VOCALinc

Massachusetts Institute of Technology, Lincoln Laboratory, 244 Wood Street, Lexington, MA, 02420

Approved for public release; distribution unlimited
VOCALinc

Automated speaker recognition software provides an accurate, objective, consistent, and efficient tool for conducting speaker comparisons.

As part of a growing trend in biometric identification, speaker recognition is becoming increasingly important to the nation’s forensic laboratories, intelligence agencies, and military branches. Voice, along with other biometrics, such as fingerprints and iris patterns, is a unique characteristic that can be used to verify or identify individuals. Until recently, only spectrogram-based speaker recognition techniques were available. These techniques rely on electronically recorded speech signals, graphically represented as a function of time and frequency (i.e., a spectrogram), that analysts must manually evaluate—an error-prone, subjective, inconsistent, and time-consuming process.

VOCALinc, a stand-alone automated speaker recognition software developed by MIT Lincoln Laboratory, overcomes many of the issues inherent in manual techniques. Incorporating recent advances in speaker recognition technology, the software consists of a suite of speaker recognition algorithms that conduct the speaker comparisons and a graphical user interface that allows analysts to specify known data about uploaded voice samples and apply the algorithm(s) of their choice. VOCALinc outputs two pieces of vital information:

- A match score between the voice samples in question provides an automatic assessment of the confidence level that the voices under evaluation were produced by the same speaker.
- A comprehensive report on the voice samples’ signals (e.g., speech duration, signal-to-noise ratio, and speech clipping percentage) informs analysts of signal conditions that could potentially discount the results (e.g., short voice samples).

Combined, this information equips analysts with sufficient data to make an informed decision about speaker identity.

A Suite of Algorithms

VOCALinc computes speaker comparisons through a suite of automated speaker recognition classifiers, or algorithms, which can be divided into two main classes: legacy algorithms and state-of-the-art algorithms. Two legacy algorithms—Gaussian mixture model and support vector machine—enable analysts to assess results from VOCALinc in the context of results obtained from older speaker recognition techniques. These classifiers work well when the voice samples under analysis are in milder conditions, such as in low-noise environments and on telephone channels, but do not work well in cross-channel (e.g., microphone versus telephone) and cross-language (e.g., English versus Spanish) conditions, both of which are common in forensic and investigative cases. To address these channel variability issues, VOCALinc incorporates a newly developed class of interoperable techniques: inner product discriminant functions, joint factor analysis, and i-vector algorithms. Each algorithm is trained to reduce or eliminate

VOCALinc’s graphical user interface features checkboxes, buttons, and drop-down menus for analysts to input speaker gender, channel type, prior probability, and other metadata regarding the uploaded voice samples (labeled Audio A and Audio B above), as well as to choose the desired algorithm(s) for comparing and scoring the samples.
Benefits of VOCALinc

VOCALinc was developed utilizing U.S. government operational data and with direct access to end-user feedback. As a result, its algorithms and interface were iteratively designed to ensure a best-practice approach. VOCALinc benefits its users by

- Reproducing consistent results among different laboratories
- Easing analyst workload and casework backlog
- Expediting open-case resolution
- Minimizing the number of false leads and the costs associated with pursuing them
- Being customizable to classified conditions
- Including an automated installation validator to calibrate speaker comparison equipment for quality-assurance purposes
- Supporting the ANSI/NIST-ITL Type-11 Record Standard

With applications in intelligence missions concerning national security (e.g., terrorism), forensic investigations within the field of law enforcement (e.g., drug trafficking), and military deployments (e.g., force protection), VOCALinc has the potential to become a valuable tool in a variety of settings. In fact, the software is already in use by several entities. Future versions of VOCALinc could reduce memory usage and enhance robustness to unseen devices such as body microphones and multirecording systems.

Technical Points of Contact

Dr. Joseph P. Campbell
Human Language Technology Group
jpc@ll.mit.edu
781-981-7522

Dr. Pedro A. Torres-Carrasquillo
Human Language Technology Group
ptorres@ll.mit.edu
781-981-5581

For further information, contact

Communications and Community Outreach Office
MIT Lincoln Laboratory
244 Wood Street
Lexington, MA 02420-9108
781-981-4204