AUTOMATIC SPEECH RECOGNITION USING PERIPHERAL AUDITORY MODELLING AND A PDP APPROACH TO CLASSIFICATION

Stephen Renals [1] and Mark Terry [2]

[1] Department of Artificial Intelligence,
University of Edinburgh,
80 South Bridge,
Edinburgh, EH1 1HN.

[2] CSTR,
University of Edinburgh,
80 South Bridge,
Edinburgh, EH1 1HN.

ABSTRACT

Recent work in the field of human-computer interaction has highlighted the need for general and sophisticated speech recognition systems. Such systems require an initial analysis which adequately represents perceptually important features and a classification component that can cope with intra- and inter-speaker variability. In this study, a prototype isolated word recogniser was constructed, with an auditory-based analysis component and a pattern classification module based on a parallel distributed processing paradigm [1].

The auditory model [2, 3] used was a BPNL configuration using 128 channels, incorporating both the effects of phase locking and lateral suppression. The features used are based on the energy and the temporal fine structure output from the auditory channels.

Pattern classification was performed by a layered, feed-forward neural network [4], consisting of an array of input nodes representing the binary features output by the auditory model, a set of hidden nodes (in effect, the system's internal representation of the problem) and an array of output nodes representing the word to be recognised. A suitable internal representation was learned by the method of back-propagation of errors by gradient descent, using the generalised delta rule.

This prototype recogniser was implemented on a MASSCOMP 5700, within the AUDLAB interactive speech and signal processing package [5]. The network was trained to recognise English digits spoken by male and female speakers and an evaluation of the model in terms of speaker and time normalisation is presented.

REFERENCES: