Mason, John S. D.; **Evans, Nicholas W. D.**; **Stapert, Robert**; **Auckenthaler, Roland**
Data-model relationship in text-independent speaker recognition. (English) Zbl 1095.68668
EURASIP J. Appl. Signal Process. 2005, No. 4, 471-481 (2005).

Summary: Text-independent speaker recognition systems such as those based on Gaussian mixture models (GMMs) do not include time sequence information (TSI) within the model itself. The level of importance of TSI in speaker recognition is an interesting question and one addressed in this paper. Recent works has shown that the utilisation of higher-level information such as idiolect, pronunciation, and prosodics can be useful in reducing speaker recognition error rates. In accordance with these developments, the aim of this paper is to show that as more data becomes available, the basic GMM can be enhanced by utilising TSI, even in a text-independent mode. This paper presents experimental work incorporating TSI into the conventional GMM. The resulting system, known as the segmental mixture model (SMM), embeds dynamic time warping (DTW) into a GMM framework. Results are presented on the 2000-speaker SpeechDat Welsh database which show improved speaker recognition performance with the SMM.

MSC:

68T10 Pattern recognition, speech recognition
68T50 Natural language processing

Keywords:

speaker recognition; segmental mixture modelling

Full Text: DOI