



Speech and Natural Language Processing and Synthesis AI Software

Background

The global voice recognition and synthesis software market is forecasted to reach 31.82 billion USD by 2025, at an annual compound growth rate of 17.2%. This market is influenced by the rapid adoption of artificial intelligence and Internet of Things. Particularly, intelligent personal assistants on smartphones and the increase in the commercial value of voice user interfaces in smartphone design, architecture and manufacturing.

Companies across the telecommunication, automotive, banking, healthcare, and military industries have growing interest in speech and natural language processing and synthesis software. Voice activated systems including voice-enabled devices and virtual assistants are readily adopted in the automotive industry and call centres. These technologies have become increasingly valuable in countering fraudulent activities and enhancing security in the banking industry through the adoption of voice biometrics for user authentication. Voice recognition and synthesis is also highly valuable in the healthcare sector as it enables efficiencies and cost savings in patient interaction and clinical documentation.

Description of the invention

Waterloo's Speech Processing and Natural Language Understanding AI Software offers a novel way of replicating digital sound files as a smooth and continuously (i.e. "analog") function in time and sound frequency.

Current competitive tools synthesize speech by taking pre-recorded phonetic fragments and words and then reordering them into the desired sentence generated by AI software responding to human queries. The resulting product lacks the phrasing of natural human speech thereby degrading the emotive connection during human interaction.

Waterloo's software addresses all of these issues by adjusting and transforming the functional representation of human generated phonetics so that they can be used to generate words and sentences in the manner of natural human speech. This technology can be used in many applications.

Advantages

- Near-lossless sound compression.
- Minimize amount of computer memory needed to store sound files.
- Replicates natural human speech.

Potential applications

- Speech synthesis, speech processing and voice recognition in the banking, automotive, telecommunications, healthcare, military and manufacturing industries.

Reference

10163

Patent status

Patent pending

Stage of development

Prototype developed, looking for industry partners to validate results.

Contact

Scott Inwood
Director of Commercialization
Waterloo Commercialization Office
519-888-4567, ext. 33728
sinwood@uwaterloo.ca
uwaterloo.ca/research