EE 5359 MULTIMEDIA PROCESSING
SPRING 2013

INTERIM REPORT

MAV-ID card processing using camera images

Under guidance of
DR K R RAO
DEPARTMENT OF ELECTRICAL ENGINEERING
UNIVERSITY OF TEXAS AT ARLINGTON

Presented by
Aditya Sharma
1000803349
aditya.sharma@mavs.uta.edu
**List of acronyms in alphabetical order:**

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHE:</td>
<td>Adaptive histogram equalization</td>
</tr>
<tr>
<td>ASCII:</td>
<td>American standard code for information interchange</td>
</tr>
<tr>
<td>CLAHE:</td>
<td>Contrast limited adaptive histogram equalization</td>
</tr>
<tr>
<td>HT:</td>
<td>Hough transform</td>
</tr>
<tr>
<td>MAC:</td>
<td>Maverick activities center</td>
</tr>
<tr>
<td>OCR:</td>
<td>Optical character recognition</td>
</tr>
<tr>
<td>ROI:</td>
<td>Region of interest</td>
</tr>
</tbody>
</table>
Overview

There are a lot of places in the university campus where one has to swipe the magnetic strip of the MAV-ID card for various reasons e.g. checking out a book from the library, take a print out, check out the sports equipment at Maverick activities centre (MAC) etc… Hence at every place there is a need of an external hardware to read out the information encoded on the magnetic strip. With this proposed approach the dependency on the external hardware to read the strip can easily be avoided and hence the purchase and installation cost can be reduced.

Proposal

Since most of the smart phones are packaged with an in-built camera, the proposal of this project is to develop an application that can take images of MAV-ID card and extract student ID and student name. With this application, there is no need for any external hardware interface.
Methods and Goals [4]

1. Figure 1 shows the image of the MAV-ID card taken from a camera which will then be used as the input image.

![Fig 1. Sample Image](image)

2. Importing image in MATLAB for processing.

3. Applying various image processing algorithms [7, 8, 10] to extract the MAV-ID and student’s name.
**Procedure:**

- **Binarization to separate background from foreground information**

![Original Image](a) ![Binarized Image](b)

Fig. 2(a) - Original image, (b) - Binarized image [1]

To find the license plate region, firstly smearing algorithm [6] is used. Smearing is a method for the extraction of text areas on a mixed image. With the smearing algorithm, the image is processed along vertical and horizontal runs (scan-lines) [1]. The method will detect the spaces in between each character on the MAV-ID card. Figure 2(a) and 2(b) shows an example of original image and its binary version.

- **Segmentation to identify different regions in the image**

![Segmented Image](a)

Fig. 3 - Segmented image [1]

Segmentation subdivides an image into its constituent regions or objects that have similar features (intensity, histogram, mean, variance, energy, texture) according to a set of predefined criteria. [2] Figure 3 shows an example of a segmented image of image shown in Figure 2(a). Segmentation is required to determine the regions of interest, which are then processed individually for text retrieval.
OCR or Morphological image processing to extract text and other features from the image

OCR (optical character recognition) [3] is the recognition of printed or written text characters by a computer. This involves photo scanning of the text character-by-character, analysis of the scanned-in image, and then translation of the character image into character codes, such as ASCII, commonly used in data processing.

In OCR processing, the scanned-in image or bitmap is analyzed for light and dark areas in order to identify each alphabetic letter or numeric digit. When a character is recognized, it is converted into an ASCII code. Special circuit boards and computer chips designed exclusively for OCR are used to speed up the recognition process [18].

OCR is being used by libraries to digitize and preserve their holdings. OCR is also used to process checks and credit card slips and sort the mail [16]. Billions of magazines and letters are sorted every day by OCR machines, considerably speeding up mail delivery [17].

Open source engine available OCR


Morphological Processing

After segmentation, there may be some noise in the image such as small holes in the candidate regions. This can be resolved with morphological processing [14]. Mathematical morphology operations are based on the shape in the image and not pixel intensities. The two basic morphological operations available in MATLAB are dilation and erosion [7]. Dilation allows objects to expand while erosion shrinks the objects by eroding the boundaries. These operations can be modified by proper choice of the structuring element which determines how many objects will be dilated or eroded [14]. Structuring element is simply a matrix of 0s and 1s that can be of any arbitrary shape and size. In MATLAB one can define neighborhood of desired size for the structuring element such as square, rectangle, diamond etc. Preferably rectangle is used as the neighborhood for the structuring element of size 6x4 which is obtained by a trial and error method. In the project, closing operation is used which is dilation followed by erosion. Removal of small holes plays an important role in obtaining the info from MAV-ID card.
Some of the applications of this project are:

- Library book/laptop checkout.
- MAC sports equipment checkout.
- Any student related process where manual entry of student information is required.
Fig 4- Different stages an image goes through before extraction of the required information
Fig. 5 - Original image

Figure 5 is the original image taken from a camera which is then processed using histogram equalization.
Contrast limited adaptive histogram equalization (CLAHE) [8] is an improved version of AHE, or adaptive histogram equalization. Both overcome the limitations of standard histogram equalization.

CLAHE was originally developed for medical imaging and has proven to be successful for enhancement of low-contrast images such as portal films [15].

The CLAHE algorithm partitions the images into contextual regions and applies the histogram equalization to each one. This evens out the distribution of used grey values and thus makes hidden features of the image more visible. The full grey spectrum is used to express the image.

Since the input image being used has a low contrast, using CLAHE instead of other histogram equalization techniques produces better results. Image shown in Figure 6 is contrast limited adaptive histogram equalized image which is then converted in to a gray scale image.
A grayscale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest.

Figure 7 shows the converted grayscale image which is then binarized as binarization is required for edge detection.
Image binarization converts an image of up to 256 gray levels to a black and white image. Frequently, binarization is used as a pre-processor before OCR. In fact, most OCR packages on the market work only on bi-level (black and white) images.

The simplest way to use image binarization is to choose a threshold value, and classify all pixels with values above this threshold as white, and all other pixels as black. The problem then is how to select the correct threshold. In many cases, finding one threshold compatible to the entire image is very difficult, and in many cases even impossible. Therefore, adaptive image binarization is needed where an optimal threshold is chosen for each image area. Figure 8 shows the image after binarization which is then given as an input to the edge detector algorithm in the next step.
Several edge detection algorithms (Sobel, Prewitt, Roberts, Canny, Laplacian of Gaussian and Morphological operators) [13] were tested. After experimenting with various parameters and methods, it was found that Sobel, Prewitt and Roberts [13] gave very similar results in comparable runtime. Canny edge detector typically resulted in more number of edges, but at the expense of longer computation time. Since the Canny edge detector [13] resulted in more number of edges which is required as an input to the bounding box algorithm to produce desired results, it was chosen to detect the edges in the binarized image. Figure 9 shows the output of the Canny edge detector algorithm.
Steps to be followed

The Hough transform (HT) [11] was originally suggested as a technique for detecting straight-line segments in an input image. It maps straight lines in the input image to points in the Hough space. Thus the HT of an input object will have points at locations corresponding to the parameters (e.g., normal distance from the origin and angle) of each line in the object. Several variations of the HT exist, depending on the two parameters (coordinates of the Hough space) used to describe a line.

The reason why Hough transform based line detection and geometric line processing has been chosen to be used is to detect the bounding box [12] of the card in a translational, rotational and scale invariant manner. In the next step, the image has to go through to find the ROI [region of interest] of the image.

The output from the ROI identification step shows the selected ROIs and the final image is used in OCR. The OCR engine is supposed to then give the desired results i.e. MAV-ID number and name from the MAV-ID card.
References:

3. https://code.google.com/p/tesseract-ocr/ [Definition of OCR and free OCR engine can be downloaded]
4. http://www.stanford.edu/class/ee368/Project_12/Proposals/Datta_Credit_card_processingusing_cell_phone.pdf [Proposal about how to extract and process information from the credit card, the information will be used in the project proposed]