سومین کنفرانس فناوری اطلاعات و دانش
۶ تا ۸ آذرماه ۱۳۸۶

Class-Dependent PCA Optimization Using Genetic Programming for Robust MFCC Extraction

Houman Abbasian
Research center for Information Technology,
Computer Engineering Department,
Iran University of Science and Technology
TEHRAN, IRAN
hooman_abbasian@comp.iust.ac.ir

Babak Nasersharif, Ahmad
Research center for Information Technology,
Computer Engineering Department,
Iran University of Science and Technology
TEHRAN, IRAN
{nasser_s, akbari}@iust.ac.ir

***Abstract*—Principal component analysis (PCA) is commonly used in feature extraction. It projects the features in direction of maximum variance. This projection can be performed in a class-dependent or class-independent manner. In this paper, we propose to optimize class-dependent PCA transformation matrix for robust MFCC feature extraction using genetic programming. For this purpose, we first map logarithm of clean speech Mel filter bank energies (LMFE) in directions of maximum variability. We obtain the mapping functions using genetic programming. After this, we form class-dependent PCA transformation matrix based on mapped LMFE and use this matrix in place of DCT in MFCC feature extraction. The experimental results show that proposed method achieves to significant isolated word recognition rate on Aurora2 database.**

I. INTRODUCTION

Feature extraction is a crucial step of speech recognition process which greatly affects the performance of speech recognition systems. Speech features must represent the temporal evolution of the speech spectral envelope. Some examples of common speech features are: LPC, PLP and MFCC. Among these features, MFCC feature are more commonly used for speech recognition. The MFCC features are obtained by applying discrete cosine transform (DCT) to logarithm of Mel filter bank energies (LMFE). There are several techniques that improve MFCC features from different points of view. Some approaches attempt to make MFCC more robust to channel and additive noises using weighitng or compression of Mel sub-band energies [1][2][3]. In other group of methods, we try to overcome to disadvantages of DCT in clean or noisy conditions [4][5]. DCT is a non-adaptive procedure that projects LMFE in the direction of global variance which achieves only partial decorrelation of features.

In order to overcome the partial decorrelation, several methods have been proposed to replace DCT and decorrelate LMFE. Some examples of such methods are: Principle Component Analysis (PCA) [4][5][6] and Independent Component Analysis (ICA) [5]. PCA, based on the principle of minimum reconstruction error, projects the data (LMFE) in the direction of maximum variability [4][5]. But, there is no guarantee that variability explained by PCA is useful for speech recognition. While PCA removes the second order dependencies of the feature vector components, ICA removes also higher order dependencies and minimizes the mutual information between the feature vector components.

One limitation of PCA is that it does not model nonlinear relationships among feature vector components efficiently. Several generalizations of PCA have been proposed to address this limitation. Two examples of such approaches are: nonlinear PCA (NLPCA) [7][8] and kernel PCA (KPCA) [9][10][11][12]. NLPCA generalizes the principal components from straight lines to curves. This can be done by using neural networks with an auto-associative architecture [7]. In KPCA, a nonlinear map is used to translate nonlinear structure of features into linear ones in a feature space with a higher dimension. After this, linear PCA applied to mapped features [9][10][11[12].

While NLPCA and KPCA optimize PCA transform to overcome non-linearity in feature space, , we propose to optimize PCA transform using Genetic Programming (GP) in order to project feature vector component in a space which they have maximum independence. We name this method as GPCA. The feature vector components in our work are LMFE. When we apply optimized class-dependent PCA transformation to LMFE, they will be more uncorrelated. Therefore, their covariance matrix will be more diagonal. This causes better HMM training based on these uncorrelated features. In addition, we expected that these uncorrelated features are less affected by noise due to the class-dependent transformation.

Fig. 1 shows the block diagram of different MFCC extraction and proposed methods.The rest of this paper is organized as follows. In section 2, we describe MFCC extraction based on PCA. Section 3, explains the used method for optimization of PCA transformation matrix for