RAPID SPEAKER ADAPTATION USING MULTI-STREAM STRUCTURAL MAXIMUM LIKELIHOOD EIGENSPACE MAPPING

Bowen Zhou; John Hansen, University of Colorado, United States

In this paper, we extend our previously proposed algorithm entitled Structural Maximum Likelihood Eigenspace Mapping (SMLEM) for rapid speaker adaptation. The SMLEM algorithm directly adapts Speaker Independent (SI) acoustic models to a test speaker by mapping the mixture Gaussian components from a SI eigenspace to Speaker Dependent (SD) eigenspaces in a maximum likelihood manner, with very limited adaptation data. In previous SMLEM paper, we presented encouraging results for SMLEM by adapting only the static feature components. In this paper, we propose a multi-stream approach where the static and dynamic feature streams are adapted. For small amounts of adaptation data ranging from 15 to 50 seconds, superior performance is demonstrated over both standard MLLR and block diagonal MLLR.

MINIMAX APPROXIMATION OF MINIMUM PHASE FIR FILTERS

Yasunori Sugita; Naoyuki Aikawa, Nihon University, Japan

This paper presents a directly design method of FIR filters with minimum phase characteristics using successive projections (SP) method. We use Rouche's theorem for coefficients update algorithm in SP method to restrict the roots of transfer function inside the unit circle and on the unit circle. The proposed method is an iterative optimization technique and its algorithm is as simple as the conventional SP method. The usefulness of the proposed method is verified through the examples.

A NEW NOISE REDUCTION METHOD USING LINEAR PREDICTOR AND ADAPTIVE FILTER

Arata Kawamura, Tottori University, Japan; Kensaku Fujii, Himeji Institute of Technology, Japan; Yoshio Itoh; Yutaka Fukui, Tottori University, Japan

A technique that uses a linear prediction error filter (LPEF) and an adaptive digital filter (ADF) to achieve noise reduction in a speech degraded by additive background noise is proposed. Since a voiced speech can be represented as the stationary periodic signal over a short interval of time, most of voiced speech is predicted by the LPEF. On the other hand, when the input signal of the LPEF is a background noise, the prediction error signal becomes white. Assuming that the background noise is represented as generate by exciting a linear system with a white noise, then we can reconstruct the background noise from the prediction error signal by estimating the transfer function of noise generation system. This estimation is performed by the ADF which is used as system identification. Noise reduction is achieved by subtracting the noise which is reconstructed by the ADF from the speech degraded by additive background noise.