A NEW MODEL OF ADAPTIVE CONTAINER IMAGE PREPROCESSING

Yajie Zhu[1,a], Qikun Zhang[1,b] and Weiming Zeng[1,c]

[1]School of Information Engineering, Shanghai Maritime University, Shanghai 201306,China;
[a]769877925@qq.com, [b]1911774061@qq.com, [c]zengwm86@sina.com

Keywords: Container image; preprocessing; color gradient; connected domain; weighting.

Abstract. The preprocessing of container image is the key to the whole container code intelligent identification system, and the result of treatment directly affects the accuracy of the code area positioning, the character segmentation and recognition. Based on some features of container images such as the color, the gradient and the connected domain we proposed a novel and adaptive container image preprocessing model. In traditional methods,there are many parameters should be optimized,and multi-parameter optimization is a difficult point.Yet,our model not only solve the problem skillfully,but also can remove a large number the background color of container image adaptively when the color difference of background and target is small.Therefore,our model can better on container area segmentation,character segmentation and recognition.A lot of experimental results showed that the proposed model can effectively extract the more effect.

1. Introduction

Container intelligent recognition system generally includes image acquisition,image preprocessing, positioning of case number area, case number character segmentation and case number character recognition and so on several major parts, each part is closely linked, as shown in figure 1.Container image preprocessing is the base and key to container character recognition[1,3].
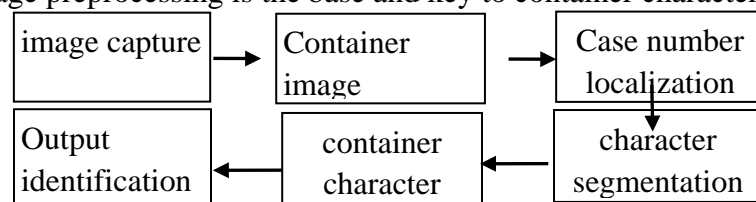
image capture → Container image → Case number localization → character segmentation → container character → Output identification

Fig.1 Flow chart of container number intelligent recognition system

In traditional methods,there are many parameters should be optimized,and multi-parameter optimization is a difficult point.Yet,our model not only solve the problem skillfully,but also can remove a large number the background color of container image adaptively when the color difference of background and target is small.Therefore,our model can better on container area segmentation,character segmentation and recognition[4,5].By the traditional methods and new model of container image preprocessing to pretreat,and the effect is verified by the successful number of the container area segmentation and time.

2. Theory of new model

Detailed steps of algorithm are as follows:

Step 1, the container image is repaired and removed the noise of image $f(x, y)$.and calculate the gradient $G(x, y)$ of the $f(x, y)$.Because the container has the color feature different from the surrounding the image, so we can filter to the matrix of the gradient image according to the feature,immediate, $G(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{m} (G(x, y) - \overline{G(x, y)})$.

Step 2, firstly, through the binarization of gradient image[7], get connected domain set.Due to the gradient image processing is more sensitive, and through the gradient filter,and many

independent noise exist in the image.So the filtering of connected domain is a very important method to removing noise.Due to the next step will make the combination of the color gradient to each connected region , so the connected domain expansion.Image expansion [8]refers to: the sets A and B of integral space Z^2 ,and A is the original image, then B is the structural image. The inflated operations of B to A that is regarded as $A \oplus B$ and defined $A \oplus B = \left\{ z \mid (\hat{B})_z \cap A \neq \varnothing \right\}$.

Step 3, due to the gradient and connected domain processing can only retain the area of high color of similar case number, the area is not necessarily a case number, it is possible that enclosure of the effects of other tags or complicated environment.So we according to the features of container color is white or black, through the analysis of the original image color of connected domain of gradient image to get strengthen and weakened.Specific process is as follows:

a.The original image of connected domain for various gradient image is the filter smoothing, attribute value of image RGB are calculated, remove the point of pixel that all the mean variance is greater than the mean square, filtering, image color non sensitive area.

$$f(x,y,3) = \sum D(k) \sum_{i,j \in D(k)} \overline{std(f(x,y,3))} - std\left(f(i,j,3)\right) \tag{6}$$

a.The filtered image is graying to get $f(x,y)$.Because the color value of case number area in the connected domain are similarity , therefore the case number of smooth, iltering and the connected domain is the area of the same color , the value of color of between the minimum and maximum is divided into 25 space $C(p)$, region and calculate the area $E(q)$ of space of color, and get color and space sets.

b. The formative reasons of gradient color of this model is implemented by weighting, the calculation of weight is finished through the color space and regional relationship of connected domain .Formula (7) for the calculation of weight of color space, because the case number is black or white, so this article by the middle value of the color space and the mean values of between maximum value and minimum value gradually to increase power than for bonus value at both ends.

$$h = \begin{cases} \dfrac{median\left(C(p)\right) - f(x,y)}{\left(Min\left(C(p)\right) - median\left(C(p)\right)\right) \div 2}, & median\left(C(p)\right) - f(x,y) \geq 0 \\[4mm] \dfrac{f(x,y) - median\left(C(p)\right)}{\left(Max\left(C(p)\right) - median\left(C(p)\right)\right) \div 2}, & median\left(C(p)\right) - f(x,y) < 0 \end{cases} \tag{7}$$

Due to the case number area within the connected domain is relatively large, so the color area and regional value than you can in addition case number area to weaken the noise area, formula (8) for the final and weighted formula:

$$l = h \times \frac{E(f(x,y))}{median(E(q))} \tag{8}$$

c. The ultimate model of gradient is got by multiplication of the value of gradient and the weight , formula (9).This model not only based on connected domain of image to do fine processing, but also it can avoid the influence of the case code area.Remove the nonsensitive area of the color , and through the analysis of color space and regional space to strengthen or weaken the image gradient value of weight.

$$G(x,y) = G(x,y) \times l = G(x,y) \times \frac{E(f(x,y))}{median(E(q))} \times h \tag{9}$$

3. Experimental results and analysis

3.1 Effect of container image preprocessing

a. Firstly,the pretreatment method container traditional image is the gray scale of the original image, and then the contrast enhancement to the gray image, median filtering, binarization and the processing of connected domain with binary image.Beacuse most of these operations need to set up its own parameters, so the effect of a large number of different container image processing will definitely influenced, when these methods are combined in container image pretreatment,several parameters of them are required to adjust, it is possible to some container image preprocessing , but they may not very good to many other images.Want to do a lot of different container image preprocessing, let the parameters of these methods are easy to optimize, the likelihood is very difficult.Taken together, these so-called classic methods together to be done that container image preprocessing is not feasible.Such as shown in figure 2: it is pretreatment and container code lost the characters, although the effect of the connected domain parameters for many container image preprocessing is very good, but there are also many container image are not enough, this is one of the serious situation.
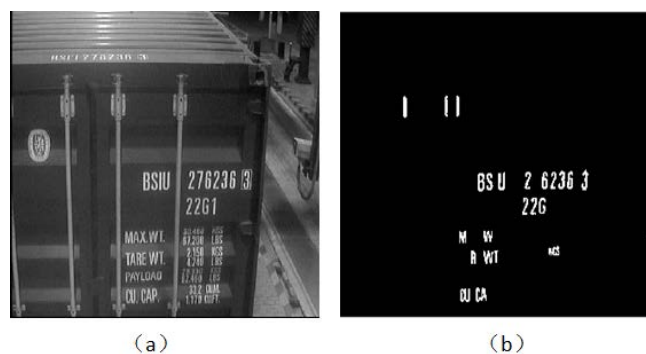

（a）　　　　　　　　　（b）

Fig.2 Traditional methods pretreatment effect of container images

b. In traditional methods,there are many parameters should be optimized,and multi-parameter optimization is a difficult point.Yet,our model not only solve the problem skillfully,but also can remove a large number the background color of container image adaptively when the color difference of background and target is small.Therefore,our model can better on container image preprocessing.The effect of container image preprocessing as shown in figure 3: image (b) is the color and the gradient image of container original image (a) and raised colour represents the size of the color gradient value of the container original image (a).From the gradient value distribution of gradient image (b) to know the gradient noise of container (a) original image is more, although we can see the red area of the image is a case number, but it is affected by the surrounding more high gradient.Image (c) is the effect of the model of the original image (a) preprocessing , figure (d) is the color and the gradient image of figure (c), through gradient distribution (d) of the model preprocessing, a large number the background color of container image has been removed.
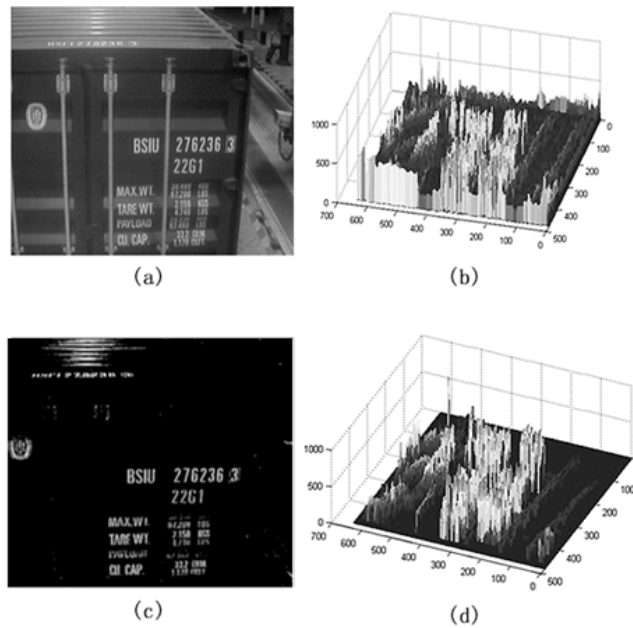
Fig.3 New model of container image preprocessing effect

3.2 Effect of container segmentation

Among the image of positioning, the character of the interferential parts has been eliminated and the character of the case number is more prominent and obvious, but between the character and character of the case number area exist the gap, when we scan the row of the case number area from left to right, the black and white pixels jump more frequently[9,10].Make it get obvious effect by new model of container image preprocessing, a large number of background is removed so that the container is much easier to locate.As shown in figure 4:
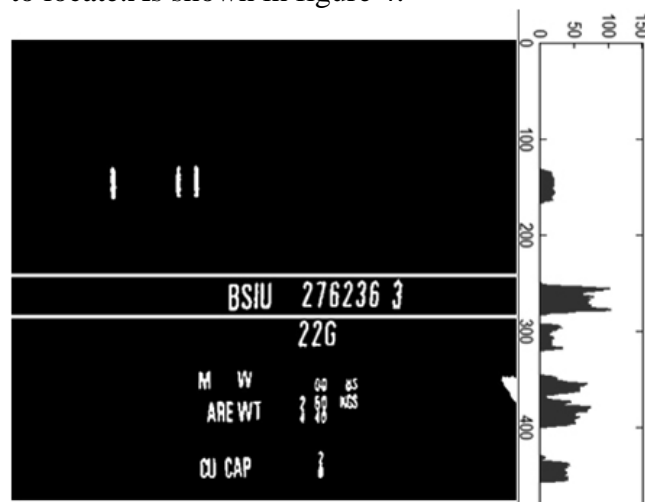

Fig.4 the effect image of localization of upper and lower boundary of case number area

According to the positioning of upper and lower boundary location of the case number area and an independent division and extraction to case number area .As shown in figure 5 (a) : through the figure 5 (a) the case number area of segmentation and extraction , there are still a lot of characters of the left and right sides of the case number area.Therefore, need to segment the area of the case number for further fine segmentation, namely fixing to the case number area around boundary location.Firstly the image 5 (a) of case number of segmentation and extraction as a vertical projection (as shown in figure 5 (b) : for projection histogram), and then locate the left and right boundary of case number area.As shown in figure 5 (c) : the result of fine segmentation for the case number area[11,12]
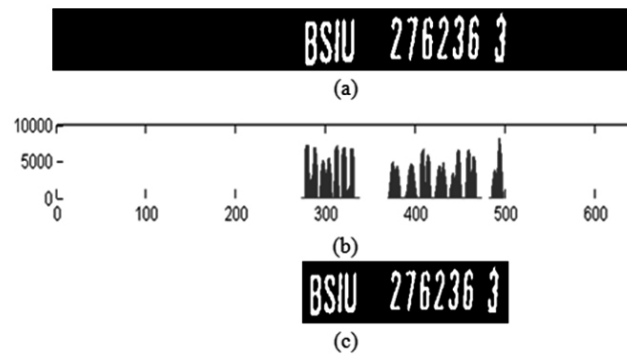
Fig.5 Case number region segmentation effect and process diagram

Through traditional methods and new model of random 50 of container image preprocessing, then use the same kind and classic method (projection) of container segmentation to case number area segmentation of image after pretreatment, through the number of successful case number region segmentation and segmentation time to verify the effect of new model on the container image preprocessing.The experimental results are shown in table 1:

Table 1 The effect of traditional method and new model of container image preprocessing

Pretreatment method	preprocessed image	Number of successful segmentation of container coder area	Rate of successful segmentation	sliced time
traditional method	50	36	72%	0.615 000
new model	50	47	94%	0.365 212

4. Summary

This is a new model of adaptive container image preprocessing.In traditional methods,there are many parameters should be optimized,and multi-parameter optimization is a difficult point.Yet,our model not only solve the problem skillfully,but also can remove a large number the background color of container image adaptively when the color difference of background and target is small.Therefore,our model can better on container area segmentation,character segmentation and recognition.A large number of experimental results and theoretical analysis show that the model is reliable, stable and has very good application prospect.

References

[1] Ting Hu . Container character recognition based on neural network research [D]. Wuhan university of technology, 2012.

[2] Linghong Sun . Container intelligent identification algorithms [D]. Wuhan university of technology, 2012.

[3] Mingshen Fei. The application of automatic identification system in container [J]. Transportation standardization, 1999 (2) : 27-29.

[4] Yuangang Li. Face detection method based on color model research [D]. National university of defense technology, 2007.

[5] Guoxiang Huang. RGB color space and its application research [D]. Central south university, 2002.

[6] Cheng Liang . connected domain markers of based on pixel and Binarization method[J]. Computer & digital engineering, 2013, 41(5) : 823-826.

[7] Tao Wang, Junfeng Wang, Qing Bai , Wei Zhang . TV model of image inpainting algorithm [J]. Computer system application, 2013, 22 (3) : 121-125.

[8] Gao Hai,Siu Wanchi,Hou Chaohuan.Improved techniques for automatic image segmentation[J].IEEE Transactions on Circuits and Systems for Video Technology,2001,11(12):1273-1280.

[9] Tai S C,Chang T Y,Wang S Y.A method of container codes detection and segmentation by using improved projection algorithm[C].Intelligent Control and Automation (WCICA),2011 9th World Congress on.IEEE,2011:587-592.

[10] Wu W,Liu Z,Chen M.An automated vision system for container-code recognition[J].Expert Systems with Applications,2012,39(3):2842-2855.

[11] He Z,Liu J,Ma H.A new automatic extraction method of container identity codes[J].IEEE Transactions on Intelligent Transportation Systems,2005, 6(1):72-78.