AUDIOJOT: EXPRESSIVE SPEECH SYNTHEIS FOR TAMIL

Mithushan Chrisjayan Jalangan

A Dissertation Submitted in Partial Fulfilment of the Requirement for
Bachelor of Engineering (Honours) Degree in Software Engineering

Department of Computing
Informatic Institute of Technology, Sri Lanka
in Collaboration with
University of Westminster, UK

2020

Abstract

Speech synthesis or commonly known as text-to-speech system converts text input into speech output that is production of speech or voice in an artificial manner. Speech Synthesis has been the crucial games changer for such form of computer to human communication.

Speech is the most effective way of communication amongst human beings. Speech has fundamentally impacted on the evolution of human civilization based on the exchange of idea.
The goal of expressive speech synthesis is to produce speech that is able to express various kinds emotions, speaking styles and attitudes.

This research proposed a solution for generated expressive speech synthesis using neural text to speech approach for Tamil, which is considered as a low resource language. The solution is a web-based platform used to create audio content that can be distributed across platforms opening ways to democratise the consumption of the audio content in Tamil.

The research open pathways to different other application of the neural speech based such speech to speech translation. The system also implements such an application based on a cascading flow.

Keywords – Speech Synthesis, Text to Speech, DCTTS, Tamil, Indian Languages, Convolutional Neural Networks