An algorithm of the recognition of inferior barcode

He Chuanliang\textsuperscript{1,a}, Xia Xin\textsuperscript{2,b}, Lv Yingjie\textsuperscript{3,c}, Liu Yongping\textsuperscript{4,d}
\textsuperscript{1,2,3,4}Beijing Electric Power Science & Smart Chip Technology Co., Ltd Being, China
\textsuperscript{a}e-mail bigone@aliyun.com, \textsuperscript{b}e-mail xiaxin1@sgitg.sgcc.com.cn,
\textsuperscript{c}e-mail lvyingjie@sgitg.sgcc.com.cn, \textsuperscript{d}e-mail lyp_zlyp@126.com

Abstract—Objective: Bar code recognition is widely used. This paper introduces a method of using digital image algorithm to complete the identification of stain bar code. Methods: firstly, the barcode image was binarized and positioned; secondly, the pixels in the barcode image were stacked on the abscissa axis by projection technology to remove the barcode stains; finally, the barcodes were effectively identified by the center distance algorithm. Results: the experiments show that the barcodes with different degrees of pollution are well recognized by our method, which is better than the traditional recognition algorithm. Conclusion: This paper proposes a new barcode recognition algorithm, which can quickly and accurately identify the bad barcode

1. INTRODUCTION
Bar code is a graphic identifier that arranges black and white bars with different widths according to certain coding rules to express a group of information. Among them, EAN-13 bar code is more widely used in various commodities in supermarkets and books in bookstores. There are a lot of barcode recognition software on the market, which can be divided into two categories: one is scanning by infrared barcode gun; the other is using camera to collect image and then identify through software processing. Both of these methods can perform high recognition work on complete and lossless barcodes. However, some barcodes such as damaged, altered and water-soaked barcodes cannot be effectively identified. This paper studies the coordinate positioning technology and image projection technology of inferior barcodes, and realizes the effective recognition of the code through the central distance measurement algorithm.

2. BARCODE INTRODUCTION
This paper will focus on the identification process of EAN-13 barcodes. Other types of barcode processing methods are similar. EAN-13 barcode is a barcode symbol that represents EAN / ucc-13 commodity identification code. It is composed of left blank area, start character, left data character, middle separator, right data character, check character, termination character, right blank area and human identification character (as shown in Fig.1). Each barcode data character is composed of two black bars and two white bars. Each black bar or white bar is composed of 1-4 modules. The total number of modules of each bar code character is 7, and all characters are 95 modules. Binary "1" is used to represent the module of black bar, and binary "0" is used to represent the module of white bar. Among them, the starting character is fixed as 101, the middle character is fixed as 01010, and the ending character is fixed as 101; the left blank area and the right blank area are used for the start detection of the whole barcode character, and the length is more than 4 modules.
3. RECOGNITION ALGORITHM

3.1. Binary image
The image captured by the camera is presented in RGB format, use the average gray method
Gray=(R+G+B)/3 to process the color image to obtain a 256-level grayscale image (as shown in Fig.2).

The barcode image has obvious brighter areas (white bars) and darker areas (black bars, stains), the
histogram of this type of grayscale image has strong bimodality, as shown in Fig.3. Choose the central
point a value of the two peaks as the threshold for binarization.

Use the threshold formula to convert the 256-level grayscale image into a binary image, as shown in
Fig.4.

\[ f_{ij} = \begin{cases} 0 & (f_{ij} > \alpha) \\ 1 & (f_{ij} \leq \alpha) \end{cases} \quad (i = 0, 1, 2 \ldots m - 1; j = 0, 1, 2 \ldots n - 1) \]
3.2. Coordinate Positioning
Create a two-dimensional array (web form) consisting of elements 0 and 1. The two-dimensional array elements have the distribution characteristics shown in Fig.5.

![Fig.5 Two-dimensional array](image)

Multiply each element in the network table with the pixel point corresponding to the barcode binarization map to obtain a new two-dimensional array and count the number of elements in the array as 1 to obtain $\text{SUM}(0\theta)$. Then the web table is rotated by the unit angle $\Phi$ and then multiplied by the barcode to obtain $\text{SUM}(\Phi)$. By analogy, add up the rotation angle to get $\text{SUM}(2\Phi)$, $\text{SUM}(3\Phi)$...$\text{SUM}(N\Phi)$ $(0\leq N\Phi < 90\degree)$.

Obviously when the vertical axis of the web form is parallel to the black bar of the barcode, $\text{SUM}(N\Phi)$ takes the maximum value: $\text{SUM}(\text{MAX})$, at this time, the vertical axis of the web form is parallel to the coordinate y axis (as shown in Fig.6), thus completing the positioning of the barcode, as shown in Fig. 7.

![Fig.6 Multiplication](image)

![Fig.7 Coordinate](image)
3.3. Image Projection

The common method of barcode recognition is to scan a line, and then calculate the width of the black and white bars to identify the barcode. There is a problem in this way. After the preprocessing of the inferior barcode image, the width of the black and white bars corresponding to each line is different, so accurate identification cannot be performed. As shown in Fig. 8, the results of barcode scanning to line a and scanning to line b are different. This paper uses image projection technology to solve this problem.

![Inferior barcode](image)

First scan the coordinate image data line by line to form a two-digit array, and store 0 or 1 in the array.

\[
m \times n \text{ indicates the number of image pixels}
\]

Then the two-bit array is accumulated vertically to get the number \( A_i \) of element 1 in each column. Sort the resulting \( A_i \) to get the projection effect shown in Fig. 9.

![Projection](image)

Finally select a valid barcode segment: According to the characteristics of EAN-13 barcode, there are 95 modules in total. Scan each line of image from left to right, when there are ninety five 0 to 1 changes in a line, it indicates that the line has complete barcode characteristics. Start scanning from the first row of pixels \( Y=0 \) to \( Y=m-1 \) pixel row, record the first row \( Y_a \) that meets the requirements, and the last row \( Y_b \) that meets the requirements, take the middle row \( Y_m=(Y_a+Y_b)/2 \) as the effective identification row, the \( Y_m \) row is more representative.

![Scan line](image)

3.4. Center Ranging Recognition Algorithm

The actual bar code generated due to water immersion, etc. shrinks and expands in the width of the black and white bars (A1, A2, A3, A4) after the projection process, as shown in Fig.11. Barcode
cannot be translated by directly identifying the width of black and white bars. However, the color of black or white bars spreads evenly to both ends, so the distance that the black or white bars spread to both ends after projection processing is the same. Based on the above principle, it can be concluded that the center position of the black bar and white bar (A1, A2, A3, A4) of the theoretical bar code and the actual bar code does not change. This paper uses this feature to propose a center ranging recognition algorithm to solve the identification of low-quality barcodes.

The digital symbol barcode is composed of 4 black bars and white bars A1, A2, A3, A4 (as shown in Fig.9), and the unit module width is D. Take A1, A2, A3, A4 center points M1, M2, M3, M4. Calibration distance D1 from M1 to M2, distance D2 from M2 to M3, distance D3 from M3 to M4. Fig.12 shows the barcode corresponding to the number character 0 in C arrangement, so D1=2.5*D, D2=1.5*D, D3=1*D.

Since each number character A1, A2, A3, A4 has a different width, each number character has its corresponding D1, D2, D3. According to Table 1, the numeric character table lists the center distances of numeric characters.

| Num | A ARRGT(*D) | B ARRGT(*D) | C ARRGT(*D) |
|-----|-------------|-------------|-------------|
|     | D1 | D2 | D3 | D1 | D2 | D3 | D1 | D2 | D3 |
| 0   | 2.5 | 1.5 | 1   | 1  | 1.5 | 2.5 | 2.5 | 1.5 | 1  |
| 1   | 2  | 2   | 1.5 | 1.5| 2  | 2   | 2  | 1.5 | 1  |
| 2   | 1.5 | 1.5 | 2   | 2  | 1.5 | 1.5 | 1.5 | 1.5 | 2  |
| 3   | 2.5 | 2.5 | 1   | 1  | 2.5 | 2.5 | 2.5 | 2.5 | 1  |
| 4   | 1  | 2   | 2.5 | 2.5| 1  | 1   | 1  | 2   | 2.5 |
| 5   | 1.5 | 2.5 | 2   | 2  | 2.5 | 1.5 | 1.5 | 2.5 | 2  |
| 6   | 1  | 1   | 2.5 | 2.5| 1  | 1   | 1  | 1   | 2.5 |
| 7   | 2  | 2   | 1.5 | 1.5| 2  | 2   | 2  | 1.5 | 1.5 |
| 8   | 1.5 | 1.5 | 2   | 2  | 1.5 | 1.5 | 1.5 | 1.5 | 2  |
| 9   | 2  | 1   | 1.5 | 1.5| 1  | 2   | 1  | 1   | 1.5 |

The numbers 1 and 7, 2 and 8 in the arrangement of A, B, and C in Table 1 have the same center distance table, so it is necessary to sort the numbers with the same center distance. Due to the changes in the width of the black and white bars of the barcode, secondary recognition cannot be achieved by
reading the width of the black and white bars. The distance between the number character M1 and the leftmost starting point Z0 of the number character is specific. As shown in Fig.13.

![Fig.13 C arrangement number character](image)

The Z0-M1 distance of the number 1 arranged in A is 1.0*D, and the Z0-M1 distance of number 7 is 0.5*D;
The Z0-M1 distance of the number 2 arranged in A is 1.0*D, and the Z0-M1 distance of the number 8 is 0.5*D;
The Z0-M1 distance of the number 1 arranged in B is 0.5*D, and the Z0-M1 distance of the number 7 is 1.0*D;
The Z0-M1 distance of the number 2 arranged in B is 1.0*D, and the Z0-M1 distance of the number 8 is 1.5*D;
The Z0-M1 distance of the number 1 arranged in C is 1.0*D, and the Z0-M1 distance of the number 7 is 0.5*D (as shown in Figure 3); The Z0-M1 distance of the number 2 arranged in C is 1.0*D, and the Z0-M1 distance of the number 8 is 0.5*D.

According to the above different distances, the same distance table is identified twice.

**Calculation of unit module width D:** The center position of the first module of the record start character (101): Mstart, the center position of the last module of the record terminator (101): Mend. Calculate the distance between the two Mall=Mend-Mstart, and the width of a single module D=Mall/(95-1).

**Calculation of Z0 point position:** Z0 refers to the starting point of the ideal number character. The width of the black and white bars changes due to image diffusion, so the Z0 position cannot be accurately located by detecting edge changes (1 to 0 or 0 to 1). Calculate ZO by using the last module center position Mend, theoretical width Dend, and the known unit module width D of the previous character of the detected character. Specific method: Z0=Mend+(Dend/2)*D.

### 4. CONCLUSION

Compared with the traditional width measurement method, the recognition algorithm in this paper has a high recognition rate, and can correctly recognize the soaked, damaged, and stained barcodes.

### AUTHOR INTRODUCTION

He Chuanliang (1983), senior engineer, Beijing Electric Power Science & Smart Chip Technology Co., Ltd, main research direction of measurement and control automation system integration.

### REFERENCES

[1] Zhang Zhiling, Liang Xiao, Yang Tianyu, et al. Application of bar code location and recognition in image [J]. Computer knowledge and technology, 2019,15 (14): 140-142
[2] Chen Chao, Yang Liu. Research on Preprocessing Algorithm of barcode recognition based on digital image processing [J]. Journal of Henan Institute of Engineering (NATURAL SCIENCE EDITION), 2015 (03)
[3] Liu Yanyang, Cao Yudong. Identification of EAN-13 barcode image [J]. Computer technology and development, 2015 (06)
[4] Zamberletti A, Gallo I, Albertini S, et al. Neural 1D Barcode Detection Using the Hough Transform. Ipsj Transactions on Computer Vision.&.Applications, 2014
[5] Feng Guowen. Research on SVM based barcode detection algorithm for complex background [D]. Zhejiang University, 2017
[6] Zhang Jin. Research on one dimensional barcode recognition technology based on image processing [D]. China University of Petroleum (East China), 2014

[7] CHEN C, HE B W, ZHANG L W, et al. Autonomous recognition system for barcode detection in complex scenes. The 4th Annual International Conference on Information Technology and Applications, 2017

[8] Zamberletti A, Gallo I, Albertini S, et al. Neural 1D barcode detection using the Hough transform. Information and Media Technologies, 2015

[9] Zhang Jinguang, Chen Xianyi, yuan Zhongkui, et al. Method and data processing equipment for gray image partition binarization [P]. Chinese patent: cn101826159a, 2010-09-08

[10] Raphael C. Gonzalez. Digital image processing [M]. Electronic Industry Press, 2017

[11] Li Fang. Research and design of embedded meter bar code recognition system [D]. Hubei University of technology, 2017

[12] Zhang Chenghai, Zhang duo, Zhao Shouxiang, et al. Barcode technology and application [M]. Beijing: Tsinghua University Press, 2010