E-Blind Examination System

Akshay Naik
Department of Computer Engineering
PVPPCOE Mumbai-400022.

Kavita Patil
Department of Computer Engineering
PVPPCOE Mumbai-400022.

Ajinkya Tandel
Department of Computer Engineering
PVPPCOE Mumbai-400022.

Vishal Patil
Department of Computer Engineering
PVPPCOE Mumbai-400022.

Manjiri Pathak
Professor & Guide
Department of Computer Engineering
PVPPCOE Mumbai-400022.

Abstract

The Online-Examination for Blind People is basi-cally an online examination system. It reflects the justification and objectivity of examination. It helps to release the workload of teachers and students. The current online examination systems do not allow blind people to appear for the examination. Hence, the proposed application is designed for the blind people so that they can appear for examination; it is also useful for the other handicapped people with upper limb disability. This research work provides the speech user interface for blind people to interact with the application. The proposed application will dictate the questions to the candidate with the help of Speech Synthesis. The application will accept the answers from candidates through voice commands using Speech Recognition. The proposed system also allows users (examiners/teachers) to add their questions to conduct the examination. The questions can be added easily by uploading a file. As the proposed system is online, result generation will be quick.

Keywords: Attribute aware data aggregation, Dynamic routing protocol, Packet driven timing control algorithm, wireless sensor network

I. INTRODUCTION

A. Current System:

In current system, the blind people who want to take the examination require a writer. The writer writes the answers which the blind dictate them. The traditional way for conduct-ing examination for the blind people requires scribe/writer. These candidates face difficulties in using the scribes for examinations. It is hard for them to dictate the answers to the scribes. Because, the scribes might be of lower qualification hence they cannot interpret their words and write them down as they are.

B. Proposed System:

Online Examination for Blind is a software solution, which allows a particular company or institute to arrange, conduct and manage examinations via online environment. This can be done through the Internet or Local Area Network Environment. Candidate can answer his/her examination paper on the computer and submit answers. The Examination Software evaluates the submitted answers and the results will be avail-able immediately after completion of the examination. The proposed application avoids main drawback of the current online examination system by providing the facility for the blind people to interact with the system.

II. SYSTEM ARCHITECTURE

The basic architecture of the system is shown as follows

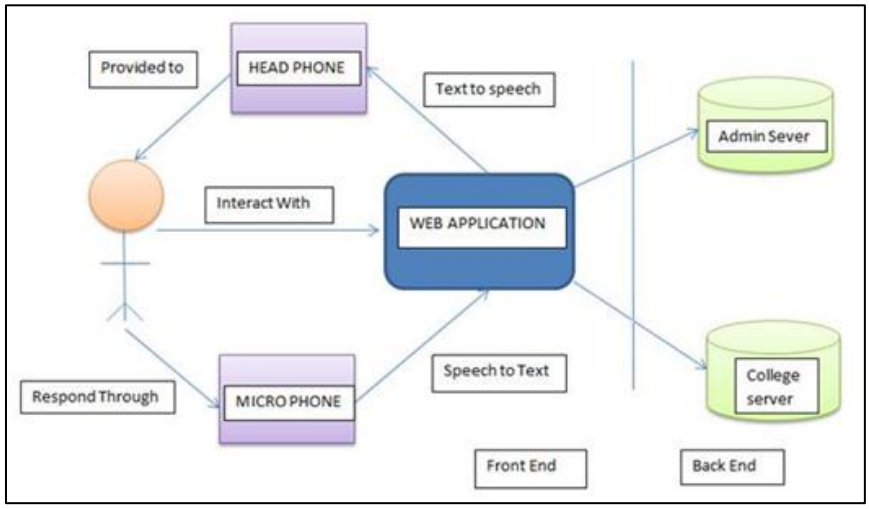

Fig. 1: Basic Architecture

A. Web Application:

The institute which wants to conduct examination will register through web application. Registering institute's ad-ministrator will fill basic information of the institute at the time of registration process. After Registration he will get the login credential of application which can be used for examination purpose.

B. Databases:

There are two databases; first at the admin server and other at the college server. Admin server will store all the details about the institutes who are using the proposed system. And the college server will store all the accounts of teachers, questions, results, etc. Because of this type of architecture the databases will be secure. The institute need not reveal its private data like teachers information and questions.

C. Web Application:

Blind user will login with his/her credential and start exam or may be start book reading. Application will dictate questions to handicapped candidates and accept answers through voice commands. Normal candidate can give examination through Graphical User Interface.

D. Speech User Interface:

The main feature provided in the proposed system is the Speech User Interface for the handicapped people. This is achieved by a series of steps like Speech Recognition (i.e. speech to text conversion) and Speech Synthesis (i.e. text to speech conversion). The description of these steps is explained in following sections.

1) Designing Web Application:

For the proposed system HTML5 is preferred. HTML5 adds many new syntactic fea-tures. These include the new video, audio and canvas elements, as well as the integration of scalable vector graphics (SVG) content (that replaces the uses of generic object tags). These features are designed to make it easy to include and handle multimedia and graphical content on the web without having to resort to proprietary plugins and APIs. [1]

2) Designing Databases:

The database contains three tables. One table contains all the student details who are appearing for the examination. Other table contains uploaded questions for the examination same as in excel file. The last table contains answers of the question given by the student and true answers of the questions for generating result of the exam. For the Proposed system PHP MySQL is preferred. For the development of robust and dynamic web application PHP MySQL is highly preferred. PHP begins with server side programming language and MySQL is open source relational database Management system software when combined, capable of delivering the unique solution. Another advantage of using this is easy to implement with simplicity.

[2]By combining the feature of MySQL, back end database connectivity to the developed web application through PHP XAMPP server must be installed on your system. XAMPP is a free and open source cross-platform web server solution stack package, consisting mainly of the Apache HTTP Server, MySQL database and interpreters for scripts written in the PHP and Perl programming language. [3]

3) Accepting Questions:

Examiner can upload questions by using Microsoft Excel (.xls) file. Import and export data between MySQL and Excel has become much easier in Excel 2007 than its previous versions. To explore the feature, you need to install MySQL for Excel. You can install the MySQL For Excel component while installing MySQL Community Server 6. Or you may install it as an add on upon your existing MySQL Server installation [4]. The questions will be read and stored with the options and the correct answer

in the database. These questions and answers will be used for conducting the examination and for generating result. The advantage of this is that the examiner will be able to set question paper easily. He can add the questions in the database simply by uploading a file dynamically irrespective of its size.

4) Designing Speech User Interface:

a) Speech Synthesis:

The speech synthesis process has a series of steps. [5] They are shown in the following figure.

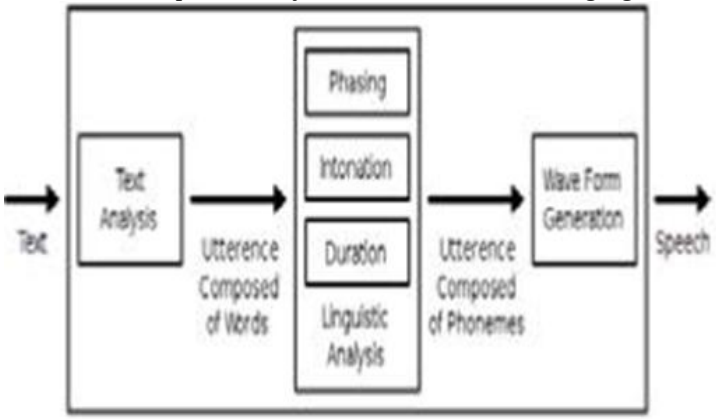

Fig. 2: Speech Synthesis

1) Text Analysis:

This contains string tokenization and normal-ization. It breaks a string into tokens which could be words, phrases or symbols. [5] Considering that punctuation marks are removed from words, there will be an issue with tokenization. In the normalization phase, [6] text is transformed into a more reliable way to be processed means it will become consistent so that it will be easier to deal with. An example of normalization would be dont and do not. Furthermore, an important step in normalizing text is expanding abbreviations. For example, if the string WWW is encountered, it will be normalized to ”World Wide Web”. [5]

2) Linguistic Analysis:

The normalized tokens are assigned with phonetic transcriptions. A phonetic transcription of a word is the visual representation of speech sounds. For ex-ample, the word elephant has the phonetic transcription el uh fuh nt. This process is called text-to-phoneme or grapheme-to-phoneme conversion [5]. A grapheme is a unit in a written language such as a letter, a number or a punctuation mark. On the other hand, a phoneme is a unit of sound to represent the utterance of a grapheme such as the letter x which is uttered eks. When the process of text-to-phoneme or grapheme-to-phoneme is finished, the text is divided into prosodic units [7]. These units determine whether the text is a clause, phrase or sentence. To elaborate more, prosody is the rhythm, stress, and intonation of speech. Therefore, when the text is divided into these kinds of units the form of utterance of the text can be determined whether it is a statement, question, command, emphasis and so on [7].

3) Waveform Generation:

After completing the previous steps, the text will be in a Symbolic Linguistic Representation form which is phonetic transcriptions and prosody merged together. The main purpose now is converting this text into sound [5]. A phone is a speech sound and a diaphone is an adjacent pair of speech sounds. Based on this there are four synthesizer technologies that are mostly used. 1) Unit Selection Synthesis: 2) Diaphone Synthesis: 3) Domain-Specific Synthesis: 4) HMM-based Synthesis: The application uses an open source speech synthesizer, Google Translator. Google uses the open-source speech synthesizer eSpeak for speech synthesis. [8] eSpeak is a compact open source software speech synthesizer for English and other languages for Linux and Windows. eSpeak uses a ”formant synthesis” method. This allows many languages to be provided in a small size. The speech is clear and can be used at high speeds but is not as natural or smooth as larger synthesizers which are based on human speech recordings. eSpeak is available as: 1. A command line program (Linux and Windows) to speak text from a file or from stdin. 2. A shared library version for use by other programs. (On Windows this is a DLL). 3. A SAPI5 version for Windows, so it can be used with screen-readers and other programs that support the Windows SAPI5 interface. 4. eSpeak has been ported to other platforms, including Android, Mac OSX and Solaris. The questions stored in the database by the examiner will be retrieved and shown to the user. The question and the options will be dictated using Google translator for handicapped people.

b) Speech Recognition:

The speech recognition process has a series of steps [5] as shown in the following figure.

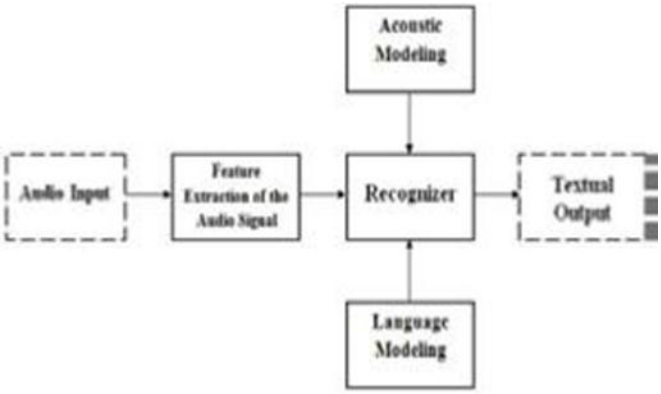

Fig. 3: Speech Recognition

1) Audio Input:

User will interact with the application via speech commands using a microphone.

2) Feature Extraction of the Audio Signal:

After obtaining the input audio signal it has to be processed and its features have to be extracted [9]. This step takes place when the signal is divided into small intervals (in terms of milliseconds) and a feature vector is generated for each interval. The input utterance is represented as a sequence of these feature vectors where each vector consists of coefficients called cepstral coefficients.

3) Acoustic Modeling:

Hidden Markov Models (HMM) are used for acoustic modeling process. Such models are probabilistic models used for pattern recognition, a field in which Speech Recognition is. Every known phoneme has its own HMM represented by a graph with different states (nodes) and transition probabilities [1] from one state to another. These HMMs are obtained after extracting the features of the recorded audios that were used to prepare and train the acoustic model. Every word is made up of a number of phonemes and the HMM of a certain word is made up of the concatenation of the HMMs of the phonemes that constitute the word. These HMMs are present in the acoustic model [1] which the recognition procedure relies on. The obtained features vectors are then matched to the available HMM phonemes to be assigned to the one that most match it (with the highest probability). Then the combined assigned HMMs are used to find the HMM of the whole word to be recognized. Combining phonemes HMMs together is also a probabilistic process. Following figure shows an example of HMMs matching. Matching HMMs, on both phoneme and word levels, is done by the Viterbi algorithm. [9] [10]

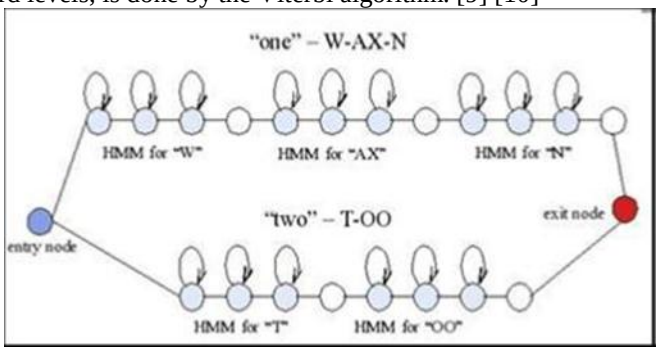

Fig. 4: Acoustic Modeling

4) Language Modeling:

In a similar manner to the one above, the HMM of a meaningful sequence of words is a concatenation of the HMMs of the words included in this sequence. To handle this phrase or sentence HMM formation, a language model is needed that knows how to group different words together to provide the desired output. [9]

5) Textual Output:

After the recognizer completes the acoustic modeling and the language modeling, the obtained textual output is returned. Again the proposed system uses Google translator for speech recognition [11]. It is client server-based. This means that the client just has to generate an audio file and then send it to Google's remote voice recognition servers. Google then responds with a formatted string containing the text recognized from the provided audio file. The whole speech recognition engine, including voice models and dynamic algorithms are placed in Google's servers. Thus, we can do speech recognition almost from any lightweight device capable to record audio.

III. CONCLUSION

The proposed system seems to be far better and efficient in terms of technology and integration point of view. The accuracy of the speech recognition system was among the top challenges. The proposed system will provide a better option for Blind people to appear for the examination.

REFERENCES

[1] http://en.wikipedia.org/wiki/HTML5.
[2] http://www.slideshare.net/brainvireinfo/doc-php-my-sql-development services11102013.
[3] http://en.wikipedia.org/wiki/XAMPP.
[4] http://www.w3resource.com/mysql/exporting-and-importing-data-between-mysql-and-microsoft excel.php.
[5] M. Hussein, J. Karim, and F. Marwan, "Multi-purpose speech recogni-tion and speech synthesis system," IEEE MULTIDISCIPLINARY ENGI-NEERING EDUCATION MAGAZINE, vol. 6 no.4, December 2011.
[6] S. Kumar, "Study of various kinds of speech synthesizer technolo-gies and expression for expressive text to speech conversion system," (IJAEST) INTERNATIONAL JOURNAL OF ADVANCED ENGINEER-ING SCIENCES AND TECHNOLOGIES, vol. 8 no.2, pp. 301–305.
[7] S. Thomas, "Natural sounding text-to-speech synthesis."
[8] http://espeak.sourceforge.net.
[9] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," Proc. of the IEEE, vol. 77 no.2, pp. 257–286, 1989.
[10] M. Nilsson and M. Ejnarsson, "Speech recognition using hidden markov model performance evaluation in noisy environment."
[11] http://panstamp.blogspot.in/2011/04/speech-recognition-from-google new.html.