Abstract

Speech conveys several levels of information. On a primarily level, speech conveys the words or message being spoken, but on a secondary level, speech also reveals information about the speaker. The Information Systems Technology Group at MIT Lincoln Laboratory has developed and experimented with approaches for automatically recognizing the words being spoken, the language being spoken, and the topic of a conversation. In this article, we present an overview of our research efforts in a fourth area -- automatic speaker recognition. We base our approach on a statistical speaker-modeling technique that represents the underlying characteristic sounds of a person's voice. Using these models, we build speaker recognizers which are computationally inexpensive and capable of recognizing a person regardless of what is being said. Performance of the systems is evaluated for a wide range of speech quality, from clean speech to telephone speech, by using several standard speech corpora.