International Journal of Progressive Sciences and Technologies (IJPSAT)
Vol. 2 No. 2 April. 2016, pp. 74-75

http://ijpsat.ijsht-journals.org

Intelligent Voice Browser

Kartheek Javvaji[1],Vamsi Krishna Reddy Bandi[2]

[1,2] 1SCOPE, VIT university,
Vellore, Tamil Nadu, India

ABSTRACT: The Internet has realized an unfathomable change in human access to learning and data. Voice programs permit individuals to get to the web, utilizing discourse combination, pre-recorded sound, and discourse acknowledgment. This can be supplemented by keypads and little shows. Voice may likewise be offered as an assistant to ordinary desktop programs with high-determination shows, giving an open different option for utilizing console or touch screen, for case in cars where hands/eyes free operation is crucial. Voice association can get away from the physical constraints of keypads and showcases as cell phones are additionally to be utilized now and again where the client can't utilize his hands or eyes. Voice collaboration can get away from the physical confinements of keypads and showcases as cell phones, similar to wearables turn out to be considerably littler. The program will have a coordinated content extraction motor that examines the substance of the page to build an organized representation. The interior hubs of the structure speak to different levels of reflection of the substance. This aide in the simple and adaptable route of the page in order to quickly home into objects of hobby. At last, the program is coordinated into a programmed content to discourse interpretation motor that yields the chose content as discourse. There are 10 times the same number of telephones on the planet as associated PCs. Telephones will turn into the real entryway to the Web. Discourse acknowledgment is not yet broadly connected with the 'visual Web', but rather this will change as gadgets proceed to psychologist and make consoles unrealistic, and as mobile phones turn out to be more predominant in areas with low education rates. Requesting headings while driving and listening to the reaction through discourse combination represents how down to earth "without hands" applications can be to portable clients.

KEYWORDS: Voice browsing, Fuzzy Logic, Universal Windows Platform.

1. INTRODUCTION

Voice Browsing refers to using speech to navigate a browser application. These applications make use of the user's voice and so some basic web browser actions. There is already a technology in this area, named VoiceXML. In much the same way that Web applications are written in HTML and are rendered in a Web browser, speech applications are written in VoiceXML and are rendered via a Voice Browser.

Real-world voice-driven Web applications abound, though people may not always realize they are interacting with a Web service; examples include airline departure and arrival information, banking transactions, automated phone appointment reminders, and automated telephone receptionists. By one estimate, over 85% of Interactive Voice Response (IVR) applications for telephones (including mobile) use W3C's VoiceXML standard.

We can easily understand voice browsing and the voice web by imagining the Web itself. Everyday we visit websites and receive information visually. The voice web has voice sites in which the information is conveyed through speech. The most basic example is calling up an airline or financial portal, where speech recognition software gives the caller a series of options - buy or sell stocks, book airline tickets - and the caller interacts with a computer system.

Corporate use of voice browsing is primarily used at banks and airlines. SpeechWorks, a voice application provider, has over 70 partners including United Airlines, Continental Airlines, AOL, FED-Ex, and Hewlett & Packard. Internationally, SpeechWorks partners with E-Plus in Germany, Singapore Telecom, and Credit Lyonais in France. "We haven't scratched the service of corporate adoption," promises Pete Settles, Media Manager of SpeechWorks [1].

Corresponding Author: Vamsi Krishna (bandivamsi95@gmail.com)

2. METHODS

Many browser activities can voice automated. Opening a new tab or bookmarking a website or even starting a private browsing tab, etc. These are the basic features generally expected from a voice browser. Apart from these, we can implement some complex features too. We use Bing speech to text translation for translating our user's speech. It is readily available in all the windows devices and the other devices can use the Bing Api. Bing voice to text translate is the top notch translator as it can even recognize the other accents of a single language. For example, there are multiple accent options in one English language itself (i.e.) English(USA), English(UK), English(India), etc.

After translating the speech, we can separate the words and store them in an array. All the letters are translated to small letters, to remove all the capitals, and the periods too. The array of the words is then subjected to fuzzy logic to determine the particular command the user meant to execute.

Each of the commands are associated with a particular set of key words. The process consists of the following steps:

- Check if any words match with the KEY-WORD database.
- Apply fuzzy logic and find the weights of the individual commands.
- If any one of them crosses the trigger weight, they will be executed.
- If two opposing commands are triggered, none will be executed.

Along with the inbuilt commands, even the user can add them. There are two ways of adding new commands. One of them is manually adding the key words and the action the browser should be doing. The other is when no command is determined, the user will be prompted to add an action for the voice command given. All the words in the particular voice command will be added to the KEY-WORD database and will be associated with the command. In this way, the voice browser can even learn new commands and the user can get a personalized experience.

This can also increase the security. Instead of saving passwords in the browser in which case anyone with the user's phone can access his account, we save the username and password in a database and the user can modify the command associated with that particular site login. All the settings where the command key words can be modified is password protected.

Since Universal Windows Platform is available for even IOT without any modifications, the same can be implemented in it. The web pages are injected with JavaScript functions that are generated with .NET and the web actions like pressing a button will be done virtually. As soon as a page is loaded all the text items and links, with the help of the html tags, are loaded up temporarily and are made available for the user.

3. CONCLUSION

Therefore, we can expect a fully voice automated voice commanding browser in the near future. This can even be personalized by using the Microsoft's Speaker recognizer API [2] made available under the Microsoft's Cognitive services.

REFERENCES

[1] Voice Browsing: How Two Great Ideas Go Great Together by C. J. Kennedy: http://www.wirelessdevnet.com/channel/voice/features/voicebrowser.phtml

[2] Microsoft Cognitive Services | Speaker Recognition API : https://www.microsoft.com/cognitive-services/en-us/speaker-recognition-api