Recognition of Human Emotion through effective estimations of Features and Classification Model

Publisher: IEEE

Satyajit Pangaonkar ; Reena Gunjan ; Virendra Shete **All Authors**

Abstract:
Voice Emotion Recognition (VER) is a dynamic and has implications on a wide range of research areas. Use of a computer for voice emotion recognition is a way to study the... **View more**

Metadata

Abstract:
Voice Emotion Recognition (VER) is a dynamic and has implications on a wide range of research areas. Use of a computer for voice emotion recognition is a way to study the voice signal of a speaker, as well as is a process that is altered by inner emotions. Human Machine Interface (HMI) is very vital and opted to implement this effectively and an innovative way. To develop new recognition methods, this research paper evaluates the basic emotions of human. Accurate detection of emotional states can be further used as a machine learning database for interdisciplinary experiments. The proposed system is an algorithmic method that first extracts the audio signal from the microphone, preprocesses it, and then evaluates the parameters based on various characteristics. The model is trained through the Mel Frequency Cepstral Coefficient (MFCC) and PRAAT (Speech Analysis in Phonetics) coefficients. By creating a feature map using these, Convolutional Neural Networks (CNN) effectively learn and classify the attributes of perceived signals of basic emotions such as sadness, surprise, happiness, anger, fear, neutral and disgust. The proposed method provides good recognition rate.

Published in: 2021 International Conference on Computing, Communication and Green Engineering (CCGE)

Date of Conference: 23-25 September 2021 **DOI:** 10.1109/CCGE50943.2021.9776405

Date Added to IEEE *Xplore*: 24 May 2022

ISBN Information:

Electronic ISBN:978-1-6654-1509-5

Print on Demand(PoD) ISBN:978-1-6654-3042-5

Publisher: IEEE

Conference Location: Pune, India

Contents

I. Introduction

For decades, research into emotional recognition in speech has been the focus of research, because including emotions in human-computer interaction can improve the user's interactive experience. VER research is an active research field. Once a researcher decides to study an emotional phenomenon in the context of information use, it is advisable to take the time to understand the differences between emotions, feelings, reactions, excitements, or other types of expressive phenomena. The most significant work of speech evaluation for sentiment analysis is the appreciated Ekman and Fox model [1]. Ekman's model is based on six basic emotional models, while Fox's model is a multi-level emotional model. Each emotion corresponds to a different part of speech. However, it is difficult to determine the boundaries between these parts. This article attempts to explain a diverse approach for detection of emotional state from human speech. As a measure of intelligence, the audio signal contains a lot of information. This information can be identified as a feature. It turns out that these features are typical landscapes of a study and, if correctly identified, can be modeled in a unique way and define different outcomes. Also, in many of the studies, hybrid method of feature selection and classification for better accuracy and recognition rate is experimented [2]. In this study, spectral features (low level features) of speech such as MFCC is identified. Other attributes like successive magnitude and time deviation like Jitter, Shimmer, and harmonics contents in reference to noise (HNR) has been extracted from PRAAT. These two different approaches of extraction for the common cause of training and classification by forming the feature map for CNN had proved a better perception of emotion.

Authors

Figures

References

Keywords

Metrics

Modified Convolutional Neural Network Architecture Analysis for Facial Emotion Recognition
2019 International Conference on Systems, Signals and Image Processing (IWSSIP)
Published: 2019

Show More