Digital Audio Authentication Forensics System using Microphone and Environment

[1] H. Ateeq Ahmed [2] Chintagunti Praveen
[1] Associate Professor, [2] P.G.Scholar
[1,2,]DEPARTMENT OF CSE

[1,2]Dr. K.V SUBBA REDDY INSTITUTE OF TECHNOLOGY
Dupadu Railway Station, Lakshmipuram Post, Kurnool, Andhra Pradesh, India - 518218

ABSTRACT

A wide range of digital audio authentication applications are emerging as a preventive and detective control in real-world situations, such as falsified evidence, infringement of copyright protection, and illegal data access, as a result of the ongoing rise in sophisticated forging. This research introduces a revolutionary automatic authentication method that can distinguish between legitimate and fake audio in order to examine and validate. Three psychoacoustic principles of hearing, which are used to mimic the human sound perception system, serve as the foundation for the design philosophy of the proposed system. The suggested method can also distinguish between sounds from various surroundings that was recorded with the same microphone. The computed features based on psychoacoustic principles of hearing are dangled to the Gaussian mixture model to generate automatic judgments in order to authenticate the audio and environment categorization. It is important to note that the proposed approach authenticates a speaker without regard to the audio content, i.e., without the need of a narrator or text. Audios in multiple environments are fabricated in such a way that a human cannot recognize them in order to assess the efficacy of the suggested system

.

1. INTRODUCTION

SCOPE OF THE PROJECT

Industrial accidents are often fatal and result in significant financial loss. Those that take place at work may be harmful to the environment, the workers, or the equipment. Data on accidents, illnesses, and fatalities caused by industrial activities show that efforts must be made to reduce these occurrences while also taking effective precautions. According to the International Labor Organization, a worker dies from an occupational injury every three minutes, and at least four other workers are injured every second (ILO).

OBJECTIVE

The major goal is to use a wide range of digital audio authentication applications as a preventative and investigative control in situations where forgery is a serious threat, such falsified documents, copyright violations, and illegal data access. This research introduces a revolutionary automatic authentication method that can distinguish between legitimate and fake audio in order to examine and validate.

EXPLANATION

With the recent unprecedented proliferation of smart devices such as mobile phones and advancements in various technologies (e.g., mobile and wireless networks), digital multimedia is becoming an indispensable part of our lives and the fabric of our society. For example, unauthentic and forged multimedia can influence the decisions of courts as it is admissible evidence. With continuous advancements in ingenious forgery, the authentication of digital multimedia (i.e., image, audio and video) is an emerging challenge. Despite reasonable advancements in image and video, digital audio authentication is still in its infancy. Digital authentication and forensics involve the verification and investigation of an audio to determine its originality (i.e., detect forgery, if any) and have a wide range of applications. For example,the voice recording of an authorized user can be replayed or manipulated to gain access to secret data. Moreover, it can be used for copyright applications such as to detect fake MP3 audio. Audio forgery can be accomplished by copy-move,deletion, insertion, substitution and splicing . The applications of copy-move forgery are limited compared with other methods as it involves moving a part of the audio at other location in the same liaison. On the other hand, the deletion, insertion, substitution and splicing of forged audio may involve merging recordings of different devices, speakers and environments. This paper deals with a splicing forgery (i.e., insertion of one or more segments to the end or middle), which is more challenging. The primary objective of the proposed system is to address the following issues with high accuracy and a good classification rate: _ Differentiate between original and tampered audio generated by splicing recordings with the same microphone and different environments. Environment classification of original and forged audio generated through splicing. Identify forged audio irrespective of content (i.e., text) and speaker. Reliable authentication with forged audio of a very short duration (i.e., _5 seconds). In the past, audio authentication has been achieved by applying various algorithms .One of the basic approaches is the visual investigation of the waveform of an audio to identify irregularities and discontinuities. For example, the analysis of spectrograms may reveal irregularities in the frequency component during the investigation. Similarly, listening to audio may also disclose abrupt changes and the appearance of unusual noise. These methods may help to decide whether the audio is original or tampered. However, one of the prime limitations of These approaches is that they are human-dependent, where judgement errors cannot be ignored. Moreover, the availability of sophisticated manipulation tools makes it convenient to manipulate audio without introducing any abnormalities. Consequently, it becomes very difficult to identify those abnormalities.

2. LITERATURE SURVEY

2.1 An Overview on Image Forensics

The aim of this survey is to provide a comprehensive overview of the state of the art in the area of image forensics. These techniques have been designed to identify the source of a digital image or to determine whether the content is authentic or modified, without the knowledge of any prior information about the image under analysis (and thus are defined as passive). All these tools work by detecting the presence, the absence, or the in congruence of some traces intrinsically tied to the digital image by the acquisition device and by any other operation after its creation. The paper has been organized by classifying the tools according to the position in the history of the digital image in which the relative footprint is left: acquisition-based methods, coding-based methods, and editing-based schemes.

2.2 Authenticationof Scalable VideoStreams With Low Communication Overhead

The large prevalence of multimedia systems in recent years makes the security of multimedia communications an important and critical issue. We study the problem of securing the delivery of scalable video streams so that receivers can ensure the authenticity of the video content. Our focus is on recent scalable video coding (SVC) techniques, such as H.264/SVC, which can provide three scalability types at the same time: temporal, spatial, and visual quality. This three-dimensional scalability offers a great flexibility that enables customizing video streams for a wide range of heterogeneous receivers and network conditions. This flexibility, however, is not supported by current stream authentication schemes in the literature. We propose an efficient and secure authentication scheme that accounts for the full scalability of video streams, and enables verification of all possible substreams that can be extracted from the original stream. In addition, we propose an algorithm for minimizing the amount of authentication information that need to be attached to streams. The proposed authentication scheme supports end-to-end authentication, in which any third-party entity involved in the content delivery process, such as stream adaptation proxies and caches, does not have to understand the authentication mechanism. Our simulation study with real video traces shows that the proposed authentication scheme is robust against packet losses, incurs low computational cost for receivers, has short delay, and adds low communication overhead. Finally, we implement the proposed authentication scheme as an open source library called svcAuth, which can be used as a transparent add-on by anymultimedia streaming application.

3. SYSTEM ANALYSIS

3.1 Existing System:

In the past, audio authentication has been achieved by applying various algorithms. One of the basic approaches is the visual investigation of the waveform of an audio to identify irregularities and discontinuities. For example, the analysis of spectrograms may reveal irregularities in the frequency component during the investigation. Similarly, listening to audio may also disclose abrupt changes and the appearance of unusual noise. These methods may help to decide whether the audio is original or tampered. However, one of the prime limitations of These approaches is that they are human-dependent, where judgment errors cannot be ignored. Moreover, the availability of sophisticated manipulation tools makes it convenient to manipulate audio without introducing any abnormalities.As a result, it becomes quite challenging to spot those irregularities.

3.1.1 Disadvantages Of Existing System:

Thevisual inspection of the waveform and spectrogram of the tampered audio depicted does not provide any clue of irregularity and hearing is also quite normal.

3.2 Proposed System:

The proposed system is able to classify between the audio of different environments recorded with the same microphone. To authenticate the audio and environment classification, the computed features based on the psychoacoustic principles of hearing are dangled to the Gaussian mixture model to make automatic decisions. It is worth mentioning that the proposed system authenticates an unknown speaker irrespective of the audio content i.e., independent of narrator and text. To evaluate the performance of the proposed system, audios in multi environments are forged in such a way that a human cannot recognize them. Subjective evaluation by three human evaluators is performed to verify the quality of the generated forged audio. The proposed system provides a classification accuracy of 99.2% ± 2.6.

Furthermore, the obtained accuracy for the other scenarios, such as text-dependent and text-independent audio authentication, is 100% by using the proposed system.

3.2.1Advantages Of Proposed System:

The system is also evaluated by using each recording of the developed forged.

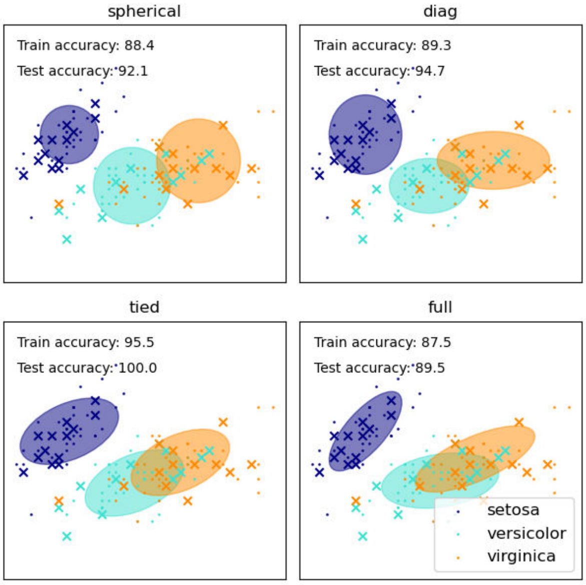
MODULES

- **Upload Forge/Real Audio Dataset**
  Using this model user Upload dataset
- **Preprocess dataset**
  Using this module dataset is preprocess here.
- **Feature Extraction**
  Using this module data is normalized here.
- **Load & Build Gaussian Mixture Model**
  Using this module data is loaded and Build GaussianMixture Model.
- **Audio Authentication**
  Using this module audio authentication here.

4. ALGORITHM

A Gaussian mixture model is a probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. One can think of mixture models as generalizing k-means clustering to incorporate information about the covariance structure of the data as well as the centers of the latent Gaussian's.Scikit-learn implements different classes to estimate Gaussian mixture models, that correspond to different estimation strategies, detailed below.The GaussianMixture object implements the expectation-maximization (EM) algorithm for fitting mixture-of-Gaussian models. It can also draw confidence ellipsoids for multivariate models, and compute the Bayesian Information Criterion to assess the number of clusters in the data. A GaussianMixture.fit method is provided that learns a Gaussian Mixture Model from train data. Given test data, it can assign to each sample the Gaussian it mostly probably belongs to using the GaussianMixture.predict method.

The GaussianMixture comes with different options to constrain the covariance of the difference classes estimated: spherical, diagonal, **tied or full covariance.**



5. SYSTEM DESIGN

5.1 SYSTEM ARCHITECTURE:

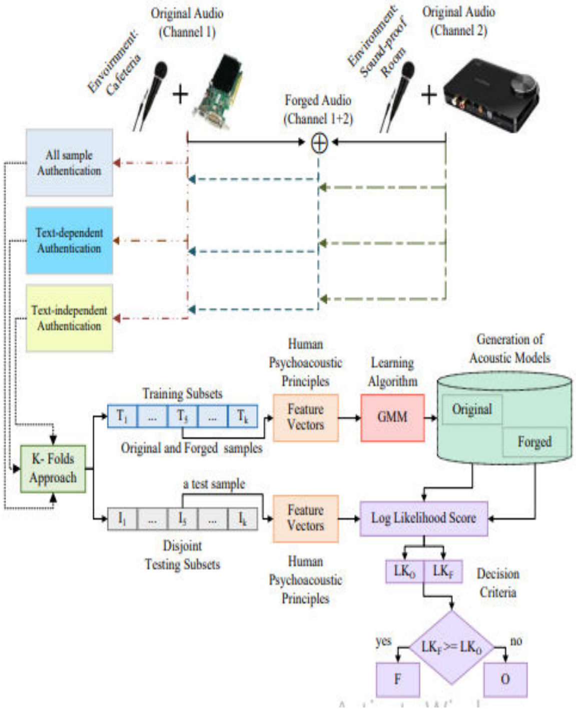

Original Audio (Channel 1)
Envoirnment: Cafeteria
Environment: Sound-proof Room
Original Audio (Channel 2)
Forged Audio (Channel 1+2)
All sample Authentication
Text-dependent Authentication
Text-independent Authentication
Human Psychoacoustic Principles
Learning Algorithm
Generation of Acoustic Models
Training Subsets
T1 ... T5 ... Tk
Original and Forged samples
Feature Vectors
GMM
Original
Forged
K- Folds Approach
a test sample
I1 ... I5 ... Ik
Disjoint Testing Subsets
Feature Vectors
Log Likelihood Score
Human Psychoacoustic Principles
LKO LKF
Decision Criteria
yes
LKF >= LKO
no
F
O

Fig:1. System Architecture

5.2 DATA FLOW DIAGRAM:

- The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.
- The data flow diagram (DFD) is one of the most important modeling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.
- DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.

- DFD is also known as bubble chart. A DFD may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow and functional detail.

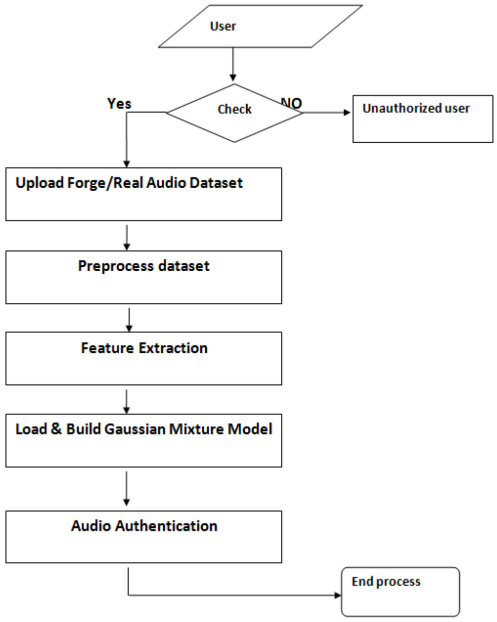
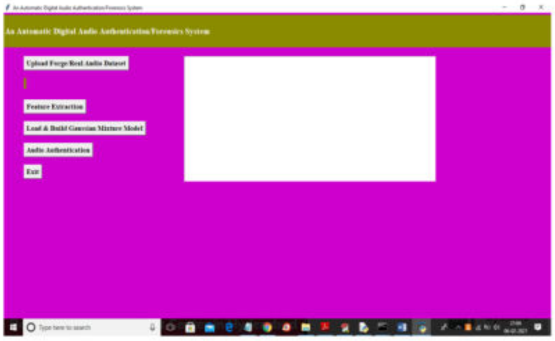
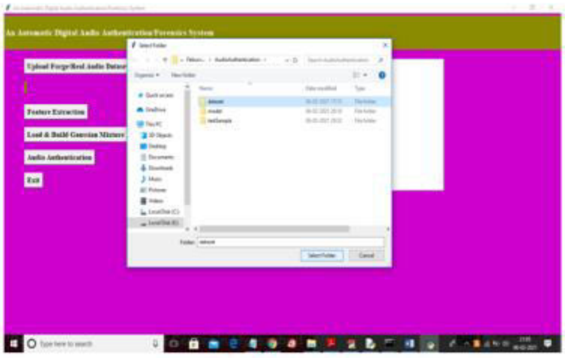


Fig:2. Data Flow Diagram

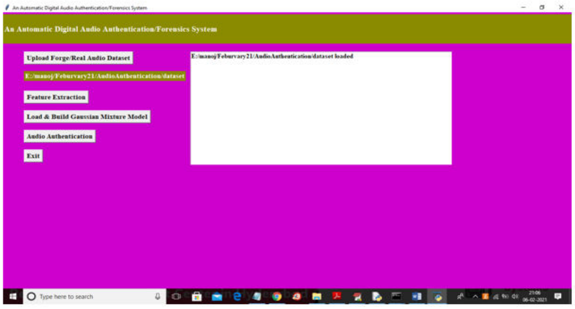
6. SCREENSHOTS


An Automatic Digital Audio Authentication/Forensics System
An Automatic Digital Audio Authentication/Forensics System
Upload Forge/Real Audio Dataset
Feature Extraction
Load & Build Gaussian Mixture Model
Audio Authentication
Exit

In above screen click on ‘Upload Forge/Real Audio Dataset’ button and then upload dataset folder.

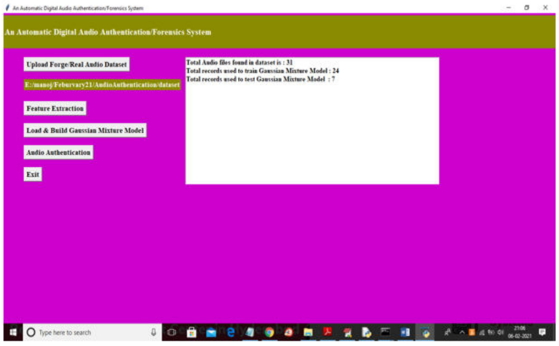

An Automatic Digital Audio Authentication/Forensics System
Upload Forge/Real Audio Dataset
Feature Extraction
Load & Build Gaussian Mixture Model
Audio Authentication
Exit
Select Folder
dataset
model
testSample
Select Folder
Cancel

In above screen selecting and uploading 'dataset' folder and then click on 'Select Folder' button to load dataset and to get below screen.


An Automatic Digital Audio Authentication/Forensics System
Upload Forge/Real Audio Dataset
E:/manoj/Feburvary21/AudioAuthentication/dataset
Feature Extraction
Load & Build Gaussian Mixture Model
Audio Authentication
Exit

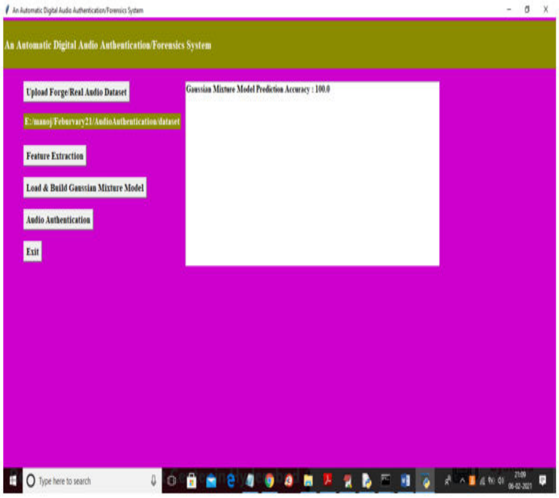
E:/manoj/Feburvary21/AudioAuthentication/dataset loaded

In above screen dataset loaded and then click on 'Feature Extraction' button to read audio files to extract features.


An Automatic Digital Audio Authentication/Forensics System
Upload Forge/Real Audio Dataset
E:/manoj/Feburvary21/AudioAuthentication/dataset
Feature Extraction
Load & Build Gaussian Mixture Model
Audio Authentication
Exit

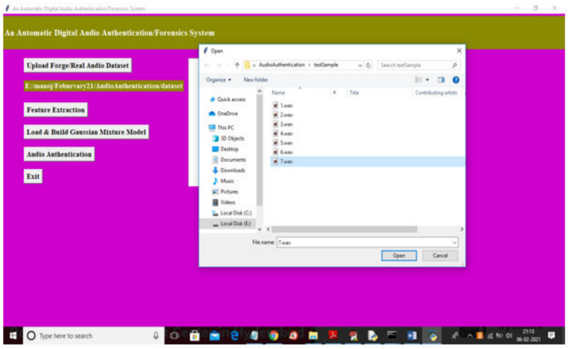
```
Total Audio files found in dataset is : 31
Total records used to train Gaussian Mixture Model : 24
Total records used to test Gaussian Mixture Model  : 7
```

In above screen all audio files features extracted and then application find total 31 audio files and the using 24 files to train GMM and 7 to test GMM. Now dataset ready with train and test records and how click on 'Load & Build Gaussian Mixture Model' button to train GMM model and calculate prediction accuracy.

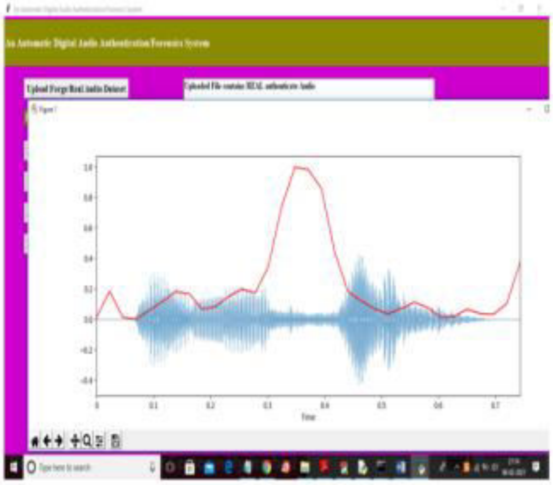

An Automatic Digital Audio Authentication/Forensics System
Upload Forge/Real Audio Dataset
E:/manoj/Feburvary21/AudioAuthentication/dataset
Feature Extraction
Load & Build Gaussian Mixture Model
Audio Authentication
Exit

Gaussian Mixture Model Prediction Accuracy : 100.0

In above screen GMM model generated and its prediction accuracy is 100% on test data and now click on 'Audio Authentication' button to upload new test file and perform prediction.


An Automatic Digital Audio Authentication/Forensics System
Upload Forge/Real Audio Dataset
E:/manoj/Feburvary21/AudioAuthentication/dataset
Feature Extraction
Load & Build Gaussian Mixture Model
Audio Authentication
Exit
Open
1.wav
2.wav
3.wav
4.wav
5.wav
6.wav
7.wav
File name: 7.wav
Open
Cancel

In above screen selecting and uploading '7.wav' file and then click on 'Open' button to get below result.


An Automatic Digital Audio Authentication/Forensics System
Upload Forge/Real Audio Dataset
Figure 1
Time

Uploaded File contains REAL authenticate Audio

In above screen text area we can see predicted result as ‘uploaded file contains REAL audio’ and you can see difference in above 2 graphs for forge and real. In forge graph due to tamper lots of fluctuation is there in red line and in second graph many fluctuations not there.

Similarly you can upload other files and test.

7. CONCLUSION

This paper proposed an automatic audio authentication system based on three human psychoacoustic principles. These principles are applied to original and forged audio to obtain the feature vectors, and automatic authentication is performed by using the GMM. The proposed system provides 100% accuracy for the detection of forged and audio in both channels. The channels have the same recording microphone but different recording environments. Moreover, an accuracy of 99% is achieved for the classification of the three different environments. In automatic systems based on supervised learning, the audio text is vital. Therefore, both the text dependent and the text-independent evaluation of the proposed system is performed. The maximum obtained accuracy is 100%. In all experiments, the speakers used to train and test the system are different (i.e., speaker-independent) and the obtained results are reliable, accurate and significantly outperform the subjective evaluation. The lower accuracy in the subjective evaluation also confirms that the forged audios are generated so sophisticatedly that human evaluators are unable to detect the forgery.

FUTURE SCOPE

In Future Work we will study other feature selection methods combined with more machine learning algorithms applied to real-time data from IoT devices.

8. REFERENCES

[1] Zulfiqar Ali, Muhammad Imran and Mansour Alsulaiman, “An Automatic Digital Audio Authentication System,” IEEE Access, 2017, vol 5, pp: 2994-3007.

[2] B. B. Zhu, M. D. Swanson, and A. H. Tewfik, "When seeing isn't believing [multimedia authentication technologies]," IEEE Signal Processing Magazine, vol. 21, pp. 40-49, 2004.

[3] A. Piva, "An Overview on Image Forensics," ISRN Signal Processing, vol. 2013, p. 22, 2013.

[4] A. Haouzia and R. Noumeir, "Methods for image authentication: a survey," Multimedia Tools and Applications, vol. 39, pp. 1-46, 2008.

[5] K. Mokhtarian and M. Hefeeda, "Authentication of Scalable Video Streams With Low Communication Overhead," IEEE Transactions on Multimedia, vol. 12, pp. 730-742, 2010.

[6] S. Gupta, S. Cho, and C. C. J. Kuo, "Current Developments and Future Trends in Audio Authentication," IEEE MultiMedia, vol. 19, pp. 50- 59, 2012.

[7]R. Yang, Y.-Q. Shi, and J. Huang, "Defeating fake- quality MP3," presented at the Proceedings of the 11th ACM workshop on Multimedia and security, Princeton, New Jersey, USA, 2009.

[8] Q. Yan, R. Yang, and J. Huang, "Copy-move detection of audio recording with pitch similarity," in 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015, pp. 1782-1786.

[8] X. Pan, X. Zhang, and S. Lyu, "Detecting splicing in digital audios using local noise level estimation," in 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012, pp. 1841-1844.

[9] A. J. Cooper, "Detecting Butt-Spliced Edits in Forensic Digital Audio Recordings," in 39th International Conference: Audio Forensics: Practices and Challenges,

[10] D. Campbell, E. Jones, and M. Glavin, ``Audio quality assessmenttechniques-A review, and recent developments," *Signal Process.*, vol. 89,pp.1489-1500, Aug. 2009.

[11] R. C. Maher, ``Overview of audio forensics," in *Intelligent MultimediaAnalysis for Security Applications*, H. T. Sencar, S. Velastin, N. Nikolaidis, and S. Lian, Eds. Berlin, Germany: Springer, 2010, pp. 127-144.

[12] B. E. Koenig and D. S. Lacey, ``Forensic authentication of digital audiorecordings," *J. Audio Eng. Soc.*, vol. 57, pp. 662-695, Sep. 2009.

[13] Audacity Team. (2016). *Audacity(R): Free Audio Editor and Recorder. Version 2.1.2 Retrieved*. [Online]. Available: http://www.audacityteam.org/

[14] GoldWave Inc. (2016). *GoldWave: Digital Audio Editing Software. Version 6.24 Retrived on November 25, 2016 From*. [Online]. Available: https://www.goldwave.com/goldwave.php

[15] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NYUSA: Springer-Verlag, 2006.