Mixed-Excitation Linear Predictive (MELP) Vocoder Algorithm

VOCAL Technologies, Ltd. software libraries include a complete range of ETSI / ITU / IEEE compliant and other standard and proprietary algorithms, optimized for execution on ANSI C and leading DSP architectures (ADI, AMD, ARM, DSP Group, LSI Logic ZSP, MIPS and TI).

The MELP (Mixed-Excitation Linear Predictive) Vocoder Algorithm is the new 2400 bps Federal Standard speech coder. The United States Department of Defense (DoD) Digital Voice Processing Consortium (DDVPC) selected it after a multi-year extensive testing program.

The selection test concentrated on four areas: intelligibility, voice quality, talker recognizability, and communicability. The selection criteria also included hardware parameters such as processing power, memory usage, and delay. MELP was selected as the best of the seven candidates and even beat the FS1016 4800 bps vocoder, a vocoder with twice the bit-rate.

MELP is robust in difficult background noise environments such as those frequently encountered in commercial and military communication systems. It is very efficient in its computational requirements. This translates into relatively low power consumption, an important consideration for portable systems.

Terminology:

- Traditional pitched-excited LPC vocoders use either a periodic pulse train or white noise as the excitation for an all-pole synthesis filter. These vocoders produce intelligible speech at very low bit rates, but they sometimes sound mechanical or buzzy and are prone to annoying thumps and tonal noises.
- These problems arise from the inability of a simple pulse train to reproduce all kinds of voiced speech. The MELP Vocoder uses a mixed-excitation model that can produce more natural sounding speech because it can represent a richer ensemble of possible speech characteristics.

Features:

- The MELP Vocoder is based on the traditional LPC parametric model, but also includes four additional features. These are mixed-excitation, aperiodic pulses, pulse dispersion, and adaptive spectral enhancement.
  - The mixed-excitation is implemented using a multi-band mixing model. This model can simulate frequency dependent voicing strength using a novel adaptive filtering structure based on a fixed filterbank. The primary effect of this multi-band mixed-excitation is to reduce the buzz usually associated with LPC vocoders, especially in broadband acoustic noise.
  - When the input speech is voiced, the MELP vocoder can synthesize speech using either periodic or aperiodic pulses. Aperiodic pulses are most often used during transition regions between voiced and unvoiced segments of the speech signal. This feature allows the synthesizer to reproduce erratic glottal pulses without introducing tonal noises.
  - The pulse dispersion is implemented using fixed pulse dispersion filter based on a spectrally flattened triangle pulse. This filter has the effect of spreading the excitation energy with a pitch period. This, in turn, reduces the harsh quality of the synthetic speech.
  - The adaptive spectral enhancement filter is based on the poles of the LPC vocal tract filter and is used to enhance the formant structure in the synthetic speech. This filter improves the match between synthetic and natural bandpass waveforms, and introduces a more natural quality to the speech output.

VOCAL Technologies, Ltd.

© 2003 VOCAL Technologies, Ltd.

Custom Product Design Division
200 John James Audubon Parkway
Buffalo, New York 14228
716-688-4675

MELP-01

http://www.vocal.com