C11 Proposed Analytical Framework for Electronically Frequency/Pitch-Modified Voices

Eliud Bonilla, BS, EB Technologies, LLC, PO Box 102, Kensington, MD 20895; Catalin Grigoras, PhD, 1020 15th Street, Ste 8I, Denver, CO 80202; and Jeff M. Smith, MS, National Center for Media Forensics - CU Denver, 1150 10th Street, Ste 177, Denver, CO 80217*

After attending this presentation, attendees will better understand the main challenges that electronically frequency/pitch-modified voices pose to speaker recognition efforts and the approach the National Center for Media Forensics is proposing in dealing with this issue in forensic environments.

This presentation will impact the forensic science community by providing a holistic analytical framework to use when dealing with electronically modified voices within investigative scenarios. Academic researchers and law enforcement personnel will also benefit since the framework will provide guidance in addressing a common core of questions.

The concept of the human voice as a reliable biometric signal is rapidly being accepted and implemented in today's society. State-of-the-art call centers are increasingly incorporating Automated Speaker Recognition (ASR) technologies in an effort to enhance customer service and minimize identity theft. In the forensics realm, ASR has been accepted in many European courts and may also be accepted in United States federal courts in the not-too-distant future; however, ASR is relatively fragile due to its dependency on the frequency dimension of the voice while not incorporating higher layers of information such as prosody and accents. It can be degraded by a number of inter-/intra-speaker characteristics, in addition to multiple variables along the signal chain.

Purposely modifying the frequency/pitch of a voice by electronic means is an effective counter-forensic measure. It is most often implemented to mask the identity of an individual. Its use can be considered legitimate when used to protect the identity of a witness in a television interview or as part of a law enforcement investigation; however, it can also be used to protect the identity of individuals committing crimes ranging from classic scenarios such as phone calls for ransom requests to recording video/audio messages inciting violence.

Electronic frequency/pitch modification impacts various common forensic analysis methods. Vowel spaces are distorted, thus neutralizing their use by phoneticians. Moderate changes degrade Likelihood Ratio (LR) scores in ASR systems while aggressive changes induce additional errors by distorting the format-frequency relationships outside of normal expected ranges.

The proposed analytical framework addresses four key questions that need to be answered in a satisfactory manner if the analysis is to have practical impact on investigative efforts. First, has the voice actually been electronically modified? Second, is the modification method or algorithm known? Third, are the settings of the method or algorithm known? Fourth, are the modifications reversible in order to enable conventional forensic voice comparison methods?

Attendees will be presented with a flowchart and guidelines that will facilitate incorporating future research in a holistic manner.

Voice Modification, Automated Speaker Recognition, Forensic Voice Comparison