A DataMatrix-based Mutant Code Design and Recognition Method Research

Leng Biao
College of Civil Engineering, Southwest Jiaotong University, Chengdu 610031, China 
bleng@126.com

Abstract

Based on DataMatrix 2D barcode, a DataMatrix mutant barcode which was also a matrix 2D barcode was designed. A distorted image of a mutant barcode was obtained by simulating actual environment. The barcode area in the image was acquired by barcode localizing. After the barcode area being analyzed, these barcode characteristic parameters were abstracted and the distorted image fitting formula was gained by exercise, which set up the correspondence between distorted barcode image point coordinates and the standard one. Thus the 01 code of mutant barcode of the distorted image can be abstracted. The implicit information in the 01 code can be decoded with DataMatrix decode method. From the actual testing result, the barcode recognizing speed is faster than others by this way, and this method was robust.

1. Introduction

Barcode[1] is information that consists of a group of bars and space characters arrayed in term of some coding rule and is used to express characters, number and sign composition. Barcode technology congregates coding, press, recognition, data acquisition. It is a new developed technique based on computer and information technology. Presently, 1-dimensional barcode is the mostly used one in china. 1-dimensional barcode is often stuck in the book verso, all kinds of commodities sold in supermarkets and books in libraries. The applications of 2-dimensional barcode are not often seen in china. The mostly used barcodes are PDF417 code, QRCode code, DataMatrix and so on. These are international universal standard barcode, which can be recognized by the corresponding barcode detecting instrument. However, for some special requirements, such as product Anti-counterfeiting and physical distribution management, some particular 2-dimensional barcodes need to be designed. Given this situation, we designed a new barcode based on DataMatrix. It is also a DataMatrix mutation.

2. DataMatrix mutation design

2.1. The introduction to DataMatrix barcode

DataMatrix[2] 2-dimensional barcode is a matrix one, whose characters are similar to algebra matrix in structural elements and arrangements. Each DataMatrix symbol consists of data regions which contain nominally square modules set out in a regular way. Based on computer imaging technique, it has specific character function graph. These models distributes in matrix elements location and expresses data information are foursquare. Fuscous models denote binary “1”, while adlittoral ones express binary “0”. Data code bit stream is presented by different combination of unit models distributed at matrix element places. DataMatrix 2-dimensional barcode is shown in figure 1.

Figure 1. DataMatrix 2-dimensional barcode
Two adjacent sides, the left and lower sides, are solid dark lines. Both of them form the L boundary. They are used primarily to determine physical size, orientation and symbol distortion. The two opposite sides are made up of alternating dark and light modules, which are named railway line. These are used primarily to define cell structure of the symbol, but also can assist in determining physical size and distortion.

2.2. DataMatrix mutation

As is shown in figure 2, In order to position images and extract characteristic point coordinate more conveniently, railway line and L boundary are added in DataMatrix mutation.

From figure 2, the four solid boundaries of the DataMatrix mutation form boundary line. The neighbor four boundaries are railway line. The out and neighbor boundaries are used to position images, and they don’t denote barcode information. The inner barcode area is the same with the one of DataMatrix. Because the bottom boundary is different from the others boundaries, when the mutation is used in commodities, it is necessary to put commodity check code or other letter information to the mutation’s bottom so as to know which orientation is the barcode’s bottom.

3. DataMatrix mutation image pre-processing

In real environment, the DataMatrix mutation barcode shot is affected by various factors, which cause the barcode to form some images with uneven brightness, besmirches and gauss noise. So, these factors affect the quality of barcode seriously, as is shown in figure 3. If the mutation barcode image is directly recognized at this time, the discrimination ratio will be very low. Therefore, the raw DataMatrix mutation barcode image is needed to be pre-processed so as to analyze and recognize the image.

These sizes of DataMatrix mutation barcode image snapped in real environment may be different. Some of them are large, and the others are small. For the sake of better pre-processed outcome being gained to different mutation barcode images, USM(Unsharp Mask) algorithm is considered to process the mutation barcode image.

The essential of USM algorithm pre-processing images is convolution operation to images with a larger image convolution template. The cell number of the convolution template \( n = (r \times 2 + 1) \times (r \times 2 + 1) \). \( r \) means the template radius. All elements’ power is \( p = -\frac{(a-1)}{(n-1)} \) except the center point, in which \( a \) denotes the center point power. After being tested, when \( r=11 \) and \( a=25 \), better pre-processed effect can be obtained using the template convoluting the raw image. Figure 4 is the result of figure 3 convoluted by the template.

As can be seen in figure 4, USM algorithm has balanced the asymmetry brightness in figure 3. In addition to, the bottom-right corner of the image in figure 3 is much blur because of the camera defocusing, which has been improved in figure 4 after image pre-processing.

4. DataMatrix mutation positioning

To improve image processing speed, the barcode area should be separately analyzed. The 01 code should
not be extracted by analyzing the whole image. Therefore, finding and positioning barcode area are needed.

For finding barcode area easily, the pre-processed image will be binaryzed. Figure 5 is the binaryzation result of figure 4.

![Figure 5. the binaryzation image](image)

In figure 5, the barcode area is surrounded by the four out boundaries, at the outside of which there is a wide white bar that separates the barcode area from the outside area. At the same time, the barcode length and width are nearly equal, so the barcode area can be found and positioned in term of the barcode connectivity and the ratio of the length and width in the binary image. As a result, the barcode area image is analyzed separately.

5. The railway line boundary center coordinate extraction

The extraction of the railway line boundary center coordinate is a key part. If the center coordinates are well extracted, the recognition effect are perfect, vice versa.

5.1. The barcode image edge detection

Generally speaking, edge detection[3] is gray image detection, which is applied to binary images here. After being processed, the edge lines in the barcode area can be acquired, as is shown in figure 6.

![Figure 6. the edge detection result to the binary image](image)

5.2. barcode horizontal and vertical edge lines

In order to extract the railway line boundary center, it is necessary to do with the barcode boundaries further in figure 6, which will be processed horizontal image and vertical image separately. For the barcode vertical edge line image being gained by processing, the processing of figure 6 edge line image is as follows:

A horizontal template was designed in which there are two elements, as is shown in figure 7. The cell marked with “×” is regarded as the template center, which is aligned with current processed pixel in the image. When the edge line image is processed, the horizontal template is moved in the image. If the gray values of two pixels aligned with the template are 255, which denotes the two pixels form a white horizon line, then the gray value of the pixel corresponded with the template cell marked with “×” is set as 0, which means the current pixel is black. After all pixels in the edge line image are processed, there will only be vertical edge lines, as can be seen in figure 8.

![Figure 7. the horizontal template](image)

In a corresponding way, the horizontal edge line image can be obtained from the image shown in figure 6.

5.3. Characteristic point coordinate extraction

After the vertical edge line image in figure 8 being analyzed, these broken lines that closely adjoin the left and the right vertical line are edges of the barcode left and right vertical railway lines. These center coordinates of these short vertical lines are required extracting. These left railway line character points are
extracted according to this method, which is shown in figure 9.

Figure 9. vertical railway line finding

Figure 10. cell center mark in distorted barcode image

At first, the pixel whose gray value is 255 in the vertical railway line image is queried from left to right. The query course is up-down. If the current pixel gray value \( g_i = 255 \), and the distance between \( g_i \) and \( g_j \) \( D(g_i, g_j) < \kappa D_{\text{max}} \), in which \( D_{\text{max}} \) is the barcode cell width and \( \kappa \) is the power, then \( g_2 \) is taken as a point in the railway line edge, whose coordinate is recorded. The current queried row is halted at this time. The next row is queried. If all pixel gray value in the current row are 0, the query course is stopped. As can be seen from these coordinates of all recorded points, these Y-coordinates of all pixels in the same vertical railroad line are monotone increasing. As a result, all recorded points can be classified formats such as \{A(\( c_{0b}, c_1, ..., c_a \)), B(\( c_{a+1}, c_{a+2}, ..., c_b \)), ...Z(\( c_{a+y}, c_{a+y}, ..., c_{a+z} \))\}. All points in every class are in the same vertical railroad line. These lines are named line A, line B etc.

The coordinate average value \( A(x, y) \) of class \( A(c_0, c_1, ..., c_a) \) is set as the center coordinate of line \( A \), in which

\[
\begin{align*}
x_A &= \frac{1}{a+1} \sum_{i=0}^{a} c_i, \\
y_A &= \frac{1}{a+1} \sum_{i=0}^{a} c_j
\end{align*}
\]

These center point coordinates from line B to line Z are also set in the same way.

All center coordinates of all lines in right edge line are similarly extracted from right to left. Similarly, these center coordinates of all lines in top and bottom line are extracted in the same way.

6. Geometric correction of imagery

Geometric correction[4] that corrects images using GCP (Ground Control Point) is a mathematic module that is used to describe image geometric distortion course. And the geometry distortion module is gained by using some corresponded points between the distorted image and the standard image. Then the module is used to correct geometric distortion. The correction corrects images only by making use of the distortion module, which doesn’t depend on specific reasons. The geometric correction course often contains geometry location transform and all image primitive lightness confirmation after being transformed. Geometric correction may be divided into four steps. They are as follows:

(1)The original image coordinate system and the corrected image one are built. For the corrected image, the origin of coordinates (the beginning row and column), the image primitive size and the image size (the row number and the column one) are demanded to ascertain.

(2)The control point pair is required finding out between the original distorted image space and standard space.

(3)The distortion mathematics module needs to be selected, and the module unsure parameters are figured out in term of these GCP data. Then the original distorted image is corrected by means of this distortion module.

(4)Accuracy analyses of the geometry correction. If the GCP selection is not accurate, the GCP number is low, the GCP distribution is unreasonable and the distortion mathematics module may not well reflect the geometry distortion process, the accuracy of the geometry correction will slow down.

The arithmetic correction module is:

\[
\begin{align*}
x' &= a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 + a_6x^3 + a_7x^2y + a_8xy^2 + a_9y^3 \\
y' &= b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + b_6x^3 + b_7x^2y + b_8xy^2 + b_9y^3
\end{align*}
\]

(2)

In the module, \((x', y')\) denotes the point coordinate in the distorted barcode image. \((x^*, y^*)\) represents the standard barcode point coordinate before the image is distorted. The corresponding relation between the center coordinates of all lines in the real railway line edge and these in standard barcode is established, which is imported to equation (2). These values of \(a_{0b}, a_{1b}, ..., a_{ob}, b_{1b}, ..., b_y\) can be figured out from equation
All image primitive center coordinates in the standard barcode image are inducted to equation (2), then these coordinates in the distorted image corresponded with those in the standard image. In figure 10, black and white points express corresponded points in the distorted image which are corresponded with the center of each unit in the standard barcode image. Whether the gray value is 0 or 255 is judged. If the gray value is 0, the code of this point is set as 1; If the gray value is 255, the code is set as 0. All 01 codes obtained are arrayed top-to-down and left-to-right, then the 01 code of the barcode is formed. Thus the mutation barcode denotation can be explained according to DataMatrix decode method.

The module can reach 0.2% statistics error precision. When it is operated to pixels, it is nearly impossible to produce pixel error. Besides, this method avoids detecting edge lines and rotating barcode images[5], so it greatly improved recognition speed.

7. Arithmetic test

Many DataMatrix mutation images are photographed without belt micro distance Nokia 6300 cellphone. The angles of these mutation barcodes are limited in -30°~+30°. Those images are inducted to the system that is researched and developed in view of the arithmetic given in this paper. The test result is shown in table 1.

| The rotated angle | Recognition Ratio |
|-------------------|-------------------|
| -10°~+10°         | 95.47%            |
| [-20,-10)U[+10,+20) | 84.62%            |
| [-30,-20)U[+20,+30) | 62.86%            |

After being tested, we can find that the barcode rotated angle is more closed to 0, the recognition result is much better. With the rotated angle being increased, the recognition ratio is gradually decreased. The mutation barcode recognition ratio is more than 90% when mutation rotated angle is from -10° to +10°. If the rotated angle is less than -10° or more than 10°, the DataMatrix mutation recognition ratio is very low. At the same time, for these images photographed in different angles, even though these images are distorted, they still can be well recognized.

8. Conclusion

A new DataMatrix-based 2-dimensional mutation barcode is designed in this paper. It can be concluded that mutation barcodes are easy to be analyzed through arithmetic analysis. The actually DataMatrix mutation barcode image is analyzed by using USM arithmetic which has good effects on processing and is robust to actual photographed mutation images. Through characteristic extraction, the length and width of all connected area that include mutation in mutation barcode images are gained. The mutation barcode location in the distorted images is positioned in term of pattern recognition method. As can be seen from the actually testing result, this method has a good effect on positioning the location. Through edge detection and railway line foundation, a method of extracting all line center coordinates of railway line edge is designed. Furthermore, that using distorted face fitting method fits distorted DataMatrix mutation barcode images avoids testing image angles and rotating images, which can greatly improve the speed of image recognition.

9. Reference

[1] XU-Jie min, XIAO Yun, “2-dimension Barcode Technological Actuality and Prospset”, JISUANJI YU XIANDAIHUA, Jiangxi, China, 2004.12, pp. 141-142.
[2] INTERNATIONAL ISO/IEC STANDARD 16002. Information technology-International symbology specification-Data matrix.
[3] KENNETH R.CASTLEMAN. DIGITAL IMAGE PROCESSING, Prentice-Hall International, Inc, 2002.
[4] Ning-Shu nian, Nu-Song tang, Yang-Xiao qing etc, Remote Sensing Image Processing and Applicatio, Earthquake Publish House, 1995.
[5] Cheng-Yuan yuan, Shi-Peng fei. “2-Dimension Barcode Recognition And Application”, Measure and Control Technology, 2006, 25(12), pp. 17-19.