ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
IJARSMT
Volume 1, Issue 2, July 2015

# Voice Enabled Active Web browser together with Included OCR pertaining to Aesthetic Impaired Clients.

**Sandip Vende[1], Asst.Prof.Krishnakant Kishore [2]**

PG Student, Dept. Of CSE, Jagadguru Dattatray College of Technology, Indore, M.P., India[1]

HOD, Dept. Of CSE, Jagadguru Dattatray College of Technology, Indore, M.P., India[2]

***ABSTRACT***— Information technology plays the vital role in trendy civilization. The various facility provided by this technologies improving the growth of globalization and also information sharing is easier than older days, mostly current accessing info. via hand-held devices like Tablets, Mobiles and desktop computers.The real piece of our general public is that visual debilitated or visually impaired persons can't ready to get to an extraordinary wellspring of data that is web. The way is accessible to such individuals is just an extraordinary printed books with Braille. it is extremely restricted wellspring of data.In this paper we propose an answer that naturally examine the pages W3C and channel it content from interactive media pages rich with Graphics and converts that content discourse with TTS framework.

KEYWORDS- Talking Internet Browser, Speech Recognition, Speech Synthesis, Active Web Browser, Smart browsers, Voice enabled Browser.

## I. INTRODUCTION

Surfing web is intelligent method for Learning, Expressing and communicating with world for ordinary individuals, however modern age data innovation Facilities frequently are not suitable for visual impeded or visually impaired individual's, that such individuals are perusing with just Braille coding. Braille client can read PC Screen with other electronic backing like as refreshable Braille display or read out from print out taken from Braille embosser, for such individuals without utilization of such assistive innovation they won't ready to get information from PC.

The present pattern of sites are rich of enhancements of visuals that require a mouse click with visual components which makes information of pages are more alluring and impressive also the visual impaired persons have no distress about it .Thereafter all the more piece of website pages is content data and that is not yet accessible by these individuals. The present Web program can't channel the text independently from interactive media pages and does not introduce content arranged in Links, headings, and paragraphs.

A proposed voice enabled dynamic program will break down such pages of world wide web and present that pages in reformatted links, headings and paragraphs by concealing the pictures according to requirement and permits client to explore through links accessible on the site page and additionally heading and sections by using voice orders and read out that content loudly in virtual sound space environment. We realize that visual debilitation imply that individuals with low vision, so a large number of them has the capacity to perused the content without utilizing assistive innovation implies that they find themselves able to peruse specifically with their eyes with some trouble like not able to peruse little size letters, partial colour blindness and many more .In such cases the program will give the facility of like zoom the content, changing content shading or changing content background shading by which their trouble of perusing content from site pages will be diminished. On account of anxiety in perusing expansive

ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
IJARSMT
Volume 1, Issue 2, July 2015

passage program will give the provision of transforming the chose content to sound records progressively with determined speaker, volume and rate by which they can capable to listen bigger sections rather than reading it. This sort of highlights will likewise help to ordinary individuals to changes over expansive writings to sound document by which they can move the sound record to their iPods or mobiles and listen it at whatever point they need, it will lesser the surfing time.

Propose arrangement will likewise incorporate the content extraction highlight from downloaded or accessible records available on local storage.The File extensions like PDF and rich content formats (RTF) and convert the extracted content to sound documents that are played by PC or whatever the other sound handling electronic gadgets. This will be open another window of information for such Visual Impaired persons or Low vision persons.

The routine procedure of perusing content for visual hindered persons is we know just Braille. Visual debilitated persons are not able to peruse printed content on diverse things like marks,books covers or any ink printed content. For this proposed arrangement will give backing of optical character recognition to peruse out such content.

The Integration of Optical Character Recognition (OCR) will likewise help the visual hindered client's to diminish the trouble of perusing content from pictures or pictures taken out from the live camera.

## II. LITERATURE SURVEY

Speech synthesis is artificial human voice generation that is played through loudspeakers it can be implemented in Software or Hardware products.

The soonest speech synthesis effort was in 1779 when Russian Professor Christian Kratzenstein made a device in view of the human vocal tract to show the physiological contrasts included in the creation of five long vowels (A,E,I,O,U) sounds.

The primary completely practical voice synthesizer VODER (Voice Operating Demonstrator) was made by Homer Dudley's with help of fellow Engineer Robert Riesz. A console from which an administrator could make expressions of discourse controlling a VOCODER with a keyboard and foot pedals. In any case, it was considered as troublesome in work. The VODER was shown at Bell laboratory displays at both in the 1939.

When we consider human to human interface is basic however human to machine collaboration is not all that simple. PC is gets info in Text User Interface (TUI) or in Graphical User Interface (GUI) .And it is awesome test for us to make voice as interface.

The paper concentrates on advancement of such programming that taking into account client interface in that computer will accept speech input through mouthpiece and gives speech output yield through speakers associated with PC and it portrays the model isolated into Two subsystems, specifically first is programmed Automatic Speech Recognition (ASR) and Second one Speech Understanding (SU). The objective of ASR is to decipher normal speech while SU is to understand the meaning of the transcription [16].

The handling of Analogue Speech Signal is the First step in speech recognition process. The paper [11] describes the paradigm of Signal model that is divided in four sub operation spectral shaping, spectral analysis, parametric trans-formation, and statistical modeling.First three operations are phase of digital signal processing and statistical modeling is divided in signal modeling and speech recognition system. Signal modeling systems is consist of first parameterizations is salient aspect of speech signal that is analogous to human auditory system. Second parameterizations are desired that, are robust to variations in channel, speaker, and transducer. We refer to this as the robustness, or invariance, problem. Finally the most recent parameters that captures spectral dynamics or changes of the spectrum with time are desired. We refer to this as the temporal correlation problem. With the presentation of Markov demonstrating systems that are able to do factually displaying the time course of the sign, parameters that consolidate both supreme and differential estimations of the sign range have turned out to be progressively basic.

ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
IJARSMT
Volume 1, Issue 2, July 2015

The Summarization of content from website pages is exceptionally dull assignment the paper [8] proposes a system architecture where client will highlights or select a content, then framework will discover proper articulation in speech library and give it back to speech synthesizer and loudly read it with yield gadgets.

The working of speech recognition framework utilizes a fundamental two model, oneis acoustic model and the another is language model. Fundamentally mouthpiece converts voice to analogue signal this is further prepared by sound card or sound processing unit of PC. The data taken from client is likewise called as utterance for speech application.An utterance is single word or expression that is accessible in configuration of "0" and "1". Acoustic model process the utterance by matching patterns and produce statistical representations of the sounds give it to language model,then language model generates a text for input utterance. The working of speech recognition system is addressed in paper[9].

Microsoft introduces Speech Application Programming Interface (SAPI) in 1995.This API provides the ultimate interface for dynamic speech synthesis and recognition. Microsoft released various versions of SAPI to use of speech recognition and speech synthesis within Windows. The paper [15] demonstrate the Windows application design with SAPI it uses a voice XML for recognition of the utterance given by user, Voice XML it is one kind the voice database stored in XML format which helps to add word to dictionary of speech library .

The paper [12] presents a complete Optical Character Recognition (OCR) system for camera captured image/graphics embedded textual documents for handheld devices. The paper involves various Phases of Character recognition Text Region Extraction, Skew Correction, Binarization, Text Region Segmentation, and finally Character Recognition. In Text Region Extraction partition image in m number of blocks Bi i=1, 2, 3----, n .Each Bi block is Information Block (IB) from where we want to extract text or it is either Background Block (BB). The images captured from camera have suffered from skew and perspective distortion, so in Skew Correction Phase Skew Correction Techniques are applied to the images. After the binarization the process of Text Region Segmentation is carried out to find the character depend on pixel matrices.

## III. PROBLEM DEFINITION

To build up a system that will help disable individual. For example, visual weakened persons to get data from World Wide Web.Such persons are not ready to utilize web due to their physical incapacity. By comprehension this issue there is huge need to make a system that works for these persons. They are all the more agreeable with customary method that is Braille to learning but switching to the Braille to typical PC with some key combination is another troublesome them.

The Web program introduces the Text, Audios, Videos, and Animations in pages from diverse locales, for it program can comprehend the mark up and scripts composed on pages. Executing the mark up and script in program is extremely mind boggling assignment for any one and isolating the connections, headings, and passage is likewise a troublesome occupation. The issue will emerge with the investigation of this content detachment from mark up and explore through the connections present on the page. Some of the issues which remain unaddressed after studying the related articles of existing methodologies are:

Problem I: Creating a proper interface to Visual impaired user by which they face least difficulties by fine GUI and VUI (Voice User Interface) to interact with Computer.

Problem II: Analyzing the web pages to grab the text in format of links, Headings, Paragraphs individually.

Problem III:Processing the text present on Web pages to TTS loudly in meaningful way by which user can make sense over it .

Problem IV: Extraction of text from web pages to convert it in Audio Files with proper parameters like Speed, Speaker, Volume andit will be saved in Profile Folder.

Problem V: Implementation of proper integration of Optical Character Recognition technique to the system to efficiently recognized text from images taken out from live Camera.

ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
IJARSMT
Volume 1, Issue 2, July 2015

So in view of this above weakness this paper proposes System Architecture with diverse modules to assemble such novel web browser software that will assist to Visual Impaired persons to open the colossal portal of data like ordinary persons.

## IV. PROPOSED SOLUTION

In this paper different modules are depict to actualize a novel program for Visual disabled individual .Presently a few program are accessible for Visual impeded persons yet they required a different Screen Reading Software to peruse out the Screen in which the Screen Reader system can read out the all superfluous data. But this proposed methodology will be not requires a Screen Reader Program to peruse out the Screen.The program will permit to clients to peruse out content loudly with help of TTS and it will read out just restricted data that is truly expected to the client.

The Web Pages may contain the data that is truly being not important to read out loudly like Advertisements and so forth. The program will isolate the Links, Heading, Paragraphs and other data encoded in Text Tags of (W3C) HTML .By which client will have choice to peruse that data or not.

### Text Extraction and Summarization

The Text Extraction and Summarization is important earlier stage. Since the at whatever point the client gives voice information to the framework it is handled by listener where Speech Synthesizer can change over that utterance to appropriate phrase, and this converted phrase will treated as input for request processor. The input is will be Voice command or either it is bookmarked address. When it is as address confirmed by Request processor then it is passed to internet Gateway to demand requested page, then it transferred to Text summarization and extractor algorithm to get a formatted text available in format of links, Heading, Paragraphs according to client demand TTS will read text loudly.

A request processor is an essential part in general procedure, on the grounds that at whatever point listener get the voice info. And passes it toward Request Processor (RP). This section will be decide what kind of service is requested by user. It will check its database without fail.

Figure 1 describes a module that when user request a Web Page but in case of

Voice command it will check its command database and when the command is valid then it perform fitting activity for it. The Command database is fixed and made at time of coding a project however though Bookmarked URL in URL database is adaptable. It will be enhance the later or there will be likewise a provision of to evacuate Bookmark URL and its relay upon client. This sort of provision will enhances the framework's convenience with adaptability in database. The whole database will made by client or client who having ordinary vision. User will have to choose the decisive word for voice quest to open specific URL for instance, rather than pronouns entire URL www.indiantimes.com he or she pronouns "indiatimes" just. This will easier for clients and agreeable to pronouns. This ability will permit to clients to utilize the non-lexicon words for to specific WebURL.

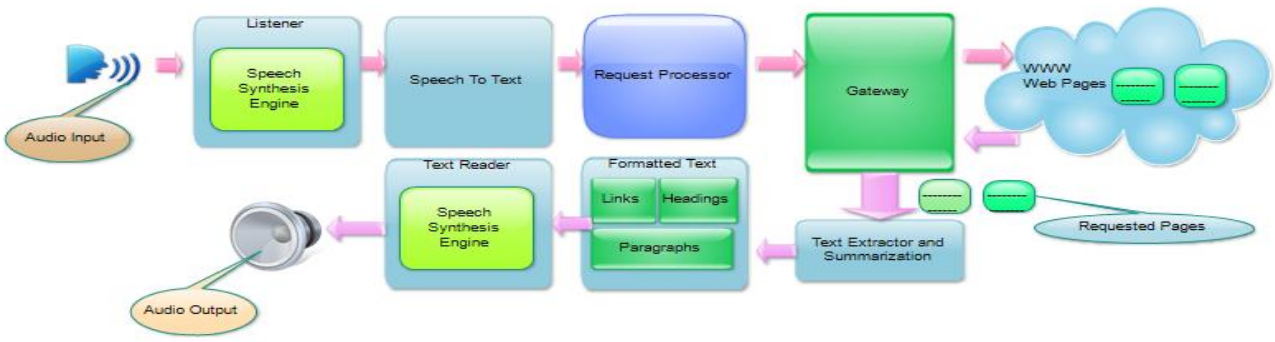


Figure 1: Proposed Module of text extraction and summarization

ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
IJARSMT
Volume 1, Issue 2, July 2015

**Voice Command Processing**

The command processing in framework is portrayed in points of interest in Figure 2 .Where how the genuine command approval and execution procedure completed is clarified. The listener will listen Voice order and offer it to Speech synthesizer then Speech synthesizer will discover fitting expression for given information after that Request processor will check it in database and when it affirm information expression is as Command then it perform its activity written in system. What's more, in the event that it is comes up short as invalid charge then it will deliver an exemption that 'Command not found 'in framework.

It is simple task than Text Extraction and Text summarization form web pages .Its job only to search command and perform the bit of code composed for it or to execute a code that are composed in the Exception.as say prior in this paper the Voice command database is settled in light of the fact that a voice command is tied with code composed at time when system is being created.

**Integration of OCR Technique**

The Optical Character Recognition method integration with program will help to client to extract the content from ink printed material like book covers, labels and etc. It just not extract content from pictures as well as will read message noisily for the client .The Figure 3 demonstrate the how OCR Plugin will help to concentrate content from pictures and in what manner will Speech synthesizer speak loudly what is written on the ink printed material from Images taken from live Camera. The Figure 3 shows a basic working structure of incorporated OCR plugin. It is similar to the utility that will encourage client to peruse out the content from ink printed material yet assume to little level just .It doesn't imply that this utility is configuration to peruse content from ink printed books .Because it is entirely hard to peruse content from books from live camera .System will skilled just to takes preview from live camera and procedure it to peruse a content.

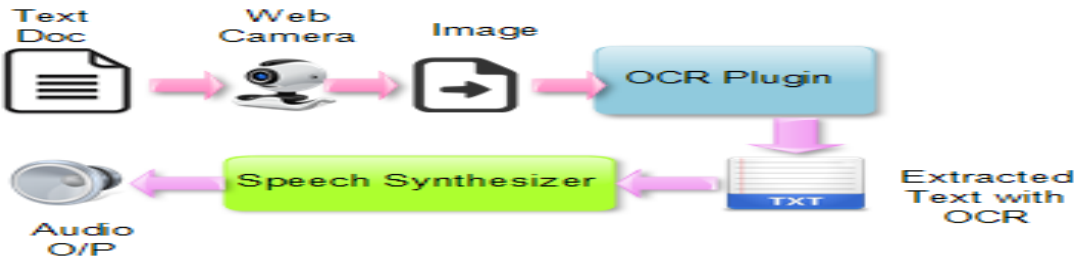


Figure 3: Working of Integrated OCR

**Text to Audio File (TAF)**

The propose framework will likewise incorporate a module appeared in Figure 4 that ready to change over the chose Text from web program to Audio File in " .WAV " design with indicated Volume, Reader and Speed in the wake of getting the best possible summon from the client .The framework TAF module will likewise permit to client to separate the content from other content document configuration like PDF, DOC and RTF and capable to convert extracted text to Audio File. This will get a little time in seconds according to measure of information. This feature gave by framework will likewise help to ordinary people groups to lessen the perusing time of substantial pages or huge content documents. We all know that hearing the information rather than reading is more pleasant.

ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
IJARSMT
Volume 1, Issue 2, July 2015

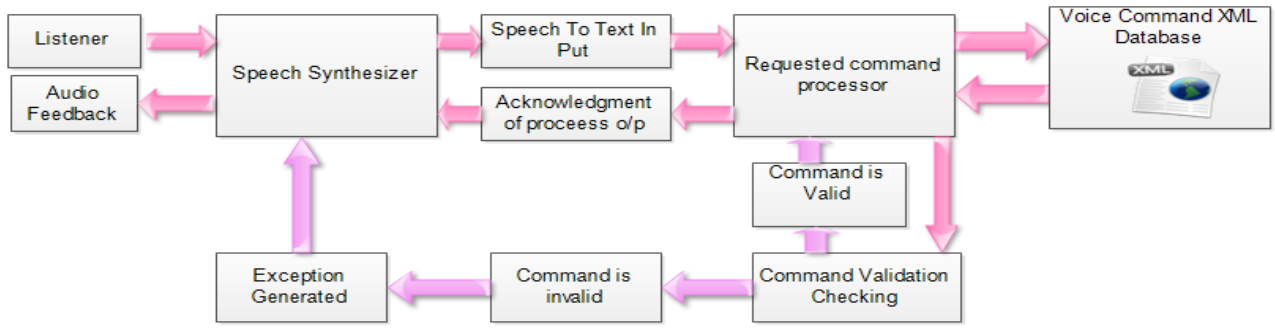


Figure 2:Voice Command Processing

By using this module the Visual impaired or normal user are able to create their own audio library by converting the selected text from browser or extracted text from other text formats. So this module will treated as more functional module in the system. Here the Figure 4 shows basic Structure of TAF.

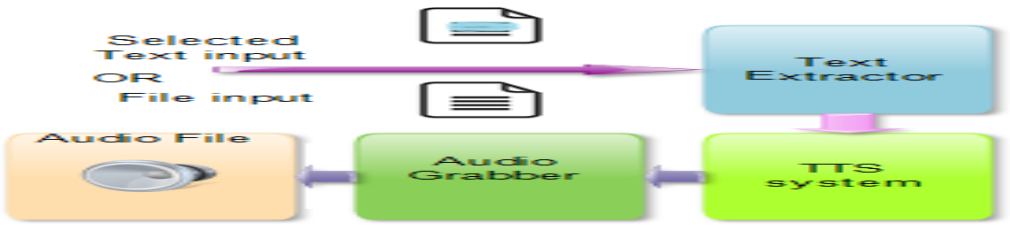


Figure 4 : Basic Structure of TAF

## V. EXPECTED RESULTS

This Proposed system will help the Visual Impaired users by reducing their difficulty of gaining knowledge from such giant source of knowledge that is Internet. The outcomes that are expected from the system is as follows

(i) The System with friendly environment
To take Voice as Input with use of some minimum key stokes to operate system.

(ii) Meaningful Presentation of Web Page
data in manageable formats such as individual Links, Heading, Paragraphs to narrate it effectively for Visual Impaired person to make a perfect sense in virtual sound space.

(iii) Accuracy in getting the Voice commands
The accuracy in getting voice command is more important issue in the system so the accuracy is maintained by use of standard rule define for to write speech grammar.

(iv) Accuracy in Character Recognition with OCR
The Accuracy OCR is always considered as Compromising because it is global issue because of distortion, improper pattern of input images.

(v) Rapid conversion Of Text File To Audio File
The conversion of selected Text available in Clipboard or The text available in Text file should be easily convert into Audio File in real time.

ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
IJARSMT
Volume 1, Issue 2, July 2015

## VI. CONCLUSION

Surfing the net is essential part of our daily need because it connects us with current World. The utilization of Internet makes our life more convenient and comfortable. But our current Browser does not satisfy the need of Visual debilitated individual to pick up information from Internet. Subsequent to concentrating on the different articles, this paper propose such novels Software that will help to peruse the Internet and also to help to gain a most extreme data from limitless wellspring of information that is Internet. It will give a complete help to such individuals by the utilization of different utilities that will furnished alongside the product. This paper discusses the various modules in proposed Model of Software by studying a challenges faced by Visual impaired or Low Vision persons by existing facilities available for them.

## REFERENCES

[1] Patrick Roth, Lori Petrucci1, André Assimacopoulos and Thierry Pun1 "AB – Web : Active audio browser for visually impaired and blind users" ICAD'98
[2] Sachin Kapoor, Kanika Mittal and Adit Gupta "Hands Free Browser-„An Interactive Speech Browser for Visually Handicapped‟ International Journal of Advanced Computer ResearchVolume-4 Number-1 Issue-14 March-2014.
[3] L.Santhoshi, K. Mahesh, Dr.K.Kuppusamy "Web Browsing For Visually Impaired People Through Handheld Devices "International Journal of Engineering Research and Applications (IJERA) Vol. 2, Issue 3, May-Jun 2012, pp.3024-3028
[4] KwangMyungJeon and Seung Ho Choi "HMM-Based Distributed Text-to-Speech Synthesis Incorporating Speaker-Adaptive Training" International Journal of Multimedia and Ubiquitous Engineering Vol.9, No.5 (2014),
[5] AntónioRamiresFernandes and Jorge Ribeiro Pereira and José Creissac Campos "ACCESSIBILITY AND VISUALLY IMPAIRED USERS" ICEIS 2004 - HUMAN-COMPUTER INTERACTION.
[6] AnandArokiaRaj ,TanujaSarkar,SatishChandraPammi1, SanthoshYuvaraj MohitBansal,KishorePrahallad1 and AlanWBlack "TextProcessingforText-to-Speech Systemsin Indian Languages" 6 th ISCA Workshop on Speech Synthesis,Bonn,Germany,August 22-24,2007.
[7] Ummuhanysifa U and NizarBanu P K "Voice Based Search Engine And Web Page Reader" International Journal of Computational Engineering Research (IJCER) May, 2013.
[8] AbhishekKawle, AmanOjha, RahatKazi and SeemaYadav "Text to Speech Web Plugin with Text Summarization " Proceedings of National Conference on New Horizons in IT - NCNHIT 2013.
[9] Parwinder pal Singh and Er. Bhupindersingh "Speech Recognition as Emerging Revolutionary Technology" International Journal of Advanced Research in Computer Science and Software Engineering Volume 2, Issue 10, October 2012.
[10] Suma Swamy1 and K.V Ramakrishnan "ANEFFICIENT SPEECH RECOGNITION SYSTEM" Computer Science & Engineering: An International Journal (CSEIJ), Vol. 3, No. 4, August 2013.
[11] JOSEPH W. PICONE "Signal Modeling Techniques in Speech Recognition" PROCEEDINGS OF THE IEEE. VOL. 81, NO. 9, SEPTEMBER 1993.
[12] AyatullahFarukMollah, Nabamita Majumder, SubhadipBasu and MitaNasipuri "Design of an Optical Character Recognition System for Camera- based Handheld Devices" IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 1, July 2011
[13] RavinaMithe, SupriyaIndalkar and NilamDivekar "Optical Character Recognition" International Journal of Recent Technology and Engineering (IJRTE) Volume-2, Issue-1, March 2013
[14] Adam Coates, Blake Carpenter, Carl Case, SanjeevSatheesh, Bipin Suresh, Tao Wang, David J. Wu and Andrew Y. Ng "Text Detectionand Character Recognitionin SceneImages with Unsupervised Feature Learning".
[15] Hao Shi and Alexander Maier "Speech-enabled windows application using Microsoft SAPI" IJCSNS International Journal of Computer Science and Network Security, VOL.6 No.9A, September 2006
[16] UrmilaShrawankar and Anjali Mahajan "Speech: A Challenge to Digital Signal Processing Technology for Human-to-Computer Interaction ".