ISSN 2229-5518

# Robust Algorithm for Developing Barcode Recognition System using Web-cam

Sampada S. Upasani, Adarsh N. Khandate, Ankita U. Nikhare, Rupali A. Mange, R.V.Tornekar

**Abstract—** This paper presents an algorithm to develop a Supermarket billing system, which recognizes the barcode using image processing. MATLAB software is used to test the algorithm. There are various types of barcodes in use today. These include Code 128, Code 39, EAN 13, etc. Our algorithm decodes EAN 13 barcode that is used universally. In this algorithm, the database of products is created in an excel sheet. Images of products are taken using USB webcam, which is to be installed manually in MATLAB Image Acquisition Toolbox. Preprocessing is done that includes processes such as binarization; spatial filtering and edge detection on these images to extract the barcode region and the barcode number is decoded using the decoding algorithm. This barcode number is further compared with the database, and the corresponding information of the scanned product is displayed in a different excel sheet. A final bill of all scanned products is created after scanning of all products. Proposed System provides a convenient method of extracting information from the barcode at a lesser cost as compared to typical electronic barcode scanners. This system can be used by businesses requiring accurate results at low cost.

**Index Terms**—Barcodes, Barcode localization, Contrast enhancement,EAN-13,Edge detection, Threshold, UPC-A

## 1 Introduction

Barcode is a visual depiction of information in the form of bars and spaces on a surface. The bars and spaces can be of different widths and consists of characters, numbers and symbols such as colon, dot, and others. Set of various combinations of these alphanumeric characters is used to depict information. There are different types of barcodes in use today. These include Code 128, Code 39 [1], and EAN 13.The barcodes are read using scanners that use laser beams or cameras. In the supermarket, there are many hand-held scanners [2] that automatically scan the barcode on the merchandises. This paper discusses an algorithm for barcode recognition system by using a web camera or digital camera for capturing barcode image and then display the barcode information for the user. Nowadays most of the barcode scanners are using infrared to scan a barcode. These types of scanners may lead to the cost issues as those scanners are expensive and unaffordable to the common user. These cost issues can be solved by using the camera-based system for reading barcodes.

It involves two steps for automatic scanning of all the barcodes:

1. To find the position of a barcode
2. To decode the barcode.

Unfortunately, images taken by a web camera or digital camera [3] are often of low quality. In this system, the proposed algorithm detects the barcode region using gradient method and this region is then cropped. Image-preprocessing module transforms the cropped barcode image into a preprocessed grayscale image, then reduces noise, and enhances the contrast between bars and spaces in this preprocessed image.

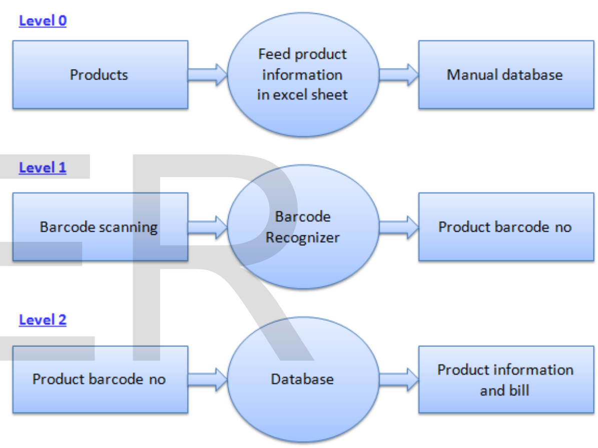



Fig. 1.Conceptual analysis diagram

This preprocessed image is then given to the decoding algorithm to extract the barcode number. Comparison of this number with the database extracts the product information.

## 2 EAN-13 Overview

An EAN-13 barcode was originally called European Article Number but now renamed as International Article Number, although the abbreviation EAN has been retained. It is a 13 digit (12 data and 1 check) bar coding standard. It is a superset of the original 12-digit Universal Product Code (UPC) [4] system developed in the United States. Fig. 2 shows the EAN-13 barcode. As EAN-13 is considered as a superset of UPC-A, any software or hardware that is capable of reading an EAN-13 symbol can also read a UPC-A symbol.

ISSN 2229-5518

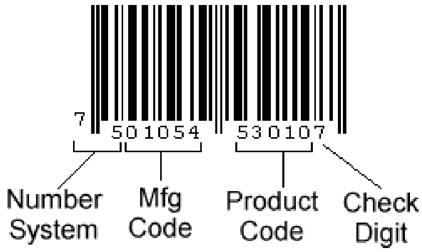



Fig. 2.EAN-13 Barcode

The following process is followed for binary encoding of data digits into EAN-13 barcode: Firstly, the digits are split into three groups such that the first digit, the first group of 6 and the last group of 6, form three distinct set of numbers. The first group of six is encoded using a scheme where each digit has two possible encodings, one that has even parity and one that has odd parity. The first digit is encoded by selecting a pattern of choices between these two encodings for the next six digits, as shown in Table 1 and Table 2. (As opposed to other digits, the first digit is not represented by a pattern of bars). Every digit in the last group of six digits is encoded using a single set of bar-patterns that are the same patterns that are used for UPC. If the first digit is zero, all digits in the first group of six are encoded using the patterns that are used for UPC; therefore, a UPC barcode is also considered as an EAN-13 barcode with just the first digit set to zero.



TABLE 1
ENCODING OF EAN-13

| First digit | First group of 6 digits | Last group of 6 digits |
|---|---|---|
| 0 | LLLLLL | RRRRRR |
| 1 | LLGLGG | RRRRRR |
| 2 | LLGGLG | RRRRRR |
| 3 | LLGGGL | RRRRRR |
| 4 | LGLLGG | RRRRRR |
| 5 | LGGLLG | RRRRRR |
| 6 | LGGGLL | RRRRRR |
| 7 | LGLGLG | RRRRRR |
| 8 | LGLGGL | RRRRRR |
| 9 | LGGLGL | RRRRRR |

TABLE 2
ENCODING OF L,G,R CODES

| Digit | L-code | G-code | R-code |
|---|---|---|---|
| 0 | 0001101 | 0100111 | 1110010 |
| 1 | 0011001 | 0110011 | 1100110 |
| 2 | 0010011 | 0011011 | 1101100 |
| 3 | 0111101 | 0100001 | 1000010 |
| 4 | 0100011 | 0011101 | 1011100 |
| 5 | 0110001 | 0111001 | 1001110 |
| 6 | 0101111 | 0000101 | 1010000 |
| 7 | 0111011 | 0010001 | 1000100 |
| 8 | 0110111 | 0001001 | 1001000 |
| 9 | 0001011 | 0010111 | 1110100 |

## 3 PREVIOUS WORK

Barcode reading has been studied for decades. And after continuous improvement, it is now represented by a well-established industrial standard. Until recently, barcode reading was performed almost exclusively with dedicated hardware. Even though there is growth in the use of camera-based readers, most of the challenges posed by this new method are yet to be solved. Commercial scanners that are used in supermarkets shine a stripe of pulsed light on the barcode and record the intensity of its reflection. As it makes use of active illuminations, it makes barcodes virtually insensitive to changes of ambient illumination. Additionally, their design often requires the barcodes to be relatively close to the scanner to scan it successfully. In general, these dedicated devices produce a high-quality signal that allows for robust barcode reading. On the other hand, reading barcodes with cameras present new challenges. The major issue faced is with the poor quality of the image due to noise or low contrast.

Orazio Gallo and Roberto Manduchi[5] proposed a new approach to barcode decoding that bypasses binarization. Their technique relies on deformable templates and exploits all of the gray-level information of each pixel. They have used UPC-A symbol of barcode and presented a new algorithm for barcode decoding (localization and reading) that can deal with blurred, noisy and low-resolution images.

Luping Fang and Chao Xie [6]present an algorithm, which can detect barcode region even in a complex background with the help of region-based image analysis. Being different from traditional region-based image analysis, the algorithm is processed with an integral image. It can deal with images that are blurred, contain obliquely positioned barcode regions and are shot under different illumination.

Xianyong Fang[1]has proposed a fast and robust recognition method for noisy Code 39 barcode. The

ISSN 2229-5518

proposed method follows two steps: search and decoding. In the first step, all asterisks in the image are found with evenly defined scan lines and then those lines with the same directions are matched together to get a valid barcode region. In the second step, a local de-noise method is applied for eliminating noise in the barcode region. Further, a middle band filter is used for decoding the barcode. Their method also has certain shortcomings. For images whose resolution are quite different from 2400x3500, this approach cannot successfully recognize the barcode automatically as many threshold values need to be set again, e.g. the local de-noise threshold.

## 4 Barcode Localization

We found out that the simple and fast algorithm presented by Orazio Gallo and Roberto Manduchi [4] provides excellent results even in challenging situations. This localization algorithm assumes that the image of the barcode is captured by the camera-oriented system such that its vertical axis is approximately parallel to the bars as shown in Fig. 3. This image is further converted to a grayscale image as shown in Fig. 4. Thus, with respect to the barcode, one should expect an extended region represented by strong horizontal gradients and weak vertical gradients. Accordingly, we compute the horizontal and vertical derivatives, Ix(n) and Iy(n), at each pixel n. Then we combine them in a nonlinear fashion given as

$$Ie(n) = |Ix(n)| - |Iy(n)|$$

It is acceptable to assume that many points within a barcode should have a large value of Ie(n). We run a block filter of size 31x 31 over Ie (n), obtaining the smoothed map Is(n).

Fig. 3.Original Image

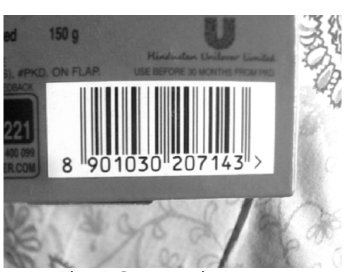


Fig. 4.Gray scale Image

The size of the filter was selected based on the range of the size of the input barcode images and the minimum size of the barcode readable by our method. Note that block filtering can be implemented efficiently so that only a few operations per pixel are required. In the end, we binarize Is(n) with a single threshold that is selected using the method proposed by Otsu [7], and the output image is as shown in Fig. 5. This binarization is only used to localize the barcode, while the decoding is performed on the grayscale image. Due to thresholding, the map Is(n) may contain more than one blob. Instead of computing the connected components of the thresholded map, we just select the pixel,No that maximizes Is(n), with the assumption that the correct blob (the one corresponding to the barcode) contains such a pixel. In our research, this assumption was always found to be correct. We then expand a vertical and a horizontal line from No, and form a rectangle with sides parallel to the axes of the barcode image and containing the intersections of these lines, with the edge of the blob. Note that a quiet zone, a white region around the barcode that facilitates localization, borders the leftmost and rightmost bars of a barcode. The quiet zone and the large size of the block filter, ensures that the vertical sides of the rectangle fall outside the area of the barcode by at least a few pixels as shown in Fig. 6.


Fig. 5.Gaussian filtered Image

Fig. 6.Barcode region Image

ISSN 2229-5518

## 5 Image Pre-processing

### 5.1Cropping

When an image is taken for scanning, along with the barcode, other unnecessary information like letters and words surrounding the barcode also gets included. Thus, it is very crucial to remove the noise and get only the barcode region for proper scanning and decoding. After barcode localization, only the barcode region is visible and surrounding information is eliminated by making all other pixels value to 0. Then, we perform cropping by observing the intensity of each pixel and extracting the rows of a barcode, i.e. the pixels having the intensity greater than 0. Fig. 7 depicts the cropped image of barcode form Fig. 6.

Fig.7.Cropped Image

### 5.2 Contrast Enhancement

In a real-time image of a barcode, the contrast between white and black bars is poor due to blurriness.Hence, it is necessary to improve the contrast of the image to distinguish between bars. This is done by making the black bars one shade darker on the grayscale compared to white bars as shown in Fig. 8.

Fig. 8.Contrast improved Image

### 5.3Converting to ideal barcode image

The contrast-enhanced image is firstly binarized. But due to the real-time acquisition, this image is distorted. To convert into an ideal image, every column of an image is scanned and checked for a maximum number of pixel intensity of either 0 or 1. If a particular column has more number of pixels with intensity 1 than 0, then the whole column is converted to pixels with intensity 1. Fig. 9 shows an ideal barcode image.

Fig. 9.Ideal Barcode Image

### 5.4Edge Detection

An edge is a pixel at which there is a sudden intensity change. An array of such pixels is created that forms the edges of the bars of the barcode image. Bar widths are calculated by subtracting consecutive elements of the array of edges. These bar widths are given as input to the decoding algorithm.

## 6 Decoding Algorithm

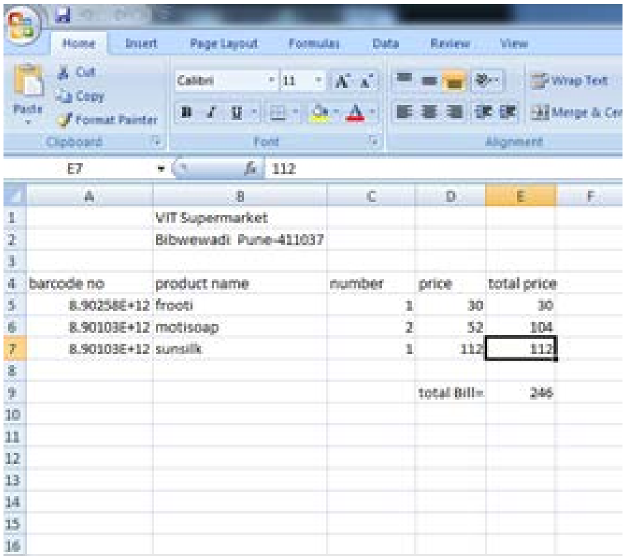
The barcode number is decoded using the array of bar widths. The first digit is always 8 as it is the first digit of the country code and is fixed. As from the Table 1, the first group of 6 digits is encoded as LGLGGL and the last group of 6 digits is encoded as RRRRR. Each digit is represented by 4 bars, such that black and white bars are alternate. For the first 6 digits the first bar is always a white bar and next black and so on and for last 6 digits it is exactly reverse. A bar can have four widths from 1 to 4. From the array of bar widths, the first and last three widths are always one indicating the starting and ending of the barcode. The remaining widths are grouped 4 at a time as a particular digit is represented using 4 bars. In Table 2, 0 represents white, and 1 represents black. For example, if a group of 4 widths (representing a number from the first 6 digits) is 3112, it is considered as 0001011 (3 is the width of a white bar hence represented by three zeroes). Here, 3 is the width of a white bar, 1 is of black and so on. Thus from the Table 1, this first group of 4 widths that is encoded under the L-code is searched for in Table 2 under the L-code and gets matched with the code of digit 9. Hence, the first digit is decoded as 9. Similarly, other digits are decoded using the table.

## 7 Display of Product Information

Initially, a database of products is created in an excel sheet using MATLAB R2009b. It is very easy to import and export data from an excel sheet to MATLAB. Barcode data can be entered into and read from an excel sheet in MATLAB conveniently. The database consists of fields such as barcode number, product name, manufacturer's name and product price. Other fields can also be added as per requirement. (Note that in an excel sheet, the 13 digit barcode number is not displayed fully in the column but is visible in the form of an exponential number. To read the whole number it is necessary to click on that particular cell.)

After the whole number is decoded, it is compared with the database and when a match is found the corresponding information is displayed in another excel sheet. As more number of barcodes is scanned their information is added to the same sheet and a bill is created (by adding the price of each product) as shown in Fig. 10.

ISSN 2229-5518



| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | | VIT Supermarket | | | |
| 2 | | Bibwewadi Pune-411037 | | | |
| 3 | | | | | |
| 4 | barcode no | product name | number | price | total price |
| 5 | 8.90258E+12 | frooti | 1 | 30 | 30 |
| 6 | 8.90103E+12 | motisoap | 2 | 52 | 104 |
| 7 | 8.90103E+12 | sunsilk | 1 | 112 | 112 |
| 8 | | | | | |
| 9 | | | | total Bill= | 246 |

Fig. 10.Final Bill

## 8 Conclusion

We have presented an algorithm for decoding (localization and reading) barcodes through a webcam, using image-processing techniques such as filtering, cropping, and contrast enhancement and edge detection. This localization technique is fast and accurate even for cluttered images. Barcodes having a flat surface and a width to height ratio between 2 to 3 are successfully decoded. We have implemented this algorithm in MATLAB R2009b and can work on any of the later versions of MATLAB. We have scanned multiple products by executing this algorithm in a loop, and a final bill of products is created in an excel sheet.

This is an efficient algorithm, which can be further improved by adding angle correction method and can be provided to small grocery shops that cannot afford costly barcode scanners. This saves a lot of time than manual calculation and also removes the possibility of human error. We have extensively tested our algorithm on challenging images and the accuracy is found to be above 60 percent.

## References


[1] Xianyong Fang, FuLi Wu, Bin Luo, Haifeng Zhao and PengWang,"Automatic Recognition of Noisy Code-39 Barcode," *Artificial Reality and Telexistence–Workshops*, 2006. (ICAT '06), pp.79,82, Nov. 29 2006-Dec. 1 2006,doi: 10.1109/ICAT.2006.45H.

[2] A.S.Bhaskar Raj, *Barcodes Technology and implementation*,pp.1-15.

[3] Chai D, Hock F, "Locating and Decoding EAN-13 Barcodes from Images Captured by Digital Cameras," *Information, Communications and Signal Processing, 2005 Fifth International Conference*,pp.1595, 1599,doi: 10.1109/ICICS.2005.1689328

[4] Thomas, V.M., "A Universal Code for Lifecycle Management of Products," *Electronics & the Environment, Proceedings of the 2007 IEEE International Symposium*, pp.180,183, 7-10 May 2007 doi: 10.1109/ISEE.2007.369390.

[5] Gallo, O.,Manduchi, R., "Reading 1D Barcodes with Mobile Phones Using Deformable Templates," *Pattern Analysis and Machine Intelligence, IEEE Transactions*, vol.33, no.9, pp.1834,1843, Sept. 2011 doi: 10.1109/TPAMI.2010.229

[6] Luping Fang; Chao Xie, "1-D Barcode Localization in Complex Background," *Computational Intelligence and Software Engineering (CiSE), 2010 International Conference*, pp.1,3, 10-12 Dec. 2010 doi: 10.1109/CISE.2010.5677072

[7] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, Jan. 1979.