ABSTRACT

The objective of this thesis is to study some of the existing algorithms of deterministic classifiers for speech recognition and to propose new algorithms. In particular this thesis addresses to the recognition of isolated spoken Marathi (Language spoken in the state of Maharashtra, India) words using minimum distance classifiers, discriminative weighting classifiers and neural network classifiers.

In this thesis five algorithms are proposed which are interesting extensions of some of the existing algorithms. These algorithms are used for the recognition of 46-mosyllabic Marathi alphabets and/or 10-Marathi digits.

DTW algorithms are used for mapping of two utterances in computing the dissimilarity between them. By improving the time warping constraints of well-established DTW algorithms a nonlinear time alignment algorithm (TAA) is proposed. The response of TAA to speaker dependent and independent modes is compared with well-established DTW algorithms and found that TAA shows better accuracy, needs less number of computations and shows robustness to endpoint detection and noise.

For confusing words recognition, an alternative approach (DW_PRO) to discriminative weighting method proposed by Rabiner and Wilpon (DW_R) based on Fisher’s linear discriminative analysis is proposed. In DW_PRO, a set of lowest distance words is selected and the frame-to-frame local distances are weighted with weighting coefficients derived from training data, using Fisher’s linear discriminant method. The performance of DW_PRO is tested with various speech parameters such as linear predictive coefficients (LPC), mel frequency cepstrum coefficients (MFCC), linear frequency cepstrum coefficients (LFCC) and regression coefficients of MFCC and LFCC and found that DW_PRO shows better performance than DW_R.
A modular fuzzy min-max neural network is proposed which uses the modified general fuzzy min-max neural network (GFMM). Words are segmented and individual segment is used for training the different modules of modular fuzzy min-max neural network. In the testing of input word, the average of a set of best fuzzy membership values of hyperboxes of all classes for each segment is computed. Then, the average of number of segments of fuzzy membership values of each class is computed. The class showing largest fuzzy membership value is assigned to input word. A modular fuzzy min-max neural network has shown better average recognition accuracy than GFMM.

A modified fuzzy hypersphere neural network (MFHS) with two learning algorithms namely two-stage learning algorithm and centroid-based learning algorithm is proposed. MFHS is a modification of a fuzzy hypersphere neural network (FHS) proposed by Kulkarni and Sontakke. The transfer function of the output layer neurons of FHS is modified. In FHS, the center of hypersphere is not the centroids of the patterns contained by the hypersphere. In MFHS, hyperspheres grow around the centroids of the patterns associated with the hypersphere; hence MFHS shows better performance than FHS.

A general fuzzy hypersphere neural network (GFHS) that uses supervised and unsupervised learning within a single training algorithm is proposed. It is an extension of FHS and can be used for pure classification, pure clustering or hybrid clustering/classification. The performance of GFHS is compared with GFMM for classification and unsupervised clustering of patterns of spoken Marathi digits and found that GFHS exhibits comparable response.