Proceedings of APIIT Business, Law & Technology Conference, 2017
July 20th, 2017 | Colombo, Sri Lanka
ISBN978-955-7678-02-3

English Character Recognition of an Image and Voicing System

G. M. N. Thilakarathna[1]
Department of Physics, Faculty of Applied Sciences
University of Sri Jayewardenepura
Nugegoda, Sri Lanka

W. K. I. L. Wanniarachchi
Department of Physics, Faculty of Applied Sciences
University of Sri Jayewardenepura
Nugegoda, Sri Lanka

***Abstract* - In this research, an attempt has been made to develop a recognition strategy for printed English character (A to Z) to help visually impaired people to overcome their disability in reading. Images in specific formats such as JPEG, BMP, JPG etc. were took as input images. Then the noise which comes to the image in acquisition was filtered. After that the coloured image was converted to a binary image using a suitable threshold. Then segmentation technique was used to identify connected areas while bounding boxes were used to identify individual letters. Finally recognition was done using a Statistical algorithm which uses a pre stored example database. Database is a collection of English alphabet letters saved as images. Apart from the recognition a complete system of image acquisition and a text to voice conversion strategy was developed to make the system more useful which helps visually impaired person to understand a the context of a writing.**

Keywords—Image Processing; Charcater Recognition; Statistical Correlation; Visually Impaired People; Text to Voice

I. INTRODUCTION

The recognition of words from scanned images or from photo images of documents has been a problem that has received much attention in the fields of image processing, pattern recognition and artificial intelligence [1,2]. Image processing can be considered as an improvement of pictorial data to information which human can predict easily, and a process of data storing, transmitting and representing. This research can be used in wide range of applications such as license plate recognition of a moving vehicle, security applications, postal address recognition, and document reading. Also, this is very useful to a visually impaired person that he /she can read a book by his/her own and can be used to read banners and names of the shop outlets while walking. Moreover an improved version can be used in reading credit card imprints for billing purposes.

This system comes under optical character recognition (OCR).OCR is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image (for example from a television broadcast) [3]. Hence, various methods can be implemented to recognize characters.

The Recognition of characters is known to be one of the applications of Artificial Neural Networks (ANN) which partially emulate human thinking in the domain of artificial intelligence [4]. Character recognition is the process to classify the input character according to the predefined character class. With the increasing interest of computer applications, modern society needs that the computer should read the text. The text may be in the form of scanned typed document or an image of typed text in various fonts or a combination of both. The character recognition system helps in making the communication between a human and a computer easy.

Recognition of visual characters may not perfect due to various reasons such as the same characters differ in sizes, shapes and styles from font to font. The source of confusion is the high level of abstraction. There are thousands styles of type in common use and a character recognition program must recognize most of these [5].

Like any image, visual characters are subject to spoilage due to noise. Noise consists of random changes to a pattern, particularly near the edges. A character with much noise may be interpreted as a completely different character by a computer program [1,6].

II. METHODOLOGY

The developed system of character recognition and voicing can be described by the following steps.

- Image Acquisition
- Preprocessing
- Segmentation
- Character Database

[1]nirashathilakarathne1217@gmail.com

ISBN978-955-7678-02-3

- Character Recognition
- Text-to-Voice Conversion

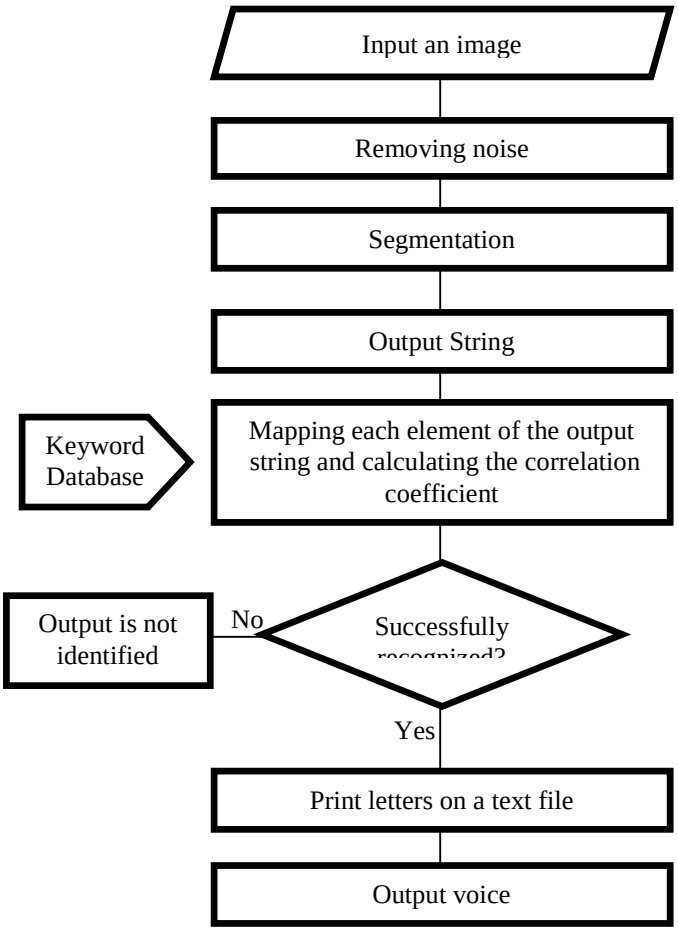


Fig.1 Process flow chart

A. Image Acquisition

Fig. 1 shows the flow charts of the developed system. Initial step was to select an input method. For this an image was taken in different ways as

- Images consist with letters and strings created using computer software.
- Scanned images.

Most of the results were tested using images created by Microsoft Power Point and the rest with the scanned images.

B. Pre-Processing

Pre-processing is one of the most important part when considering character recognition systems because the output of this stage (segmented image) was used in other stages, so better the pre processing easier to extract features (better output will be produced). The input original true RGB color image converted to a binary image (black and white image) by considering a global threshold. Noise reduction was carried out by removing of objects less than 30 pixels. Fig.2a shows the original true color image (RGB image) which was created using computer software. Fig.2b shows the complement image of the noise removed binary image.

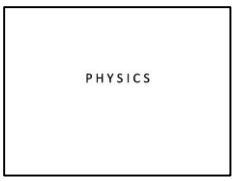


Fig.2a Original RGB image

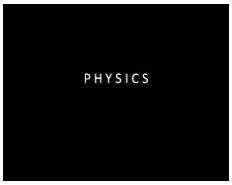


Fig.2b Negative binary image

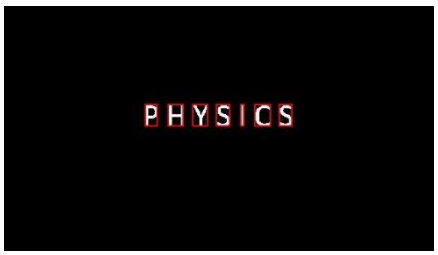


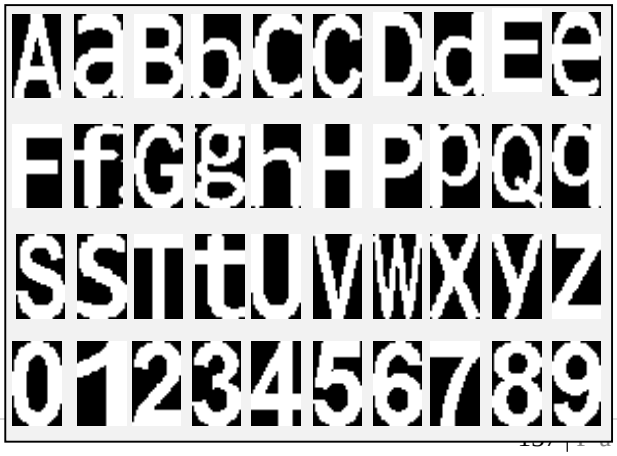
Fig.3 Character Segmantation

C. Segmentation

In this step characters were grouped from the binary image. Blob detection was used to separate edges of the letters from the background and rectangular boxes in red colour were created around the connected areas. Through these bounding boxes the connected areas were separated (Fig.3) and resized into binary image of 42x24 (no. of rows x no. of columns).

D. Database Creation

To identify the characters a database was created using English alphabet letters (both simple and capital letters) of Times New Roman and Calibri font types. The database consists with total of 100 images. For that an algorithm was developed to extract the icons of images as the same procedure explained above. After creating bounding boxes the letters were cropped along the bounding box and they were automatically saved in a folder numbering from 1: n (n=any number). These images were then subjected to a resizing of 42x24 pixel size.



ISBN978-955-7678-02-3

$\overline{A}$ and $\overline{B}$ – average of matrix elements

Fig.4 Sample images of 42x24 pixel size characters

E. Character Recognition

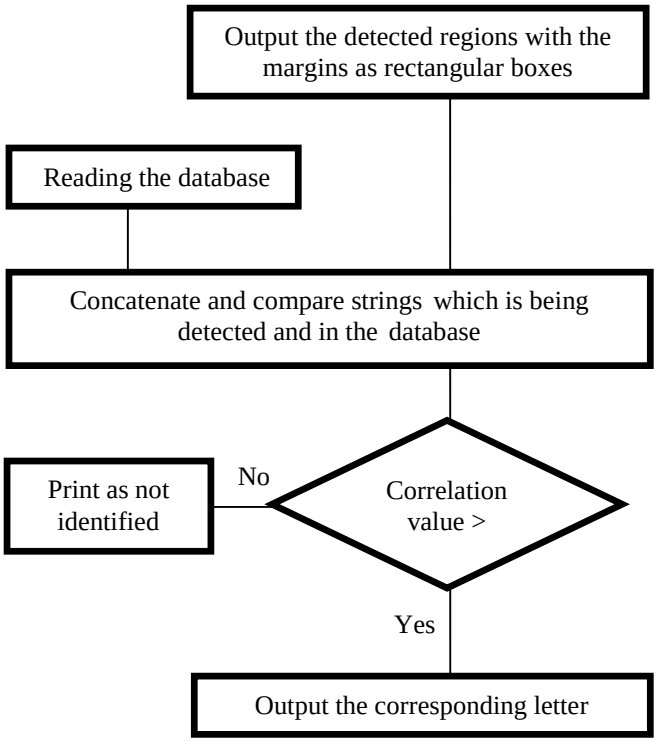


Fig.5 Flow chart of character recognition

The detected areas using bounding boxes were subjected to a comparison with the images of the created character database. Images were compared and a relationship was found using correlation coefficient. Each image was compared with the template images and then a correlation coefficient value is given as the output. Each value is stored in an array and the maximum correlation coefficient points were detected. Then the system automatically detects the corresponding letter and the letter is printed on the created .txt file.

The equation that is used to calculate the correlation coefficient is as below[ref]. Here 'A' is the template image and 'B' is the image to be recognized. The images are in m x n (42x24) format.

Correlation coefficient

$$\frac{\sum_m \sum_n (A_{mn} - \overline{A})(B_{mn} - \overline{B})}{\sqrt{\left(\sum_m \sum_n (A_{mn} - \overline{A})^2\right)\left(\sum_m \sum_n (B_{mn} - \overline{B})^2\right)}}$$

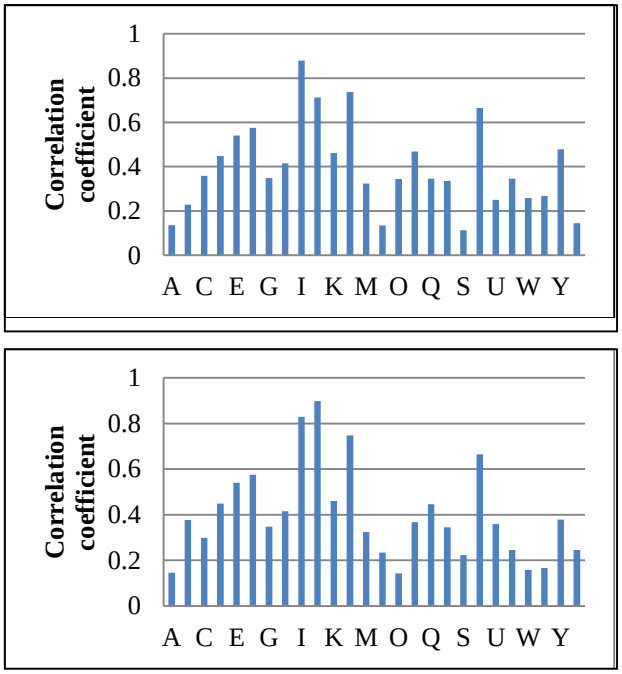


Fig.6 Correlation coefficients for comparison of letters 'I' (top) and 'J' (bottom)

F. Text to Speech Conversion

Text-to-speech (TTS) is the generation of synthesized speech from text [7]. The synthesized speech should be intelligible, natural and pleasant to listen, as human speech. During synthesis very small segments of recorded human speech are concatenated together to produce the synthesized speech. In this work a PC based text to speech synthesizer has accessed using MATLAB algorithms. Here, the accessed text to speech synthesizer consists a database required for character to voice conversion with recorded alphabets (A-Z), digits (0-9) and words in the form of wave files (.wav).

The first step in converting text to speech is to create a text file(.txt) to print the identified characters . Once the file is created, it is opened and read in MATLAB. For every element read, corresponding wave file is played so as to output the sound of that character. We can read as big file as possible in character wise or in word wise. [7,8]

Text to Speech Algorithm

STEP1: Create a text file (.txt)

STEP2: Open and Read the .txt file in MATLAB.

STEP3: Play the corresponding wave (.wav) file for corresponding letter.

ISBN978-955-7678-02-3

III. RESULTS AND DISCUSSION

As mentioned in methodology the RGB images were converted to binary images for character recognition. Suitable threshold for the conversion may depend on image intensity distribution and it is differ from one to another. Therefore, global threshold value was considered for a given image where the program set the suitable threshold itself. Object consists with pixels lesser than 30 were removed from the binary image to reduce noises. In scanned images the edges were filled with unnecessary pixel values due to defects of the machine and due to malfunctioning of image acquisition methods, so that not only the characters every connected element is taken as separate objects and created bounding boxes. To overcome the issue the images were cropped and resized. Filtering methods like median filters were not used as from our investigations, where we found that to a black and white image the filtering methods add noise rather than minimizing.

Fig.7 Segmentation of capital letters

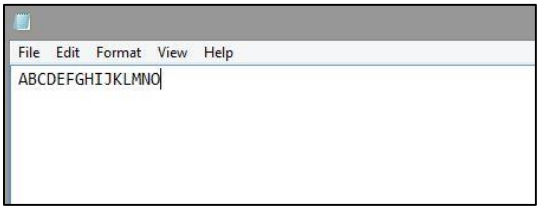


Fig.8 Text file of recognized capital letters

The developed computer code as explained above was successfully identified the capital letters of unknown image consist with Time new Roman and Calibri font type. Character recognition of images consists with other font types cannot identify accurately by the developed method. Therefore, it is need to expand the character database for different font style for better character recognition. When recognizing of 'i' and 'j' make some complications. The developed program identifies i (and j) as two characters as there is a gap with in the letter itself. Other simple letters were recognized correctly including the numbers. Fig.9 shows that the there are two bounding boxes for each letter 'i' and 'j'. For this particular case letter 'i' was identified as letter 't' and letter 'l'. Fig.10 shows the text file of the identified characters for simple letters. For the case letter 'j', it was identified as letter 'j' and letter 'i'.

Fig.9 Character segmentation for simple letters

Fig.10 Text file of recognized simple letters

Other than that the developed program not capable separately identifying the simple and capital letters for the cases of 'O', 'S', 'V', 'W', 'X' and 'Z' where both simple and capital letters have similar features.

The system has the ability to convert all the meaningful words to speech without voicing letter by letter. So this has become not only a character recognition system but also word recognition and voicing system.

The speed of the recognition mostly depends on the size of the database and with the length of the word. Machine specifications also influence for process speed of the system. Following table show results of process speed in different sizes of databases and different length of words.

TABLE I. TIME TAKEN TO IDENTIFY LETTERS

Word Used	**Time taken to print letters (s)**	
	100 letter database	***50 letter database***
A	1.423654	1.494532
AB	1.514552	1.383118
FOX	1.447655	1.414765
DEAR	1.410674	1.460191
FLOWS	1.448465	1.397348
FLOWER	1.458298	1.418548
NIRASHA INDUNIL	1.703936	1.650331
UNIVERSITY OF SRI JAYEWARDENEPURA	1.813692	1.828469

IV. FURTHER IMPROVEMENTS

The developed system need to improve for recognition of simple letters such as 'i' and 'j'. The system can be implemented to detect Sinhala characters, by making a database using Sinhala character letters for different font types. Also this can be further improved to read number plates of vehicles and cutouts of the shops. There is a possibility to use the proposed work as a language translator in Sinhala or English.

ACKNOWLEDGEMENT

Authors wish to acknowledge university research grant ASP/01/RE/SCI/2017/13 supported for this project.

REFERENCES

1. R.C. Gonzalez, R.E. Woods, S.L. Eddins, Digital image processing using Matlab, Prentice Hall, ISBN 81-297-0515-X(2004).
2. Rosenfeld A; King-Sun F, Pattern Recognition and Image Processing.,IEEE transactions on computers,Vol c-25,No.12,(1976).

ISBN978-955-7678-02-3

3. Mithe R; Indalkar S; Divekar N, Optical character recognition International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-2, Issue-1, (2013)

4. Pohankar P; Karmare S; Taralkar N; Kulkarni S, Character Recognition using Artificial Neural Network, International Journal of Electronics Communication and Computer Engineering Volume 5, Issue (4) July, Technovision-2014, ISSN 2249–071X

5. Ambler, B. A; Proctor R; The familiarity effect for single-letter pairs.Journal of Experimental Psychology: Human Perception & Performance,Vol 2, 222–234, **(1976).**

6. Kauleshwar P; Devvrat C. N; Lakhotiya A; Dheeren U, Character Recognition Using Matlab's Neural Network Toolbox, International Journal of u- and e- Service, Science and Technology Vol. 6, No. 1, February, 2013

7. Swetha N; Anuradha K, Text-to-Speech Conversion, International Journal of Advanced Trends in Computer Science and Engineering, Vol.2 , No.6, Pages : 269-278 (2013)

8. Rajesh R., and Seththilkumaran N, A Note on Image Segmentation Techniques. International J. of Recent Trends in Engineering and Technology, Vol. 3, No. 2, (2010)