
Abstract

This paper presents high performance speaker identification and verification systems based on Gaussian mixture speaker models: robust, statistically based representations of speaker identity. The focus domain is for unconstrained speech, although the systems can equally be used for text-dependent tasks. The identification system is a maximum likelihood classifier and the verification system is a likelihood ratio hypothesis tester using background speaker normalization. The systems are evaluated on three widely used speech databases: TIMIT, NTIMIT and Switchboard. The different levels of degradations and variabilities found in these databases allow the examination of system results for different task domains. An identification accuracy of 99.7% was obtained for a 168 population on TIMIT, 76.2% for NTIMIT and 82.8% for a 113 population on Switchboard. Global threshold equal error rates of 0.3%, 5.4% and 7.0% were obtained in verification experiments on TIMIT, NTIMIT and Switchboard, respectively.