



ON THE RECOGNITION OF KEY WORDS IN UNCONSTRAINED CONVERSATION

Michael Carey, Amanda Howe and Roger Tucker

Enigma Ltd  
Archway House, Welsh Street, Chepstow, UK

ABSTRACT

For many applications it is not necessary or even desirable to recognise every word in an utterance. A more satisfactory approach may be to try to identify the occurrence of a number of key words within the utterance string. In this paper we describe a system which identifies key words occurring in unconstrained conversation.

The system consists of a feature extractor followed by a pattern matching algorithm. The feature extractor performs a filter bank analysis of the incoming speech and the outputs of the filter bank are transformed into melcepstral coefficients and transitional mel cepstral coefficients. These coefficients comprise the observation vector, which is fed to a pattern matcher based on the Viterbi algorithm. The Viterbi algorithm attempts to match the incoming sequence of observation vectors to a group of continuous density hidden Markov models. The Markov models are re-estimated on utterances selected from a database of speakers. The utterances are chosen so that the system will work in a speaker independent mode.

When the log probability output of the Viterbi algorithm exceeds a threshold there is a potential match between the incoming speech and one of the models. A traceback procedure is then initiated. The traceback evaluates the number of vectors of the incoming speech which have been matched to each state in the model. If the distribution of vectors between states is not within specified bounds, then the overall match is rejected and no output is made. Otherwise the output of the occurrence of a valid word is indicated.

The performance of the algorithm is dependent on the matching threshold, but over a useful range of thresholds, between 25% and 75% of the incoming key words are detected. The false alarm rate varies between one false alarm every thirty minutes and one false alarm every three minutes of speech. The algorithm works in near real-time on a Western Electric DSP32C digital signal processor.