboratory of Information Network Security, The Ministry of Public Security, The Third Research Institute Of Ministry Of Public Security, Shanghai 201204*

Abstract
PCB technology is particularly important in electronic industry. However, the increasing of technology complexity makes it difficult to inspect the quality of PCB. AOI system greatly improves the PCB inspection efficiency, and it has become one of the most instructive subjects that how to identify the text information in the chip image rapidly and accurately. Since the chip images have complex natural background, and are greatly affected by light, shadow, noise, font and size, texture, color, position and arrangement., it is often difficult to inspect, extract and identity the text.

This paper presents an effective algorithm of scene character segmentation and recognition in natural scene images. The algorithm segments lines of character, then segments every line of character into individual words for further processing, such as feature extraction and character recognition according to the known features of character. After the segmentation of character, we use an advanced BP algorithm to recognize the character. It improves BP mainly through restructuring gradient in the SIGMOID. The experiment shows a significantly improvement in PC: the recognition accuracy rate of this system is above 96% and the response time is 5ms/100 words.
Keywords Scenes Character Segmentation, Feature Extraction, BP Algorithm, and Image Processing

1 Introduction

As the optical character recognition (OCR) technology comes in vogue, many scholars begin to research on the character extraction from the document images. Till the 1990s, during the rapid development of computer technology and multimedia technology, the content-based multimedia retrieval has become a research hotspot. Then, the character extraction in the natural scene images has again gradually become one of the research hotspots. Usually, the natural scene image character has substantial changes in font, size, color, alignment and arrangement; also with the complex character background, low resolution and high noise. Furthermore, many systems also require the algorithm has higher processing speed in the application. All of these causing difficulties to effectively extract the characters from the natural scene images, especially for the natural

scene images based on the videos. Many domestic and foreign scholars have made beneficial explorations and attempts in this field. Y.Zhong[1] firstly put forward the resolution that position the characters in the complex images. However, this resolution mainly aims at the character positioning in the image scanning for the colorful disc covers, and fails to be directly applied into the natural scene images. A.K.Jai[2] and others proposed a kind of character positioning method, which applies to the newspaper, web pages and general images and video frames, but not ideal to identify the small character fonts. M.A.Smith[3] and others developed a kind of method detecting the characters on the images. However, because of the limitation in size, this method can only detect the characters within the specific font scope. It can not use the feature that the same characters will appear in multi frame to further enhance the character detection performance. Also, the word segmentation prepared for the OCR is not done. Sato[4] and others developed a character segmentation and recognition system aiming at the static low resolution headlines. The system will take advantage of the method mentioned in literature[3] to recognize the character. Then magnify as 4 times as the recognized characters, which will use the smallest time-based image pixed value to conformify the headline. Such system has a good effect on news program. R.Lienhart[5-6] and others have developed systems in character detecting, identifying, recognizing successively. The color-based algorithm system of inter infiltration segmentation adopted in literature[5], only deals with image segmentation and recognition in single frame, without consideration in successive images in multi frame. However, the system in literature [6] takes further consideration on texture feature in title character, and tracking and conforming in successive images in multi frame, which makes it a better recognition in OCR.

Character recognition has a close connection with character segmentation which is one of critical factors in character recognition[7]. Since the threshold process in natural scene image tends to bring about noise and low quality problems, and probably results in character conglutination. Such traditional character segmentation will not be satisfied with people. In order to identify characters in the natural scenes fast and accurately, localizing the chunks should be taken as first priority, and later extracting smaller chunks. Therefore, the small chunk could be used respectively to sharpen the noise tremendously, without the complex impact from other parts in images.As a result, this paper presents a algorithm of scene character segmentation and recognition in natural scene images, based on feature feedback, which will segment Chinese, English and adherent character. After the segmentation of character, an improved BP algorithm is afterwards used to recognize the character.

2 Feature Feedback Based Algorithm of Character Segmentation & Recognition between Chinese, English and Adherent Character

Several Traditional Methods in Character Segmentation & Recognition[8]

① Image- based Analysis. Search for rational segment point between characters, chiefly adopt static projection analysis;

② Recognition- based. Select various kinds of current segmentations via identifying ability before actual segmentation.

③ Synthetical based on image analysis and recognition. Reduce and filter vertical segmentation hypothesis by image analysis.

④ Unity Recognition. Use the whole word as recognizing object, according to the features of the whole word, to avoid segmentation harm on character.

However, in the case of word conglutination, those methods can not segment adherent characters accurately. This paper deals with this kind of situation by an algorithm of feature feedback- based character segmentation and recognition between Chinese, English and adherent characters. First segment each line of character, and then search and extract the adherent character in each line, finally transfer small segmenting module to discover the most reliable segment point, so as to achieve the segmentation purpose.Key points as follows:

1.Identify the Category of Adherent Characters

After initial identification, the recognition and length of each image as follows:

① Suppose the length of each character image is w, and the average length of each character image is w. When w is much wider than wv, and pick up a threshold via experiment, we could identify whether it is adherent character or not.

② After identifying adherent character image, according to English character image length is wider than the Chinese character image length, and the separation distance is small between the adjoining characters, we could identify whether the adherent character image is Chinese or English; according to average English adherent character length is shorter than the Chinese one, we could identify those character images with wrong structure as adherent character image.

2.Identify the Category of Adherent Characters

For the adherent Chinese character, we can segment the Chinese character by judging the potential character image length. We use identifying module to judge the border to segment the character. In this process, we confirm the most reliable segmenting point by right-to-left and left-to-right searching methods.Algorithm as follows:

Step1, Record the value of left border bL, right border bR and line height HL.

Step2, Cobfirm X_0 value (X_0 = HL)from the left border, choose one threshold value T0, the segmenting point should be picked up from (X_0-T_0 , X_0+T_0), modulate it according to steplong.

Step3, Transfer identifying module, and judge whether the result is in the range or not. If not , do Step2; if is , do step 4.

Step 4, Save segmenting result bR1, Take bR1 as the left border , and then do step2 and the following steps successively, till the right border bR. To compare the validity of the final segmenting result, do all steps in turns from the right border, and save the best segmenting result.

3.Segment Adherent English Number

For the adherent English number, first use Border Searching Algorithm to segment these adherent character with gaps in between, which can not be segmented by projection method. Then suppose the potential height of each adherent character string is X, according to X, transfer the identifying module to segment the adherent character string so as to locate the segmenting point. Border Searching Algorithm as follows:

Step1, Take note of left rect urlLi ,Tiand right rect urlRi ,Bi, and pick up a threshold value T,Find a P point whose gray-scale value is 1, and take it as initial point of the contour line.

Step2, Start with the fund point and continue with the searching job. If the gray-scale value is 1, please carry on to the left; if the gray-scale value is 0, please carry on from the right turn. Once meet the point whose gray-scale value is 1, these points would be contour point.

Step3, Repeat Step 2 till the contour line point coincides with the initial point;

Step4, Confirm the right border via contour line searching. If bR1 coincides with the right border, there is no adherent character; if not,For identified character we can pick up a suitable threshold via experiment, and change it to the left or eight. Then use it as the segmenting position and save the result.If character width to height ratio is large, the character is adherent; if the ratio is relatively small, we could transfer the identifying module to check it is or not, after it being segmented.

If it is adherent we can calculate its potential height value, according to whose height we adjust a threshold value and use it as its possible width. Then we take advantage of identifying module to segment the adherent character and confirm the segmenting position via identified result. Save the correct final segmenting result. The English character baseline feature as follows:

For the adherent English number, if the line height is close to 5/6 of Chinese

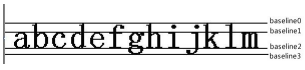


Fig.1 English character baseline feature

line height, there are up-protruding letters (e.g. t) as well as down-protruding letters (e.g. p). We confirm the baseline approximately as much as 1/4 as adher-

ent character height, and height X is approximately as much as 1/2 of adherent character height;

If the line height is close to 2/3 of Chinese line height, there are up-protruding letter or down-protruding letter. We confirm the baseline approximately as much as 1/3 as adherent character height, and height X is approximately as much as 2/3 of adherent character height;

If the line height is close to 1/3 of Chinese line height, there are neither up-protruding letter nor down-protruding letter. We confirm the baseline approximately as much as 1/2 as adherent character height, and height X is approximately as much as adherent character height. And we can adjust the potential segmenting position to meet height X via threshold-adjusting. And the same method could be used in English number adhesion and Chinese adhesion.

Feature feedback based algorithm of character segmentation & recognition between Chinese, English and adherent character identify result as Fig.2 and Fig.3.

Fig.2 English character segmenting result

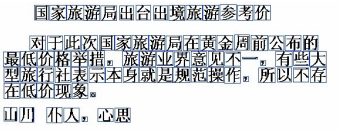


Fig.3 Chinese character segmenting result

3 Character Size Normalization

After being segmented,each character is in different size; therefore the segmented character region is different.In order to extract the character feature,we need to normalize the segmented character region.

Suppose Height and Width to represent respectively normalized character vertical and horizontal length. And H and W respectively represent height and width of the character rectangle. Therefore, the ratio of vertical to horizontal HR and WR would be:

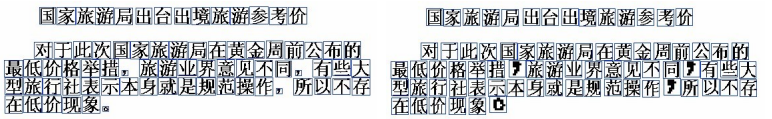
$$H_R = \frac{Height}{H}, W_R = \frac{Width}{W}$$

Then we make coordinate mapping, and we suppose coordinate (i_new,j_new) is the mapped coordinate (i,j) after it has been normalized. Similarly we map the following coordinates:

$$i\_new = top + (i - top)/H_R, j\_new = left + (j - left)/W_R$$

Note: Top is the vertical coordinate of the POINT a in the mornalized character region, and left is the horizontal coordinate of the POINT a in the mornalized character region. The character size normalized result is shown in Fig.4.

Assign the gray-scale value of the character coordinate (i,j) before being normalized to the normalized character coordinate (i_new,j_new) to accomplish the character normalization.



(a) Before Normalization (b) After Normalization

Fig.4 Contrasts two normalization results

4 Improvement of BP Algorithm

Base on the traditional BP Algorithm, this paper presents an improved algorithm which improves BP mainly through restructuring gradient in the SIGMOID. If the value is supposed too much, the input 0 or 1 of each layer would be scattered, and the study result is worse. Otherwise, if too little, the system linearity is enhanced whereas non- linearity is weakened. Therefore, the best value is between the above both. The most important point is to amend the gradient of each allnodes in the SIGMOID so as to find a best value. In the improved BP Algorithm, SIGMOID is turned into binary driving function on the connection of gradient and the potential value of the junction point:

$$\varphi(\alpha, \upsilon) = \frac{1}{1 + e^{-\infty}}$$

We can easily deduce that after all the amendment, adjusting method of the weight and threshold is the same with BP Algorithm. Suppose the actual output vector $O = \{O_0O_1, \ldots, O_L\}$, teacher vector $T = \{T_0, T_1, \ldots, T_L\}$, mean square error E:

$$E = \frac{1}{2}\sum_k (T_k - O_K)^2$$

Gradient α is changed according to $E^{'}$ negative gradient change, where mean square error E reduces the least. Hence the improved gradient algorithm would be: $\Delta\alpha_{\infty\text{-}} = \frac{\partial E}{\partial \alpha}$

The node point of the output layer is: $O_k = \frac{1}{e^{-\alpha_k(\mu_k-\theta_k)}}$,Then

$$\Delta\alpha_k = -\eta\frac{\partial E}{\partial \alpha_k} = -\eta\frac{\partial E}{O_k}\cdot\frac{\partial O_k}{\partial \alpha_k} = -\eta(T_k - O_k)\frac{\partial}{\partial \alpha_k}\frac{1}{1+e^{-\alpha_k(\mu_k-\theta_k)}}$$
$$= \eta(T_k - O_k)(\mu_k - \theta_k)O_k(1-O_k)$$

η is studying step length,$O_k$is the actual output of output layer neuron K,$T_k$is teacher signal,$\mu_k$is the Input signal linear combination of the output layer node, then

$\mu_k = \sum_k H_j W_{kj}$,$W_{kj}$ is weight;$\theta_k$is threshold;$H_j$is the input signal of output layer nodepoint, that is the output of intermediate layer neuron. For the node point of

$$\Delta\alpha_j = -\eta\frac{\partial E}{\partial \alpha_j} = -\eta\frac{\partial E}{\partial H_j}\cdot\frac{\partial H_j}{\partial \alpha_j}$$

hidden layer

$$\frac{\partial H_j}{\partial \alpha_j} = \frac{\partial}{\alpha_j}\cdot\frac{1}{1+e^{-\alpha_j(\mu_j-\theta_j)}} = (\mu_j-\theta_j)H_j(1-H_j)$$

$$\frac{\partial E}{\partial H_j} = \sum_k \frac{\partial E}{\partial \mu_k}\cdot\frac{\partial \mu_k}{H_j} = \sum_k \frac{\partial E}{\partial \mu_k}\cdot W_{kj}$$

$$\frac{\partial E}{\partial \mu_k} = \frac{\partial E}{\partial O_k}\cdot\frac{O_k}{\mu_j} = -(T_k - O_k)\alpha_k O_k(1-O_k)$$

Substitute in turns would be:

$$\Delta\alpha_j = \eta(\mu_j-\theta_j)H_j(1-H_j)\sum_k (T_k - O_k)\alpha_k O_k(1-O_k)W_{kj}$$

Using driving function via the improved gradient, we can adjust the gradient of SIGMOID to its best in the process.

To study the photo sample, we memorize the character feature in the sample and find the weight and threshold, and then we establish and improve BP neural network. Then we can input a photo, using BP neural network, to find a closest character to the input Chinese character in the sample. The steps as follows:

(1)Network TopologyMost characters to be identified are Chinese characters. Because the chinese character structure is quadrel, we divide them into 26×26 unit. Each unit corresponds to one input; hence the output nodes in the neural

network would be 26×26=676. The numbers of the output nodes in the neural network is the same with the characters to be identified.

(2)Output Node Confirmnesscalculation of node output is equivalent to calculation of the function SIGMOID output. Because of the functions nature, we would be aware of its value range [0,1]. The closer the value is to the limits, the less sensitive the function value changes to its independent variable. Therefore, we can enhance the nodes precision by its output-to-input sensitivity. We suppose T is 0.9, and F is 0.1; for input, if with stroke, then input 0.9, if not, we input 0.1. And we deal with output the same way.

(3)Edit of training sampleThe more samples the collection owns, the more corresponding sample each Chinese character has. Also the better identifying ability the neural network is after being trained, however, the more time the training takes. Here we only take the sized 18 thick Song Ti character as sample. There are English characters, numbers, 3,775 Chinese characters of one-level, and basic punctuation in the sample collection.

(4)Experiments and analysis The identifying result of BP-Algorithm-based character identifying algorithm is shown in Fig.5. This system is designed to deal with images including character, whose recognition rate is above 96%, and its response time is 5ms / 100 words. Compared to the traditional recognition method it has been improved, whose recognition result is better. However, the worse the images quality will result in the worse the recognition result. In the image treatment phase, the wiping-out of useless information is partial, which has an effect on character recognition and the whole recognition result.

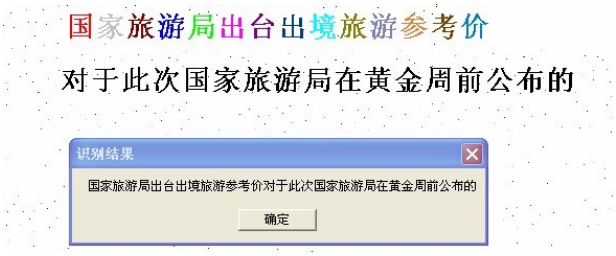


Fig.5 Results of character identifying algorithm

5 Conclusion

According to the above discussion, the identifying result of BP-Algorithm-based character identifying algorithm can improve the efficiency of IC identifying in AOI system.

References

[1] Yu Zhong, K.Karu, A.K.Jain. (1995), "Locating text in complex color images", *In Proceedings of the Third International Conference on Document Analysis and Recognition*, Vol.1, No.14, pp.146-149.

[2] A.K.Jain, B.Yu. (1998), "Automatic text location in images and video frames", *Pattern Recognition*, Vol.13, No.12, pp.2005-2076.

[3] M.A.Smith, T.Kanade. (1995), "Video skimming for quick browsing based on audio and image characterization", *3Technology Report CMU-CS-95-186*, pp.13-15.

[4] T.Sato, T.Kanade. (1995), "Video OCR: Indexing digital news libraries by recognition of superimposed caption", *Multimedia Systems*, Vo.l7, No.5, pp.385–395.

[5] H.J. Zhang, C.Y. Low, S.W. Smoliar, and J.H. Wu. (1995), "Video parsing, retrieval and browsing: An integrated and content-based solution", *Proc. ACM Multimedia* , pp.15-24.

[6] R.Lienhart, W.Effelsberg. (2000), "Automatic text segmentation and text recognition for video indexing", *Multimedia Systems*, Vo.l8, No.2, pp.69-81.

[7] Huiping Li, Omid Kia, David Doermann. (2004), "Text enhancement in digital videos", *Proc. SPIE99-Document Recognition and Retrieval*, pp 3651:2-9

[8] Casey R G, Lecolinet E. (1996), "A survey of methods and strategies in character segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.18, No.7, pp.690-706.