Emotion Recognition using Fuzzy Clustering Analysis

Tejal Udhan
Georgia Southern University, t.p.udhan@gmail.com

Follow this and additional works at: https://digitalcommons.georgiasouthern.edu/research_symposium

Part of the Engineering Science and Materials Commons, and the Materials Science and Engineering Commons

Recommended Citation
https://digitalcommons.georgiasouthern.edu/research_symposium/2016/2016/37

This presentation (open access) is brought to you for free and open access by the Conferences & Events at Digital Commons@Georgia Southern. It has been accepted for inclusion in Georgia Southern University Research Symposium by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.
Emotion Recognition using Fuzzy Clustering Analysis

Tejul Udhan1, Shonda L. Bernadin, Ph.D.1 and Earl Smith, Ph.D.2
1FAMU-Florida State College of Engineering, Department of Electrical and Computer Engineering, Tallahassee, Florida
2Georgia Southern University, Department of Mechanical Engineering, Statesboro, Georgia

Abstract

This research project investigates using fuzzy clustering algorithms for emotion recognition. Emotion recognition has gained significant attention in recent years in applications such as artificial intelligence, human-computer interaction, speech and voice recognition. The ability of a computer or machine to understand human emotion and respond to users in a more human way can lead to significant advances in conversational speech recognition systems, improved quality of life in persons with speech disorders, such as Parkinson’s disease and even in voice response systems, such as Google Voice or Apple’s Siri. Experimental results in this area can inform discovery and innovation of machine intelligence and actionable response algorithms that use physiological methods for characterizing speech. Human emotion is a complex signal that is difficult to characterize analytically. One proposed method for characterizing emotion is to use fuzzy clustering techniques to partition the data into classifications of emotions based on feature similarities. Fuzzy clustering provides a method for organizing data into groups either in unsupervised fashion or based on the selected feature and classifying each group as a different emotion. In this work, an emotional prosody speech dataset is prepared input to a fuzzy clustering toolbox to explore underlying structures in the dataset and perform data reduction for optimal feature extraction. The emotion dataset includes fifteen different categories of emotions: happy, elation, sadness, despair, boredom, interest, shame, pride, contempt, disgust, panic, anxiety, hot anger, cold anger, and no emotion. The primary goal of this research project is to identify a fuzzy clustering technique that will partition the dataset into different categories of emotions.

Background and Objectives

- The interaction between human beings and computers will be more natural if computers will be able to perceive and respond to human verbal communication or non-verbal communication. Verbal mode has speech i.e. acoustic signals where as non-verbal mode contains facial expressions.
- Past research shows that use of either of the modes of communication for emotion recognition is not efficient (only between 50% to 70%). Neural networks, hidden Markov models (HMMs) and support vector machines (SVMs) are the few types of algorithms that has been used.
- Since Fuzzy clustering is one of the powerful tools in data mining and has supervised and unsupervised algorithm availability, it performs well in situations where there is large variability in the data clusters. Consequently, it may should provide a good platform for emotion recognition algorithm.

The main objectives of this project are two-fold:
(1) To determine an acoustic feature that can characterize human emotion efficiently.
(2) To develop a clustering algorithm performing emotion recognition within the dataset that has closely related emotions.

Methods

To accomplish the objectives, the following process is used:

Step 1: Identify speech utterances for 15 different emotions
Step 2: Segment a single speech utterance for each emotion
Step 3: Derive different features for each emotion
Step 4: Selection of the feature representing each emotion uniquely
Step 5: Fuzzy clustering algorithm for emotion recognition

First three steps of the method are completed in the analysis. The results are presented in the following section.

Results and Discussion

Step 1: Out of eight speakers, having three females and five male speakers, all the utterances from one male speaker are taken into account. A single utterance for each of the emotion has been chosen. The selection was based on time duration for each utterance.

Step 2: The desired speech utterances are segmented. Segmenting the data provides better understanding. All the speech utterances lie between 1.00 seconds to 1.14 seconds; divided in six different groups with time duration 1, 1.02, 1.03, 1.04, 1.07 and 1.14 seconds.

Step 3: Different time and frequency features of the speech segments are analyzed. The analysis includes first four frequency formants, spectrogram, pitch and intensity listings for each emotion. The features are shown for four dominantly different emotions: hot anger, happy, sadness and neutral (no emotion).

The emotions are color coded; the color of the figure name describes the emotion it represents. First row of the figures indicate acoustic signal on channel 1 and 2, frequency formants, duration of the utterance and pitch contour whereas the second row indicates spectrograms and intensity levels. Figures show that, acoustic signals for each of the emotion are fairly different, which is depicted in the distribution of the frequency formants. The spectrograms show the distribution of frequency component distribution for each emotion. The intensities listed on the spectrograms are between 40 dB to 100 dB.

Implications and Future Research

- The achieved results for intensity and frequency formants are distinctive but cannot be independently used. Hence, desirable feature extraction still remains a challenge for the use of speech signals in emotion recognition. A feature vector that can represent each emotion differently is of prime importance to apply fuzzy clustering algorithms.
- The ultimate goal of the research is to design an emotion recognition system which can use acoustic and electromyographic (EMG) data as an input and provides better accuracy in recognizing emotions than most of the systems which use just one of these parameters but are less efficient.

References