A REVIEW ON VOICE ASSISTANT FOR BLIND AND PARTIALLY SIGHTED PEOPLE

Dr.S.Rajeswari[1], R.Srimathi[2]

[1]Assistant Professor, PG Department of Computer Science,

SDNB Vaishnav College for Women, Chrompet, Chennai-600 044,

[2]Student, PG Department of Computer Science,

SDNB Vaishnav College for Women, Chrompet, Chennai-600 044,

Email:vrajee2008@gmail.com[1],p21cs022@sdnbvc.edu.in[2]

ABSTRACT

Visual debilitations individuals are 31% more averse to get to the web than people without incapacities. Those individuals can't ready to peruse text straightforwardly from the book without the utilization of Braille framework. They need assistance from third individual. To defeat the trouble, utilizing advances like NLP and text-to-speech (TTS) framework can be created which helps the outwardly weakened pay attention to a sound read-back of any text from image This audit paper represents about the execution of programming that give voice help to the outwardly debilitated for getting to the web.

KEYWORDS: Natural language processing, Text to speech, Image Processing

I. INTRODUCTION:

Outwardly impeded individuals deal with a ton of issue in their regular routine and as opposedto being trustworthy on someone else, so the advances like NLP and Text-to-speech which canassist with helping the outwardly hindered individual. Voice help have a long history that goesback more than 100 years, which is amazing since Alexa, the first that we know about , was just delivered decade prior and presently we can see computer based intelligence ,AI at work consistently as voice partners. The text-to-speech works virtually every individual advanced gadget, including PCs, cell phones and tablets. A few TTS devices likewise have an innovationcalled optical person acknowledgment (OCR) .This OCR permits TTS instruments to peruse text resoundingly from pictures. Furthermore, NLP assumes a significant part in this voice help,which is the driving innovation that permits machines to comprehend and communicate with human discourse, however isn't restricted to voice connections. NLP joins computational etymology with factual, AI and profound learning models. So in this paper we shows

70

about the voice help for outwardly weakened individuals in view of NLP and Text-to-speech advancements.

TEXT-TO-SPEECH CONVERSION METHODOLOGY

TTS is a cycle where incoming text by the client is examined first, after at that point, it handled and comprehended, the last text of the specific client is switched over completely to computerized sound and afterward spoken

- Text Processing: The text is examined first, handled and standardized (It will deal with certain abbreviations and shortened forms and similar texts are matched) then translated into phonetic portrayal.

- Speech Synthesis: A portion of the discourse blend procedures of the text are

i) Articulator Synthesis:

Involves instinctive and audio model as discourse age. It creates understandable manufactured discourse however it is nowhere near regular sound and consequently not generally utilized.

ii) Formant Synthesis:

Thereby, portrayal of single discourse sections also put away on a parametric premise. There are two essential designs in formant union, equal and fountain, however for superior execution, some sort of mix of those two designs is utilized. A fountain synthesizer-formant comprises of band-pass filter associated in this series. The result of resonator-formant is registered to the contribution of the progressive one.

iii) Concatenate Synthesis:

Here strategy blends sound which is linking is the short examples of sound are units. It is utilized in discourse combination to produce client explicit succession of sound from a data set worked from the recording of different grouping

II. LANGUAGE TRANSLATION:

In India, we have different dialects communicated. The 2001 Evaluation audiotaped 30 dialects which was communicated by in excess of 1,000,000 local presenters and hundred and twenty two speakers which was communicated by in excess of 10,000 individuals, this is the reason it will exceptionally important to possess some applications, cycles that switch message starting with one language over completely then onto the next, protecting the sacredness of the text. Machine Interpretation is a Man-made consciousness also a Normal Language Handling that manages interpretation starting with single

language then into the next utilizing machine interpretation framework.

Rule Based Machine Translation (RBMT):

Interpretation exists produced based on semantic, syntactic and morphological investigation of the two source and the objective dialects. A framework comprise of assortment of rules: Punctuation controls-fundamentally comprise of examination about dialects as far as language structures (grammar, semantic, morphology, grammatical form labeling and orthographic highlights)

Statistical machine translation (SMT):

This is portrayed by utilization of AI strategies. And this is also an information motivated approach which utilizations equal adjusted corpus and treats interpretation as one numerical thinking issue. From that, each line in the objective languages are an interpretation along likelihood from the source language.

Example based machine translation (EBMT):

EBMT depends on possibility of similarity. From this methodology, the particular corpus which is utilized one, consists texts that have previously been deciphered. From the specified sentence that will be interpreted, particular sentences from this are chosen that contain comparable sub-sentential parts. The comparative sentences are then used to interpret the sentential parts of the first line into the objective language, and those expressions are assembled to shape a total interpretation. This Relationship interpretation utilizes different phases; coordinating, adaption, also recombination

Matching

The incoming text is divided, trailed via look for models from information base which intently matching the info of SL piece characters and the important parts are chosen. It includes TL parts comparing into significant sections are removed.

Adaption

Assuming the similar matches are definite, the parts are combined to shape TL yield, otherwise find TL piece of the applicable similar match compare to explicit piece and adjust them.

Recombination

Blend of pertinent TL sections to frame legitimate linguistic objective text.

III. LITERATURE REVIEW

Yi-Ching Huang, Lung-Chuan Liao [1] proposed a project to investigate the impacts which includes the advanced material integrated into Text-to-speech framework for understudies English

vocabularies. A computerized detail was made based on the vocabulary bee jargon lists (around three hundred words) gave by chosen school. Twenty one third graders from a confidential school in the Taiwan were chosen for mentioned paper. The review utilized four information assortment strategies, includes poll, pre- testing and post-testing, casual perception and interviews, semi-organized individual meetings.

Ahmed Tammaa et al [2] had projected an idea of visual impedance and visual deficiencies are viewed as one of the most difficult openness spaces for PC designers. Text-to-speech (TTS) is viewed as one of the arrangements that can assist on making with bettering availability for PCs for everybody. Notwithstanding, the TTS sounds unnatural and awkward for the clients. In certain dialects, for example, Arabic, the elocution is off-base. Subsequently, Generative skilled disposed Organizations can naturalize and further develop the result through High Loyalty GAN (HIFI-GAN)

Jeisson A. Rodríguez Bonces [3] completed a gathering of 10 rudimentary degree understudies from expansion streams in a confidential college in Bogota and Colombia. The principal aim of particular study was to decide the impact of (TTS) innovation utilization on perusing familiarity with authors of English like an unknown dialect. The outcomes display that students foster vocal perusing familiarity with regions, for example, connecting sounds, articulating precisely, and understanding timing.

Ayushi Trivedi,Navya Pant Et.al [4] have proposed a system that siri a shrewd robotized colleague executed on an electronic gadget, to work with client connection with a gadget, and to help the client all the more really draw in with neighborhood or potentially remote administrations utilizes Subtlety Correspondences voice acknowledgment and text-to discourse (TTS) innovation. In this paper, the author investigates the various kinds of discourse, discourse acknowledgment, discourse to message change, message to discourse transformation and discourse interpretation.

Et.al Nwakanma Ifeanyi [5] implemented a voice/speech combination is a field of software engineering that arrangements with planning PC frameworks that incorporate composed text. An innovation permits a PC to change over a text into speech through a receiver or phone. As the new age of registering innovation, it comes as the following significant development in man-machine collaboration, after usefulness of Discourse acknowledgment (TTS), supporting Intelligent Voice Reaction (IVR) frameworks.

Et.al Itunuoluwa Isewon [6] introduced a TTS synthesizer is one developed application that believers input text into expressed word and by investigating also handling the input text utilizing Normal Language Handling (NLP) and afterward utilizing Computerized Signal Handling (DSP) innovation to change over this handled input text into integrated discourse portrayal of the particular

International Journal of Multidisciplinary Research in Arts, Science & Commerce
(IJMRASC) ISSN Online: 2583-018X

Vol. 3(1), January 2023, pp. 70 - 77

input text, Nowwe fostered a valuable TTS synthesizer in the type of a basic application that changes over inputted stored text into integrated discourse and peruses to the client which can save as mp3 file.

S. Venkateswarlu et al [7] have proposed to Optical character recognition (OCR) is an interaction that translates filtered or copied text including images, manually written input text into modified text for additional handling. The author has introduced a vigorous methodology for input text extrication and switching it over completely to discourse. Verifying of gadget was finished on raspberry-pi stage.Raspyis at first associated with the web through VLAN.

Table 1 Review of Various Techniques for Speech to Text Conversion

| AUTHOR &TITLE	METHODOLOGY	FUTURE PERSPECTIVE	DATASET	ACCURACY
Yi-Ching Huang, LungChuan Liao, A Study of Text To Speech(TTS) In Children's English Learning	Self-directed learning, data collection techniques	Interactive teaching and learning activities	Questionnaire, pre-test and post-test	83.3%
Ahmed Tammaa, Ahmed Ramy, Neural Text-ToSpeech Synthesis Literature Review	Natural Language Processing, Deep Learning	Changing the engineering of the TacoTron2	Audio files, Total of 13,100short audio clipfor the same person.	78%
Jeisson A. Rodríguez Bonces, Implementing Text-To-Speech Technology As A Means Of Enhancing L2 Reading Fluency	Information and Communication, Oral reading fluency	For those teachers who love designing New	Field notes, voice audio recordings.	80.93%

74

Ayushi Trivedi,Navya Pant Et. al, Speech to text and text to speech recognition systems-Areview	Dynamic time warp, Neural Network models	Output is far from naturalvoice.	Text image	56.87%
Et.al Nwakanma Ifeanyi, Oluigbo Ikenna, TexttoSpeech Synthesis	Object Oriented Analysis and Development Methodology	By using of Information	Images of plain text	85.98%
Et.al Itunuoluwa Isewon, Jelili Oyelade, Design, and implementationof Text to Speech Conversion for Visually Impaired People	Analysing and processing of the text is done	To make motors for confined Nigerian	Containing symbols like	86%
S. Venkateswarlu, D. B. K. Kamesh, J. K. R. Sastry,and Radhika Rani, Text to SpeechConversion	Optical Character Recognition(OCR) and Text to Speech Synthesizer (TTS)	This convenient gadget, doesn'tneed web association	Collection of texts from colorimages	82.22%

IV. ANALYSIS OF TECHNIQUES

Based on the above review articles various techniques were taken over for the speech to text conversion is identified. These techniques are analyzed based on the accuracy obtained during the training and testing of datasets. It is observed that the techniques such as Optical Character Recognition (OCR), Natural Language Processing , Deep learning, Neural Networks, Object Oriented Analysis and Development produce more than 80% of accuracy for analyzing the concerned text and also conversion of speech in their respective languages. Therefore the proposed system is composed of these techniques for developing the voice based text to speech conversion in Tamil language.

V. PROPOSED SYSTEM

The TTS framework gets the text as the info and afterward a PC calculation which called TTS motor examinations the text, pre-processes the text and combines the discourse for certain numerical models.
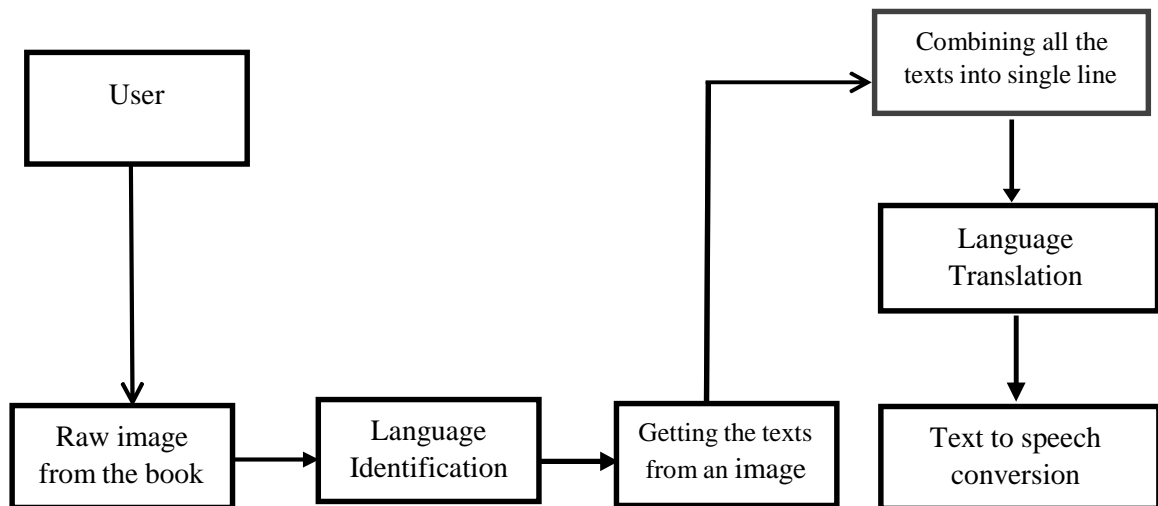
Fig.1 Architecture of TTS System

The raw image is captured by the user and getting the texts presented in the image then the texts are divided into several parts by using Natural Language Processing, After that, all the texts are combined in a correct way, and then by using google API the given text can be converted into speech.

VI. CONCLUSION

TTS combination remains quickly developing part of PC innovation and is progressively assuming more significant part in the manner of communication with framework and connection   points over various stages. In future many new designs are implemented for an attempt to the area of execution of text-to-discourse framework on different stages. For example, communication framework ATMs, computer games whatever other stages where TTS innovation could be an additional benefit and increment usefulness.

REFERENCES

[1] Yi-Ching Huang and Lung-Chuan Liao, "A Study Of Text-To-Speech (TTS) In Children's English Learning", *Teaching English with Technology*, 15(1), 14-30 (2015).

[2] Ahmed Tammaa, Ahmed Ramy, "Neural Text-To-Speech Synthesis Literature Review", *Natural Language Processing Phase 1* ©2022 IEEE (2022).

[3] Jeisson A. Rodríguez Bonces, "Implementing Text-To-Speech Technology As A Means Of Enhancing L2 Reading Fluency", *The International Education and Learning Review* (2022).

[4] Ayushi Trivedi, Navya Pant, Pinal Shah, Simran Sonik, and Supriya Agrawal," Speech to text and text to speech recognition systems-Areview", *IOSR Journal of Computer Engineering*, Volume 20, Issue 2, Ver.I (2018).

[5] Nwakanma Ifeanyi, Oluigbo Ikenna and Okpala Izunna, "Text-To-Speech Synthesis(TTS)", *IJRIT International Journal of Research in Information Technology,* Volume 2, Issue 5 (2014).

[6] Itunuoluwa Isewon, Jelili Oyelade, Olufunke Oladipupo," Design and Implementation of Text To Speech Conversion for Visually Impaired People" *International Journal of Applied Information Systems* 7(2):25-30 (2014).

[7] S. Venkateswarlu1, D. B. K. Kamesh1, J. K. R. Sastry and Radhika Rani," Text to Speech Conversion", *Indian Journal of Science and Technology*, Vol 9(38)(2017).