

# Active threshold design of PDF-417 two-dimensional bar-code

La Yeon An\* , Hong Chae Woo\*\* , Han Yong Kim\*\*\*

\*A College of Information and Communication Daegu University , [lyan@daegu.ac.kr](mailto:lyan@daegu.ac.kr)

\*\* A College of Information and Communication Daegu University , [hcwoo@daegu.ac.kr](mailto:hcwoo@daegu.ac.kr)

\*\*\*AD Information and Communication , [manager@adinc.ac.kr](mailto:manager@adinc.ac.kr)

**Abstract** - In this paper, an algorithm to extract bar-space area is suggested . In a section of bar-code space area the threshold value is computed, and bar and space are extracted according to threshold value. PDF417 is used everyday life and printed in many different materials. The printed PDF417 is especially influenced by various light source. The decision of bar and space is very hard under the change of illumination. The fixed threshold value to distinguish the bar and space can not be applied. in these cases, The proposed algorithm is developed to investigate variable threshold. The variable threshold can be obtained by simple calculation.

**Keywords:** two-dimentional bar-code, edge location detection

## 1 Introduction

Bar-code that is commonly used on modern society often can see in many surroundings. We can observe bar-code which is composed of black lines and various numbers, most of goods that is exhibited in the market. These work as a key to store data such as code and information of goods briefly as most one dimensional bar-code or to approach to DB of goods and to approach to relevant information. However, we need to store more data or information. The one- dimensional bar-code has a limited capacity. Two dimensional bar-code can store information tens or hundred times more then one dimensional bar-code [1].

These two-dimension bar-code can have data information of Hangul, Japanese, Chinese character. as well as English and number. It has high density, so it is to restoration when data was destroyed more than 30% [2].

In this paper an algorithm to convert bar-code into binary image in various environment is investigated. It is focused on the process which a captured image is changed into a binary image using an active threshold value to de-code PDF417, two dimensional bar-code.

## 2 Symbol structure of PDF-417

PDF417 as character set has available structure of symbol character formed 17 modules that is expressed by each 4 bars and space arrays. So it has named as PDF417. Signal has 3~90 rows with 1~30columns. One of the column in figure 1 is divided into start pattern, stop pattern, left and right row indicator code words and data region. Although capacity of maximum code word is 928 and maximum data code word capacity is 925, the ratio of

width and vertical length in PDF417 symbol can be variable according to the conditions of printing space.

PDF417 is consisted of rows more than 3 minimum code words. Data region must include at the least 1 symbol character except beginning character, stop character and row indicator code words. PDF417 must have top margin, bottom margin, right margin and left margin. With beginning pattern and end pattern, the direction and bar-code can be recognized. As bars and spaces are exist alternately, start pattern is known as 81111113(bsbsbs'cs) and stop patterns is 711311121. [3].

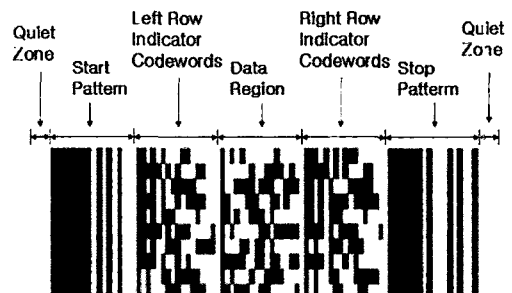


Figure 1. Symbol structure of PDF 417 code.

The left and right row indicator code words has information of total number of a row code word as symbol length descriptor. These indicator code words has information about row number and code error correct on level. The code words of data region are arranged in the order of symbol length descriptor, data code word, pad

code word from left to right and from top to bottom. The symbol length of data region is computed except error detection code word.

### 3 PDF-417 two-dimensional bar code reading

When a bar-code is used on the spot, it can be disturbed of various noise since it attached to goods or to package unit such as box, palette, and it is recognized through image processing. To remove these noisy factors, the proposed algorithm developed. This algorithm can produce a variable threshold and can result a binary image bar-code.

#### 3.1 Extraction of Bar-code Region

Two-dimensional bar-code has special code type for detecting a position of bar-code and its standard. In case of PDF417, a start pattern and a stop pattern is one of the special code types. These patterns guide to read PDF417 bar-code.

After acquiring an image of the bar-code, the start pattern is searched. In this procedure it computes overall length of bar-code form starting black stripe to ending black stripe. The vertical length of a bar-code is computed from bottom point to top point. In the next step, a start pattern is searched. The start pattern as shown in figure 2 is fixed as 81111113 and it is a guideline to detecting remaining patterns. In figure 3 a structure of start pattern is given. It has a special sequence of numbers in which the number 8 pattern has the most wide black stripe and the number 1 pattern has the most narrow black stripe. The mix of number 8 and 1 patterns can be easily detected [4].

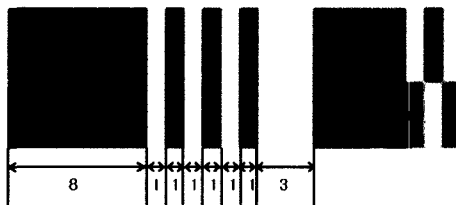


Figure 2. Structure of Start Pattern

#### 3.2 Extraction of Dimension

A dimension in PDF417 is defined as the width of number 1 pattern in figure 2. The width of number 1 pattern is not fixed because the image of PDF417 pattern can have many different sizes.

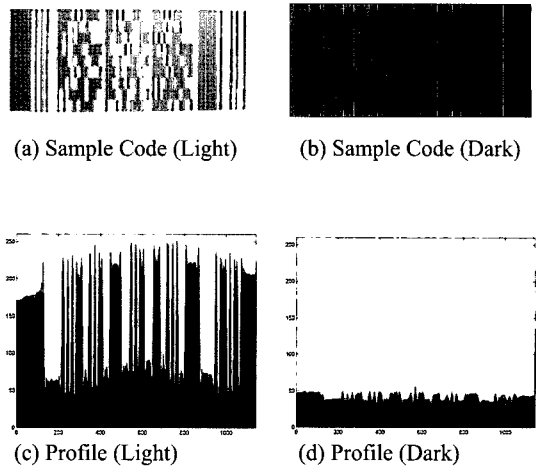


Figure 3. Sample code of PDF417 and Histogram

The PDF417 image has 8 bit resolution which is 256 gray levels. The bar-code image is quite different according to surrounding environment. The biggest impact is illumination from light source. The image figure 3 has dark intensity or light intensity. In figure 3 (c), the profile is obtained from the top row of figure 3 (a). In figure 3 (c) the profile has a large gap between the highest point and the lowest point so that bar and space patterns can be distinguished easily. In figure 3 (d) the profile has a small gap so that it is hard to distinguished bar and space pattern. [5].

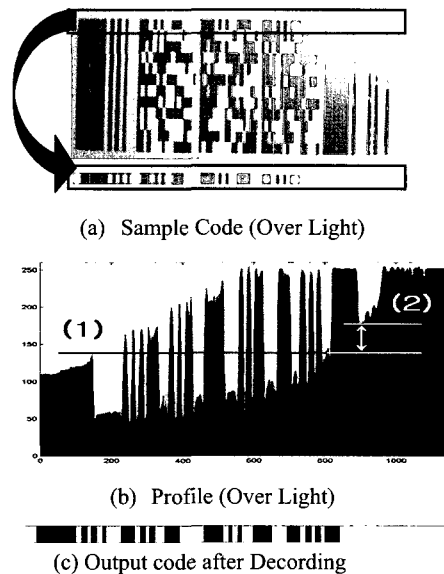


Figure 4. Sample code of PDF417 and Histogram

In case of figure 4 (a) the intensity of image is not flat. The intensity becomes high as it goes right-side. The irregularity of intensity produces the profile in figure 4 (b). In figure 4 (c) the legend (1) shows the highest point in the dark region and the legend (2) is the lowest point of the bright region. Since the legend (2) is lower than the legend (1), it is hard to distinguish bar and space pattern.

### 3.3 Extraction of variable Threshold Value

We propose a variable threshold approach to extract codeword from irregular intensity bar-code images. As shown in figure 5 one dimension pattern is divided into 3 zones. When the average of the first zone is greater than the one of the second zone, the largest average among 3 zones is decided as the maximum value of the dimension. When the average of the first zone is smaller than the one of the second zone, the smallest average among 3 zones is decided as the minimum value of the dimension. In each dimension, the threshold is decided as the average of current dimension minimum or maximum, and each of 3 future dimension minimum or maximum value in figure 6. It looks like a type of moving average.

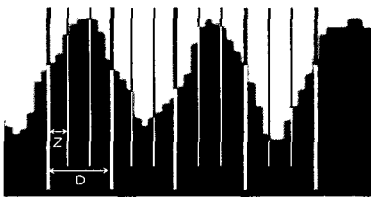


Figure 5. Dimension and Zone ( Z = Zone, D = Dimension)

The variable threshold can overcome the irregularity of light intensity so that correct codeword can be extracted.

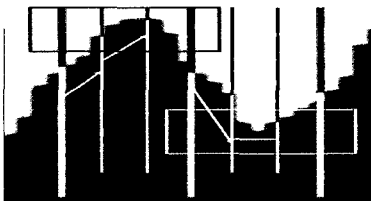
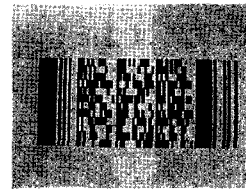


Figure 6. extract of threshold value

## 4 Test result

The bar-code image is quite different according to surrounding environment. The image figure 7(1) is highest point in original sample image of PDF417. The image figure 7 (2) and (3) has dark intensity or light intensity highest point in sample image. The image figure 7 (4) is irregularity of intensity, figure 7 (5) has dark region and bright region. The image figure 8 is these image's profile. The different of intensity produces



(a) normal sample image



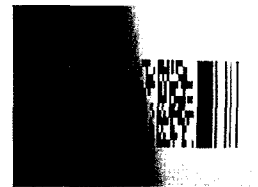
(b) Bright image



(c) Dark image



(d) abnormal illumination image



(e) two regions in sample

Figure 7. Data Image and Sample Image

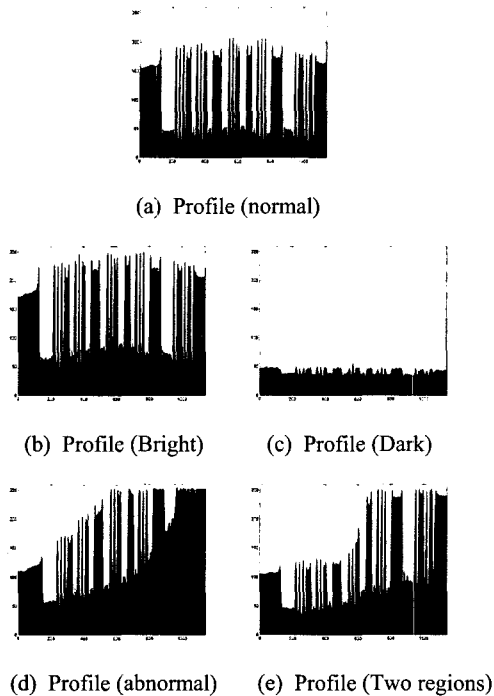


Figure 8. Profile of sample image

Values of table 1 are result of 2-way test. One test used fixed threshold and other test used fixed threshold. These result show result of success rate with variable threshold better then result of success rate with fixed threshold. Exceptionally result of abnormal illumination image has lower success rate, because this sample image has a lot of white zone why narrow and bright light. But table 1 shows improve success rate of abnormal illumination image

Table 1. Test result

( /100)	Success rate with fixed threshold	Success rate with variable threshold
(a) Normal	97	100
(b) Bright	100	100
(c) Dark	63	98
(d) abnormal	12	89
(e) Two regions	57	100

## 5 Conclusions

The proposed algorithm can produce a effective variable threshold for PDF417 and can result a clean bar and space image. The 100 test image were a irregular intensity image, dark or bright images, and abnormal illumination image, in each case , Success result is higher than test of fixed variable value.

## References

- [1] Understanding 2D Symologies a detailed overview and technical introduction , AIM USA
- [2] Roger C. Palmer, The Bar Code Book
- [3] Uniform Symbology Apecification PDF417,AIM USA
- [4] H.I. Han and J.K. Joung, "Implementation of Algorithm to Decode Two-Dimensional Bar-code PDF-417", ICSP, 2004
- [5] Gonzalez E. Woods, Digital Image Processing