International Journal of Latest Trends in Engineering and Technology
IJLTET Special Issue- ICRACSC-2016 , pp.197-200
e-ISSN: 2278-621X

A MULTICLASS SVM APPROACHES FOR AUTOMATIC EMOTION RECOGNIZATION USING SPEECH

P.Jayaram., M.Tech[1], Dr.J.Sreedhar.,M.Tech.,PhD[2] & E.S.Madhan.,M.E.,(PhD)[3]

Abstract-Automatic Emotional Recognization (AER) is one of most research issue in speech processing.The applications such as lie detection, medicine, e-learning and so on. In this paper, we propose a multiclass SVM classification, which is used to recognition the emotions. However, we consider emotions like anger, fear, happy and neutral. The features extract from speech signals plays a prominent role to identify the emotion. The combination of both spectral features like (MFCC, SKEWNESS) and prosodic features like (pitch, formants,zero-crossings) provides good Accuracy in recognition of emotions. Moreover, we use the database to consider different combinations are both male and females speakers. Our results show the stability and good rate of recognization for emotions.
Keywords – MFCC,Feature Extraction, prosodic features, spectral features, SVM, speech emotion recognition.

1. INTRODUCTION

Speech emotion recognition is one of the new research scope in the era of human machine interaction. [1,6] It can taken speech as input and identify the corresponding emotion. The application of emotion recognition are monitoring of psycho physiological state of individuals, Educational institutions, software engineering etc.

Feature extraction:

The rich amount of information presents in the signals, we can extract the relevant or robust features from speech using feature extraction techniques. The quality of emotion recognition is depends upon the features extracted from the speech. [8] The features can be prosodic and spectral features. Only consideration of prosodic or spectral is not enough to identify the emotion in very efficient manner. So in this we propose the combination of both spectral and prosodic features which gives the recognize rate better than the previous workout.

The features consider here are prosodic features like fundamental frequency, formants, energy and spectral features are MFCC, Skewness [2]. The performance of the emotion recognition system depends on the features that are used. The multiple classification algorithms used for the speech emotion recognition are Support Vector Machine, Artificial Neural Network, [3] Gaussian Mixture Model, and Hidden Markov Model.

From the fig 1The speech samples are taken as input. Preprocessing is the first step to be done with the speech samples are preprocessing from which noise can be removed. [4] Next from the noise free samples the features are extracted. These features are further passed to the classifier algorithm. The classifier then classifies the emotions accordingly and outputs the emotions.

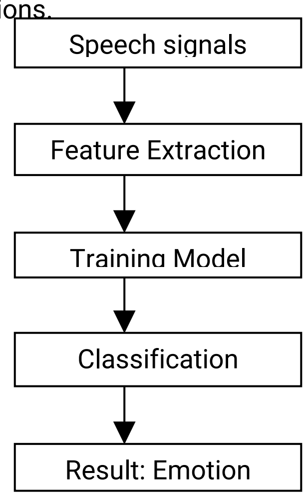


Fig1. Working Model Of Speech Emotion Recognition System.

[1] *Department of ComputerScience and Engineering, SVEC,TIRUPATHI.*
[2] *Department of ComputerScience and Engineering, SVEC,TIRUPATHI.*
[3] *Department of ComputerScience and Engineering, SVEC,TIRUPATHI.*

This paper is organized as follows. Section II Contains the database that used in this proposed system. Section III discuss about the classification algorithm, that used classify the emotions. The section IV is display the results that are classified by the algorithm. Section V all about conclusion and future scope.

2. DATABASE

The database used in this paper is, IITKGP-SESC. It was record at All India Radio India. All the artists had the professional experience of 8–12 years.[5] They consider 15 sentences for interpreting the emotions 10 sentences have been spoken by each artist in the chosen emotions. Finally 2400 utterances have been uttered in grand total. The number of words and syllables in the sentences were varying from 3–6 and 11–18 respectively. The duration of the database was around 36 minutes. The four emotions considered for collecting the proposed speech corpus were: Anger, Happy, Fear, and Neutral.

After feature extraction the database is divided into training and testing data sets. The database includes both male and female voice speech samples. The tabular results shows that the different combination of training and testing data. The training duration considered is 30 seconds and the testing duration of 3 seconds.

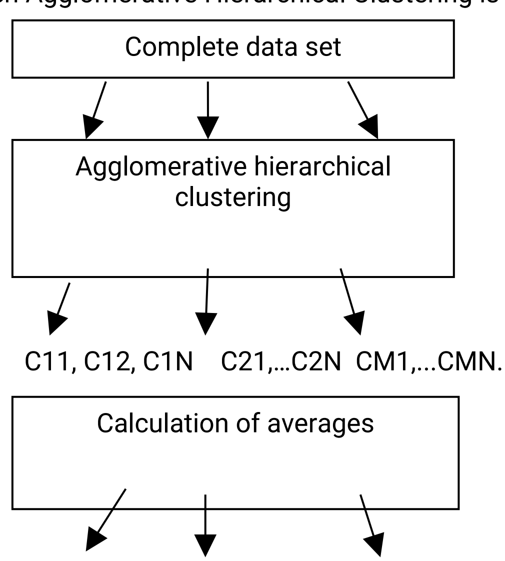
3. CLASSIFICATION ALGORITHM

SVM is one of the widely used classification algorithm for pattern recognition and machine learning areas. Under limited training data, it produces the more effective outcomes compared to all other classifiers. So we consider the support vector machine to classify the emotions in this paper. SVM is basically implemented for binary classification problems. And it provides good competence in classifying binary class problems and identifies the labels.

In this paper we consider four emotion classes like, fear, anger, neutral and happy. Then we have chosen multiclass SVM methodology for classification the emotions. Already several approaches have been listed out for multiclass classification.

The popular approaches are one against all and one against one approach. In one against all approach one emotion is selected as one individual class and remaining all emotions are grouped together as second class. SVM classification draws a hyper plane between this one class and remaining all classes. This process continues until all classes are classified. In one against one method choose the two classes in random manner and apply binary classification between the classes. And this process continues till all the combinations of the classes are taken and apply then binary classification between them. Still this is an open and broadly emergent area of research. The classes are in any machine learning algorithms or pattern recognition is generally big in number. We need much capable and precise multiclass classification methodologies.

In the midst of existing methods we have chosen an approach called Dendogram Based Support Vector Machine (DSVM)[12].The proposed DSVM model have the working out of the both **Agglomerative Hierarchical Clustering** (AHC) of classes and provides the efficient binary classification in SVM. The Training phase of DSVM needs (p-1) SVMs for p-class problem and the testing phase of DSVM require an optimal set of SVMs selected from the root of the classification through the chosen class between the leaf nodes. The major two steps in DSVM method are first computation of clusters for classes known. And then second is associating a support vector machine at each node of the classification obtained by clusters from first step. Consider a set of samples c1, c2, c3. . . ,cn labeled each one by yiϵ {x1, x2, . . . , xk}, where k≤n.[12] DSVM method consists of calculates k gravity centers for the k known classes in the first step. Then Agglomerative Hierarchical Clustering is applied over these k centers.



Y11, Y12, Y1K Y21, Y2K YM1...YMK

Fig 2 Dendogram Based Support Vector Machine Classification Model

The above Figure2 shows that classification completed by Agglomerative Hierarchical Clustering algorithm over the p classes. Then each SVM is related to a node and trained with the elements of the two subsets of this node. For example, we consider clustering of 4 classes as SVM1.And then categorize the elements of {c3, c4} as positives and elements of {c1, c2,} as negatives. This process is repeated for every SVM related to a node in the classification. Finally, we will train the ($p-1$) classes SVM for p-class problem. Dendogram Based Support Vector Machine to classify the pattern query then presents it to the root SVM which provides an output either left or right on the classification. This process repeated for every selected node in the system of the classification until incoming to a leaf which to end with related class for our pattern query.

4. RESULTS

Table 1:Male Training + Male Testing:

Emotions	Anger	Fear	Happy	Neutral
Anger	**89.7**	**1.78**	**2.76**	**3.67**
Fear	**74.5**	**0**	**1.98**	**25.4**
Happy	**55.9**	**4.56**	**94.5**	**1.09**
Neutral	**87.2**	**0**	**98.7**	**15.6**

Table 2:Male Training +Female Testing:

Emotions	Anger	Fear	Happy	Neutral
Anger	**2.67**	**1.34**	**3.56**	**94.8**
Fear	**1.80**	**5.67**	**0**	**89.7**
Happy	**0**	**13.5**	**6.98**	**97.6**
Neutral	**1.54**	**0**	**0**	**100**

Table 3:Female Training + Female Testing:

Emotions	Anger	Fear	Happy	Neutral
Anger	**74.2**	**7.48**	**28.3**	**3.84**
Fear	**25.7**	**47.0**	**17.8**	**8.04**
Happy	**16.8**	**56.7**	**0**	**25.6**
Neutral	**64.8**	**5.64**	**6.89**	**20.3**

Table 4:Female Training + Male Testing:

Emotions	Anger	Fear	Happy	Neutral
Anger	**5.40**	**1.48**	**67.8**	**1.58**
Fear	**2.48**	**0**	**34.5**	**0**
Happy	**0**	**7.64**	**56.2**	**7.86**
Neutral	**3.67**	**0.83**	**45.7**	**25.7**

From the results we observed that the emotion recognition rate is increased when the training and testing set sample data from same gender and compare to the different gender samples data.

5. FUTURE SCOPE:

In this field of research to identify the emotion recognition from speech signals depends upon the robust features extraction and selection. So with these robust features to apply various classification algorithms to produce the results very effective accuracy.

6. REFERENCES

[1] Rao, K. S., Kumar, T. P., Anusha, K., Leela, B., & Bhavana, I. (2012). Emotion Recognition from Speech, 3(2), 3603–3607.
[2] Koolagudi, S. G., & Rao, K. S. (2012). Emotion recognition from speech using source , system , and prosodic features. http://doi.org/10.1007/s10772-012-9139-3
[3] A, F. S., A, R. S., & P, B. A. (2009). Automatic Emotion Recognition from Speech using Artificial Neural Networks with Gender-Dependent Databases, 162–164. http://doi.org/10.1109/ACT.2009.49
[4] Yun, S., Member, S., & Yoo, C. D. (2012). Loss-Scaled Large-Margin Gaussian Mixture Models for Speech Emotion Classification, 20(2), 585–598.

[5] Anna, P. (2013). Evaluation of influence of spectral and prosodic features on GMM classification of Czech and Slovak emotional speech, 1–22.

[6] Albahri, A., Lech, M., & Cheng, E. (2016). Effect of Speech Compression on the Automatic Recognition of Emotions, 4(1), 55–61.

[7] Kamel, M. S., Fellow, L., Karray, F., & Member, S. (2014). Multiview Supervised Dictionary Learning in Speech Emotion Recognition, 22(6), 1056–1068.

[8] Song, P., Jin, Y., Zha, C., & Zhao, L. (2015). Speech emotion recognition method based on hidden factor analysis, 51(1), 1–2. http://doi.org/10.1049/el.2014.3339

[9] A, S. N. R., Gadhe, R. P., Deshmukh, R. R., Waghmare, V. B., & Shrishrimal, P. P. (2015). Automatic emotion recognition from speech signals: A Review, 6(4), 636–639.

[10] Yoon, W., & Park, K. (2011). Building Robust Emotion Recognition System on Heterogeneous Speech Databases, 747–750.

[11] Zong, Y., Zheng, W., Zhang, T., Huang, X., & Member, S. (2016). Cross-Corpus Speech Emotion Recognition Based on Domain-adaptive Least Squares Regression, 9908(c), 1–9.

[12] Cedex, L. (n.d.). Dendogram Based SVM for Multi-Class Classification Khalid Benabdeslem, 173–178.