Method for pitch-synchronous speech parameterization with applications

Technology #cu16133

Speech parameterization is used in speech recognition and synthesis to convert between audible speech and a manipulatable digital representation. However, current methods to parametrize voice signals can be corrupted by variations in pitch, leading to inaccurate speech recognition and robotic-sounding speech generation. This technology is a speech parameterization method that is able to completely separate and accurately quantify unique elements of human speech such as pitch and timbre, enabling more accurate speech recognition and more realistic-sounding speech generation.

Pitch-synchronous parameterization allows for more accurate modeling of human speech

In order to establish a link between audible speech and written text as well as speech coding, speech parameterization is performed to convert voice signals into a parametric representation. Most current commercial systems use pitch-asynchronous methods such as mel-frequency cepstral coefficients and linear predictive coding, which suffer from low accuracy and may lead to inaccurate speech recognition or robotic-sounding voice transformation and speech synthesis. In contrast, this technology enables pitch-synchronous parameterization using timbre vectors, which serve to decouple the pitch and the timbre of speech signals. This method of speech parameterization is thus sensitive to variations in tone and prosody, and better models natural human speech. Pairing this technology with vector quantization and scalar quantization will lead to more accurate speech coding and allow for generation of CD-quality audio from low-bite-rate coded signals.

Lead Inventor:

C. Julian Chen, Ph.D.
Applications:

• Speech recognition (speech-to-text)
• Speech synthesis (text-to-speech)
• Voice transformation
• Speech coding

Advantages:

• More accurate modeling of human speech for speech recognition and speech synthesis
• Can be used with tonal languages (e.g. Mandarin Chinese)
• Compatible with any speech database used in traditional parameterization
• Can achieve higher voice quality than traditional speech coding methods at same bit-rate

Patent Information:

Patent Issued (US8,744,854)
Patent Issued (US8,719,030)
Patent Issued (US8,886,539)
Patent Issued (US8,942,977)
Patent Issued (US9,135,923)

Tech Ventures Reference: IR CU16133

Related Publications:


Inventors

Julian Chengjun Chen