Audio Forgery Detection Techniques: Present and Past Review

Prabhu R. Bevinamarad, M.S.Shirldonkar
Department of Computer Science and Engineering
BLDEA'S V.P. Dr. P.G. Halakatti College of Engineering and Technology, (Affiliated to Visvesvaraya Technological University, Belagavi-590018)
Vijayapur-586103, Karnataka, India
prabhu.rb@bldeacet.ac.in

Abstract—**The increase in low cost digital audio recorders, mobile applications and accelerating growth of the audio-based Internet of Things (IoT) has initiated ease in obtaining speech and audio data which will be later used for the purpose of identifying human traits, implementing voice authentication system and development of voice based embedded systems. On the other hand, the availability of free advanced mobile applications and audio editing software like Adobe, Audition CC, etc. enabling people to edit easily the meaningful content of audio recordings data for getting benefit from e-services or producing it in a courtroom for the purpose digital proof. Perhaps, most of the people do it for fun as well as the strong intention of hiding reality present. Moreover, the audio recording captured in a real-life today does not contain digital watermarking and signature content for authentication because of expensive procedure. Therefore, in recent years the researches focused more on developing active audio forgery detection techniques for copy-move and audio splicing forgeries to authenticate and verify for its genuineness. In this paper, put an effort to describe past and present developments in audio copy-move and audio splicing forgery techniques. The paper also presents an overview of audio forgeries, its classification, various post-processing operations used and available audio dataset for forgery sample preparation for benchmark testing.**

Keywords— Digital watermarking; copy-move forgery; Audio splicing;Post processing operations;

I. INTRODUCTION

In today's digital world multimedia data have gained significant importance due to the accelerating growth of cost-effective, smart, and tiny recording devices as well as the availability of free mobile applications. As per the statistics, at least more than 300 million multimedia data get circulated per day in social media to disseminate information and replace a traditional system with an automated one. On the other hand, protecting multimedia information from tampering by unauthorized users is a critical issue. Since many researchers have implemented different methods for detection and localizing image, video forgeries[19], and audio recording forgery detection techniques for the protection and authentication of multimedia information. Whereas very few number of literature have been focused on audio forgery detections techniques such as audio signal copy move forgery and splicing etc. Of course, digital watermarking and signature are widely used techniques [1] for image, video and audio data protection and authentication. Despite, most audio captured day today in a real life is random and does not embed digital watermarking or signature information to protect its modification. Moreover, the speech recordings which are captured in real-life scenarios are very long in length with background noise and acoustics artifacts embedded. While the tampered, copy-move and spliced part in audio recordings may be very short and hearing to this over and over again and identifying the forged part is difficult and need lots of time. Sometimes even the detection yields lower accuracy rates due to various post audio forgery operations used to hide forgery traces. The purpose of writing this paper is to give an insight on following concepts.

1. Overview of an audio forgery.
2. Classification of the digital audio forgeries.
3. Existing techniques to detect copy-move and splicing audio forgeries.
4. Audio forgery detection framework.
5. Available audio dataset and audio forgery samples preparation for benchmark testing.
6. Conclusion and future research issues.

The rest of the contents are arranged as follows: The section II focuses an overview of audio tampering. Section III, describe the existing techniques. Section IV, contains audio forgery detection framework. Section V, includes description of audio forgery detection technique. Section VI, consist of discussion on audio dataset and tampering samples preparation. Section VII, explore post processing during forgery, and finally Section VIII, comprise the conclusion and further research issues need to be addressed.

II. AN OVERVIEW OF AUDIO TAMPERING

Audio tampering methods are broadly classified into two categories. (i)Based on container (ii) Based on content [20]. Alteration of the structure of the audio file, file metadata, and related description is referred to as container-based tampering, while the modification of actual audio contents referred to as content-based tampering. Audio splicing, audio copy-move, and audio cutting are the most generally used techniques that can alter the contents of audio. In the case of audio splicing, an extrinsic audio source is added in between another audio source to modify the source content. In the latter two cases, no extrinsic audio sources are used to modify its contents instead; the part of the same audio content is copied/cut and pasted at some location within the same audio. The audio tampering artifacts used are usually hardly noticeable for human hearing. So it is essential to develop an audio Integrity verification techniques that can automatically detect various audio forgeries and expose artifacts used during audio forgery. The Fig. 1 depicts classification of audio forgery techniques. The Table I: contains the definition of audio tampering and types.

Fig. 1. Hierarchy of audio forgery

TABLE I. DEFINITION OF AUDIO TEMPERING, CONTAINER BASED AUDIO FORGERY AND CONTENT BASED AUDIO FORGERY

Terminology	Definition
Audio tampering	Computing technique that spread or edit a digital audio recording/speech recording.
Container based audio forgery	The modification of audio file structure, metadata and its description. For example time stamp, file format, and hex data of an audio file.
Content based audio forgery	Modification of actual audio contents i.e. bits and bytes of audio wave forms. such as time frequency, enhancement, copy-move, and splicing etc.
Copy move audio forgery	In this forgery, the semantic information of audio can be changed by copying some parts of audio segments and pasting at other locations within the same audio file. Here no extrinsic audio sources are being added.
Audio splicing forgery	Most popular and simple to-do forgery. Here forged audio is generated by assembling spliced segments obtained from other audio recordings.

The Fig. 2 depicts audio copy move forgery where the audio signal of word present at second position i.e. “two” is copied and moved at third position within same audio signal. The Fig. 3 depicts audio splicing forgery where the audio segment from another audio source is combined to form a target audio signal.

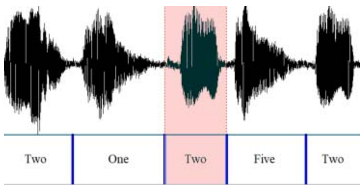

Fig. 2. Audio copy-move forgery sample

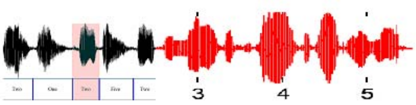

Fig. 3. Audio splicing forgery sample

III. RELATED WORK

Many steganography techniques are widely available to ensure protection and authentication of any type of digital content [1]. However, to apply these techniques for real-world captured audio recordings it is essential to have inbuilt support for hardware/software configurations that most non-professional audio recording devices lack. In recent years, many researchers have developed various active audio forgery detection techniques. The aim of this section is to give brief insight into past and present audio forgery detection techniques developed that are relevant to Content-based audio forgery.

The author in [2], proposed a method for microphone classification and environment classification and how these parameters are used to digital audio forgeries in [2]. One the other hand in [5], proposed method for microphone classification and identification of more than one microphone exist within a recorded audio content. The audio forgeries such as removal, addition, substitution and splicing detection based on verification of frame offset have been discussed in [3]. In [4, 8, 14, and 18] discussed techniques such as estimation of magnitudes of acoustic channel impulse response and local noise levels in audio signals to detect audio splicing forgeries. The author in [16], proposed a method based on channel response multi-feature to detect audio splicing localization. The Discrete wavelet packet decomposition and singularity points analysis of audio signals are described in [6 and 9] respectively to detect time-domain audio forgeries such as audio detection, insertion, substitution and splicing. The author in [7] proposed a method for estimation of reverberation time for detecting traces of editing in audio recording. Copy-move forgery is usually hardly noticeable since the forged audio segments are taken from the same speech recordings. In the paper [10 and 17] the similarities of pitch and formant sequences are used to detect copy-move forgeries in audio recording. Computation of histogram via LBP and comparison technique is used to identify locations of copy-move forgery in [11]. Discrete cosine transforms (DCT) and dimensionality reduction technique is presented in [12]. Discrete Fourier transform (DFT) transform and Pearson Correlation Coefficient (PCCs) measure to verify the similarity between audio segments to detect copy-move forgery is discussed in [13]. The author in [15], detailed a method based on analysis of electric network frequency (ENF) and an autoregressive (AR) model for audio tampering detection. The Table II summarizes various audio tampering techniques and issues addressed along need of further improvement.

TABLE II. SUMMARY OF EXISTING TECHNIQUES AND FUTURE ISSUES.

Ref#	Objective	Methodology	Detection Accuracy	Remarks/Issues to be addressed
[2]	Classification of multiple environment and microphone	Classify categories of recording environment and microphones	75.9%	Intra-room classifications, identification, room and different locations within a room and time of recordings made.

[3]	Audio segment deletion, insertion, substitution and splicing	frame offset checking	94%	Need to test with different bitrates, additive noise, and MP3 encoding
[4]	Audio Splicing	local noise level estimation method	0.06(SDO) 0.23(SFP)	This method reported robust and effective for both artificial and realistic audio splicing
[5]	Audio Tampering detection	Microphone classification	95%	During preprocessing de-reverberation algorithm could be considered for input file.
[6]	Time domain forgeries detection and localization	singularity points analysis of an audio signals	84.20%	Need to explore the dissimilarity between the intrinsic singularity points.
[7]	Detection of audio editing traces	Analysis of reverberation time	85%	Technique to transitory event and speaker moves.
[8]	Audio splicing	Analysis of magnitudes of acoustic channel impulse response and ambient noise	100%	This method yields better result for selected state of art.
[9]	Detect and locate time domain audio forgeries	Discrete wavelet packet decomposition and singularity analysis	87.80%	Need a method to choose designing parameters automatically such as inherent features of an audio file and low-pass filtering
[10]	Audio copy move forgery	Pitch tracking method	56.20%	Accuracy and efficiency of detection need to be improved by considering varying pitch sequence and more audio as characteristics.
[11]	copy-move audio forgery in detection and localization	Boundary point analysis and histogram computation	96.56%	Voice activity detection need to be refined to address wrong inference of the boundary points.
[12]	Audio Copy-Move forgery detection	Syllable division and comparison	98%	Robust and efficiency against common audio forgery attacks.
[13]	Audio Copy-Move forgery detection	Syllable division and comparison	99.40%	Effective features and optimize methods need to be developed to improve the accuracy and achieve quicker detection and better performance.
[14]	Audio Splicing	Analysis of environmental signature	96%	New techniques to be implemented to detect very small audio insertions, robustness against noise and the performance issue under anti-forensics
[15]	Audio Splicing, cutting and insertion	analysis of electric network frequency (ENF)	96.9%	Robust against noisy conditions and MP3 compression. The supervised learning framework can automatically render the parameters or the thresholds without the need for manually tuning.
[16]	Audio splicing detection	Estimation of channel response and multi-feature	85.54%	The precision as well as detection rate is improved for varying frame sizes.
[17]	Audio copy-move forgery detection	Similarity evaluation of pitch and formant frequencies	99.12%	Need to develop method to detect forgery in unvoiced speech segment and robust against post-processing operations for changing pitch and formant frequency.
[18]	Heterogeneous audio splicing detection	Local noise level estimation and similarity comparison	0.12(SDO) 0.21(SFP)	This method can accurately locate heterogeneous syllables

a. SDO and SFP represent sample detection omission rate and sample false positive rate respectively.

IV. AUDIO FORGERY DETECTION FRAMEWORK

The basic steps of the audio forgery detection process are depicted in Fig. 4. The main functions of the detection framework are shown in a dashed rounded rectangle: Initially, a queried audio signal is preprocessed using different frequency representations and scaling techniques. The preprocessed audio is divided into overlapping frames to extract features using MFCC, MDCT, DCT and logarithmic spectrum coefficients techniques for audio splicing detection. The prosodic features of the audio signals are extracted using voice activity detection (VAD) module and various discrete transform functions for audio copy-move forgery detection. Finally, the correlation between the feature vectors is calculated using the Pearson Correlation Coefficient (PCCs) and Euclidean distance measure along with the appropriate threshold value to detect suspicious part of an audio signal.

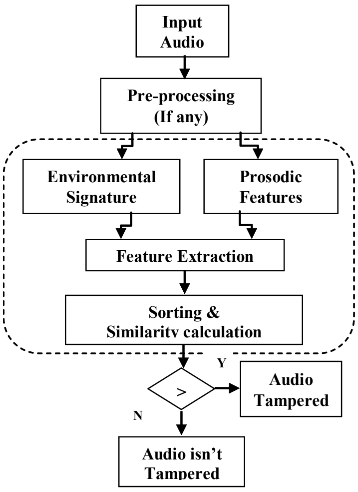

Fig. 4. General flowchart of audio forgery detection framework

V. DESCRIPTION OF AUDIO FORGERY DETECTION TECHNIQUE

This section provides a detailed description of audio splicing and audio copy-move forgery detection techniques.

A. *Audio forgery detection in time domain*

The following steps describe audio forgery detection such as audio splicing, deletion, insertion, and substitution using singularity point's analysis and discrete wavelet packet decomposition [6].

1. Initially, the queried audio speech is divided into(2^{n+1}-2) subbands using 'n' level wavelet packet decomposition technique.
2. Ensure decomposed wavelet subband has same length as of the queried audio speech signal so that local peak value of a subband can indicate forged location in time domain.
3. Obtain unprocessed wavelet subband and calculate the the corrosponding absolute value denote as Abs_subband using following equation.
4. Define Sub_Arr of with size equal to the number of samples in the speech signal and initialize to zero. And store decomposed wavelet subband for further processing. The Sub_Arri denote ith wavelet subband of Sub_Arri.
5. Find the mean of Abs_subband and set parameter 'N' to 5 using MS=mean(Abs_subband).
6. Compute maximum value 'M1' of the Abs_subband and maximum value of LEMR value using LEMR=max (Abs_subband/MS).
7. Check and go to step 11; if LEMR value is less than N; otherwise, if the maximum LEMR value is greater than kxN, then change the parameter N to half of the maximum LEMR value.
8. Find the maximum value (M2) of the Abs_subband in the ranges [n1; n2] and [-n2;-n1].
9. The ratio of M2 to M1 denoted by MLER less than the parameter P, the element present in the array SubArri that represents the position of M1 is set to "1" to claim that singular point corresponding to the M1 is forged.
10. Set points in the fluctuation region [-n1, n1] as zeros to avoid repeated detection of forged regions. Go back to step 6.
11. Go to step 12, if all wavelet subband of an array is completely processed. Otherwise, return to step 3 to process remaining elements of a Sub_Arr
12. The values '1' in SubArr indicate the forged region detected from the ith wavelet subband. Finally, perform 'OR' operation on SubArr to output forgery detection results..

B. *Audio Copy-move forgery detection.*

In this section, copy-move detection procedure based of discrete Fourier transform [13] technique is discussed. This method basically includes three steps: feature extraction, sorting of features and similarity comparison.

1) *Feature extraction:*

1.1. Initially input audio signal is divided into number of frames and DFT is used to transform audio sample from time domain into frequency domain and compute the short term power spectrum Pi(ω) using following equation $P_i(\omega)=|P_i(f)|^2$ where 'i' represents i^{th} frame.

1.2. Sum the energy of each filter by applying mel filterbank to power spectrun using equ."(1)" and calculate filterbank energy using equ."(2)"

$$M(f) = \tag{1}$$

$$\theta(Mk) = \sum_{k=1}^{K} |P_i(f)|^2 H_m \tag{2}$$

Where k means kth filter, k=1,2,3,.....,K and K is number of filters and H_m indicate mel filterbanks.

1.3. Calculate logarithm of all filterbank energies and take DCT of log filterbank energies and compute MFCC coefficients using equ. "(3)"

$$MFCC(n) = \sum_{k=1}^{K} X_k \cos\left[n(k-0.5\right. \tag{3}$$

Where 1<=n<=K/2.

2) *Sorting*

The segmented audio sample contains different aspect of features and almost all time the duplicated segments contains same feature as original segment. Therefore for ease of computation and reduce the time complexity the feature of segmented audio signal is sorted in a list. Also in order to achieve accuracy of detection only first 'β' points of segmented audio features are selected for sorting. Consider for an instance k-length feature sequence of one audio segment X (k)={X_0,X_1.....X_{k-1}} is extracted, then the characteristics value 'v' used for sorting is calculated using equ. "(4)"

$$v = \sum_{k=0}^{\beta-1} \jmath \tag{4}$$

Finally the quick sorting algorithm is selected for sorting feature sequences of audio signal.

3) *Similarity calculation*

The Pearson Correlation Coefficient (PCC) used to compute similarity at trend level of every feature sequences along with optimal threshold value using following equ.

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\left(\sum X^2 - \frac{(\sum X^2)}{N}\right)\left(\sum Y^2 - \frac{(\sum}{}\right.} \quad (5)$$

If PCC value 'r' is greater than appropriate threshold then the particular audio segment is classified as suspicious.

VI. Audio Tampering Datasets

One of the main challenges faced by the researcher during the evaluation of a forgery detection algorithm is the lack of publicly available and ready to use dataset. Therefore to evaluate the efficiency of the latest algorithm, one need to prepare audio tampering (e.g. audio copy-move and splicing forgery detection) samples using various standard speech recording corpus [21, 22, 23, and 24] are described in Table III. In this section, the general procedure involved in the preparation of digital audio tampering samples for copy-move and audio splicing forgery to facilitate the evaluation of audio forgery algorithms is discussed. The Fig. 5 shows a block diagram of an audio forgery making process for audio copy-move and splicing forgery.

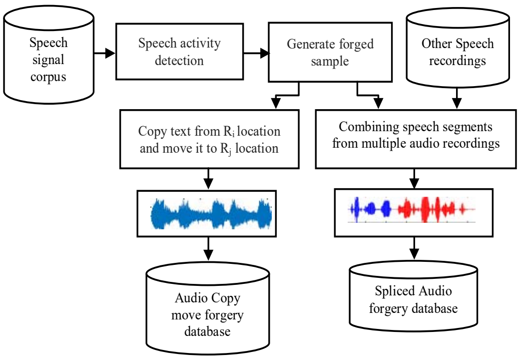

Fig. 5. Audio forgery making process

A. *Audio splicing forgery making process*

The Fig. 5. shows general steps involved in the generation of audio splicing forgery database. Currently, very few audio forgery databases are available for evaluating existing and new algorithms. The researchers are insisted to follow the general procedure for generating Audio splicing samples for the purpose of evaluation of algorithms. Initially, a speech sample is taken from any one of the speech recording corpus. The speech activity detection technique is used to find voiced and unvoiced regions of the speech signal to decide at which location the splice segment of another audio is assembled. As a result the spliced audio could be meaningful and doesn't comprise any perceptible distortion. Finally, the generated audio splicing samples are stored in a database for the purpose of evaluation.

B. *Audio Copy-move forgery making process*

Audio copy-move is another type of easy to do forgery operation commonly used during speech tampering. This kind of forgery is created using the same speech recording. The Fig. 5 shows basic steps involved in the creation of an audio copy-move forgery sample as well as a database. Initially, the speech sample from speech corpus is a preprocessed using speech activity detection module to extract pitch sequences for every syllable and words. Later generate forged regions by copying text from location 'R_i' and move it to 'R_j' location of the same speech recording.

TABLE III. Audio Database Description

Sl. No	Database Name	Description
1	Wall Street Journal database	Contains 400hr English Speech data i.e.13240 speech samples with 16 kHz sampling rate.
2	TIMIT Acoustic-Phonetic Continuous Speech Database	Contains American English and phonetically rich sentences in eight major dialects of 630 speaker's recordings.
3	Spanish Speech Database	Comprise audio recordings of 200 native Spanish speakers (100 males, 100 females).In total it includes 3,748 pronunciation lexicon headwords plus variants.
4	Arabic Speech Database (KSUD)	It contains Arabic speech data of 590 hours from 269 male and female speakers and stored as sequences of two-channel 48 kHz 16-bit FLAC compressed PCM Wav files.

VII. Post-Processing

Most of the early works of audio tampering detection has less considered different post-processing operations used during audio tampering in order to make the tampered artifact less susceptible as well as to hide the forgery traces from detection technique. The following are the different types of post processing operations used during audio tampering process such as noise addition, filtering, MP3 compressing, frequency scaling, pitch shift, and re-sampling. The Fig. 6 to Fig. 9 illustrates the audio copy move forgery with post processing operations applied. It can be observed from the Fig. 7 to Fig.9 the segmented audio waveform changes a lot in it frequency after post processing operation is performed.

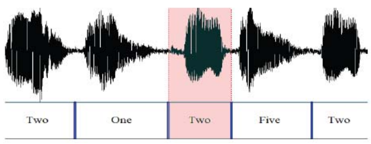

Fig. 6. Copy-move forgery

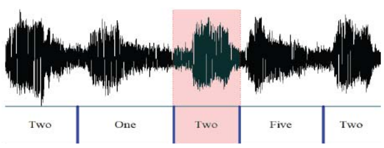

Fig. 7. Copy-move forgery with white gaussian noise addition

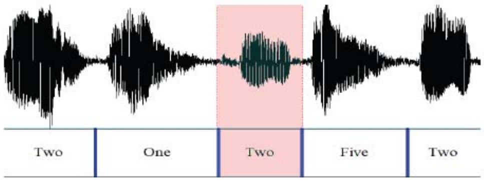

Fig. 8. Copy-move forgery with median filtering

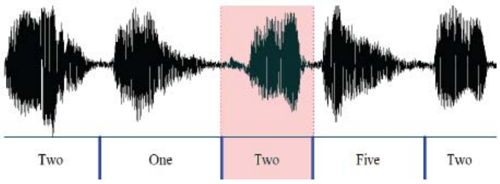

Fig. 9. Copy-move forgery with phase vcoder

VIII. CONCLUSION AND FUTURE RESEARCH ISSUES

In this survey, research works on audio tampering and its detection techniques especially copy-move and audio splicing forgeries have been reviewed. This survey details past and latest research developments in the field of audio forgery detection techniques followed by a general framework for audio forgery detection and a various standard audio dataset which can be used for the preparation of tampered audio samples. Lastly, it comprises the limitations and demand for future developments for unresolved issues in the field. This field is one of the emerging research area recently many researchers working on efficient algorithm development. I hope this survey would help many researcher, students and engineers who are currently working or interested in this field will benefit from the information presented in this paper.

The following are the key research issues to be considered and addressed for further improvement over existing system. They are,

1. Need to develop techniques to detect audio forgeries for unvoiced speech segments.
2. A novel approach need to be implemented that are robust against various audio signal post-processing operations.
3. There is a requirement of new methods for small audio segment insertions, deletion and splicing forgeries.

REFERENCES

[1] Kirovski and Malvar. Spread-spectrum watermarking of audio signals. IEEE Transactions on Signal Processing, 51(4):1020–1033, 2003.

[2] Kraetzer, Christian & Oermann, Andrea & Dittmann, Jana & Lang, Andreas. Digital audio forensics: A first practical evaluation on microphone and environment classification. MM and Sec'07 - Proceedings of the Multimedia and Security Workshop 2007.

[3] R. Yang, Z. Qu, and J. Huang, "Detecting digital audio forgeries by checking frame offsets," in Proceedings of the 10th ACM workshop on Multimedia and security. ACM, pp. 21–26,2008.

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

[5] Luca Cuccovillo, Sebastian Mann, Marco Tagliasacchi, and Patrick Aichroth. Audio tampering detection via microphone classification. In Proc. of 2013 IEEE 15th Int. Workshop on Multimedia Signal Processing (MMSP), pages 177–182. IEEE, 2013.

[6] J. Chen, S. Xiang, W. Liu, and H. Huang, "Exposing digital audio forgeries in time domain by using singularity analysis with wavelets," in Proceedings of the first ACM workshop on Information hiding and multimedia security. ACM, pp. 149–158, 2013.

[7] A. Ciobanu, T. Culda, C. Negrescu and D. Stanomir, "Analysis of reverberation time blind estimation used in audio forensics," 2014 11th International Symposium on Electronics and Telecommunications (ISETC), Timisoara, pp. 1-4,2014.

[8] Zhao, Hong & Chen, Yifan & Wang, Rui & Malik, Hafiz. Audio Source Authentication and Splicing Detection Using Acoustic Environmental Signature 2014.

[9] Jiaorong Chen, Shijun Xiang, Hongbin Huang, and Weiping Liu. Detecting and locating digital audio forgeries based on singularity analysis with wavelet packet. Multimedia Tools and Applications, 75(4):2303– 2325, 2016.

[10] Qi Yan, Rui Yang, and Jiwu Huang. Copy-move detection of audio recording with pitch similarity. In Proc. of 2015 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), pages 1782–1786. IEEE, 2015.

[11] Muhammad Imran, Zulfiqar Ali, Sheikh Tahir Bakhsh, and Sheeraz Akram. Blind detection of copy-move forgery in digital audio forensics. IEEE Access, 5:12843–12855, 2017.

[12] F. Wang, C. Li and L. Tian, "An algorithm of detecting audio copy-move forgery based on DCT and SVD," 2017 IEEE 17th International Conference on Communication Technology (ICCT), Chengdu, pp. 1652-1657.2017.

[13] Zihan Liu and Wei Lu. Fast copy-move detection of digital audio. In Proc. of 2017 IEEE Second Int. Conf. Data Science in Cyberspace (DSC), pages 625–629. IEEE, 2017.

[14] Hong Zhao, Yifan Chen, Rui Wang, and Hafiz Malik. Audio splicing detection and localization using environmental signature. Multimedia Tools and Applications, 76(12):13897–13927, 2017.

[15] Xiaodan Lin and Xiangui Kang. Supervised audio tampering detection using an autoregressive model. In Proc. of 2017 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), pages 2142–2146, 2017.

[16] Sanjay Kumar Rouniyar , Yu Yingjuan , Yongjian Hu, Channel Response Based Multi-Feature Audio Splicing Forgery Detection and Localization, Proceedings of the 2018 International Conference on E-Business, Information Management and Computer Science, Hong Kong, April 14-15, pp.46-53, 2018.

[17] Q. Yan, R. Yang and J. Huang, "Robust Copy–Move Detection of Speech Recording Using Similarities of Pitch and Formant," in IEEE Transactions on Information Forensics and Security, vol. 14, no. 9, pp. 2331-2341, Sept. 2019.

[18] X. Meng, C. Li and L. Tian, "Detecting Audio Splicing Forgery Algorithm Based on Local Noise Level Estimation," 2018 5th International Conference on Systems and Informatics (ICSAI), Nanjing, pp. 861-865,2018.

[19] Mulla, M.U.,Bevinamarad, P.R. Review of techniques for the detection of passive video forgeries. Int. J. Sci.Res. Comput. Sci. Eng. Inf. Technol, pp. 199–203, 2017.

[20] Zakariah, M., Khan, M.K. & Malik, H. Digital multimedia audio forensics: past, present and future. Multimed Tools Appl 77, 1009–1040 January 2018.

[21] Douglas B Paul and Janet M Baker. The design for the wall street journal-based CSR corpus. In Proc. of the workshop on Speech and Natural Language, pages 357–362. Association for Computational Linguistics, 1992.

[22] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, N. L. Dahlgren, and V. Zue. TIMIT Acoustic-Phonetic Continuous Speech Corpus. Linguistic Data Consortium, Philadelphia, 1993.

[23] P. Esquef, 1. Apolinario Jr., and L. Biscainho, "Improved edit detection in speech via ENF patterns," In IEEE International Workshop on Information Forensics and Security, pp.I-6, Rome, Italy, Nov. 2015.

[24] Alsulaiman, Mansour, et al. King Saud University Arabic Speech Database LDC2014S02. Hard Drive. Philadelphia: Linguistic Data Consortium, 2014.