

Web application for visually impaired Deaf User

Patil Vishal Chandrakant¹, Tambe Seema Sampat², Gade Kalayni Manikrao³

B.E, Computer Engineering, Jaihind College of Engineering, Pune, India¹

B.E, Computer Engineering, Jaihind College of Engineering, Pune, India²

B.E, Computer Engineering, Jaihind College of Engineering, Pune, India³

Abstract: People who are deaf or hard-of-hearing and Visually Impaired User also may have challenges communicating with others via spoken words for and may have challenges being aware of audio events in their environments. Blind person also face problem like deaf ,they can't see words or message. This is especially true in public places, which may not have accessible ways of communicating announcements and other audio events. In this paper, the design and evolution of a mobile sound transcription tool for the deaf and hard-of-hearing is elaborated. Transcriptions include dialog and descriptions of environmental sounds. The transcriber is a multilingual transcriber who listens to the audio and sends the converted text message to the server, the server then using the IP address of the user sends it to the user and Transcriber also message is converted into audio blind person able to listen this message. so Communication done by this System for Blind and Deaf Users .If the user is not logged into the server then the message is stored in the server database and sent when the user logs in.

Keywords: Transcription, Deaf Blind Person , Shapi, HMM algorithm, Headphone, security

1 INTRODUCTION

Communications Process The most basic form of communication is a process in which two or more persons attempt to consciously or unconsciously in uence each other through the use of sym-bols or words to satisfy their respective needs. Likewise, integrated marketing communications uses this communications process to persuade target audiences to listen and act on marketing messages. Our ability to receive, communicate, and process information from other communicators and outside stimuli enables us to perceive the advertising and promotional messages central to integrated marketing communications.

receiver basic information i.e. which category receiver may include either deaf or blind so program generate such type of output so that receiver easily understand the sender mes-sage and sent appropriate feedback. People who are deaf or hard-of-hearing may have challenges communicating with others via spoken words and may have challenges being aware of audio events in their environments. This is especially true in public places, which may not have accessible ways of communicating announcements the design and evolution of a mobile sound transcription tool for the deaf and hard-of-hearing is elaborated

1.1 PROPOSED SYSTEM

A basic type of such communication has various methods as Deaf blind manual alphabet, Block, British Sign Language (BSL), Hands of Signing and Clear Speech. We propose the basic commu-nication system for Deaf Blind people. Using these techniques Deaf blind person can understand whole or some sort of communication. Our aim is to implement simple methods which can sort many complex dependencies in proposed solution. When user apply for transaction module, user require to submit

1.2 OUTLINE OF SYSTEM

Transcriptions include dialog and descriptions of environmental sounds. The transcriber is a multilingual transcriber who listens to the audio and sends the converted text message. In similar way we can send receive in reverse communication signal, text is send toward him in the form of sound.

The objective is to provide Transcription services free which help the deaf-blind categories people to communicate normally, as behave as a essential and

supporting part of society. Application records the speech, and sends the recorded audio to the service providers automatically. And in turn receive the converted text message of the recorded audio thus helping the deaf and making them aware of what is happening around them

Application to receive the voice speech signal. The received samples are stored into memory. The microphone input port with the audio codec receives the signal, amplifies it, and converts it into 16-bit PCM digital samples at a sampling rate of 8 KHz. The codec requires initial configuration, which is performed using hardware on the board.

2 LITERATURE SURVEY

1. Telecommunications Device for the Deaf (TDD). These devices look like typewriters or word processors and transmit typed text over regular telephone lines.
2. There are several new Telecommunications Relay Service technologies including IP Relay and captioned telephone technologies.
3. Mobile textphone devices allowing simultaneous bidirectional text communication.
4. Phone captioning is a service in which a hearing person's speech is captioned by a third party, enabling a hearing impaired person to conduct a conversation with a hearing person over the phone.
5. Hearing dogs are a specific type of assistance dog specifically selected and trained to assist the deaf and hearing impaired by alerting their handler to important sounds, such as doorbells, smoke alarms, ringing telephones, or alarm clocks.

3 SYSTEM DESIGN

3.1 SPEECH TO TEXT

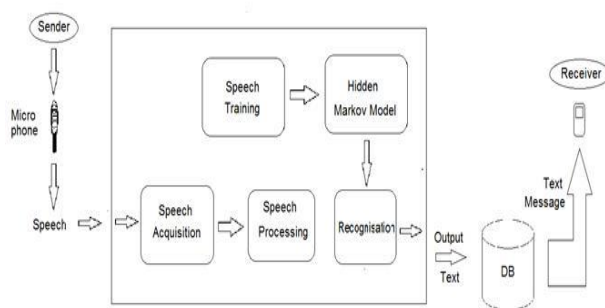


Fig 3.1: Speech To Text Conversion

3.1.1 SPEECH ACQUISITION

Speech acquisition requires a microphone coupled with an analog-to-digital converter (ADC) that has the proper

3.1.2 PREPROCESSING

Speech signal consists of the uttered digit along with a pause period and background noise. Pre-processing reduces the amount of processing required in later stages. Preprocessing involves taking the speech samples as input, blocking the samples into frames, and returning a unique pattern for each sample.

3.1.3 TRAINING

Basic or main process of STT is part of speech-to-text conversion using pattern recognition is training involves creating a pattern representative of the features of a class using one or more test patterns that correspond to speech sounds of the same class.

Model commonly used for speech recognition is the HMM, which is a statistical model used for modeling an unknown system using an observed output sequence. The system trains the HMM for each digit in the vocabulary using the Baum-Welch algorithm. The codebook index created during preprocessing.

Training is a one-time process. Due to the complexity and resource requirements, it is performed using standalone PC application software.

3.1.4 RECOGNITION

The process of comparing the unknown test pattern with each sound class reference pattern and computing a measure of similarity (distance) between the test pattern and each reference pattern Recognition The digit is recognized using a maximum likelihood estimate.

Preprocessing, feature vector extraction, and codebook generation are same as in HMM training. The input speech sample is preprocessed and the feature vector is extracted.

Then, the index of the nearest codebook vector for each frame is sent to all digit models. The model with the maximum probability is chosen as the recognized digit.

3.1.5 STORAGE

Speech-to-text conversion system can convert the recognized digit to a text for storage. The hidden Markov model would output a sequence of n-dimensional real-valued vectors (with n being a small integer, such as 10), outputting one of these every 10 milliseconds. The vectors would consist of cepstral core clients, which are obtained by taking a Fourier transform of a short time window of

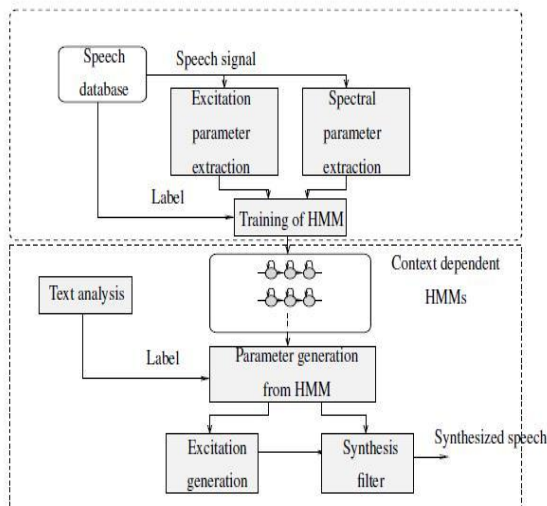


Fig 3.2 HMM Algo

The basic components of a HMM: state-to-state transition probabilities, observation probability distributions, number of states,... There is no such thing as a State class, so each state is simply designated by an integer index; a sequence of states is thus implemented as an array of integers. A hidden Markov model (HMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states. A HMM can be considered the simplest dynamic Bayesian network. Hidden Markov models are especially known for their application in temporal pattern recognition such as speech, handwriting, gesture recognition, part-of-speech tagging, musical score following, partial discharges and bioinformatics.

3.1.7 SPHINXBASE.dll AND Pocketsphinx.dll

speech and decorrelating the spectrum using a cosine transform, then taking the rest (most significant) core clients. The hidden Markov model will tend to have in each state a statistical distribution that is a mixture of diagonal covariance Gaussians, which will give likelihood for each observed vector. Each word, or phoneme, will have a different output distribution; a hidden Markov model for a sequence of words or phonemes is made by concatenating the individual trained hidden Markov models for the separate words and phonemes.

3.1.6 HIDDEN MARKOV MODEL

We use Sphinx.jar for Speech to text conversion. Sphinx.jar contain two dll library le These two dynamic libraries directly con gure with motherboard processing signal for conversion of signal state into other state using HMM algorithms. Above diagram show ow of sample word

from entry node to exit node with HMM algorithm. It is base on Theory of Computation concept. The rest phase in a compiler reads the input source and converts strings in the source to tokens. Using regular expressions, we can specify patterns to lex that allow it to scan and match strings in the input. Each pattern in lex has an associated action. Typically an action returns a token, representing the matched string, for subsequent use by the parser.

3.2 TEXT TO SPEECH

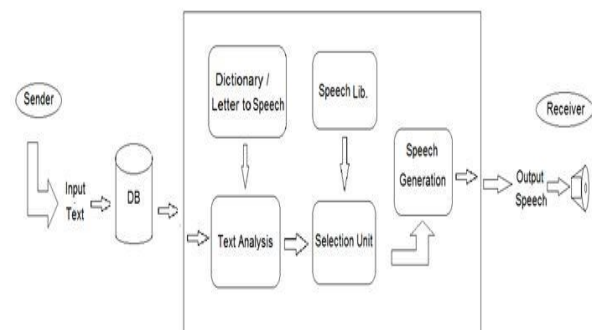


Fig3.3 Text TO Speech

3.2.1 SPEECH APPLICATION PROGRAM INTERFACE (SAPI)

SAPI (Speech Application Program Interface) is an application program interface (API) provided with the

Microsoft Windows operating system that allows programmers to write programs that offer text-to-speech and speech recognition capabilities. Interfaces are provided for the C, C++, and Visual Basic programming languages. Using Microsoft's COM (Component Object Model) architecture, SAPI is the most widely used speech application program interface used today. In the future, Microsoft plans to embed speech technology using SAPI into their operating system.

SAPI has seven main components:

1. **Voice Command:** Voice Command is a high-level interface that provides command and control speech recognition for applications. Voice Command allows a developer to create a Voice Command menu that contains voice commands, such as "new file" or "send mail to someone@anywhere.net" that a user speaks into a microphone or other audio device. The user can control the computer without needing a keyboard or mouse.
2. **Voice Dictation:** Voice Dictation allows the user to dictate into any application that supports speech recognition. An invisible or virtual edit box receives the text the user dictates and displays the text in an application window. Voice Dictation allows text formatting such as capitalization, translation of punctuation words into punctuation symbols, built-in glossary entries, and correction of the last word spoken or a selected word. Applications that use Voice Dictation classify speech by topics that use different language styles. Topics include e-mail speech, formal writing, or programming speech. Voice Dictation stores the information for each topic on your hard drive.
3. **Voice Text:** Voice Text converts text into speech that is played over computer speakers or sent over a telephone line. The speech played has several different modes, each with a different voice.
4. **Voice Telephony:** Voice Telephony uses telephony controls that are similar to Windows controls. Windows controls include buttons, list boxes, sliders and other objects that can be manipulated by a mouse or keyboard. Telephony controls are codes that recognize spoken

responses such as Yes or No, your phone number, the date, and the time. Telephony controls create a dialogue between the user and the computer. For example, a user calls a vendor to order an item. The user then answers several questions by speaking into the telephone receiver. The telephony controls recognize these responses and sends them to the application that processes responses. Telephony controls also handle error conditions (these are common with spoken numbers or when the caller does not respond) and variations of answers such as "January 4th" or "tomorrow."

5. **Direct Speech Recognition:** This is a low-level interface similar to Voice Command. The main difference is Direct Speech Recognition speaks directly to the speech engine. This gives the application more control and speed.
6. **Direct Text To Speech:** This is a low-level interface similar to Voice Text that also speaks directly to the speech engine. Audio Objects
7. **An Audio Object:** tells the speech engine where to get its audio.

The future of speech technology will include products that allow you to do such things as surfing the Internet using speech and asking your television what is showing tonight. Software developers are developing applications that understand concepts. For example, if you tell your computer to print a certain document, your application will know whether to print it on your printer or the network's printer. Speech technology is important for medical professionals, law enforcement personnel, the physically handicapped, as well as many business and home users

3.3 SECURITY

Java Cryptography Architecture Standard Algorithm public class Cipher extends Object This class provides the functionality of a cryptographic cipher for encryption and decryption. It forms the core of the Java Cryptographic Extension (JCE) framework.

In order to create a Cipher object, the application calls the Cipher's getInstance method, and passes the name of the requested transformation to it. Optionally, the name of a provider may be specified. A transformation is a string that

describes the operation (or set of operations) to be performed on the given input, to produce some output. A transformation always includes the name of a cryptographic algorithm (e.g., DES), and may be followed by a feedback mode and padding scheme. A transformation is of the form: "algorithm/mode/padding" or "algorithm"

the latter case, provider-specific default values for the mode and padding scheme are used). For example, the following is a valid transformation:

```
c=Cipher.getInstance(DES/CBC/PKCS5Padding);
```

modes such as CFB and OFB, block ciphers can encrypt data in units smaller than the cipher's actual block size. When requesting such a mode, you may optionally specify the number of bits to be processed at a time by appending this number to the mode name as shown in the "DES/CFB8/NoPadding" and "DES/OFB32/PKCS5Padding" transformations. If no such number is specified, a provider-specific default is used. (For example, the SunJCE provider uses a default of 64 bits for DES.) Thus, block ciphers can be turned into byte-oriented stream ciphers by using an 8 bit mode such as CFB8 or OFB8. Modes such as Authenticated Encryption with Associated Data (AEAD) provide authenticity assurances for both confidential data and Additional Associated Data (AAD) that is not encrypted. (Please see RFC 5116 for more information on AEAD and AEAD algorithms such as GCM/CCM.) Both confidential and AAD data can be used when calculating the authentication tag (similar to a Mac). This tag is appended to the ciphertext during encryption, and is verified on decryption. AEAD modes such as GCM/CCM perform all AAD authenticity calculations before starting the ciphertext authenticity calculations.

To avoid implementations having to internally buffer ciphertext, all AAD data must be supplied to GCM/CCM implementations (via the updateAAD methods) before the ciphertext is processed (via the update and doFinal methods).

```
GCMParameterSpec s = new
```

```
GCMParameterSpec(...);cipher.init(...s);
// If the GCMParameterSpec is needed again
cipher.getParameters().getParameterSpec(GCMParameterSpec.class));
cipher.updateAAD(...);\ADD
cipher.update(...);// Multi-part update
cipher.doFinal(...);// conclusion of operation
```

4 APPLICATION

4.1 SPEECH TO TEXT



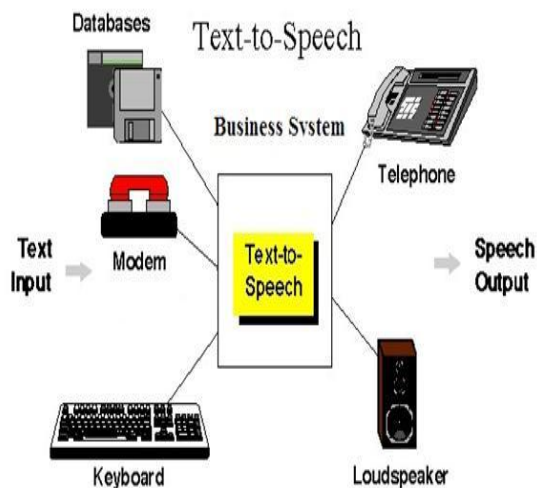
Fig:4.1 Hospital Management

When the patient visits a doctor, the doctor spends time with the patient discussing his/her medical problems. The doctor performs a physical examination and may request various laboratory or diagnostic studies; will make a diagnosis or differential diagnoses, and then decides on a plan of treatment for the patient. After the patient leaves the office, the doctor records.

the information about the patient encounter using handheld cassette recorder or into a regular telephone, dialed into a central server located in the hospital or transcription service office, which will 'hold' the report for the transcriptionist. This report is then accessed by a medical transcriptionist, it clearly received as a voice file or cassette recording, who then listens to the dictation and transcribes it into the required format for the medical record, and of which this medical record is considered a legal document. The next time the patient visits, the doctor will call for the medical record or the patient's entire chart, which will

contain all reports from previous encounters. The doctor on occasion can recall the patient's medications after seeing only the medical record, although doctors prefer to not recall prescriptions without seeing the patient to establish if anything has changed.

4.2 TEXT TO SPEECH



In the Business use of the Speech To Text Conversion. At the any notice At a spread all the information in building and wing possible to this paper implementation software.

5 CONCLUSION AND FUTURE SCOPE

5.1 CONCLUSION

As our system successfully communicate with blind deaf people using conversion models, according to receiver type. Two conversions are applied in system Speech To Text (STT) and other is Text To Speech (TTS).

Using this system physically blind deaf category communicates with normal society and they can be a routine part of current assets.

5.2 FUTURE SCOPE

As we implement our speech to text text to speech utility, in future we can try for can be enhanced to picture and video transcription. In picture and video transcription the contents can be transcribed and described. Performance wise with high speed networking the time required during server communication can be reduced, as

the current project handles only GPRS. In future it can be extended to 3G, wi- and more. And more importantly noise filters can be used for filtering the noise (unwanted sound) from the user recorded audio.

Also we can use latest technologies to improve application structure, attractive GUI well perform or user friendly operating functionality. In future we image processing and recognition algorithms dedicated for blind users are proposed. Namely the color detector, the light direction detector and the object recognition algorithm.

REFERENCES

- 1) Zdenek Bumbalek, Jan Zelenka, Lukas Kencl ?Cloud-Based Assistive Speech-Transcription Services? Computers Helping People with Special Needs Lecture Notes in Computer Science Volume 7383, 2012, pp 113-116.
- 2) Schmitt, D.G., Slowiaczek, M.L. ?An experimental study of synthesized speech intelligibility using text created by telecommunication device for the deaf(TDD) users? Global Telecommunications Conference, 1990, and Exhibition. 'Communications: Connecting the Future', GLOBECOM '90, IEEE
- 3) Bryan Bergeron ?Voice Recognition and Medical Transcription? MedGenMed. 2004; 6(3): 54. Published online 2004 August 27. PMID: PMC1435648.
- 4) Anssi Klapuri ?Signal Processing Methods for the Automatic Transcription of Music? ISBN 952-15-1147-8 ISSN 1459-2045.
- 5) Martin H. Block, Marc Okrand, Ph.D ?Real Time Closed Captioned Television as an Educational Too? American Annals of the Deaf September 1983 Vol. 128, No. 5
- 6) Martin H. Block, Marc Okrand, Ph.D ?General Packet Radio Service (GPRS)? By Peter Rysavy, Rysavy Research, for GSM Data Today online journal. September 30, 1998
- 7) Marcel Garf ? An introduction to Java Telephone API(JTAPI)? IBM@IBMCH IBM Research Division Zurich Research Lab March 2000.
- 8) T. Matthews, J. Fong, F. Ho-Ching and J. Manko , ?Evaluating non-speech sound visualizations for the deaf?, Behaviour Inf. Technol. 2006.
- 9) James Koegh ?J2ME: The Complete Reference? Second Edition McGraw-Hill Publishers.