Research on Cigarette Carton Rechecking Technology Based on Morphological Processing Algorithms

Xi Chen[1, a], Linzi Shan[2]

[1]School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

[2]School of Automation, Beijing University of Posts and Telecommunications, Beijing 100876, China

[a]572214342@qq.com

Keywords: One-dimension barcode extraction, morphological processing, sub-region screening, projection, cigarette carton recognition

Abstract. This paper puts forward a method for automatic recognition of images on cigarette cartons taking advantage of the uniqueness of the barcode on the carton. It is to solve the mismatch problem between the traditional way of manual rechecking and the high-speed automatic sorting system. Based on the square adaptive morphological structure element algorithm, it deals with monochrome images to reduce the interference caused by blank structures within the barcode to the connected domain, so as to attain multiple candidate connected sub-regions. By screening sub-regions, it positions the barcode area. Then, it locates human-readable characters, and cuts out digital items by projection. At last, barcode digits are recognized, and the recognition is done. Experiment results show that this method is effective in removing the interference caused by characters and images on cigarette cartons as well as by illumination noises.

Introduction

For the mainstream high-speed automatic cigarette carton sorting system at a rate of 15,000 or even 20,000 cartons per hour, cigarette carton rechecking is a critical step to examine its accuracy. The traditional rechecking method depending on labor, with its lower efficiency and inability to do a large amount of work, is no longer suitable for producing lines. There are existing researches which recognize one single carton by identifying its energy features[1], or features extracted by principal element analysis combined with Fisher Linear Discriminant[2]. Due to the variability of cigarette cartons, a large template library is needed to match their features. Therefore, efficiency and accuracy in the practice of extracting features of images are major problems. In addition, cigarette packages and types are updating constantly, which means the traditional feature matching modal needs to be particularly designed each time. This is not practical for the development of cigarette industry.

This paper puts forward a new method to recognize cigarette cartons, which identifies cigarette categories according to digital features extracted from one-dimension barcodes. First, locate the barcode areas on the monochrome images. Then, cut out human-readable digits from the barcode. Due to the uniqueness and versatility of the barcode, we can get rich information about cigarettes, such as brand, producer and place of production by identifying the digits on the barcode. This method can adapt to the upgrade and development of cigarette products quickly. When the feature template is simplified into 10 Arabic numerals from 0 to 9, the matching efficiency and accuracy are greatly improved.

The Overall Technology Roadmap

Cigarette carton rechecking technology based on one-dimension barcodes mainly consists of two parts: barcode area extraction, and digital segmentation and recognition of barcodes. The first step to extract barcode areas is to pre-process inputs of cigarette carton images by graying the color images, and by adaptive double peak threshold segmentation processing. Second, attain connected sub-regions by adaptive morphological processing. Third, fill up the sub-regions and cut out barcode

areas by screening these sub-regions. As for the digital segmentation and recognition, at first, use adaptive threshold processing to segment thresholds of gay-level images in barcode areas, so as to attain monochrome images. Then, locate and segment the digital part by projection. Finally, recognize digits according to weighted template matching and vague criterion. The overall technology roadmap is shown in Figure 1.
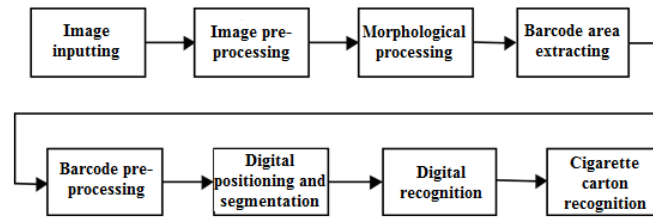

Figure 1 Technology roadmap

This paper mainly improves the barcode area positioning algorithm and the digital segmentation algorithm. Barcode area positioning algorithm, at first, processes monochrome images by square self-adaptive morphological structure element algorithm, which reduces interference to the connected regions caused by blank areas within barcodes. In this way, we get several candidate connected sub-regions. Then, cut out barcode areas according to the set size of area and length-width ratio. Barcode digital segmentation algorithm first locates the prepositive code in binarized barcode images. According to its upper and lower bounds, locate those bounds of the digits, and cut out the digital parts. Then locate each digit according to their projection minimums. Examine the position of each digit according to the fixed intervals between them, and at last accurately locate and segment digital items on the barcode.

Barcode Locating Algorithm Based on Self-adaptive Morphological Processing and Sub-region Screening

The cigarette carton photos in this paper are taken by fixed camera, and the distance between the cigarette and the camera is fixed. By analyzing cigarette carton photos, characteristics of barcode area are concluded as follows: the background is white, and the barcode area is black; barcode areas do not bear much difference in size and are at the same magnitude; the cigarette cartons are rectangular in shape with an aspect ratio at about 13:7. According to these features, barcodes are located and extracted from the background by image binarization combined with morphological processing.

Square Self-adaptive Morphological Structure Element Processing and Barcode Positioning by Sub-region Screening. In actual cigarette sorting line, interference caused by illumination variation or cigarette cartons inclined is unavoidable. These factors have great impact on structure elements of morphological processing. There has been researches[3] using linear adaptive structure elements to locate car license plate through morphological processing. However, in the process of cigarette carton barcode positioning, a lot of linear segments in blank areas within the barcode can not be eliminated by linear elements. To solve this problem, this paper puts forward a barcode positioning algorithm based on square adaptive structure elements which works efficiently in separating barcode areas from the background.

Figure 2 Threshold segmentation effect in adaptive double peak method

First, gray the original color image, and then segment thresholds with the adaptive double peak processing [4]. The effects are shown in Figure 2:

After analyzing and improving monochrome images processed by double peak segmentation, we find that though the barcode area has been separated from the background, some fragments of characters and images, as well as noises caused by uneven illumination, are still left on the carton. All these factors will interfere the barcode area positioning. Therefore, after morphological processing, we still need to screen sub-regions, remove invalid regions, and keep the barcode area only[5].

In this paper, square is taken as the shape of structure element, and the length of the square varies from short to long, which are treated as parameters of structure elements. Monochrome images are processed by morphological closing operation, and then through image analysis functions, regional structure STATS is attained including 'Area' and 'BoundingBox'. They respectively refers to the area of each sub-region and the minimum rectangle within the area. The latter provides the rectangle's length and width in the form of [x_width, y_width]; first the area size of the marked BoundingBox is screened; then its length-width ratio is examined; if the ratio is close to 13:7, barcode areas are found. Details and steps of the procedure are as Figure3 shows:
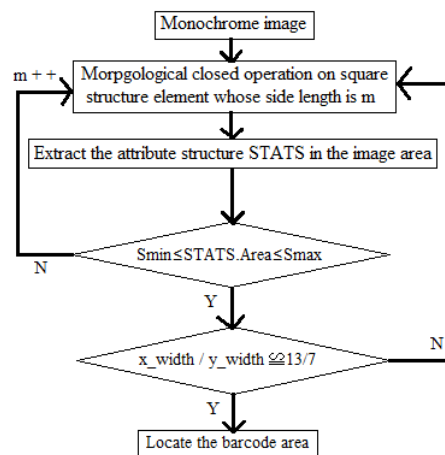

Figure 3 Procedure of adaptive morphology structural elements processing algorithm

Comparison of Positioning Effects with Linear Adaptive Morphological Structure Element Algorithm. Comparison of barcode positioning effects of the two kinds of morphological structure element algorithm (the linear adaptive structure element algorithm and the square adaptive structure element algorithm) is shown in Figure 4. The figure shows that the method proposed in this paper is effective in retaining barcode areas, and eliminating interference from illumination and the other information on the carton.
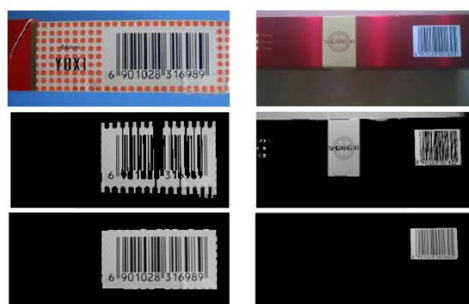

Figure 4 Comparison of barcode positioning effect of two kinds of morphological structure elements

According to the located barcode image, barcodes can be extracted from the original image, and barcode digits can be positioned and segmented.

Barcode Digital Positioning and Segmentation Based on Projection

All the barcodes on cigarette cartons are EAN-13 code including blank regions within barcodes and human-readable characters. The effects of traditional methods are less than good in recognizing barcodes inclined and barcodes of low resolution ratio. After several experiments, this paper adopts another method, that is, human-readable character recognition: locate human-readable digits on the barcode, and recognize them by cutting out every single digit[6].

Digital Positioning and Segmentation by Projection. In this paper, the blank area on top of the

barcode image is extracted. Images in this area will produce great inference to the projection process. Therefore, how to separate the digital part from the blank area is crucial. By studying barcode structures and monochrome images, the specific procedure to position and segment human-readable digits is as follows:

Step1: Binarize the barcode image by standard adaptive threshold processing after barcode positioning; scan the monochrome image horizontally; eliminate the blank areas on the left and the right sides of the barcode to attain the horizontal axis of the valid part. In the same way, scan the barcode vertically to attain the vertical axis of the valid part. Then, cut out the valid part.

Step2: By vertical projection, locate the vertical axis of the prepositive code; cut out the prepositive code; scan it horizontally to locate the horizontal axis of the prepositive barcode, that is, the horizontal axis of the human-readable digits; and then cut out the images of digits from the barcode image.

Step3: From studies on vertical projection, we know that there are a great number of white pixels on each digital area, among which are some barcode separators. According to the position of the minimum on the vertical projection image, we divide these barcode separators into single characters, and then we will find pixels of character set barcode separators gather in upper part of the image. Therefore, by examining the horizontal projection of each character, interference from barcode separators can be eliminated so as to attain a single digital image collection. Then calculate the number of the digits. If it is not 13, the segmentation procedure fails.

Step4: As the shooting distance and angels are different, the photos are in different sizes, so images are different after being segmented by projection. To improve digital recognition ratio, this paper unifies the size of segmented digital images into one whose resolution ratio is the pixel. This facilitates digital template matching from 0 to 9 to realize digital recognition. The specific procedure of the algorithm and the image processing are shown in Figure 5.
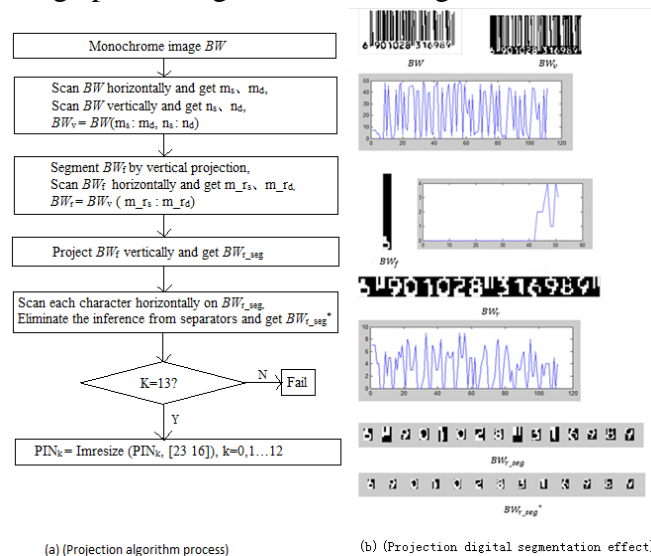

(a) (Projection algorithm process)   (b) (Projection digital segmentation effect)

Figure 5 Projection method for digital positioning and segmentation

The Experiment Result and Analysis of Cigarette Carton Recognition

To test the robustness of barcode positioning and segmentation algorithm after combining morphological filter with sub-region screening, cigarette carton images in different condition and of different types are used in simulation experiment in Matlab. Figure 6 shows the recognition effect of images inclined and images of low quality. Table 1 lists the correct rate and average time consumption in each step of cigarette carton recognition system.

The experiment result shows that the digital positioning and segmentation procedure proposed in this paper is of high efficiency. The projection adapts well to digital segmentation of low quality images and images inclined. See figure 6(a). Digital recognition algorithm, which combines weighted

template matching with vague criterion, is rather effective in recognizing low quality digits that are connected or disconnected, and confusable digits (like 3 and 8, 1 and 7). See Figure 6(b).
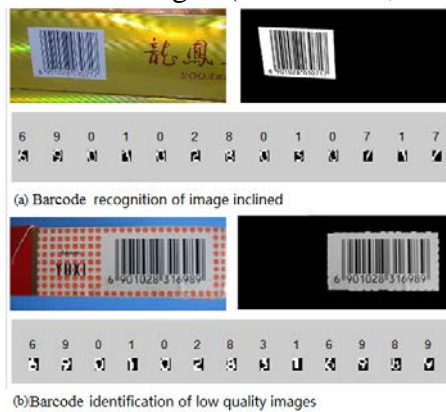

Figure 6 Effect of barcode recognition of special image

Table 1 Result statistics of cigarette recognition

NO.	Item	Test number	Correct number	Correct rate	Average time consumption
1	Barcode area positioning	500	486	97.2%	0.82s
2	Digital positioning and segmentation	500	494	98.8%	0.23s
3	Digital recognition	500	492	98.4%	4.35ms
4	Overall barcode recognition	500	476	95.2%	1.34s

Conclusion

This paper puts forward a new cigarette carton rechecking system, and makes some improvements to key technologies at present, which increases the robustness of cigarette carton rechecking system in different complex situation. In the barcode area positioning and extracting process, adaptive morphological structure element and sub-region screening are adopted to locate barcode areas. This is effective in removing interference caused by the other characters and images on the carton, as well as by illumination noises. In the digital positioning and segmenting process, digits are cut out by projection, which separates human-readable digits from blank areas within the barcode at a high speed, and finally effective digital positioning is realized. Experiments show that this method, whose procedure is simple with high accuracy and robustness, meets requirements of the system for immediacy. Therefore, this method can be used by cigarette distribution centers in cigarette carton rechecking procedures.

References

[1] Sun Dong,Ming Jun,Zang Xiaoxi.Cigarette boxes identifying by energy characteristics [J], Computer Technology and Development, 2006, 4. 131-134.

[2] Zeng Yi, Design and realization of cigarette boxes image identifying system[D], Beijing:Beijing University of Posts and Telecommunications, 2010.

[3] Fang Xinglin, A car license plate locating algorithm based on adaptive morphological structure elements [J], Computer Engineering and Applications, 2013, 23. 149-152.

[4] Wang Lei, Duan Huichuan.Applications of Qtsu method in multi-threshold image dividing [J], Computer Engineering and Design, 2008, 6. 2844-2872.

[5] Rafael C Gonzalez, Richard E Woods.Digital Image Processing Second Edition[M]. Prentice Hall, 2003. 467-469.

[6] N.R.Pal,S.K.Pal.A Review on Image Segmentation Techniquea[J].Pattern Recognition, 1993, 26: 1277-1294.