Automatic Data Entry of Passport for Security System

Mya Mya Thinn and Myint Myint Sein

Mandalay Technological University & University of Computer Studies, Yangon, Myanmar

ngenge17@gmail.com,mlm@wwlmail.com

Abstract

A passport contains the important personal information of holder such as photo, name, date of birth and place, nationality, date of issue, date of expiry, authority and so on. A system for the automatic data entry of passport is developed in this paper. In proposed system, the capable of extracting the photo, character segmenting, handwritten recognition and printer character transformation have been done automatically from the scanned image of a passport. Especially, Myanmar passport is used in experiment. Recognition algorithm is extended from the Hopfield Neural Network for character recognition and eigenface method for face recognition. It will be provide to easy and convenient to immigrant staff for registering the personal documents of foreigner. This system is also preventing the distribution of imitation passport.

Keywords: *passport, face recognition, personal documents, character recognition, eigenfaces, Hopfield Neural Network*

1. Introduction

Biometric recognition techniques can be found for methods based on different physiological and behavioral modalities. The term biometrics refers to the science of measuring identifying features or attributes of human beings. We distinguish two approaches: passive and active schemes. Example of passive biometric is face recognition. Handwritten recognition is one of an example of the behavioral biometrics. For ID documents, e.g. passport, there are several requirements regarding security aspects. The two main security aspects for classical passports are the ability to detect tampering and to prove authenticity. To achieve the security criterion of non-transferability, most ID documents contain a photograph and an image of the handwritten signature of the legitimate holder and in some cases some information on the holder's appearance, like date of birth, eye color, nationality. Face recognition has the great advantage of not requiring any sort of contact, so there are no hygienic concerns. The biometric samples here are typically taken as 2D image of the frontal section of the face (e.g. photograph).

We describe a system to recognize characters such as name, date of birth, Nationality, passport number, date of issue, place of issue, etc. For storage of this information, database system or mobile memory units can be used. There are many different storage approaches such as photograph, OCR font. This paper is organized as follows. Section 2 briefly introduces the preprocessing, while in Section 3 region segmentation. Recognition processes are illustrated in Section 4. Section 5 provides experimental results of the process. Finally, conclusion and future work are discussed in section 6.

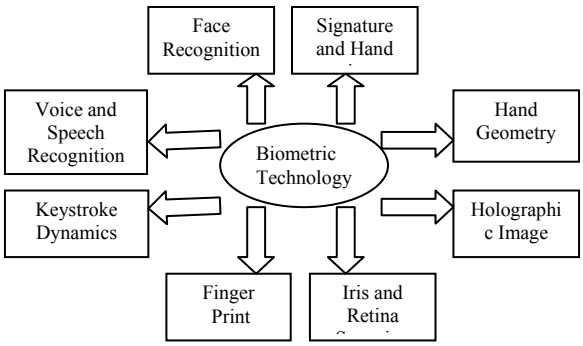

Figure1. Biometric system

2. Preprocessing

In the preprocessing step, passport is scanned by using Cannon N340P/n60P scanner with 300 DPI. An input image of passport is shown in figure 2. In the preprocessing task the system refits the input passport and extracts the content of the some printed characters. The size of the input image is resized. After resizing, the image may be appeared skewing angle or titled due to the user scanning error. So that the input image is tested for skewed detection algorithm which has been developed using equation.

$$P = x\cos\theta + y\sin\theta$$

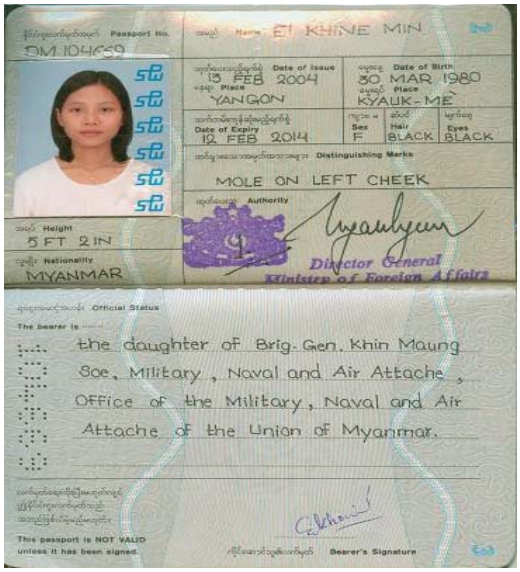

Figure2. An Input Image of Passport

The skew angle estimation and correction algorithm is:

-Read the input image

-Change the image to gray level

-Make the image sharper

-Find the radon Transform of the image

-Get the height of the image

-For I = 1 to height

-For J = 1 to 180

-Find max value

-Find Theta (skew angle) from max Value

3. Region Segmentation

The input image had to be extracted in order to segment photo and handwritten characters. After scanning the image from the scanner, preprocessing stages have been done for that image. Original passport color image is converted to the gray scale image. Then, binarylizing and thresholding are performed for deleting the background. After then, foreground region is extracted from binary image. Individual region of the photo and character are estimated by segmenting the extracted foreground region in image.

Figure3.Extracted regions from a passport image

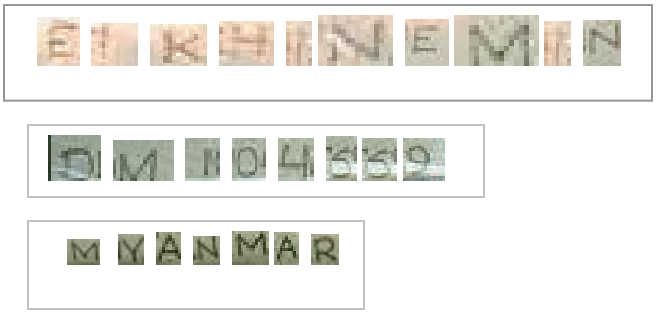

Figure4. Example of character segmentation

Segmentation methods can be divided into three groups, these are global knowledge about and edge-based segmentation and region based segmentation. Region based segmentation is applied to separate the

region of the image and character. Region based segmentation method is easy to construct regions from their borders, end it is easy to detect borders of existing regions. Then characters are segmented to transform the printed characters.

Segmentation Algorithm

```
For i=1 to width
        For j= 1 to height
        Check the flag of the pixel
        F = chk ( img [ i, j]  );
        If f = 1 (# it shows it is the start # )
             Then new [ 0, 0 ] = img [ i, j ]
        If f = 2( # it shows it is continues the character#
)
              Then new [ i, j] = img [ i, j]
        If f = 0 ( # no data at that pixels #)
        Loop j
         Loop i
         A segmented character is in the new.
```

Segmented regions from a passport and character segmentation are expressed in figure 3 and figure 4, respectively.

4. Recognition Process

Face recognition system and character recognition system are essential processes for automatic data entry by matching with database information.

4.1. Eigenfaces for recognition

The following steps summarize the recognition process using eigenfaces.

-Initialization – Acquire the training set of face images and calculate the eigenfaces, which define the face space.
-When a new face image is encountered, calculate a set of weights based on the input image and the M eigenfaces by projecting the input image onto each of the eigenfaces.
-Determine if the image is a face at all (whether known or unknown) by checking to see if the image is sufficiently close to "face space".
-If it is a face, classify the weight pattern as either a known person or as unknown.
-If the same unknown face is seen several times, calculate its characteristic weight pattern and incorporate into the known faces (i.e. learn to recognize it.)

4.2. Recognition of Printed character by using Hopfield Neural Net

A Neural Net comprises a set of weighted edges and nodes. It may be characterized as pattern association and pattern recognition, clustering and constrained optimization. Pattern recognition includes such applications as speech and character recognition for various languages, visual image recognition and classification. There are several neural networks available for printed character recognition problem. Hopfield method is one of the efficient methods of the neural networks. Hopfield neural net is a single layer feedback network operating in discrete time interval. The Hopfield neural network is a simple artificial network which is able to store certain memories or patterns in a manner rather similar to the brain.

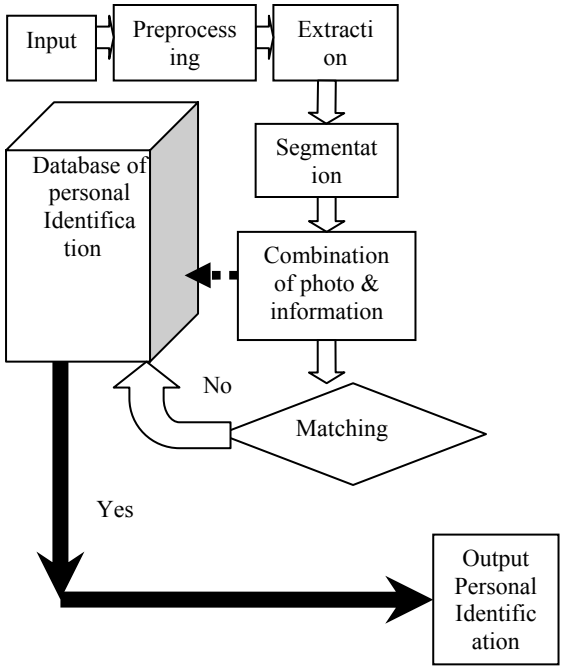

Figure5. General process of our purpose system

5. Experimental results

We first created a large database of photo images and personal information with respect to that image. Using this database, we have conducted several experiments to access the performance under known variations of lighting, scale and orientation. There are fifty persons' information are used in this experiment.

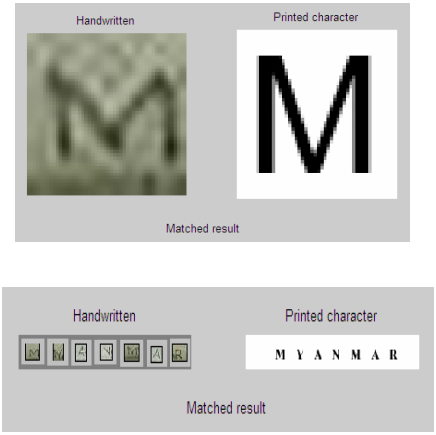

Figure6. Transforming of Handwritten character

Figure7. Personal Identification of the passport holder

In our system, MATLAB programming language version 8 is used. Transforming of the handwriting character to the printed character is as shown in Figure6. In our printed character recognition system, both well-separated printed characters, poor printed characters can be recognized. Information in the passport are recognized and matched with the database information. The output result for personal identification of the passport holder is illustrated in figure 7.

6. Conclusion and Future Work

In this paper, a system for automatic data entry of passport has been developed. The personal information of passport holder described in printed characters. Personal identification has been performed based on the eigen face for face recognition and Hopfield Neural Network for character recognition, respectively. If the information will change, the new information of the passport holder will be added by using automatic data entry system. For connecting to the network of database, we will extend to search the personal identification from every where.

References

[1] P.Y. Simard, D. Steinkraus, and J. Platt. Best Practice for Convolutional Neural Networks Applied to Visual Document Analysis. *IEEE International Conference on Document Analysis and Recogntion*, pp.958–962, 2003.

[2] BSI (Bundesamt für Sicherheit in der Informationstechnik), *BioFace – Comparative Study of Facial Recognition Systems*, *Bio-Face – Comparative Study of Facial Recognition Systems*, 2003.

[3] M.Turk and A.Pentland., "Eignefaces for Recognition", journal of cognitive Neuro science, March 1991.

[4] T. Kanadae, "Picture processing system by computer complex and recognition of human faces, " Dept. of Information Science, Kyoto University, Now, 1973.

[5] Dr. AZidouri, K.Fahd, "Automatic Text Recognition: A need in Arabization "Electrical Engineering Department, University of Petroleum &Minerals, Kingdom of Saudi Arabia.

[6] Bundesdruckerei GmbH, *All security feature of the passportdata page,* http://www.bundesdruckerei.de/en/iddok/2_1/2_1_7.html, 2004.