AUTOMATIC SELECTION OF BINARIZATION METHOD FROM IMAGES WITH SERIAL NUMBERS ON INDUSTRIAL PRODUCTS

Vít PASKER, Ondřej GRYCZ, Robert HLAVICA, Pavel FORETNÍK, Ivana BARČÁKOVÁ

VSB - Technical University of Ostrava, Ostrava, Czech Republic, EU, vit.pasker@vsb.cz

https://doi.org/10.37904/metal.2020.3636

Abstract

The article deals with the automatic selection of the binarization method using advanced methods of artificial intelligence. The input images to the algorithms are images of serial numbers from industrial environments, for example on iron and steel billets, slabs, etc. The surface of these products is in most cases severely damaged by industrial processes, such as traces of cut, rust, noise, surface roughness, etc. Text recognition is a very common topic nowadays. All investigated solutions are based on the fact that each image is binarized by a single defined method and the accuracy of recognition is given only by the quality of learning of the neural network. Especially in an industrial environment, it is difficult to create a universal method for unambiguous methods for text recognition. The innovation described in this article is the automatic selection of the binarization method (from the Bradley, Niblack, Sauvola methods etc.), which increases the accuracy already in the phase before the text recognition itself, which with the subsequent correct combination of filters leads to an overall increase in accuracy.

Keywords: Binarization, image recognition, neural network, image filters

1. INTRODUCTION

Obtaining information from text files, photos, videos and other resources, it is nowadays a very attractive sector with innumerable practical applications. Text recognition on industrial products is complicated by the large number of defects and the extensive range of font and material properties themselves. The basic defects include traces of cuts, corrosion, abrasions and damage caused by handling the product, etc. As a result, many text detection and recognition systems are absolutely unusable, or rely on the created functions for basic data, making them applicable to only one issue. The main goal of the topic is to create algorithms and methods for different types of binarization of images from the industrial environment. Based on specific properties and parameters of images, an algorithm was subsequently created using artificial intelligence methods, which can select the optimal method of binarization. Such a selected image will continue to be subjected to algorithms for character selection and text recognition in the image. The output of all applied methods is reading text from an image from an industrial environment with higher success. [1]

2. BINARIZATION METHODS

The process of converting an image to binary is called binarization. Binary images are those images whose pixels have only two possible intensity values. From a programming point of view, it can be said that the input image, which contains multiple color spectra (RGB, CMYK), is converted into a two-dimensional boolean array according to certain rules. Binary images are produced by so-called thresholding in grayscale or color image so that the object in the image is separated from the background. The color of an object (usually white) is called the foreground color. The rest (usually black) is referred to as the background color. However, depending on the image to be thresholded, this polarity may be inverted. In this case, the object is displayed with zeros and the background has a non-zero value. [2,7]

Niblack method

Wayne Niblack proposed an algorithm that calculates a binarization threshold for each individual pixel by moving a rectangular matrix across the image. The threshold value T for the center pixel of the matrix is calculated using the mean value and the variance of the intensity values in the matrix according to the equation (1). [8]

$$T = m + k \cdot \sigma \tag{1}$$

where:

m - the mean value of the pixels inside the array

k - constant in the interval <0; 1>, set to 0.2 according to Niblack

σ - the standard deviation value of the pixels inside the array

The quality of binarization depends on the parameter *k* and the size of the working matrix. The Niblack method is one of the most commonly used methods and is the basic principle of several other methods. This method can effectively detect text even at very low contrast. One of the main limitations of this method is that it generates a large amount of noise in empty matrices. It also requires that the size of the matrix and the values of the factor k be determined manually. [3-5,9]

Sauvola method

The Sauvola algorithm is a modified form of the Niblack method. The aim of the method is to solve the problem of the existence of noise depending on the impact on the standard deviation value using a range of gray level values in the image. The method finds better results under conditions where there is an obvious difference in color contrast and light structure in the image. The threshold value is calculated according to the equation (2).

$$T = m\left(1 - k\left(1 - \frac{\sigma}{R}\right)\right) \tag{2}$$

where:

m - the mean value of the pixels inside the array

k - constant in the interval <0; 1>, set to 0.2 according to Niblack

σ - the standard deviation value of the pixels inside the array

R - standard deviation dynamics, usually 128

The method works well in eliminating noise to a high level. Conversely, if the contrast between the background and the foreground is low, or if the text in the image is thin in style, the algorithm cannot process the image properly and the result is greatly distorted. The size of the matrix and the value of the factor *k* must also be set manually, as with the Niblack method. [3,4]

Bradley method

The method was designed for adaptive thresholding using image integration. The proposed method is an extension of the Wellner method, but it is more resistant to changes in luminosity in the image. At the same time, however, it remains relatively simple and easy to implement. In the first phase of the algorithm, the sum of *f(x, y)* is calculated at each location *I(x, y)* of the expressions to the left and above the calculated pixel according to equation (3).

$$I(x,y) = f(x,y) + I(x-1,y) + I(x,y-1) - I(x-1,y-1) \tag{3}$$

where:

x, y – position of pixels in matrix

After converting the given image to an integral image, the sum of the function for any rectangle with the upper left corner (x_1, y_1) and the lower right corner (x_2, y_2) can be calculated in constant time according to equation (4). [3-5]

$$\sum_{x=x_1}^{x_2}\sum_{y=y_1}^{y_2} f(x,y) = I(x_2,y_2) - I(x_2,y_1-1) - I(x_1-1,y_2) + I(x_1-1,y_1-1) \quad (4)$$

where:

x, y – position of pixels in matrix

Outputs from different binarization methods

As can be seen in **Figure 1**, the output from each method is quite different. It is not easy to say which method is the best, each is used for a different type of input image. This is why it is not advisable to always use only one type of binarization, but to dynamically select the one that shows the most suitable results for other algorithms. These outputs of binarization methods are better perceptible for algorithms than for the human eye, especially in such a low resolution. Therefore, the outputs may appear to have much worse properties than the input image. After the subsequent filtering of the image, elimination of various types of noise and other adjustments, however, the text or its parts that can be recognized will remain.

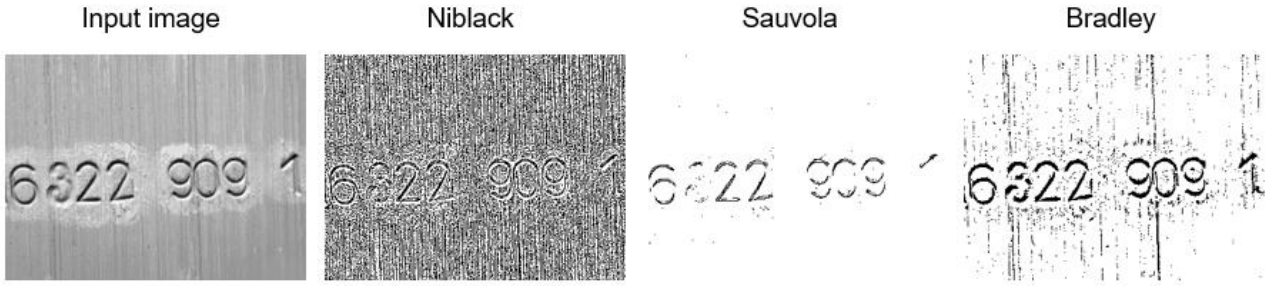

Figure 1 Results of different binarization methods

3. AUTOMATIC SELECTION OF THE BINARIZATION METHOD

There are several ways to achieve automatic selection of a binarization method. One of the possibilities is to compare the parameters of images after individual binarizations. It is assumed that the values of the parameters, their ratios, etc., at which the output is most suitable, are known. Another possibility is to use artificial intelligence methods for recognition, such as a neural network. A backpropagation neural network was used for the procedure below. This variant was chosen because it is possible to define an output for individual inputs for the learning phase. The advantage of the input images is their color spectrum, which is basically always in shades of gray. Therefore, these images can be considered as a training set of the neural network without major modifications. It can be assumed that future images will have the same properties.

Image with the serial number is taken from mobile camera, 3D scanner or a standard high-resolution reflex camera with subsequent image processing. Mostly all of these images have high resolution, which makes the whole process many times slower. Therefore, first need is to reduce the input image size. The output resolution was chosen to be 320x240 px so that the aspect ratio was maintained, the number of pixels was significantly reduced, but at the same time the data in the image was preserved. The next process is to transform the data from a two-dimensional array to a vector. At the end of this process, all pixel values

are divided by the maximum value of the color spectra to achieve values in the range from 0 to 1. The data prepared in this way can be used as input to the neural network. [6]

There were created algorithms for binarization by methods Bradley, Feng, Niblack, Sauvola and Nick. Between these methods, the neural network decides according to the internal properties of each input image. The principle of the automatic selection method is that for certain parameters of the input image, such as brightness, contrast, accumulation of pixels in parts of the image, etc., a certain type of binarization is suitable. So far, a partial disadvantage is the smaller training set used, as it is relatively difficult to find optimal images for individual binarization methods. As is generally known, a neural network needs a sufficient number of patterns with a suitable distribution, ideally one that corresponds to reality. Therefore, it is necessary to supplement the training set and set the network parameters so that it shows the most suitable results without overtraining. The neural network was learned in 1000 epochs with 40 hidden neurons and a 0.00005 learning rate. Several variants of parameter settings were used, these showed the best success on the verification set. [10]

Figure 2 shows the input images on which the text recognition method was performed. For comparison, the OCR function in the MATLAB programming environment was used, which generally has satisfactory results.

Figure 2 Test images from the industrial environment

The OCR results of the original images and those that were subjected to the automatic binarization method selection algorithm are shown in **Table 1**. It is possible to see that after the application of advanced methods, the success of OCR was increased only due to the appropriate choice of binarization. The average success rate of OCR on original images is around 18 %, after the application of the above methods it increased to almost 63 %. These outputs are filtered and modified to eliminate as far as possible interfering elements and parts of the image that are in no way associated with the text that needs to be cleaned.

Table 1 OCR results that were applied to the original images and to the images that were binarized using the automatically selected method

Test image	Text in image	Original image		Automatic binarization method		
		OCR result	Success rate (%)	Selected method	OCR result	Success rate (%)
1	14675400	- 4 - 7 5 4 - -	50	Nick	- 4 - 7 5 4 0 0	75
2	163229041	- - - - - - - - 1	11	Feng	- 6 - 2 2 9 0 - 1	67
3	734822	1 - - - - -	0	Nick	7 8 - 8 - -	33
4	734952	- - - - - 2	17	Bradley	7 - - 8 5 2	50
5	14675400	- - - - - 4 2 2	13	Nick	1 4 - 7 5 4 0 0	88

4. CONCLUSION

There are many methods by which image binarization can be achieved. The important question, however, is which method is suitable for a given type of input image. It is not possible for there to be one method that is optimal for all types of images. The article describes and applies a method for automatic selection of image

binarization method. The basis for the created algorithms is a backpropagation neural network, as it provides suitable properties for such operations as the type of learning, the required performance or the range of network settings. Part of the article is to test the developed method on five different images with serial numbers from the industrial environment, which show obvious material defects. The result is an average increase in the recognition success of individual characters by 45 %. The next process after binarization is the need to filter the image, reduce noise and eliminate other defects. These processes are another important factor for the correct recognition of text from an image.

ACKNOWLEDGEMENTS

The work was supported by the specific university research of Ministry of Education, Youth and Sports of the Czech Republic No. SP2020/18, SP2020/61 and RPP2020/136.

REFERENCES

[1] FRISCHER, R., DAVID, J., SVEC, P., KREJCAR, O. Usage of analytical diagnostics when evaluating functional surface material defects. *METALURGIJA.* 2015, vol. 54, no. 4, pp. 667-670.

[2] PASKER, V., HLAVICA, R., GRYCZ, O., SPACKOVA, H., MENSIK, M. Optical Character Detection on Industrial Products. In *METAL 2019: 28th International Conference on Metallurgy and Materials*. Ostrava: TANGER, 2019, pp. 1820-1824.

[3] GRDIET, P.B., NARESH, K.G. Binarization Techniques used for Grey Scale Images. *International Journal of Computer Applications*. 2013, vol. 71, no. 1, pp. 8-11.

[4] BRADLEY, D., ROTH, G. Adaptive Thresholding using the Integral Image. *Journal of Graphics Tools*. 2011, vol. 12, no. 2, pp. 13-21.

[5] SAINI, R. Document Image Binarization Techniques, Developments and Related Issues: A Review. *International Journal of Computer Applications*. 2015, vol. 116, no. 7, pp. 41-44.

[6] FRISCHER, R., KREJCAR, O., SELAMAT, A., KUCA, K. 3D surface profile diagnosis using digital image processing for laboratory use. *Journal of Central South University*. 2020, vol. 27, no. 3, pp. 811-823.

[7] GHOSHAL, R., ROY, A, BANERJEE, A., DHARA, B.CH., PARUI, S.K. A novel method for binarization of scene text images and its application in text identification. *Pattern Analysis and Applications*. 2019, vol. 22, no. 4, pp. 1361-1375.

[8] SAXENA, L.P. Niblack's binarization method and its modifications to real-time applications: a review. *Artificial Intelligence Review*. 2019, vol. 51, no. 4, pp. 673-705.

[9] ZHOU, M., WU, Z., CHEN, D., ZHOU, Y. An Improved Vein Image Segmentation Algorithm Based on SLIC and Niblack Threshold Method. In *2013 International Conference on Optical Instruments and Technology: Optoelectronic Imaging and Processing Technology*. Beijing, 2013.

[10] TULI, R. Character Recognition In Neural Networks Using BackPropagation Method. In *3rd IEEE International Advance Computing Conference (IACC)*. Ghaziabad, 2013, pp. 593-599.