Quest Journals
Journal of Electronics and Communication Engineering Research
Volume 6 ~ Issue 2 (2020) pp: 46-50
ISSN(Online) : 2321-5941
www.questjournals.org

Research Paper

An Overview of Digital Signal Processing Segmentation

Rekha P [1], Vishalakshi V [2]

[1] *Department of Electrical and Electronics Engineering, Government polytechnic Arakere, Karnataka, India.*
[2] *Department of Electrical and Electronics Engineering, Government polytechnic Channapatna, Karnataka, India.*

ABSTRACT
Whenever humans find a way to make life more convenient and pleasant, they often create a new era. There is a growing need for methodologies that can automatically analyze audio content, such as audio event recognition for determining object distance, depth, and the height of wells and valleys. This is because audio information is increasingly important in the vast amount of digital content available today. As an example of a non-stationary process, consider sound waves. The purpose of this report is to demonstrate how to measure the distance between two unknown objects using the physics of sound waves. We used approximated sound waves that were segmented using pyAudioAnalysis to determine fundamental frequencies. Through the use of digital signal processing and machine learning, we can filter the signals that are eco-sound wave measurements. Digital signal processing is a widely utilized method for segmentation since it can distinguish between desired echos, noise, and quiet. Additional signal analysis is required to align and translate the signals into a time-frequency graph. It is crucial to get the highest level of precision possible since the fundamental frequency selection is done on a frame basis. A summary and analysis of this technique for audio segmentation and fundamental frequency signal prediction are presented in this study.

Keywords: *Digital signal processing, Machine learning, Segmentation, Fundamental frequency estimation, Time-Frequency graph.*

I. INTRODUCTION

Digitized signal processing, sometimes known as DSP, is a numerical method of signal processing that is often used to quantify characteristics derived from compressed continuous analog data. Digital signal processing allows for the representation of signals as discrete-time, discrete-frequency, or other discrete domain signals in the form of a series of numerical symbols. In order to implement numerical techniques, digital signals were needed, for example, to create an analog to digital converter. There are subfields within signal processing, such as digital signal processing and analog signal processing. Some examples of DSP applications include processing signals involving voice and audio, processing signals involving sonar and radar, estimating spectra, doing statistical analyses, processing digital images, and processing signals for communication.

The majority of audio assessment tools rely on segmentation as a critical processing phase. Separating a continuous audio source into smaller, more manageable pieces is the goal. If the segmentation is supervised, then the input indications are categorized and sectioned using some type of supervised data. One approach to do this is by using a classifier to assign labels to restore-sized segments that are added to a set of predetermined lessons. Another option is to use a Hidden Markov Model (HMM) to accomplish combined segmentation-category.

For tasks when a supervised model isn't available, such as speaker dualization, the identified segments are clustered, making the process unsupervised. It may classify the uses of audio content analysis into two broad groups. One component is to separate a continuous audio stream into homogeneous regions, and the other is to separate a speech pattern into segments containing distinct speakers.

There are now three primary groups that audio segmentation algorithms fall under. The design of classifiers takes place in the main class. Classifiers are used to distinguish audio signals mostly according to their content once the functions are extracted in the time domain and frequency domain. Classifiers use the skills extracted from statistics in the second kind of audio segmentation. Posterior probability-based characteristics describe these types of skills. In order for the classifier to provide accurate results, a large amount of observed data is required. The implementation of efficient classifiers is the primary focus of the audio segmentation rule set. Bayesian facts criteria, Gaussian chance ratio, and a hidden Markov model (HMM) classifier are the classifiers used in this category. When given large amounts of training data, these classifiers also provide outstanding results.

Advanced performance systems are necessary for the analysis of overlaid speech, which is a complex problem. Audio segmentation is an essential preprocessing step in most audio processing applications. It also has a significant effect on how well frequency recognition works. For that reason, we provide an efficient and quick audio class and segmentation method that can work with multimedia packages in real time. Silence, natural-eco, noise, and environmental sound are the four main categories into which the incoming audio is divided. We provide a set of guidelines that, with little training data, may achieve high accuracy—that is, a low misclassification rate.

Two sliding overlapping windows and the detection of changes in signal attributes form the basis of the proposed technique of signal segmentation [2]. To improve accuracy, most studies combined segmentation methods with intelligent techniques like support vector machines, neural networks, and others. At the current stage of development of methodologies and tools for building artificial intelligence systems, voice signals processing plays a key role [3]. Methods for the analysis and synthesis of voice data often need real-time assistance. Consequently, linear prediction algorithms [5, 7] become quite popular and intriguing when applied to simulate speech signals. Transforming the signal into an autoregression model is the fundamental premise of these approaches. [4] A nonlinear object is any signal that cannot be determined. With a particular discretization period, however, it is always feasible to choose a temporal interval from the signals. A quasistationary interval is what we term this kind of interval. It is possible to construct a parametric model of the speech signal using the quasistationary interval and the qualities of the signal.

Classification and segmentation of audio has several uses. Typical applications of content-based audio retrieval and categorization include the following: commercial music consumption, surveillance, the entertainment sector, audio archive management, and so on. With millions of datasets accessible online now, audio segmentation and categorization are essential tools for audio search and indexing. Audio categorization is used in the monitoring of broadcast news programs to aid in the efficient and correct navigation of broadcast news archives.

II. MATERIALS AND METHODS

Steps for audio segmentation and categorization

An audio sample may be classified into simple statistical kinds using the suggested hybrid classification system. There is a pre-classification phase that examines the audio clip's windowed bodies independently before typing. A normalized feature vector is subsequently obtained after the feature extraction stage. The usage of the hybrid classifier technique follows feature extraction. The first stage involves using bagged SVM to categorize audio clips/frames into natural-eco and noise categories. Since silence frames are prevalent in audio signals, the frequency section is evaluated using a rule-based completely classifier to separate the quiet frames from the pure-eco parts.

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

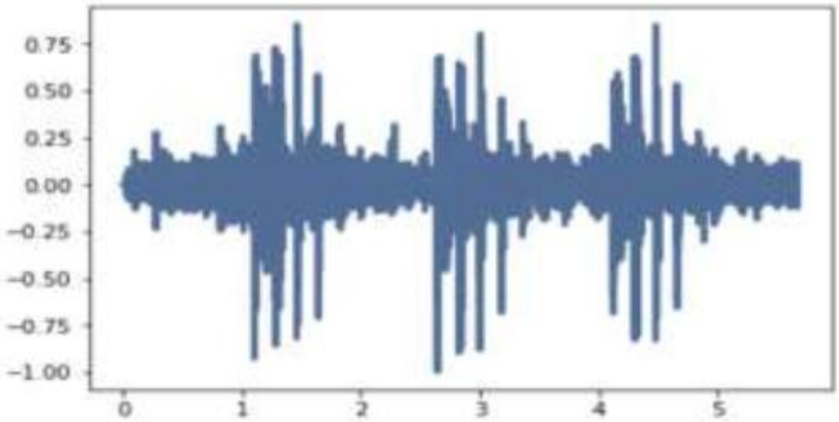

Fig. 1 : Audio clip's windowed.

Pre-classification

Because digital signals are overlaid, or in a combined form, it is possible to have a conversation in any setting with a lot of background noise and echoes. Cocktail party effect is another name for this. Blind source separation is a method used in the independent thing analysis framework to separate the source or desired parts.

In most cases, blind supply is used to decompose the combined sign into its component parts, even if the act of combination is not always acknowledged. The majority of methods for blind supply separation rely on higher-order statistics. These algorithms need to do iterative computations for more complex information. Improved order data and repeated computations are no longer desired. Separation is based on an analysis of the temporal form of warnings.

The first step is to utilize Fourier transform at short time intervals to transform the mixed sign into the time-frequency domain, which is also called the spectrogram of sign.
The hamming window is installed. Each spectrogram is processed independently to prevent their merging.

Most of these short periods are correlated. Making an optimistic prediction about supply alerts is all that a statement is. The spectrograms of the individual signs are used to complete the reconstruction process. All the frequency components that have been dissected are then merged. In order to discover the relationship between the separated indications, the permutation phase is completed at the end. A classifier is used to help make the decision.

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

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

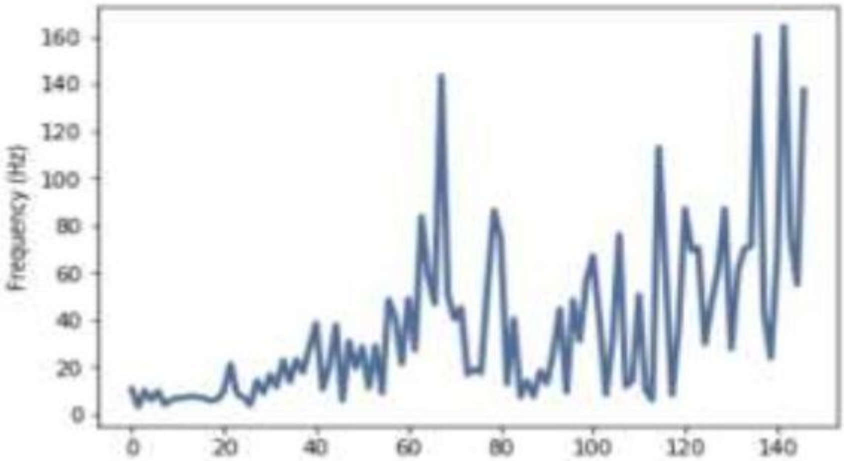

Fig. 2 : Signals in a combined form.

Segmentation
pyAudio Analysis

Breaking the silence. In addition, the library offers a semi-supervised noise reduction algorithm. Each one accepts an audio file that hasn't been interrupted as input and, after eliminating "silent" parts of the file, provides the segment endpoints that match to certain audio events. The following procedures comprise a semi-supervised technique that accomplishes this goal:

A support vector machine (SVM) model is trained to differentiate between low-energy and high-energy short-term frames. All of the recording's short-term characteristics are retrieved. For example, the support vector machine (SVM) classifier is trained using 10% of the highest energy frames and 10% of the lowest. Then, applied to the entire recording, the SVM classifier produces a series of probabilities that correspond to the confidence level that each short-term frame belongs to an audio event (rather than a silent segment). The active segments are detected via dynamic thresholding.

Pseudo code for segmentation

```
[Fs,x]= aIO.readAudioFile("1904060514208888592030.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("1904060514208888592030.wav")
for j in segments:
T1 = float(j[1])#timestamp 1
T2 = float(j[2])#timestamp 2
label = ("1904060514208888592030.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 )
```

1. Proposed Flow Model Based on Segmentation:

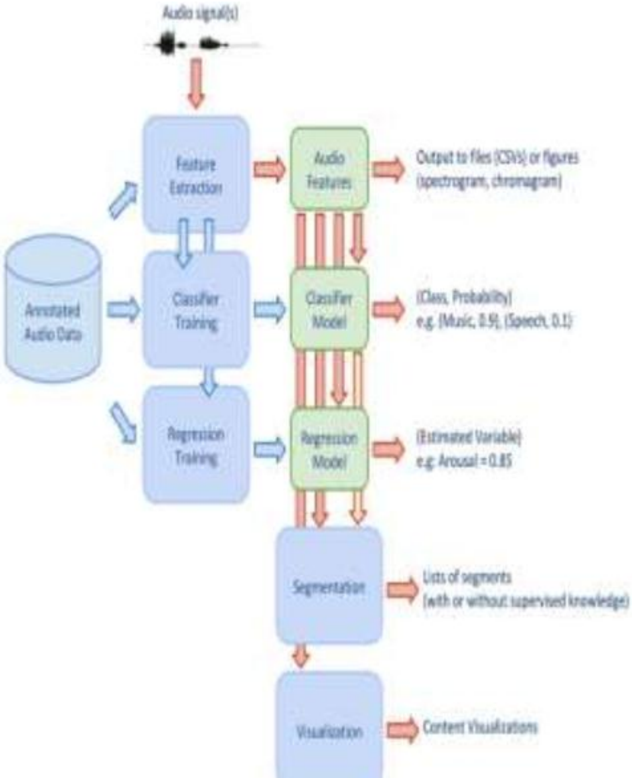

Fig. 3 : Proposed Flow Model Based on Segmentation.

III. Conclusions

In this study, several parts of audio classification are described, including the segmentation algorithm, the pre-classification phase, the feature extraction process, and the stages that are utilized for discriminating. A broad variety of audio analysis functions that are applicable to a variety of applications were exposed as a consequence of the investigation. Through the use of segmentation, it is possible to match an unidentified audio clip to a collection of predetermined categories. In addition, the process of segmenting an audio recording may be used to categorize segments that are similar to one another and eliminate quiet sections from the recording.

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 Ale˘s Proch´azka, Martina Mudrov´ a∗ , Old˘rich Vy˘sata∗ ∗ , 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.

[6]. J.R. Black , Electromigration —a brief survey and some recent results, IEEE Trans. Electron. Devices 16 (4) (1969) 338–347.