

Automatic Speaker Verification: state of the art and current issues

Dijana Petrovska-Delacrétaz, Asmaa El-Hannani and Gérard
Chollet

Speech is often the only available modality to recognize the identity of a person (over the telephone, the radio, in the dark, ...). Automatic speaker recognition has been studied for several decades. Few applications are deployed. This paper analyses the individual characteristics of human voice, its variability, possibilities of disguise and mimicry, and automatic techniques to capture robust characteristics and use them to identify speakers.

In this paper the state of the current speaker recognition research is reviewed. Basic principles of speaker recognition are first summarized. The choice of the speech features and speaker models are mostly related to the individual characteristics (variability) of the speakers' voices. Besides the speaker's variability, we are faced with other factors, such as channel variability, silence detection, or score normalization, that influence the performance of speaker verification algorithms. All these issues are illustrated on recent NIST speaker evaluation databases. The utility of open source reference (baseline) algorithms and controlled evaluation campaigns is also pinpointed.

The field of speaker recognition is also reviewed in relation to speech recognition, focussing on the usage of this new source of information for the speaker recognition task. This relationship has to be seen as an important issue in the development of new services based on speaker and speech recognition. Overview of recent developments in this field is given. With the combination of speech and speaker recognition, we introduce also the issue and challenges of multimodal biometrics (i.e. talking faces).