

USA-Japan Joint Acoustical Society Meeting, 1988.11.14,
Hawaii, USA.

Incremental Learning of Large Phonetic Neural Networks from Smaller Subnets. A. Waibel (ATR Interpreting Telephony Research Laboratories, Twin 21 MID Tower, 2-1-61 Shiomi, Osaka, 540, Japan)

Time Delay Neural Networks (TDNNs) [Waibel et al., ATR-TR-0006 (1987)] have recently been shown to achieve excellent recognition performance for difficult phonetic discrimination tasks (e.g., voiced stops). This was achieved in part by the TDNNs' ability to not only activate the correct output category, but also to inhibit all incorrect outputs. The disadvantage of this property is that training larger networks with many more output categories (e.g., all consonants) becomes non-trivial as all categories have to be incorporated in the learning process requiring excessive amounts of training. Several techniques are presented that overcome this problem by exploiting the hidden structure of previously learned smaller neural nets to train larger nets incrementally in comparatively short training runs. Experimental results show that the resulting larger networks aimed at stop consonants, at voiced stops and nasals and at all consonants achieve recognition scores (95% to 99%) as high as the smaller subnetworks from which they were constructed.

Suggested for Special Session on Speech Communication, Human-Machine Interaction

Technical Committee: Speech Communication

Method of Presentation: Prefer Lecture but Willing to Give as Poster

(PACS) Subject Classification Number: 43.72.Ne

Telephone Number: 412-268-7674 (Secretary, Karen Olack)

Send Acceptance or Rejection Notice to:

A. Waibel, Computer Science Department, Carnegie-Mellon University, Pittsburgh, PA 15213