International Journal of Computer and Communication Technology

Volume 3 | Issue 2 Article 3

April 2012

E-Services for the Visually Challenged in India:Current Scenario and the Road Ahead

Sohag Sundar Nanda
P.G.Department of CSA Utkal University, nanda.sohagsundar@gmail.com

Basanta Kumar Swain
P.G.Department of CSA Utkal University, technobks@yahoo.co.in

Sanghamitra Mohanty
Dept of Computer Sc. & Application Utkal University Bhubaneswar, Orissa, India, sangham1@rediffmail.com

E-Services for the Visually Challenged in India:Current Scenario and the Road Ahead

Sohag Sundar Nanda,Basanta Kumar Swain, Sanghamitra Mohanty
P.G.Department of CSA
Utkal University
nanda.sohagsundar@gmail.com
technobks@yahoo.co.in
sangham1@rediffmail.com

Abstract-**This paper discusses the present status of availability of e-services for the visually challenged in India. An analysis of five major Governments to Citizen (G2C) e-service initiatives is done to figure out the level of assistance for the visually challenged users. E-services are judged on a 3-point scale that includes Text to Speech (TTS) support, adherence to W3C accessibility guidelines and provision for voice based e-services. Finally, we discuss the architecture and working of *E-Prakash*, a voice based e-service delivery system..**

Keywords-**E-Prakash, Mel-frequency cepstral coefficient (MFCC), dynamic time warping (DTW)**

1. INTRODUCTION

Efficient and timely delivery of G2C e-services is an important aspect of citizen-centric e-government. For an inclusive e-government, citizen groups with special needs should be provided with equal opportunities to access e-services provided by the government [1]. Numbering more than 15 million [2], the visually challenged citizenry of India is a major special needs group. In this paper we analyze five popular and commonly used e-services provided by the government of India with respect to their suitability of use by visually challenged users.

2. METHODOLOGY USED

The selected e-services are rated on their performance in the following three tests:

a)TTS Support: We test whether the website/portal providing the e-service provides text to speech support. Speech Synthesis Markup Language (SSML) based web browsers are used for this purpose.

b) WCAG 2.0 Guidelines: We test whether the website/portal providing the e-service adheres to the W3C recommended Web Content Accessibility Guidelines (WCAG) 2.0. These guidelines are meant to provide accessible web pages for special need users.

c) Voice based e-service support: We test whether the website/portal providing e-service has support for voice

based e-services like voice passwords etc.

One point is given to a particular e-service if it passes a test. Half a point is awarded for partial success.

3. PRESENT STATUS

3.1 E-Service 1: Filing Tax Returns

This e-service enables citizens to file their annual income Tax return to the government in a convenient and hassle free manner. This e-service is accessible at https://incometaxindiaefiling.gov.in a)TTS Support: No Text to Speech support is available. b)WCAG Guidelines: WCAG Guidelines have not been followed. No accessibility option available. c)Voice based e-service support: No voice based e-service support available. Score: 0 points.

3.2 E-Service 2: Railway Ticket Reservation

This award winning [3] e-service enables citizens to book tickets on the Indian Railway network. Users have to provide source station name, destination station name, date of journey and type of seat required. If seats are available the user is taken to the payment gateway. After successful payment the user can print the ticket. This e-service is accessible at http://www.irctc.co.in

a)TTS Support: No text to speech support is available.

b)WCAG Guidelines: WCAG Guidelines have not been followed.No accessibility option available.

c) Voice based e-service support: No voice based e-service support available.

Score: 0 points.

3.3 E-Service 3: Telephone Bill Payment

This e-service enables citizens to pay their telephone and

Cell phone bills. The bill amount is generated and mapped to the user account. The user has to log in to his account and pay the bill. This e-service is available at http://www.bsnl.co.in

a)TTS Support: No Text to Speech support is available.

b)WCAG Guidelines: WCAG Guidelines have not been followed. No accessibility option available.

c) Voice based e-service support: No voice based e-service support available.

Score: 0 points

3.4 E-Service4: Bangalore One

This award winning [3] e-service portal provides multiple services to the users ranging from bill payments to traffic violation records. This e-service is available at http://www.bangaloreone.gov.in

a)TTS Support: No Text to Speech support is available.

b)WCAG Guidelines: WCAG Guidelines has been followed partially. Font enlargement for partially visually challenged is available.

c)Voice based e-service support: No voice based e-service support available. Score: 0.5 points

3.5 E-Service 5: National Portal of India

This e-service portal is the national portal of India and provides multiple services like information about various government agencies and processes,tourism information, trade information etc.This e-service is available at http://www.india.gov.in/

a)TTS Support: No Text to Speech support is available.

b)WCAG Guidelines: WCAG 2.0 Guidelines have been followed. Accessibility option is available along with an accessibility statement.

c)Voice based e-service support: No voice based e-service support available.

Score: 1 point

The above comparison clearly shows that very little has been done to provide e-services for the visually challenged in India.

4. E-PRAKASH: A VOICE BASED E-SERVICE SYSTEM

In this section we discuss the architecture and working of E-Prakash(Prakash means light in Oriya language), a voice based e-service delivery system. We use a Spinx[4] based Text to Speech engine for reading web pages. In E-Prakash, speaker identification technique is used to authenticate a visually challenged user. We use a four step approach for speaker identification. In the first step, we register a new user. The speech sample of the user along with other related data are stored in a database in an encrypted form. When a user requests authentication, some features are extracted from his voice. This forms the second step of speaker identification. In the third step speaker modeling is done using the extracted features. In the final step the models are matched with the voice database and the speaker is identified. After successful identification, the user secures access to various e-services. Figure 1. shows the architecture of E-Prakash.

4.1 User Registration

New user registration is done by the User Registration Module (URM). Voice samples of the user are saved and used for training the identification system. The user chooses a voice password and it is saved in the voice database. Other information like personal and financial data are also stored and mapped to the users account. User registration is done by taking help of volunteers who assist the visually challenged users. We have tested E-Prakash with 100 visually challenged users consisting of 69 men and 31 women.

4.2 Signal Processing Subsystem

Signal processing subsystem converts the analog signal to digital signal for subsequent analysis and manipulations. The whole corpus designed in URM was digitized at 16000Hz with A/D conversion precision of 16 bits [5, 6].

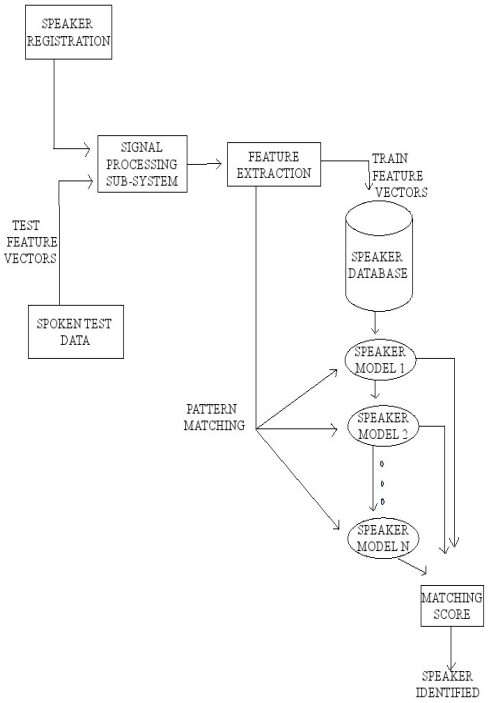

Fig1. Architecture of E-Prakash

4.3 Feature Extraction

Since speech is a highly non-stationary signal, speech analysis is carried out on short segments across which the speech signal is assumed to be stationary. We extract the voice features by using 25ms windows with 10ms shift between two consecutive windows. In E-Prakash we have used mel-frequency cepstral coefficient(MFCC) as feature vectors. The reason for choosing MFCC is that it is considered to be the best approximation of the human ear[7]. The first step in extracting MFCC from speech is pre-emphasis. Pre-Emphasis makes a speech sample depend on the next sample by a certain weight. We then divide the speech signal into frames and apply hamming window to each frame sample. This enables smooth transition between frame samples. Next, we calculate the total energy of the frame samples. The data is then passed through Mel filters. The Melspectrum coefficients are obtained and discrete cosine transform (DCT) is performed on them. The result is Melcepstral coefficients which give the MFCC. Figure 2 shown the process of feature extraction used in E-Prakash.

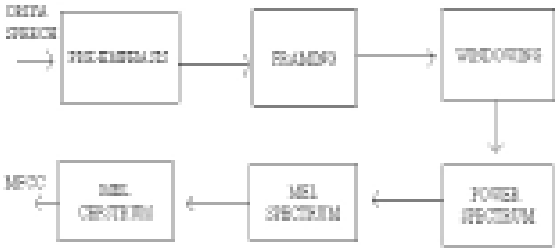

Fig.2 MFCC computation from Oriya Speech

4.4 Speaker Modeling And Identification

We use Dynamic Time Wrapping (DTW) algorithm for speaker modeling and identification. We first convert the training data into speaker models in the form of templates. We then match the incoming speech (test data) with stored templates. The template with the lowest distance measure from the input pattern is the identified speaker. We obtain local distance between feature feature vectors by using Euclidean distance metric. This is represented by

$$d(x,y)=\sqrt{\sum(xj-yj)^2}$$

where x is the feature vector of signal 1 and y is the feature vector of signal 2.We perform time normalization of different utterances of the same word because the time duration and structure of utterance of the same word varies. This is clearly shown in figure 3 which shows the spectrogram of utterance of Oriya numeral Aatha(eight) in slow and fast pace.

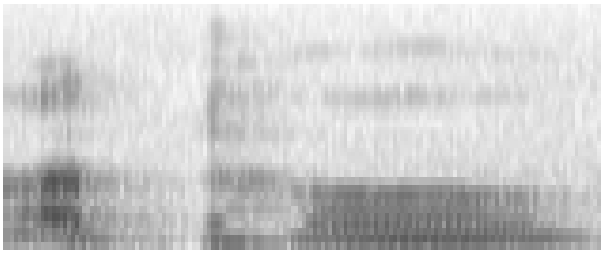

Fig.3 (a) Slow utterance of Oriya numeral Aatha

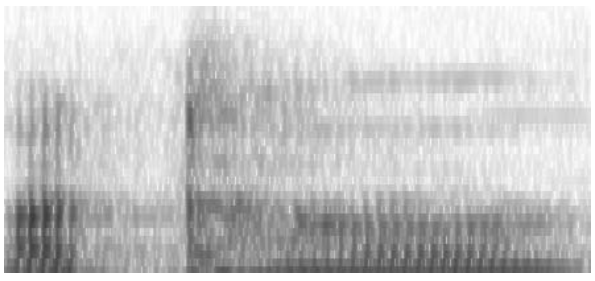

Fig.3 (b) Fast utterance of Oriya numeral Aatha

Time alignment is done to obtain global distance between two speech patterns. The final global distance is matched with the test data(incoming voice) and the lowest matching score is identified as the speaker.[8]

After the speaker is identified, voice based e-services are made accessible to him.

5. RESULT AND CONCLUSION

The following table shows the percentage of false rejection (FR) and false acceptance (FA) of the first five members in our experiment.

Table 1: Accuracy Table for E-Prakash

User	Speaker 1	Speaker 2	Speaker 3	Speaker 4	Speaker 5
Speaker 1	3%FR	5%FA	5%FA	7%FA	7%FA
Speaker 2	7%FA	5%FR	9%FA	8%FA	5%FA
Speaker 3	8%FA	9%FA	3%FR	9%FA	9%FA
Speaker 4	7%FA	4%FA	7%FA	6%FR	5%FA
Speaker 5	9%FA	5%FA	9%FA	9%FA	3%FR

It is clear that the results are encouraging. Using voice based authentication, e-services can be provided securely and efficiently to the visually challenged users. Future work includes incorporation of new techniques to increase accuracy, especially to reduce false acceptance rates.

ACKNOWLEDGEMENT

The authors would like to thank DIT,MCIT Government of India for providing financial support for this work.

6. REFERENCES

[1] Towards Faster and More Inclusive Growth- An Approach to the 11th 5 Year Plan. Planning Commission. Government of India.

[2] Times of India. Retrieved from http://timesofindia.indiatimes.com/articleshow/2447603.cms Processing, Pearson education, Third Impression-2007.

[3] Web Ratna Awards. http://webratna.india.gov.in

[4] SPINX http://cmusphinx.sourceforge.net/

[5] Quatieri T.F., Discrete-Time Speech Signal

[6] B.W.Jervis, Digital Signal Processing A Practical Approach , Pearson Education Asia, 2nd Edition.

[7] L.R. Rabiner, R.W. Schafer, Digital Processing of Speech Signals, Pearson Education(Singapore), India,2004.

[8] S.Mohanty, S.Bhattachrya. "Recognition of voice signals for Oriya Language using Wavelet Neural Network", Proceedings of symposium on Machine Translation System,IIT Kanpur,2001