Segmentation in Digital Signal Processing

Adesh Potdukhe, Prajwal Tendulkar, Prof. Prajakta Pawar

[1,2]*Student, BE Computer, Bharati Vidyapeeth's College of Engineering, Pune, India*
[3]*BE (CSE), MTech (CSE), Bharati Vidyapeeth's College of Engineering, Pune, India*

---***---

Abstract - *Mankind usually invents a new era to make existence easier and better to stay. In this era audio information plays a rather important role in the increasing digital content that is available today, resulting in a need for methodologies that automatically analyze such content: audio event recognition for calculating the distance between two objects, depth, and heights of wells and valleys. Sound waves are a non-stationary process in nature to represent the non-stationarity. In this report, we are using soundwaves properties for measuring the distance between unknown objects. To calculate fundamental frequencies, we used estimated sound waves and segmented them using pyAudioAnalysis. By filtering the signals which are eco of sound waves measure by using machine learning and digital signal processing. To implement segmentation, we used technology of digital signal processing is extensively used to distinct wish echo's, noise and silence further analysis through signal is necessary to align and translate signal into time frequency graph. The fundamental frequency selection is made on frame basis so it is important to achieve right possible accuracy. This paper analyses and summarized this algorithm of audio segments and predict fundamental frequency of signal.*

Key Words: Digital signal processing, Machine learning, Segmentation, Fundamental frequencies, pyAudioAnalysis, Fundamental frequency estimation, Time-Frequency graph.

1. INTRODUCTION

Digital signal process (DSP) a numerical manipulation of signals usually to measure features produced over compress continuous analog signals. It is characterized by the use of digital signals to represents these signals as discrete-time, discrete frequency or other discrete domain signals in the form of a sequence of numbers symbols to permit digital processing of signals. Numerical methods required digital signals such as to produce analog to digital convertor. Digital signal processing and analog signal processing are subfields of signal processing. DSP applications are speech and audio signal processing, sonar and radar signal processing, spectral estimation, statistical processing, digital image processing, signal processing for communication.

Segmentation is a completely vital processing level for maximum of audio evaluation programs. The intention is to cut up an uninterrupted audio signal into homogeneous segments. Segmentation can either be

<u>Supervised</u>: in that case some sort of supervised information is used to categories and section the enter indicators. This is both carried out via applying a classifier a good way to classify successive restore-sized segments to a hard and fast of predefined lessons, or employing using a HMM method to achieve joint segmentation-category.

<u>Unsupervised</u>: a supervised model isn't always to be had and the detected segments are clustered (instance: speaker dualization)

Applications of audio content analysis may be categorized in two categories. One part is to discriminate an audio stream into homogenous areas and the alternative categories is to discriminate a speech movement into segments, of different speakers.

Audio segmentation algorithms may be divided into 3 fashionable categories. In the primary class, classifiers are designed. The functions are extracted in time domain and frequency area; then classifier is used to discriminate audio signals primarily based on its content material. The second category of audio segmentation extracts capabilities on statistics that is used by classifier for discrimination. These styles of capabilities are known as posterior probability-based features. Large quantity of observed records is needed by using the classifier to present correct results. The category of audio segmentation set of rules emphasizes putting in effective classifiers. The classifiers used in this category are Bayesian facts criterion, Gaussian chance ratio, and a hidden Markov version (HMM) classifier. These classifiers additionally provide excellent effects whilst huge training facts is provided

The analysis of superimposed speech is a complicated trouble and progressed performance systems are required. In many audio processing applications, audio segmentation plays a critical position in preprocessing step. It additionally has a good-sized impact on frequency recognition performance. That is why a fast and optimized audio class and segmentation algorithm is proposed which can be used for real-time packages of multimedia. The audio input is classed and segmented into four primary audio types: natural-eco, noise, environment sound, and silence. A set of rules is proposed that calls for less training facts and from which high accuracy may be achieved; this is, misclassification rate is minimum.

[2] The proposed method of signal segmentation is based upon the two sliding overlapping windows and the detection of signal properties changes. [4] Most of the researches integrated segmentation approaches with some intelligent techniques such as neural network, support vector machines

and so on to enhanced accuracy. [3] Nowadays, the speech signals processing takes important place at the present development stage of methods and tools for building artificial intelligence systems. Real-time support is a common requirement that is proffered to methods of analysis and synthesis of speech information. As a result, methods of the direction of linear prediction [5, 7] during the simulation of speech signals become significantly interesting and widely used. The main idea of these methods is to represent the signal into a form of autoregression. [4] Any undetermined signal is defined as a nonlinear object. But it is always possible to select time interval from the signals with a given discretization period. Such an interval is called quasistationary interval. The quasistationary interval in combination with the speech signal quality characteristics can be used to build a parametric model of the speech signal.

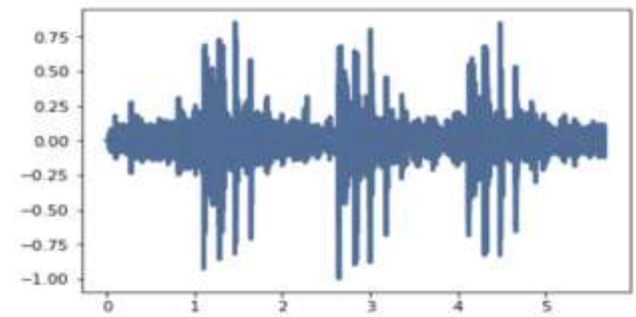
Audio segmentation and classification have many applications. Content-based audio classification and retrieval are mostly used in entertainment industry, audio archive management, commercial music usage, surveillance, and so forth. Nowadays, on the World Wide Web, millions of databases are present; for audio searching and indexing audio segmentation and classification are used. In monitoring broadcast news programs, audio classification is used, helping in efficient and accurate navigation through broadcast news archives..

2. Materials and methods:

2.1 Audio classification and segmentation steps

Hybrid classification scheme is proposed that allows you to classify an audio clip into simple statistics sorts. Before type a pre-classification step is done which analyzes each windowed body of the audio clip separately. Then the feature extraction step is carried out from which a normalized feature vector is acquired. After feature extraction the hybrid classifier approach is used. The first step classifies audio clips/frames into natural-eco and noise segments via the use of bagged SVM. As the silence frames are mostly found in audio signal so the frequency section is assessed into silence and pure-eco segments on the idea of rule-based totally classifier.

```
import thinkplot
import matplotlib
wave.plot()
#wave.show()
```



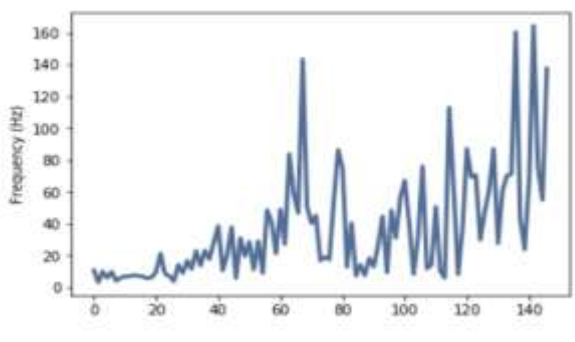
2.2 Pre-classification

Digital signal is superimposed (i.e., in combined shape) which means that a communication is held at any location or celebration where there's echoes and plenty of noise. This is likewise known as cocktail celebration impact. Separating the source or the desired segments in the independent thing analysis framework is known as blind source separation. Blind supply is usually a method used to separate the combined sign into independent sources (while the combination process isn't always recognized) [. Most blind supply separation techniques use better order statistics. For higher order facts these algorithms require iterative calculations. This does no longer want better order data and iterative calculations. The temporal shape of alerts is analyzed and the separation is done on this foundation.

The mixed sign is first of all converted to the time-frequency domain, additionally known as spectrogram of sign, via making use of Fourier remodel at short-time periods. Hamming window is used. In order to avoid blending of spectrograms each spectrogram is handled one by one. Correlation is performed on a majority of these short durations. A statement is simply a projection of supply alerts in positive direction. Reconstruction step is achieved on each separated sign's spectrogram. All the decomposed frequency components are then combined. At the end permutation step is done for finding the relation between the separated indicators. The choice is made with the aid of using classifier.

```
spectrum_1=spectrum
spectrum_1.plot(high=147.85)
plt.ylabel('Frequency (Hz)')
```

```
Text(0,0.5,'Frequency (Hz)')
```



2.3 Feature Extraction

The method of converting an audio signal into a chain of feature vectors is called characteristic extraction system. The function vectors deliver temporal as well as spectral characteristic records about the audio signal. Feature vectors are calculated on window basis. The function choice has a notable effect at the performance of audio segmentation systems. Two types of features are calculated in this proposed work: Time-domain and Frequency-domain area features. To shape a characteristic vector these normalized functions are mixed.

Feature extraction step is performed on the separated signals obtained after pre-classification step. These separated signals are divided into nonoverlapping frames. These frames are used as classification unit. On the basis of the classification results segmentation is performed.

We used following function for feature extraction:

```
def stFeatureExtraction(signal, fs, win, step):

    """
```

This function implements the shor-term windowing process. For each short-term window a set of features is extracted.

This results to a sequence of feature vectors, stored in a numpy matrix.

ARGUMENTS

signal: the input signal samples

fs: the sampling freq (in Hz)

win: the short-term window size (in samples)

step: the short-term window step (in samples)

RETURNS

st_features: a numpy array (n_feats x num Of Short Term Windows)

"""

2.4 Segmentation

2.4.1 pyAudioAnalysis:

[1] Silence removal. A semi-supervised silence removal functionality is also provided in the library. Their respective function takes an uninterrupted audio recording as input and returns segment endpoints that correspond to individual audio events, removing "silent" areas of the recording. This is achieved through a semi-supervised approach which performs the following steps:

• The short-term features of the whole recording are extracted

• An SVM model is trained to distinguish between high-energy and low-energy short-term frames. In particular, 10% of the highest energy frames along with the10% of the lowest are used to train the SVM model

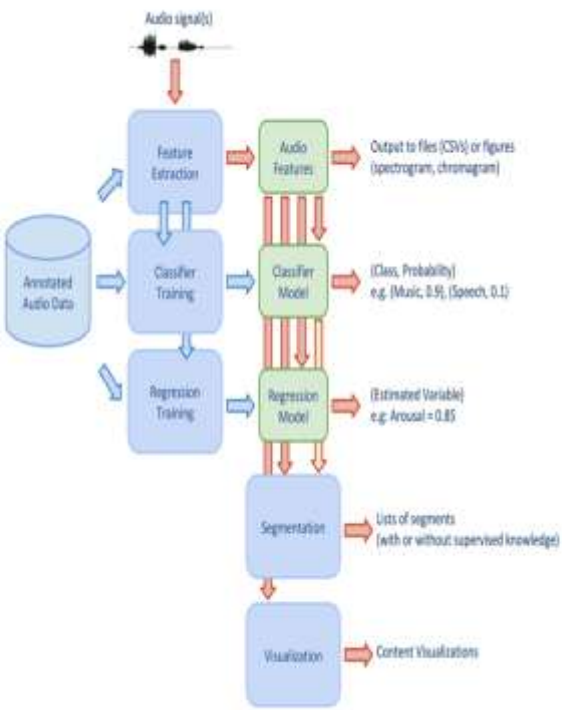
• The SVM classifier is applied (with a probabilistic output) on the whole recording, resulting in a sequence of probabilities that correspond to a level of confidence that their respective short-term frames belong to an audio event (and do not belong to a silent segment).

• A dynamic thresholding is used to detect the active segments.

2.4.2 Pseudo code for segmentation

```
[Fs,x]= aIO.readAudioFile("190406051420888859203 0.wav")

segments = aS.silenceRemoval(x, Fs, 0.025, 0.025, smoothWindow = 0.85, plot = True)

#type(segments)

    #print(Fs) Fs: the sampling rate of the generated WAV files

#adding label

n=1

for i in segments:

#indexing of data to avoid using csv files and manual editing

i.insert(0,n)

#insert appends at the beginning

n=n+1

    print("The segmented wav files written are:")

import scipy.io.wavfile as wavfile

[Fs, x] = aIO.readAudioFile("190406051420888859203 0.wav")

for j in segments:

T1 = float(j[1])#timestamp 1

T2 = float(j[2])#timestamp 2

label = ("190406051420888859203 0.wav", j[0], T1, T2)

        xtemp=x[int(round(T1*Fs)):int(round(T2*Fs))]

#time x (sapmle/time)

        print (T1, T2, label, xtemp.shape)

        wavfile.write(label, Fs, xtemp)

print("the labeled limits are:\n",segments )
```

3. Proposed Flow Model Based on Segmentation:



4. CONCLUSION

The paper presents selected aspects of Audio classification i.e segmentation algorithm, pre-classification step, feature extraction step, and steps used for discrimination are discussed. The result showed a wide range of audio analysis functionalities that can be used in several applications. Using segmentation one can classify an unknown audio segment to a set of predefined classes. Also, segmentation of an audio recording can classify homogeneous segments and remove silence areas from an audio recording.

ACKNOWLEDGEMENT

All the project members express their deep sense of gratitude to Dr. R. N. Patil (The Principal of Bharati Vidyapeeth College of Engineering Lavale), Prof. Uday Patkar (Department Head), Prof. Prajakta Pawar for her keen interest in promoting our project in Savitribai Phule university and have always been an inspiration for achieving great heights.

REFERENCES

[1] pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis by Theodoros Giannakopoulos * Computational Intelligence Laboratory, Institute of Informatics and Telecommunications, NCSR Demokritos, Patriarchou Grigoriou and Neapoleos St,AghiaParaskevi,Athens,15310,Greece

[2] Multi-Channel EEG Signal Segmentation and Feature Extraction Aleˇs Proch´azka∗, Martina Mudrov⍰a∗, Oldˇrich Vyˇsata∗∗, Robert H⍰ava∗, and Carmen Paz Su⍰arez Araujo∗∗∗ ∗ Institute of Chemical Technology in Prague, Department of Computing and Control Engineering,

[3] Automatic initial segmentation of speech signal based on symmetric matrix of distances D.Peleshko, Y.Pelekh, M.Rashkevych, Y.Ivanov, I.Verbenko Professor, Dmytro Peleshko, Department of Publishing Information Technologies, Institute of Computer Science and Information Technologies, Lviv Polytechnic National University, Lviv, Ukraine.

[4] Dorohin, O.A., Starushko, D.H., Fedorov, E.E., Shelepov, V.Y. 2000 The segmentation of the speech signal. "Artificial Intelligence".– №3. – 450- 458 pp.

[5] Performance analysis of character segmentation approach for cursive script recognition on benchmark database Amjad Rehman∗, Tanzila Saba Department of Computer Graphics and Multimedia, Universiti Teknologi Malaysia, Skudai, Malaysia.