International Journal of Engineering Research and Science & Technology

ISSN : 2319-5991
Vol. 4, No. 2
May 2015

www.ijerst.com

Email: editorijerst@gmail.com or editor@ijerst.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 4, No. 2, May 2015

Research Paper

ASSISTIVE TEXT READING SYSTEM FOR VISION IMPAIRED PERSON

R Priya[1]* and A Ashlin Jeba[2]

Corresponding Author: **R Priya ✉ Riyasri.10@gmail.com*

In this paper an assistive system has been proposed for visually impaired persons. It reads textual information on papers by using hand held camera and produces corresponding voice using OCR (Optical Character Recognition) and TTS (Text-to-Speech) system. To localize text regions in images connected component labeling approach using histogram analysis is done on binarized image. TTS system using Concatenative synthesis based on Excel application platform is used. This system is operated via a voice-based user interface and also has a user friendly GUI (Graphical User Interface) to scan the text and to control various speech parameters. The distance between the product and the camera is measured and maintained by using ultrasonic sensor.

Keywords: Neural Network, USB Camera, Image Processing, MATLAB, Speed detection, Speech synthesizer, Text to Speech converter, Ultrasonic Sensor, Optical Character Recognition

INTRODUCTION

The purpose of this project is to take handwritten English characters as input, process the character, train the neural network algorithm, to recognize the pattern and modify the character to a beautified version of the input. This project is aimed at developing software which will be helpful in recognizing characters of English language. This project is restricted to English characters only. It can be further developed to recognize the characters of different languages. Neural networks are particularly useful for solving problems that cannot be expressed as a series of steps, such as recognizing patterns, classifying them into groups, series prediction and data mining. Pattern recognition is perhaps the most common use of neural networks. The neural network is presented with a target vector and also a vector which contains the pattern information, this could be an image and hand written data. The neural network then attempts to determine if the input data matches a pattern that the neural network has memorized. A neural network trained for classification is designed to take input samples and classify them into groups. These groups may be fuzzy, without clearly

[1] ARJ College of Engineering and Technology, Mannargudi, Thanjavur, India.
[2] Assistant Professor, Department of ECE, A.R.J College of Engineering and Technology, Mannargudi, Thanjavur, India.

defined boundaries. This project concerns detecting free handwritten characters.

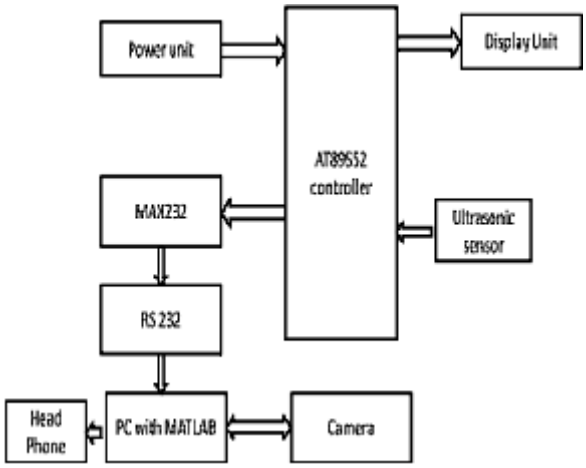
RELATED WORK

A feature extraction technique based on character geometry for character recognition describes a geometry based technique forfeiture extraction applicable to segmentation-based word recognition systems. The proposed system extracts the geometric features of the character contour. These features are based on the basic line types that form the character skeleton. The system gives a feature vector as its output. The feature vectors so generated from a training set were then used to train a pattern recognition engine based on Neural Networks so that the system can be benchmarked. As Sheng Wang described in A Review of Gradient-Based and Edge-Based Feature Extraction Methods for Object Detection. In computer vision research, object detection based on image processing is the task of identifying a designated object on a static image or a sequence of video frames. Projects based on such research works have been widely adapted to various industrial and social applications. The field to which those applications apply includes but not limited to, security surveillance, intelligent transportation system, automated manufacturing, and quality control and supply chain management. In general, we categorize those methods into to gradient-based and edge based feature extraction methods, depending on the low level features they use. In this paper, the definitions for gradient and edge are extended. Because an image can also be considered as a grid of image patches, it is therefore reasonable to incorporate the concept of granules to gradient for a review.

EXPERIMENTAL SETUP

A. USB Camera

Figure 1: System Working Setup



USB Cameras are imaging cameras that use USB 2.0 or USB 3.0 technology to transfer image data. USB Cameras are designed to easily interface with dedicated computer systems by using the same USB technology that is found on most computers. The accessibility of USB technology in computer systems as well as the 480 Mb/s transfer rate of USB 2.0 makes USB Cameras ideal for many imaging applications. An increasing selection of USB 3.0 Cameras is also available with data transfer rates of up to 5 Gb/s. Edmund Optics offers a variety of USB Cameras suited to meet many imaging needs. EO USB Cameras are available in both CMOS as well as CCD sensor types making them suitable across a larger range of applications. USB Cameras contain out-of-the-box functionality for quick setup. USB Cameras using low power USB ports, such as on a laptop, may require a separate power supply for operation.

B. Ultrasonic Transducer

Ultrasonic transducer uses the physical

characteristics and various other effects of ultrasound of a specific frequency. It may transmit or receive the ultrasonic signal of a particular strength. These are available in piezoelectric or electromagnetic versions. The piezoelectric type is generally preferred due to its lower cost and simplicity to use. The Ultrasonic wave propagation velocity in the air is approximately 340 m/s at 15°C of air or atmospheric temperature, the same as sonic velocity. The system consists of a transmitter and a receiver module controlled by a microcontroller At 89S52. A microcontroller development kit has been used for testing of the system (Figure 1). 40 KHz ultrasound sensors have been used for the experiments. Figures 2 and 3 show the photographs of the transmitter module and the receiver shows the photograph of ultrasonic transmitter and receiver sensors. An interrupt INT1 signal initiates the system. When the interrupt INT1 signal is received, MCU (microcontroller unit) starts the timer1 and simultaneously generates the controlled 40 KHz burst pulse having a train of specific number of pulses. These pulses are applied to the amplifier circuit and after amplification; the ultrasound transmitter transmits the 40 KHz ultrasound pulses in the air in the direction of the object. These ultrasonic pulses are reflected from the object and travels back in different directions. When these waves arrive at receiver, the signals received by the receiver is amplified and processed by the receiver module. The receiver module also generates an interrupt signal; INT2 at the instant the first pulse of the burst is received. Interrupt INT2 stops the timer1, and MCU calculates the time period between the generation of the wave and reception of the wave, which is proportional to the distance travelled by the waves. Using the formula, MCU calculates the distance of the Product or object and displays it or transfers it to the part of the total system where it is to be used for further control.

C. Optical Character Recognition

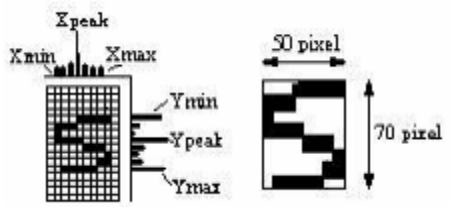
Read Image

The image is first being converted to grayscale image follow by the threshing technique, which make the image become binary image. The binary image is then sent through connectivity test in order to check for the maximum connected component, which is, the box of the form. After locating the box, the individual characters are then cropped into different sub images that are the raw data for the following feature extraction routine. The size of the sub-images are not fixed since they are expose to noises which will affect the cropping process to be vary from one to another. This will causing the input of the network become not standard and hence, prohibit the data from feeding through the network. To solve this problem, the sub-images have been resize to 50 by 70 and then by finding the average value in each 10 by 10 blocks, the image can be down to 5 by 7 matrices, with fuzzy value, and become 35 inputs for the network. However, before resize the sub-images, another process must be gone through to eliminate the white space in the boxes.

Feature Extraction

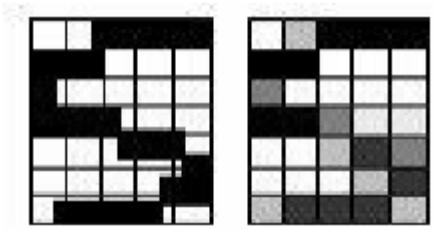
The sub-images have to be cropped sharp to the border of the character in order to standardize the sub-images. The image standardization is done by finding the maximum row and column with 1 s and with the peak point, increase and decrease the counter until meeting the white space, or the line with all 0 s. This technique is shown in figure below where a character “S” is being cropped and resized.

Figure 2: Cropped and Resized Picture



The image pre-processing is then followed by the image resize again to meet the network input requirement, 5 by 7 matrices, where the value of 1 will be assign to all pixel where all 10 by 10 box are filled with 1s, as shown below:

Figure 2.1: Image Resize Again to Meet the Network Input Requirement



Creating Vectors Data for the Neural Network (Objects)

These few line of codes creates training vector and testing vector for the neural network. This is to match the input accepted by the neural network function. The front 4 rows will be used to train the network, while the last row will be used to evaluate the performance of the network.

```
P = out(:,1:40);

T = [eye(10) eye(10) eye(10) eye(10)];

Ptest = out(:,41:50);
```

Creating And Training of the Neural Network

Create and Train NN. The "edu_createnn" is a function file to create and train the network by accepting the training-target datasets. Because of the nature of the problem, we used feed forward back propagation neural network for classification or with other words we were experimenting with Multi Layer Perceptron (MLP). Below is a picture of MLP structure with two hidden layers [35 10] neurons and sigmoid, linear activation function. Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors, or classify input vectors in an appropriate way as defined by you. Networks with biases, a sigmoid layer, and a linear output layer are capable of approximating any function with a finite number of discontinuities.

Figure 2.2: Typical Neural Network

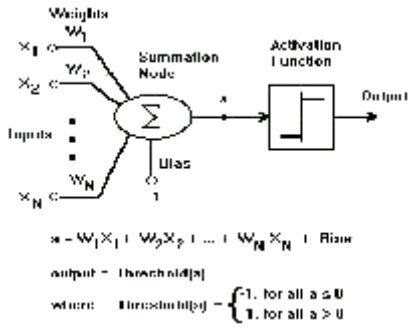
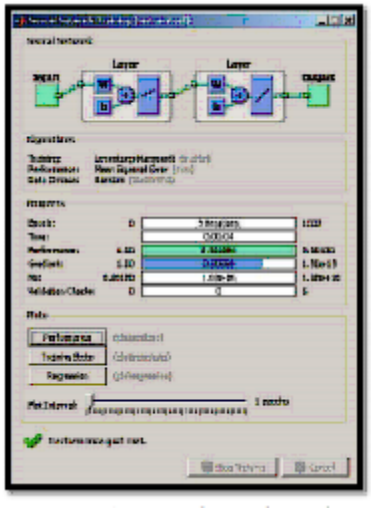


Figure 2.3: Training Neural Network



```
%Training with the help of training function
net = edu_createnn(P,T);
```

We can also experiment and test the network separately

```
net = newff(P,T,[35], {'logsig'}) %net.performFcn = 'sse';
net.divideParam.trainRatio = 1; % training set [%]
net.divideParam.valRatio = 0; % validation set [%]
net.divideParam.testRatio = 0; % test set [%]
net.trainParam.goal = 0.001;
[net,tr,Y,E] = train(net,P,T);
```

Tempelate Matching

In our recognition we use/train the net with ten different signs (characters). Five were numbers and five were characters, so input is a matrix of size 1 x 10 and [1,2,3,4,5,'A','B','C','D','E'] values. Because we cannot define network target as a character we have to use numbers – integers. So on output of the neural network we get matrix with values [1,2,3,4,5,6,7,8,9,0] instead of above, first input matrix. From that reason we have to encode/decode the output of our application. The encoding is made with simple code table, in which every index gets a new value. The text in output text file from OCR is matched with the saved product names in the Database the matched product is identified.

TEXT-TO-SPEECH SYNTHESIZER

A Text-To-Speech (TTS) synthesizer is a computer-based system that should be able to read any text aloud, when it is directly introduced in the computer by an operator. It is more suitable to define Text-To-Speech or speech synthesis as an automatic production of speech, by 'grapheme to phoneme' transcription. A grapheme is the smallest distinguishing unit in a written language. It does not carry meaning by itself. Graphemes include alphabetic letters, numerical digits, punctuation marks, and the individual symbols of any of the world's writing systems. A phoneme is "the smallest segmental unit of sound employed to form meaningful utterances".

Synthesizer Technology

The most important qualities of a speech synthesis system are naturalness and intelligibility. Naturalness describes how closely the output sounds like human speech, while intelligibility is the ease with which the output is understood. The ideal speech synthesizer is both natural and intelligible. Speech synthesis systems usually try to maximize both characteristics. The primary technology for generating synthetic speech is concatenative synthesis.

Concatenative Synthesis

Concatenative synthesis is based on the concatenation (or stringing together) of segments of recorded speech. Connecting prerecorded natural utterances is probably the easiest way to produce intelligible and natural sounding synthetic speech. However, concatenative synthesizers are usually limited to one speaker and one voice and usually require more memory capacity.

One of the most important aspects in concatenative synthesis is to find correct unit length. The selection is usually a trade-off between longer and shorter units. With longer units' high naturalness, less concatenation points are achieved, but the amount of required units and memory is increased. With shorter units, less memory is needed, but the sample collecting and labeling procedures become more difficult and complex. In present systems units used are usually words, syllables, phonemes.

CONCLUSION

The effectiveness of the method that uses feature extraction using character geometry and gradient technique from scanned images containing handwritten characters is presented. The feature extraction methods have performed well in classification when fed to the neural network and pre-processing of image using edge detection and normalization are the ideal choice for degraded noisy images. The method of training neural network with extracted features from sample images of each character has detection accuracy to a greater extent. The proposed methodology has produced good results for images containing handwritten text written in different styles, different size and alignment with varying background The system is developed in MATLAB and evaluated for a set of sample images containing handwritten text on Intel dual core computer. The method is advantageous as it uses nine features to train the neural network using character geometry and twelve features using gradient technique.

FUTURE ENHANCEMENT

As the feature extraction methods such as gradient technique and character geometry used in the method does not classify characters of different language, the method can be extended for language independent classification from the images of other languages with little modifications. The performance of the method has been tested for classifying English text written in upper case, but needs further exploration. Refinement of the segmented characters can be done in order to achieve higher accuracy rate. The performance of the neural network can be increased by adding some more features other than the existing ones. The classification rate can be increased by training the neural network with more number of test images.

REFERENCES

1. Alexander J Faaborg (2002), "Using Neural Networks to Create an Adaptive Character Recognition System", Cornell University, Ithaca NY, (May 14, 2002).
2. Ashutosh Aggarwal, Rajneesh Rani and Renu Dhir (2012), "Handwritten Character Recognition Using Gradient Features", *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 2, Issue 5, May 2012.
3. Chirag I Patel, Ripal Patel and Palak Patel (2011), "Handwritten Character Recognition Using Neural Networks", *International Journal of Scientific & Engineering Research,* Vol. 2, Issue 5, May-2011.
4. Dinesh Dileep, "A Feature Extraction Technique Based on Character Geometry for Character Recognition".
5. Kauleshwar Prasad, Devvrat C Nigam, Ashmika Lakhotiya and DheerenUmre (2013), "Character Recognition Using Matlab's Neural Toolbox", *International Journal of u- and e- Service, Science and Technology,* Vol. 6, No. 1, February, 2013.
6. Miller G A (1955), "The Magical Number Seven, Plus or Minus Two: Some limits On our Capacity for processing Information", *The Psychological Review,* Vol. 63, pp. 81-97, 1955.
7. Rahul Kala, Harsh Vazirani, Anupam Shukla and RituTiwari (2010), "Offline Handwriting Recognition", *International Journal of Computer Science issues,* Vol. 7, March-2010.
8. Sheng Wang (2011), "A Review of Gradient-

Based and Edge-Based Feature Extraction Methods for Object Detection", Computer and Information Technology (CIT), 2011 IEEE 11th International Conference.

9. Swapnil A Vaidya and Balaji R Bombade (2013), "A Novel Approach of Handwritten Character Recognition using Positional Feature Extraction", *IJCSMC*, Vol. 2, Issue. 6, June 2013.
10. Vinita Dutt and Sunil Dutt (2011), "Handwritten Character Recognition Using Artificial Neural Network", *Advances in Computing*, Vol. 1, No. 1, pp. 18-23.

International Journal of Engineering Research and Science & Technology
Hyderabad, INDIA. Ph: +91-09441351700, 09059645577
E-mail: editorijlerst@gmail.com or editor@ijerst.com
Website: www.ijerst.com