ISSN 2321 3361 © 2016 IJESC

Research Article **Volume 6 Issue No. 9**

Electronic Braille Reader and Tutor

Arshita Gupta[1], Hetu Shah[2], PrashitaPrathapan[3], AniketSaha[4], TusharChonde[5]

Abstract:
In the face of a digital world, communication on paper is being rapidly replaced by messengers, SMS services, e-book readers and the like. This primarily affects the lives of the 39 million blind people living worldwide, who have to either rely on audio books or avail the services of a visually able person, causing a grave disadvantage to their habit of reading. There is also a rising problem of increasing number of people turning accidentally blind, and hence the conventional methods of teaching Braille to such a class of visually challenged people are very less effective. The average test subject would never get acquainted to pages of Braille language texts, creating a need and hence a potential market for a singular device that performs the same action as pages of Braille characters. Thispaper, thus, presents a device: The ERT-Braille (Electronic Reader and Tutor- Braille) made to aid these differently abled persons. The two major objectives of ERT- Braille for the blind, are to permit access to digital matter (MODE 1) and also a self-tutor for learning the braille language for beginners(MODE 2) This device, consisting of a braille cell, is used to display characters from a Portable Document Format (PDF) file, as raised symbols on the cell. The self-tutor ability of the device helps to learn braille language through speech recognition using MFCC technique.

Keywords: braille cell, Speech recognition, MFCC

I. INTRODUCTION

Visually challenged individuals rely on their sense of touch for pattern perception, much as the rest of us depend on vision. It is traditionally written with embossed paper, perceived by the blind. This pattern perception can be electrically controlled by any microcontroller or embedded system, and the output pattern is recorded by the linear motion of a three by two braille character cell. The ERT Braille involves a mechatronics system where a micro-controller is programmed in such a way that it converts each character of the English alphabet, from the digital text available, to its corresponding braille character, further given to a tactile terminal. There are certain cases where people are not visually challenged since birth. To help such people understand the Braille language, this device also provides a learning method for the braille language, where using the basic principles of speaker recognition we differentiate the characters of the English language. The micro-controller simultaneously generates the respective braille character of this English alphabet, which is perceived by the blind to understand the corresponding braille character spoken by him. Thus the two modes of operation of the ERT Braille are as follows. Mode One: In this mode the device acts as a Reader Mode Two: In this mode the device acts as a Tutor

II. BRAILLE OVERVIEW

Invented in 1821 by Louis Braille, who had been blind from age three, Braille is a tactile writing and reading system for blind and partially sighted people. Braille system employs group of dots embossed on a paper or some other flat surface to represent printed letters and numbers. The system's basic "braille cell" is illustrated in Figure 1. Each character is made of up to six dots, positioned in two columns of three. For convenience, a standard numbering system has been established for the dots whereby the dots in the left column are numbered downward from one to three, and the dots in the right column are numbered downward from four to six. It is read by passing fingers over each character.

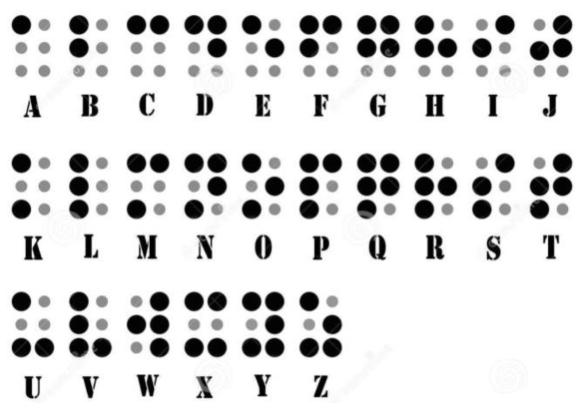

Figure: 1.MODE ONE: ELECTRONIC BRAILLE READER
The block diagram of ERT in mode 1 is shown in figure 2.

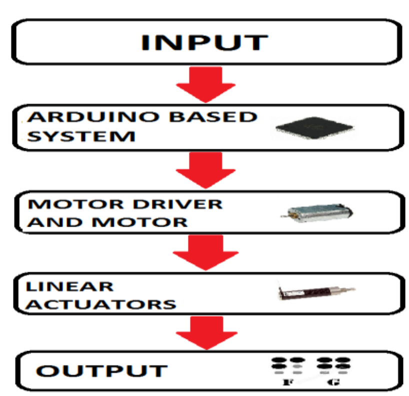

Figure 2.INPUT AND OUTPUT

The first objective of the ERT Braille is to convert any text to an output that can be perceived by a blind person through touch. The user begins with the RESET button of the micro-controller, and the first character is taken from the selected document file.

Implementation:

The character is passed to the switch table in the program code and the respective pins are set to logic HIGH as per the braille equivalent of the English character. This HIGH signal is received by the toy motors and they turn ON hence raising the mechanical pins using the respective linear actuators. The program goes in a hold state for five seconds after which the next letter is imported from the text file. This way the entire text file is read out character by character into the braille cell of the display. The reset button on the device is takes the control of the program back to the initial character of the document, and all pins are retracted when there is no source document to read out from.

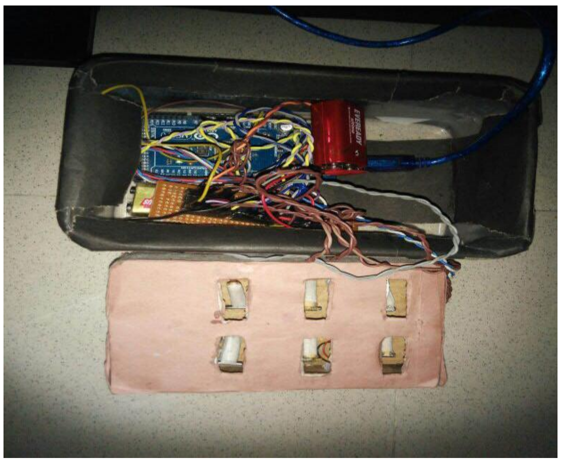

Figure: 3.IMPLEMENTATION

III. MODE TWO: ELECTRONIC BRAILLE TUTOR

The second and a very new concept on this braille system is the ability of this device to provide assistance in learning the braille language by automatically recognizing the letters spoken in isolation. The device is able to accept microphone input and process this input to provide a tactile output corresponding to the input alphabet spoken. The extraction of the best parametric representation of acoustic signals is an important task to produce a better recognition performance. The efficiency of this phase is important for the next phase since it affects its behavior. MFCC is based on human hearing perceptions which cannot perceive frequencies over 1Khz. In other words, in MFCC is based on known variation of the human ear's critical bandwidth with frequency [8-10]. MFCC has two types of filter which are spaced linearly at low frequency below 1000 Hz and logarithmic spacing above 1000Hz. A subjective pitch is present on Mel Frequency Scale to capture important characteristic of phonetic in speech. Thus the apt feature that can be extracted for this application is the MFCC or 'mel' frequency cepstrum coefficients.

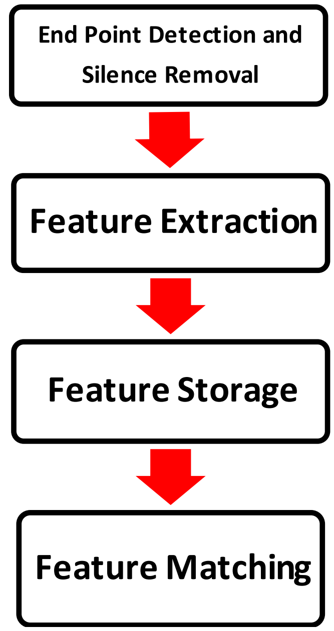

Figure: 4. at the highest level, all speaker recognition systems contain two main modules:
(i) Feature Extraction
(ii) Feature Matching

Figure 4 gives a basic flow of the source code for speech recognition for this particular application. This application has an input signal which is of a short duration due to capturing of only single letters at a time. Thus the input signal needn't be framed or windowed and is subjected to the process of feature extraction thus reducing the computation time. Figure 3 illustrates the basic flow. The steps are described as follows:

1. End point detection & Silence removal

This process removes the unwanted part of signal before the hard core processing takes place. Pre-processing is very crucial in Speech recognition and other such speech applications where silence and background noise have to be clearly removed. End point detection and Silence Removal are techniques that are used in this paper for removal of unwanted noise and blank spaces from the main sound track. This technique also helps in dimensionality reduction in speech that facilitates the system to be computationally more efficient.

2. Feature extraction

Once the main speech signal is captured, feature extraction is done, where a particular signal's main features that distinguishes it from other signals, is carried out. In this paper, the technique applied is MFCC. MFCC (mel frequency Cepstrum Coefficient) are coefficients that are distinct for every speech signal (here for every alphabet).

The steps followed for obtaining MFCC for any given particular signal are as follows:

Step 1: Pre–emphasis

The first process is the Pre-emphases process, which acts like a high pass filter. It increases the energy of signal at higher frequency thus allowing for more efficient processing.

Step 2: Framing
Framing consists of segmenting the speech signals into small frames with length within 20 to 40 msec. Also if any signal is divided into N frames, then a separation of M (M<N) is kept between adjacent frames.

Step 3: Hamming windowing
Hamming window is used as window shape by considering the next block in feature extraction processing chain and integrates all the closest frequency lines.

Step 4: Fast Fourier Transform
It is to convert each frame of N samples from time domain into frequency domain. The Fourier Transform is to convert the convolution of the glottal pulse U[n] and the vocal tract impulse response H[n] in the time domain.

Step 5: Mel Filter Bank Processing
The frequencies range in FFT spectrum is very wide and voice signal does not follow the linear scale. The bank of filters according to Mel scale as shown in is then performed.

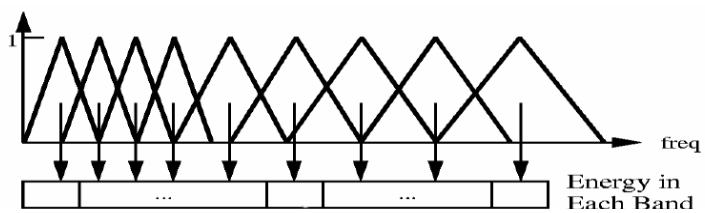

Figure: 4. Energy in each band

Step 6: Discrete Cosine Transform
This is the process to convert the log Mel spectrum into time domain using Discrete Cosine Transform (DCT). The result of the conversion is called Mel Frequency Cepstrum Coefficient. The set of coefficient is called acoustic vectors. Therefore, each input utterance is transformed into a sequence of acoustic vector.

3. Feature Storage
Once the features are extracted from the signal, the important task is to store this feature in such a way that it can be later compared with user input.

4. Feature matching
The final stage of recognition of the input speech signal consists of pattern recognition. Here, the goal of pattern recognition is to classify objects of interest into one of a number of categories or classes. The objects of interest are patterns which in our case is the acoustic vector which we found out with the help of the MFCC technique mentioned earlier. Every class corresponds to a particular letter in our case. Since we are classifying features here, this technique is also termed as feature matching. To be more precise the technique used here is supervised pattern recognition. This may be neural networks or logistic regression. In this paper, we have used neural networks to solve our problem of feature matching.

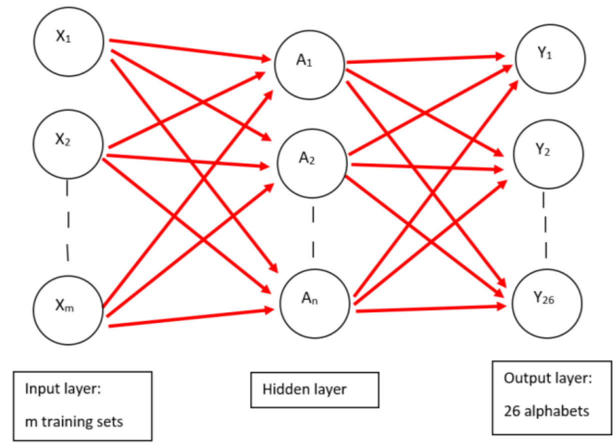

Figure :5. Setd are used to tarin our network

The different m training sets are used to train our network. Corresponding 26 different outputs for every alphabet is achieved. A hyphothesis for every letter letter is calculated.
When switched to Mode 2, the device is ready to take in an input speech signal. The signal spoken into the microphone, is processed and then the corresponding alphabet is predicted using the above networks. After the prediction, it is further passed on to the switch table and the Braille output is obtained in the same way as obtained in MODE 1.

V. CONCLUSION
The device uses the principles of actuation and speech processing to achieve two different objectives for the visually challenged. The two modes, when incorporated into a single device, acts as a standalone instrument to convert not only text but also speech signals into linear motion, serving the purpose of a reader and a tutor.

VI. REFERENCES
[1] Braille L (1829) Procedure for Writing Words, Music and Plain-song Using Dots for the Use of the Blind and Made Available to Them. Institut National des JeunesAveugles (Royal Institution of Blind Youth), Paris.

[2] Wormsley DP, D'Andrea FM (1997) Instructional strategies for braille literacy. American Foundation for the Blind, AFB Press

[3] ‘Speech Identification and Recognition Using Data Mining’ by Mr. H. D. Gadade, Miss. Manisha R. Jadhav and Miss. Shilpa V. Deogirkar

[4] ‘A New Silence Removal and Endpoint Detection Algorithm for Speech and Speaker Recognition Applications’ by G. Saha ,SandipanChakroborty , SumanSenapati

[5] ‘Review on classification based on artificial neural networks’ by Saravanan and S. Sasithra

[6]’ A Portable eBook Reader for the Blind’ by Ramiro Vel´azquez, Hermes Hern´andez, and Enrique Preza

VII. FIGURES USED

[1] Braille representations of the English Alphabet

[2] Working of the device in Mode one

[3] Speech recognition code flowchart

4] Mel scale filter bank, from (young et al, 1997)

VIII. IMAGES USED

[1] Final representation of the device working in mode one