D33 Crosscorrelation-Based Pulse Suppression for Forensic Audio Analysis

Robert Dunn, MS, and Thomas Quatieri, PhD, MIT Lincoln Laboratory, 244 Wood Street, Lexington, MA 02420; and Reva Schwartz, MA, and Michael Piper, BA, United States Secret Service, Forensic Services Division, 950 H Street NW Suite 4200, Washington, DC 20223*

After attending this presentation, attendees will learn about a new method for improving audio, specifically how to increase the clarity of audio signals affected by a class of interference consisting of repeated pulses.

This presentation will impact the forensic community and/or humanity by informing the community about forensic audio analysis and different methods for improving the clarity of speech in law enforcement recordings.

A goal of forensic audio enhancement is to combat additive background interferences in a desired signal to assist in forensic analysis and presentation in legal cases. In this presentation, the authors introduce a new software-based approach that runs on a personal computer to enhance speech. This software increases the clarity of audio signals affected by a class of interference consisting of repeated pulses. Algorithm is based on pulse detection through crosscorrelation with a prototype pulse followed by pulse scaling or subtraction.

Surveillance recordings in law enforcement are typically made using a variety of equipment. The proliferation of cell phones has carried over into the law enforcement arena and subsequently a large percentage of surveillance recordings involve cell phones. Interference between recording equipment and the transmitters in cell phones creates noise. This noise is often heard as a pulsing similar to that from a lawn sprinkler and thus is often referred to as "sprinkler pulses." This pulsing gets recorded along with the speech and seriously diminishes speech intelligibility and increases listener fatigue. Removing this particular class of pulses without damaging the speech is the subject of this presentation.

Most current algorithms for removing impulse-type noise require very short duration pulses, typically a few milliseconds, are susceptible to random wideband noise, and result in holes in the time waveform. Standard algorithms also require that the pulses be largely deterministic in nature. The authors have developed a technique to detect and suppress sprinkler pulses, which are, on the other hand, typically relatively long in duration and contain a random component. These pulses have a characteristic signature, being low frequency and deterministic at the beginning and end of the pulse but wideband and random in their mid-region. The low frequency component at pulse edges is consistent enough from one pulse to the next so that a matched filter can be used as a detector. The authors' new approach allows sprinkler pulse durations on the order of 20 ms, is robust, and avoids temporal holes. The method also provides the audio analyst with easily adjustable parameters for a detection threshold, suppression pulse duration, and suppression level.

The steps in the pulse suppression algorithm are as follows:

1. Select initial prototype pulse
2. Crosscorrelate and detect pulses
3. Refine prototype pulse by averaging detected pulses
4. Crosscorrelate and attenuate pulse over pulse duration
5. Adjust detection threshold interactively

Pulses that change over time are also allowed by performing a timevarying average in Step 3. The most effective means of removing the pulses is to apply pulse scaling after detection. First, pulses are detected and then suppressed. Therefore, suppression occurs only during pulses but speech during pulses is also suppressed. To counteract the speech suppression, perceptual continuity is exploited to improve intelligibility. Application of the algorithm to real-world signals results in significant noise suppression and improved word perception. Due to the presence of a random pulse component, less successful results using a pulse subtraction method where a prototype pulse is subtracted has been observed.

Future work in this area may include examining the possibility of reconstructing the speech lost when a pulse is suppressed. The pulse duration is short enough that in most cases it should be reasonable to perform such a reconstruction. Another possibility is the use of a multiband pulse-suppression scheme.

[1] MIT Lincoln Laboratory, 244 Wood Street, Lexington, MA 02420

[2] United States Secret Service, 950 H Street NW, Suite 4200, Washington, DC 20223

The United States Government Technical Support Working Group under Air Force contract F19628-00-C-0002 sponsored this work. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

Digital Evidence, Audio Enhancement, Speech Processing