IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, January 2023
Impact Factor: 7.301

A Recent Survey Paper on Text-To-Speech Systems

Shruti Mankar[1], Nikita Khairnar[2], Mrunali Pandav[3], Hitesh Kotecha[4],Manjiri Ranjanikar[5]
Pimpri Chinchwad College of Engineering, Nigdi, Pune, Maharashtra, India[1,2,3,4,5]

Abstract: *All of us are aware of how important knowledge is and it is also true that mostly the data or knowledge is in the form of books or online articles, or various pdfs i.e. in text format. But not all of us are privileged to read. Some are illiterate, some are blind, and some have reading difficulties. Hence, a form of adaptive technology or procedure that reads digital text aloud, which is called text-to-speech (TTS) was developed. It is occasionally referred to as "read-aloud" technology. Words on a computer or other digital device can be converted into audio using TTS. TTS is particularly beneficial for all those people who have reading difficulties or are illiterate. Also, a significant amount of research has been done and is currently being done on text-to-speech technology. Various technologies, methodologies, and algorithms are used in the various proposed approaches and solutions for TTS. This research presents a systematic review of all those methods which have been proposed and implemented by different active researchers in this field.*

Keywords: Text-to-speech conversion, text-to-speech synthesis, machine learning, neural networks, optical character recognition, etc

REFERENCES

[1]. Patil Mrunmayee and Ramesh Kagalkar, "A review on conversion of image to text as well as speech using edge detection and image segmentation", International Journal of Advanced Research in Computer Science Management Studies 2, 2014
[2]. Isewon, Itunuoluwa, Jelili, Oyelade, and OlunfunkeOladipupo, “Design and implementation of text to speech conversion for visually impaired people”, International Journal of Applied Information Systems 7, no. 2, pp. 25-30, 2014
[3]. Venkateswarlu, S., D.B.K. Kamesh, J.K.R. Sastry, and Radhika Rani, “Text to speech conversion”, Indian Journal Of Science and Technology 9, no. 38, pp. 1-3, 2014
[4]. Ma, Shuang, Daniel McDuff, and Yale Song, “Unpaired image-to-speech synthesis with multimodal information bottleneck” In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 7598-7607, 2019
[5]. Tae-Ho Kim, Sungjae Cho, Shinkook Choi, Sejik Park? andSoo-Young Lee ."Emotional Voice Conversion Using Multitask Learning with Text-To-Speech" ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)(2019).
[6]. Cong Zhou, Michael Horgan, Vivek Kumar, Cristina Vasco, Dan Darcy. "Voice Conversion with Conditional SampleRNN" Interspeech 2018, Hyderabad, India.
[7]. Kuan Chen, Bo Chen, Jiahao Lai, Kai Yu. " High-quality Voice Conversion Using Spectrogram-Based WaveNetVocoder" Interspeech, 2018.
[8]. Nagdewani, Shivangi, and Ashika Jain. "A Review On Methods For Speech-To-Text And Text-To-Speech Conversion " International Research Journal of Engineering and Technology, (2020).
[9]. Donahue, J., Dieleman, S., Bińkowski, M., Elsen, E., &Simonyan, K. (2020). End-to-End Adversarial Text-to-Speech.arXiv. https://doi.org/10.48550/arXiv.2006.03575
[10]. T. -H. Kim, S. Cho, S. Choi, S. Park and S. -Y. Lee, "Emotional Voice Conversion Using Multitask Learning with Text-To-Speech," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 7774-7778, doi:10.1109/ICASSP40776.2020.9053255.
[11]. Kuan Chen, Bo Chen, Jiahao Lai, Kai Yu, "High-quality Voice Conversion Using Spectrogram-Based WaveNetVocoder," Interspeech 2018- isca-speech.org.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, January 2023
Impact Factor: 7.301

[12]. MingyangZhang,Xin Wang2, Fuming Fang2, Haizhou Li1, Junichi Yamagishi2, "Joint training framework for text-to-speech and voice conversion using multi-source Tacotron and WaveNet," April 2019, doi: https://arxiv.org/abs/1903.12389.

[13]. Tae-Ho Kim, Sungjae Cho, Shinkook Choi, Sejik Park and Soo-Young Lee, "Emotional Voice Conversion Using Multi task Learning with Text to Speech " , November 2019, arXiv:1911.06149v2.

[14]. Cong Zhou, Michael Horgan, Vivek Kumar, Cristina Vasco, Dan Darcy, "Voice Conversion with Conditional Sample RNN", Interspeech 2018, Hyderabad, India. https://doi.org/10.48550/arXiv.1808.08311.

[15]. Łańcucki, Adrian, "Fastpitch: Parallel text-to-speech with pitch prediction" In ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 6588-6592, IEEE, 2021

[16]. A. Kain and M.W. Macon, "Spectral voice conversion for text-to-speech synthesis", Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat.No. 98CH36181), 1998, pp. 285-288 vol. I, doi: 10.1109/ICASSP.1998.674423