

Automatic Speech Emotion Recognition System using Mel-frequency Cepstral Coefficients and Fuzzy Gaussian Mixture Models

Neha Gupta

Student, Poornima College of
Engineering, Jaipur, Rajasthan
neha.pce.jaipur@gmail.com

Mahesh M. Bundele

Professor, Poornima College of
Engineering, Jaipur, Rajasthan
maheshbundele@gmail.com

Indu Gupta

Student, Sri Ram Murti Smarak
College of Engineering and
Technology, Bareilly, Uttar Pradesh

Abstract— Digital processing of speech signal and voice recognition algorithm is very important for fast and accurate automatic voice recognition technology. The voice is a signal of infinite information. A direct analysis and synthesizing the complex voice signal is due to too much information contained in the signal. Therefore the digital signal processes such as Feature Extraction and Feature Matching are introduced to represent the voice signal.

in order to improve both the performance and the efficiency of the conventional Gaussian Mixture Models (GMMs), generalized GMMs are firstly introduced by integrating the conventional GMMs and the active curve axis GMMs for fitting non-linear datasets, and then two types of Fuzzy Gaussian Mixture Models (FGMMs) with a faster convergence process are proposed based on the generalized GMMs, inspired from the mechanism of Fuzzy C-means (FCMs) which introduces the degree of fuzziness on the dissimilarity function based on distances are evaluated with a view to identify a straight forward and effective method for voice signal. The extraction and matching process is implemented right after the Pre Processing or filtering signal is performed. The non-parametric method for modelling the human auditory perception system, Mel Frequency Cepstral Coefficients (MFCCs) are utilize as extraction techniques. Since it's obvious that the voice signal tends to have different temporal rate, the alignment is important to produce the better performance. This paper present the viability of MFCC to extract features and FGMM to compare the test patterns.

Keywords— MFCC, Fuzzy C-means, Gaussian Mixture Models, Fuzzy Gaussian Mixture Models, emotion recognition, prosody modeling.