



CREATING STYLE SHEET BY USING SPEECH RECOGNITION

Preethi.S

Department of computer science & IT, Jain University, Bangalore, India

Article DOI: <https://doi.org/10.36713/epra4478>

ABSTRACT

Today's rich Web applications use a mix of Java Script and Asynchronous communication with the application server. This mechanism is also known as Ajax: Asynchronous JavaScript and XML. The intent of Ajax is to exchange small pieces of data between the browser and the application server, and in doing so, use partial page refresh instead of reloading the entire Web page. AJAX (Asynchronous JavaScript and XML.) is a powerful web development model for browser-based Web applications. This project is designed and developed keeping that factor into mind, and a little effort is made to achieve this aim. Our project is capable to recognize the speech and convert the input audio into text & style sheet it also enables a user to perform operations "save, open, exit" a file by providing voice input.

1. INTRODUCTION

1.1 Introduction

Today's rich Web applications use a mix of Java Script and Asynchronous communication with the application server. This mechanism is also known as Ajax: Asynchronous JavaScript and XML. The intent of Ajax is to exchange small pieces of data between the browser and the application server, and in doing so, use partial page refresh instead of reloading the entire Web page. This project is designed and developed keeping that factor into mind, and a little effort is made to achieve this aim.

Our project is capable to recognize the speech and convert the input audio into text & style sheet it also enables a user to perform operations "save, open, exit" a file by providing voice input.

Internet browsers are common applications any computer user likes to work with irrespective of his computer knowledge. There are various third party tools available in the market that can attach to a browser application to change the actual display of a webpage in accordance to the user requirement.

1.2 Speech Recognition

Speech Recognition (or Sometimes referred to as Automatic Speech Recognition) is the process by which a computer identifies spoken words. Basically, it means talking to a computer & having it correctly understand what you are saying.

1.3 Style Sheet

A Style Sheet is a collection of style rules that tells a browser how the various styles are to be applied to the HTML tags to Present the document. Rules can be applied to all the basic HTML elements, for example the <p> tag, or you can define your own variation and apply them where you wish. Style sheets can be external or internal to your web site. A Style sheet is a text file that is stored on any webpages or style sheet can also be included as part of your web page .

There are three types of style sheets:

- **Embedded:** the style rules are included within the HTML at the top of the web page – in the head
- **INLINE :** the style rules appear throughout the HTML of the web page – i.e in the body .
- **Linked:** The style rules are stored in a separate file external to all the web pages.

2. REVIEW OF LITERATURE

2.1 Speech Recognition Process

In humans the speech or acoustic signals are received by the ears & then transmitted to the brain for understanding & extracting the meaning out of the speech & then to react it appropriately. Speech recognition enabled computer or device too, work



under the same principle. Any speech recognition system involves following five major steps

➤ **Signal processing:**

The sound is received through the microphone in the form of analog electrical signals. These signals consist of the voice of the user & the noise from the surroundings. The noise is then removed & the signals are converted into digital signal.

➤ **Semantic Interpretation**

Here it checks if the language allows a particular syllable to appear after another. After that, there will be grammar check.

2.2 Types of speech recognition systems

Speech recognition system can be separated in several different classes by describing what types of utterances they have the ability to recognize. These classes are based on the fact that one of the difficulties of SR is the ability to determine when a speaker starts and finishes an utterance. Most packages can fit into more than one class, depending on which mode they're using.

➤ **Isolated word :**

Isolated word recognizers usually require each utterance to have quiet on both sides of the sample window. It doesn't mean that it accepts single words, but does require a single utterance at a time.

➤ **Connected word:**

Connect word systems are similar to isolated words, but allow separate utterances to be run-together with a minimal pause between them.

➤ **Voice verification/identification:**

Some ASR system have the ability to identify specific users by characteristics of their voices. If the speaker claims to be a certain identity and the voice is used to verify this claim, this is called verification or authentication. There are two types of voice verification/identification system, which are as follows

• **Text-Dependent:**

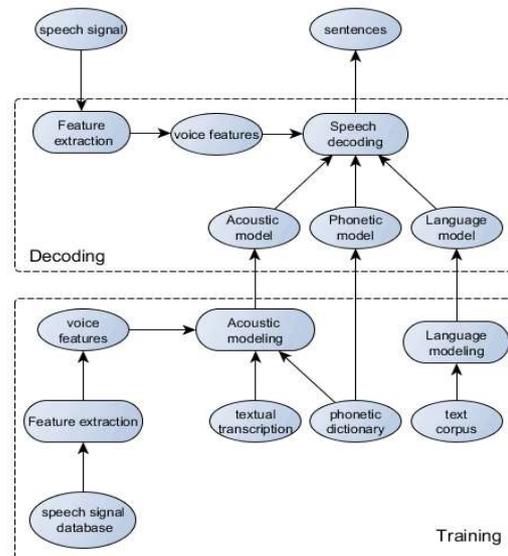
If the text must be the same for enrolment and verification this called text-dependent recognition. In a text-dependent system, prompts can either be common across all speakers (e.g: a common pass phrase) or unique.

• **Text-Independent:**

Text-independent systems are most often used for speaker identification as they require very little if any cooperation by the speaker. In this case the text during enrolment and test is different. In fact, the enrolment may happen without the user's knowledge, as in the case for many forensic applications.

3. ARCHITECTURAL DESIGN

Speech recognition is a difficult task, and is one of most important branches in digital signal processing. Figure presents the general architecture of an automatic speech recognition system, which emphasizes the two main steps of the architecture, namely the training and decoding tasks.



3.1 Architecture

Fig General architecture of an recognition also shows that the recognition does not model speech directly. Here, the feature extraction block is employed to extract specific features, which are further used in the acoustic model creation. The block is also used in recognition or decoding process for speech representation.

4 . METHODOLOGY

A connected digits speech recognition system is a limited vocabulary recognizer. This means the system will only recognize and transcribe in style sheet.

Speaker dependent system should be better at decoding speech uttered by the specific user for which it was trained. However, this demands a new system to be constructed and trained individually for each speaker, a time consuming and non-scalable task. speech development toolkit is hidden markov model toolkit

4.1 speech recording

A speech database comprising recorded audio files is required to build an acoustic model for the speech recognition system. An online speech recorder application, developed by speed research group, is used to record the audio files.

4.2 creating a language model

Current system trained with a large vocabulary for speech recognition use a statistical language model. These language models are built based on large text corpora specific for recognition task. Estimating the probability of occurrence for words and sequences of words for that task. The language models are then used in the process of decoding to select the most likely sequence of words proposed by the acoustic model.

4.3 Decoding

The three basic components of a speech recognition system are available for use in the decoding process. As a result the system can decode using the evaluation data and then compare the textual transcription of decoding process with the reference transcription.

5. RESOURCE AND TOOLS

5.1 Speech Recognition Software

There are ample of speech recognition software available in the market. These software's are available for various kinds of platforms including smart phones, PCs, Tablets etc & are designed for different operating systems as well

➤ Julius

- ✓ Open source & freeware speech recognition engine
- ✓ Developed by –Nagoya institute of technology
- ✓ Developed in C language
- ✓ Operating systems – unix, windows
- ✓ High- performance two-pass large vocabulary continuous speech recognition decoder software for speech-related researchers and developers.

➤ Google Now

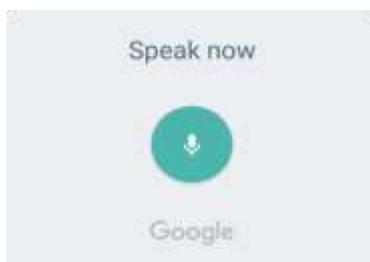


Figure 5.1.1 Google Now Interface

- ✓ An intelligent personal assistant software
- ✓ Developed by- Google
- ✓ Operating System - android & later
- ✓ Language available in –English
- ✓ Google now is implemented as an aspect of the google search application. It recognizes repeated actions that a user

performs on the device and to display more relevant information to user in the form of cards.

➤ SIRI



Figure 5.1.2 SIRI

- ✓ An intelligent personal assistant and knowledge navigator software
- ✓ Developed by apple inc.
- ✓ Operating system –ios 5 & later
- ✓ Platform –iphone, ipad touch, ipad and later
- ✓ Languages available in – English, French, german, japanese etc..
- ✓ The application uses a natural language user interface to answer questions, make recommendations, and perform actions by delegating requests to a set of web services.

➤ S Voice



Figure 5.1.3 S Voice

- ✓ An intelligent personal assistant and knowledge navigator software
- ✓ Developed by samsung
- ✓ Operating system - Android 4.0 & 4.1
- ✓ Platform- Samsung
- ✓ Languages available in – English, French, german, japanese etc..
- ✓ The application uses a natural language user interface to answer questions, make recommendations, and perform actions by delegating requests to a set of web services

➤ **Dragon naturally speaking**

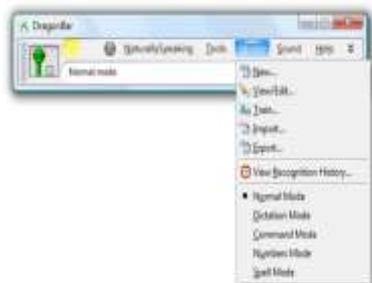


Figure 5.1.4 Dragon Naturally Speaking Interface

- ✓ A speech recognition software package
- ✓ Developed by – Nuance communications
- ✓ Operating system – Windows
- ✓ The software has three primary areas of functionality: dictation, text-to-speech and command input

➤ **Windows speech recognition**

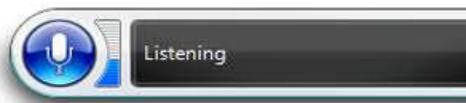


Figure 5.1.5 Windows Speech Recognition Interface

- ✓ A speech recognition application
- ✓ Developed by –microsoft
- ✓ Operating system – windows vista, windows 7 and windows 8
- ✓ Languages available in – English
- ✓ Allows the user to control the computer by giving specific voice commands.

5.2 Advantages & disadvantages

Advantages

- ✓ Increases productivity
- ✓ Able to write the text both through keyboard and voice input.
- ✓ Can help with menial computer tasks, such as browsing and scrolling
- ✓ Voice recognition of different notepad commands such as open save and clear
- ✓ Lower operational cost
- ✓ Open different windows softwares based on voice input
- ✓ Provide significant help for people with disabilities.
- ✓ Requires less consumption of time in writing text

Disadvantages

- ✓ Difficult to build a perfect system
- ✓ Involves more than just words (non verbal communication stutters)
- ✓ Every human being has differences such as their voice , mouth and speaking style
- ✓ Filtering background noise is a task that can even be difficult for humans to accomplish.

6. IMPLEMENTATION & RESULTS

The advancement of technology has led to the growth of many voice-activated products and applications. It will only become more widespread in the next several years

Speech-recognition system implementation in the studied orthopaedic division started with one physician,

7. FUTURE SCOPE

➤ **Achieving efficient speaker independent word recognition**

All the SR systems well be speaker independent and will produce the same kind out output for a particular command irrespective of user.

➤ **Ability to distinguish nuances of speech and meanings of words**

SR systems would be able to distinguish between nuances phrases & meaningful commands, & would be able to process the proper command out of the nuances phrases correctly.

➤ **Stand-alone speech recognition systems**

Presently there is no SR stand-alone systems available, all the SR systems been developed are based on one or the other pre-existing hardware and software platforms, but in near future stand alone speech recognition systems might be available in the market.

➤ **Wearable speech recognition system**

SR systems will be embedded in wearable devices or things such as wrist watch, necklace, bracelet etc. there will be no need of carrying bulky devices and the technology can be used on the go.

8. CONCLUSION

Speech recognition will revolutionize the way people conduct business over the web and will, ultimately, differentiate world-class e-businesses. VoiceXML ties speech recognition and telephony together and provides the technology with which businesses can develop and deploy



voice-enabled web solutions Today these solutions can greatly expand the accessibility of web-based self-service transactions to customers who would otherwise not have access, and, at the same time, leverage a business' existing web investments.

Speech recognition and voiceXML clearly represent the next wave of the web. In near future people will be using their home and business computers by speech not by keyboard or mouse. Home automation will be completely based on speech recognition. Speech recognition will revolutionize the way people interacted with smart devices & will ultimately differentiate the upcoming technologies. Almost all the smart devices coming today in the market are capable of recognizing speech. Many areas can benefit from this technology. Speech recognition can be used for intuitive operation of computer-based systems in daily life.

9. REFERENCES

1. JOE TEBELSKIS {1995}, *SPEECH RECOGNITION USING NEURAL NETWORKS*, school of computer science Carnegir Mellon University
2. KARE SJOLANDER {2003}, *An HMM-based system for automatic segmentation and alignment of speech*, Umea University, Department of philosophy and linguistics
3. KLAUS RLES {1999}, *HMM AND NEURAL NETWORK BASED SPEECH ACT DETECTION*, International Conference on Acoustics and Signal Processing
4. B.PLANNERER {2005},*AN INTRODUCTION TO SPEECH RECOGNITION*
5. KIMBERLEE A. KEMBLE, *AN INTRODUCTION TO SPEECH RECOGNITION*, Voice Systems Middleware Education, IBM
6. LAURE SCHINDLER {2005}, *A SPEECH RECOGNITION AND SYNTHESIS TOOL*, Department of Mathematics and Computer Science, College of Arts and Science, Stetson University
7. MLKAEL NILSSON, MARCUS EGNARSSON {2002}, *SPEECH RECOGNITION USING HMM*, Blekinge Institute of Technology