

Research Paper :

Speech synthesis

KOMAL SINGH AND GEETA KAUSHIK

Accepted : December, 2009

ABSTRACT

Speech synthesis systems are often called text-to-speech (TTS) systems in reference to their ability to convert text into speech. However, systems exist that instead render speech synthesis to artificial production of human speech. A system used for this purpose is termed as a speech synthesizer and can be implemented in software or hardware. However, systems exist that instead render symbolic linguistic representations like phonetic transcriptions into speech. A text-to-speech system is composed of two parts: a front-end and a back-end. Broadly, the front-end takes input in the form of text and outputs a symbolic linguistic representation. The back-end takes the symbolic linguistic representation as input and outputs the synthesized speech waveform. TTS software can “read” text from a document, Web page or e-Book, generating synthesized speech through a computer’s speakers. TTS can also convert text files into audio MP3 files that can then be transferred to a portable MP3 player or CD-ROM. This can save time by allowing the user to listen to reports or background materials while performing other tasks. TTS makes a critical difference to those with disabilities such as poor vision or visual dyslexia. People with speech loss can utilize specialized TTS programs to turn typed words into vocalization. TTS programs provide a valuable edge, particularly for learning new languages. This thesis aims to study the speech synthesis technology and to develop a cost effective, user friendly text to speech conversion system using Laboratory virtual instruments engineering workbench (LabVIEW) graphical programming language

See end of the article for authors’ affiliations

Correspondence to:

GEETA KAUSHIK

Maharishi Markandeshwar
University, Mullana,
AMBALA (HARYANA)
INDIA

Key words : Text-to-speech (TTS), Speech Synthesis

The aim of the speech synthesis is producing the human speech artificially either in software or hardware. The natural language text is converted into speech by text-to-speech (TTS) systems. These systems have been widely used as assistive technological tools for a long time. The pre-school kids, the people who have visual impairments or reading disabilities and the ones who suffer from severe speech impairment can get benefit from these systems. The news web sites that convert written news to audio content, entertainment productions such as games, cartoons, mobile tools, preparation of audio supplementary materials in various fields, automated question-answering systems, attaining certain information (price list, the weather forecasting report, etc.) and vocalizing e-mail, fax, sms, and daily journals for handicapped ones are only a limited number of items that can be listed as the typical application areas of TTS. The aim of the TTS is that the system converts all digital texts and printed texts into speech automatically. Commercial and non-commercial systems have been continuously developed and recent advances are promising for future applications.

Objectives:

Speech synthesis is the artificial production of human

speech through the use of computer.

A very large set of different rules and their exceptions is needed to produce correct pronunciation and prosody for synthesized speech. The main objective of this report is to:

- Study the speech synthesis technology,
- Develop text to speech module using LabVIEW software

Synthesized speech can be produced by several different methods. All of these have some benefits and deficiencies. The methods are usually classified into three groups:

- Articulatory synthesis, which attempts to model the human speech production system directly.
- Formant synthesis, which models the pole frequencies of speech signal or transfer function of vocal tract based on source-filter-model.
- Concatenative synthesis, which uses different length prerecorded samples derived from natural speech.

The formant and concatenative methods are the most commonly used in present synthesis systems. The formant synthesis was dominant for long time, but today the concatenative method is becoming more and more popular.