Enhancement of crime voice sample by using ACCUSTECK Lab+ Audio Spectrometer

S. H. Mondal[1*], T. Bera[2], C. Sarkar[3]
[1,2,3]Physics Division, State Forensic Science Laboratory, Kolkata, India, 700037

Abstract

In this work, we aimed to improve the sound quality of the speech casually spoken by the criminals during a robbery at a factory using a voice spectrometer. The audio file included mentions of two local gangsters, who were later caught by the police, which helped advance the investigation. We achieved this objective by stretching the recorded speech over time, which allowed us to understand the names better. We have accomplished this task by elongating/expanding the recorded speech in the time scale/time domain, which led to the decipherment of names.

Keywords: Voice authentication, Levelogram analysis, Histogram analysis, DC-offset analysis.

I. Introduction

This section of the paper gives a brief introduction to it. The investigation agency faces challenges in identifying the accused individuals, and the State Forensic Science Laboratory, Kolkata, plays a crucial role in interpreting their names. The technical processes involved in doing so are followed by the success in the conclusion section.

II. Brief nature of the crime

On 10th January, 2023, a requisition was sent to the Administrator of the State Forensic Science Laboratory from the office of the Deputy Commissioner of Police, Central Division of Barrackpore Police Commissionerate, West Bengal, for the purpose of clearing the illegible voice, which was seized by the police during the time of the offense, by use of a voice spectrometer at the Physics Division. Some criminals were on their way to rob an industry situated under Rahara P. S., North 24 Pgs., West Bengal. The said industry was under the surveillance of CCTV. The miscreants entered into the factory and forcefully took away a huge amount of cash, amounting to about 3,400,000 INR, along with some valuable articles. During the course of the robbery inside the said industry, voices of some miscreants were caught in the recorders of the DVR of the CCTV, which lacked clarity and was accompanied by noises and overlapping voices.

The task that was endorsed by the State Forensic Science Laboratory, Kolkata, was to decipher the said voice to get any clue. We received a 25-second '.mp4' video clip featuring two male speakers conversing in the 'Bhojpuri' dialect during the night when the crime occurred at the

*Corresponding Author Email: shamimmondal709@gmail.com
Published: 18/11/2025
DOI: https://doi.org/10.70558/IJST.2025.v2.i4.241127

specified location.

Problem:

We were tasked to decipher the 25-second-duration video clip to obtain any clue/evidentiary material from it.

Methodology:

We initially extracted the audio part from the above video file by converting the extension to '.wav' format from '.mp4' format. Then we reduced the background noise of the extracted audio sample. Next, we amplified the audio signal by 200% through voice spectrometer software. Then we reduced the playback rate by up to 25% to extend the duration of the audio signal. Now, we split the whole sound signal into seven segments to decipher the content of the audio signal. After that, we coined each of the segments and amplified it to make it more legible. Along with these, we also fine-tuned the playback rate to get better computational outputs.

We represent the above procedure pictorially in a chronological manner.

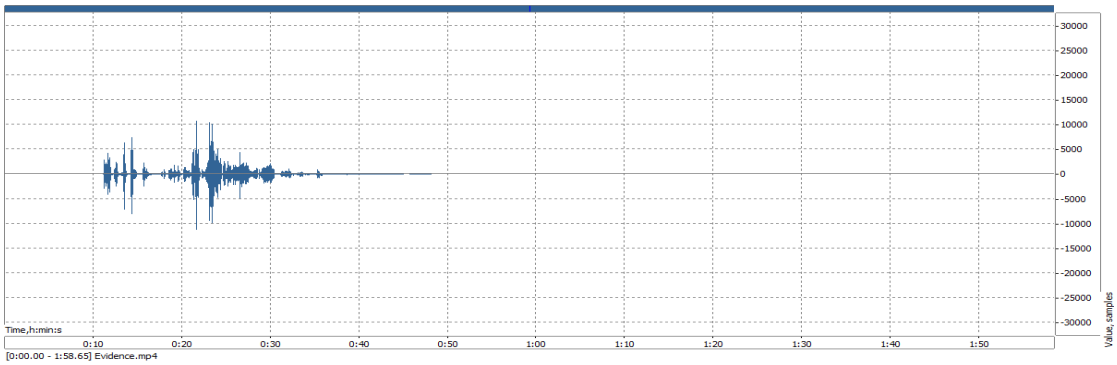


Fig.1: Plot of amplitude vs. time (in db-sec unit) of the extracted audio file.

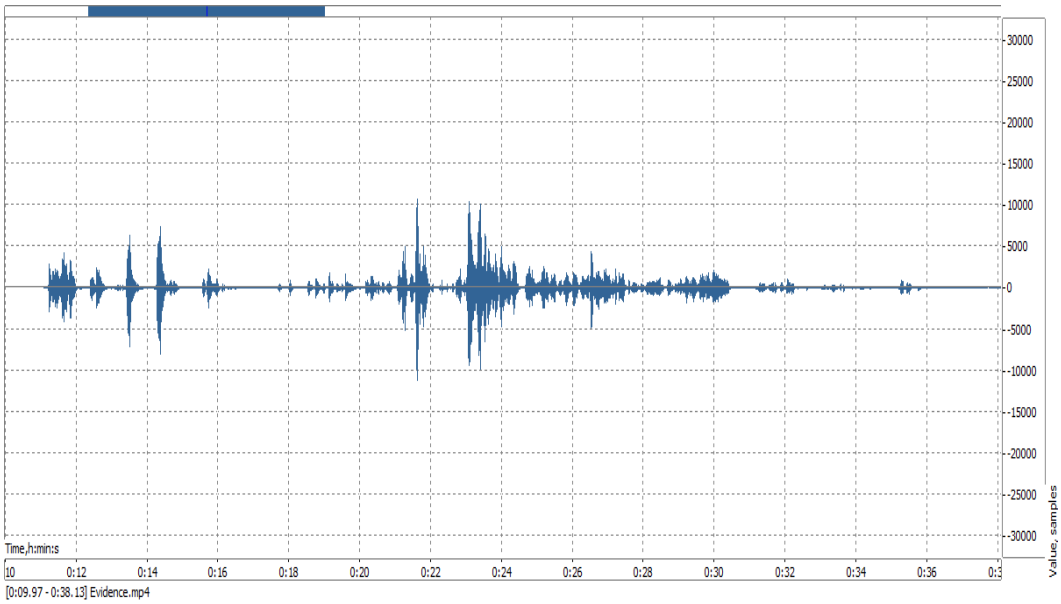


Fig. 2: Plot of amplitude vs. time (in db-sec unit) of the extracted audio file after expansion in time-scale, and filter.

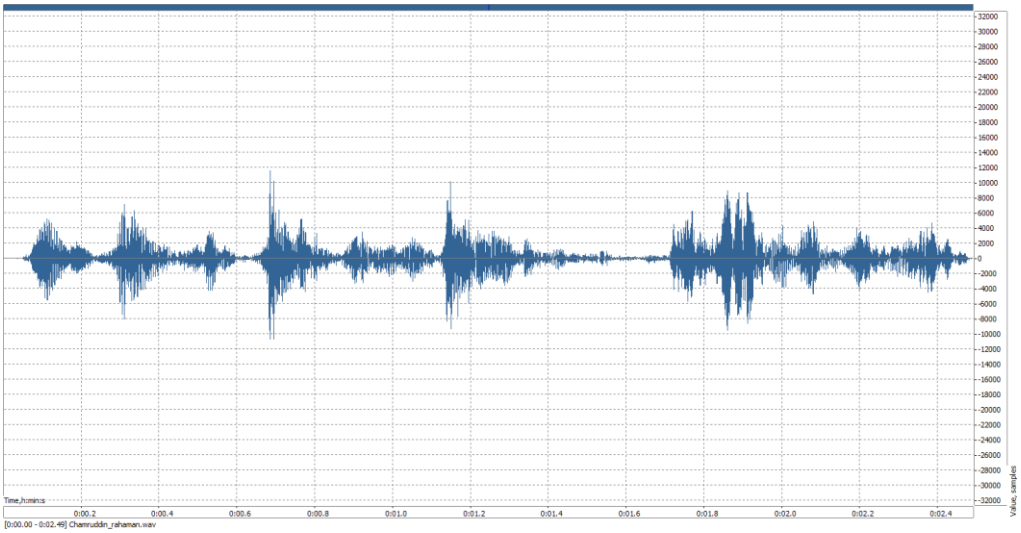


Fig. 3: Plot of amplitude vs. time (in db-sec unit) of the extracted audio file after expansion in time-scale, and f filter for the part **“SO JA LIYE FEK BHI DIYE”**

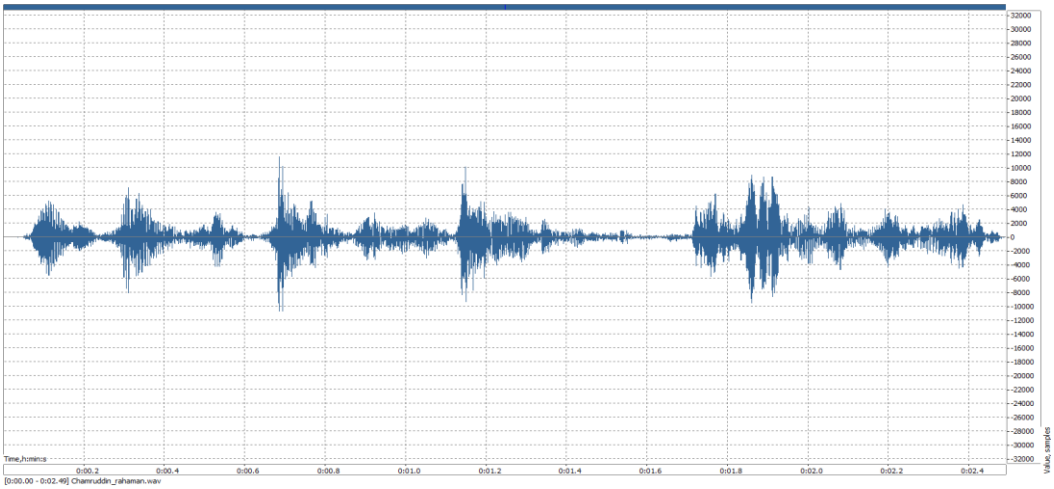


Fig.4: Plot of amplitude vs. time (in db-sec unit) of the extracted audio file after expansion in time-scale, and filter for the part **“chamruddin comission liya Rahamn bhi liya”**

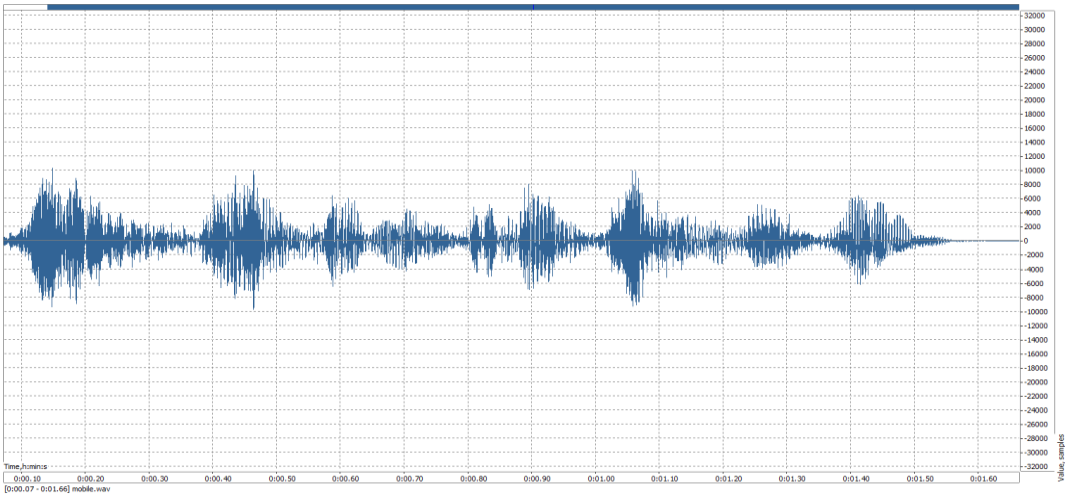


Fig.5: Plot of amplitude vs. time (in db-sec unit) of the extracted audio file after expansion in time-scale, and filter for the part **“mobile lene giya ka uha pher?”**

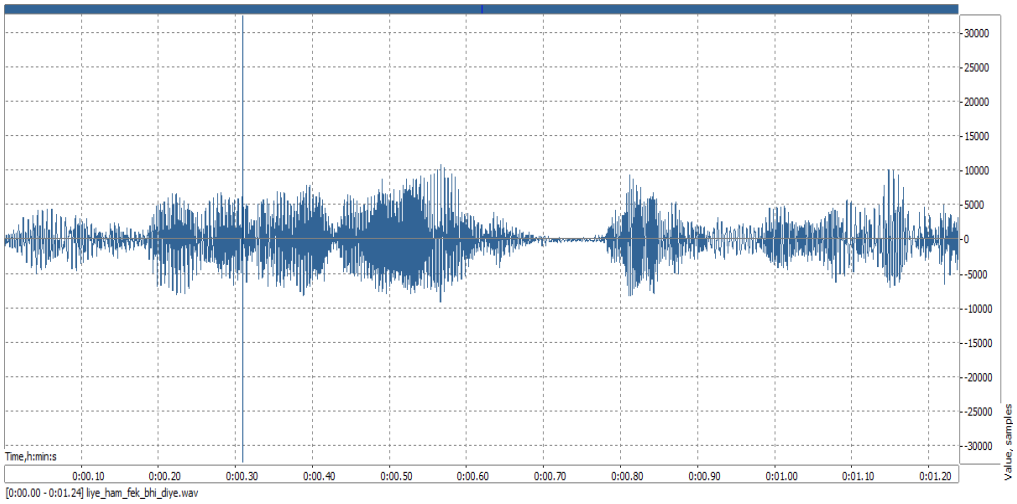


Fig.6: Plot of amplitude vs. time (in db-sec unit) of the extracted audio file after expansion in time-scale, and filter for the part "**lele hum fek vi diye"**

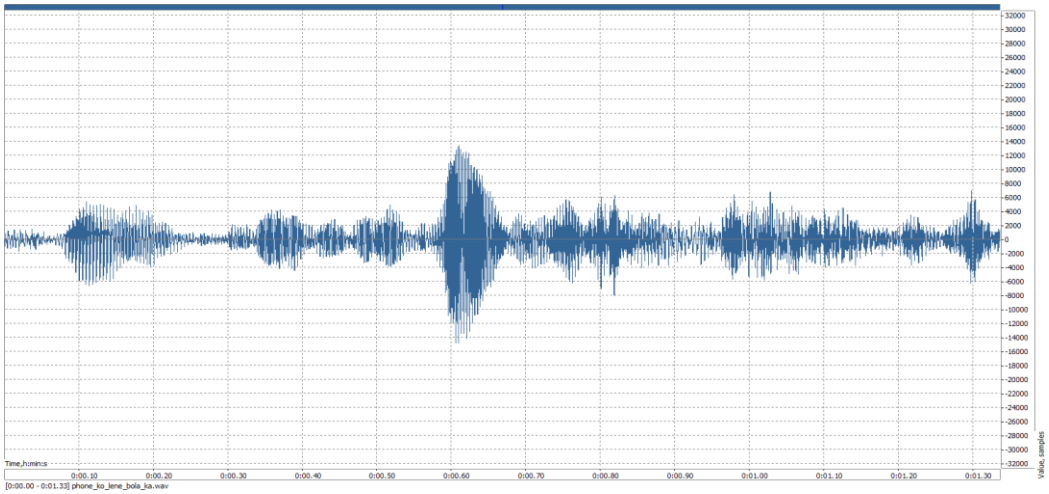


Fig.7: Plot of amplitude vs. time (in db-sec unit) of the extracted audio file after expansion in time-scale, and filter for the part "**phone ko lene bola ka"**

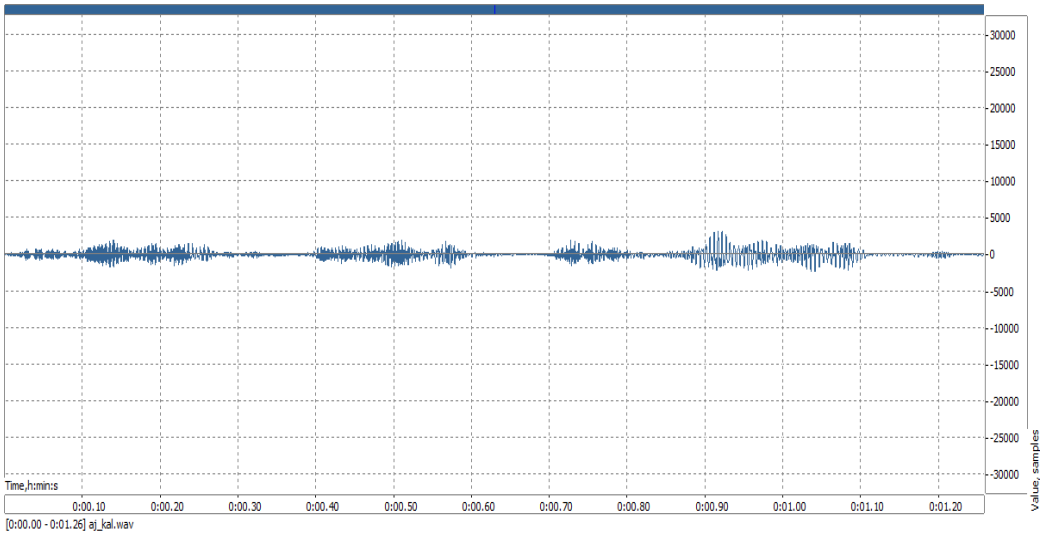


Fig. 8: Plot of amplitude vs. time (in db-sec unit) of the extracted audio file after expansion in time-scale, and filter for the part "**aj kal sabhi**"

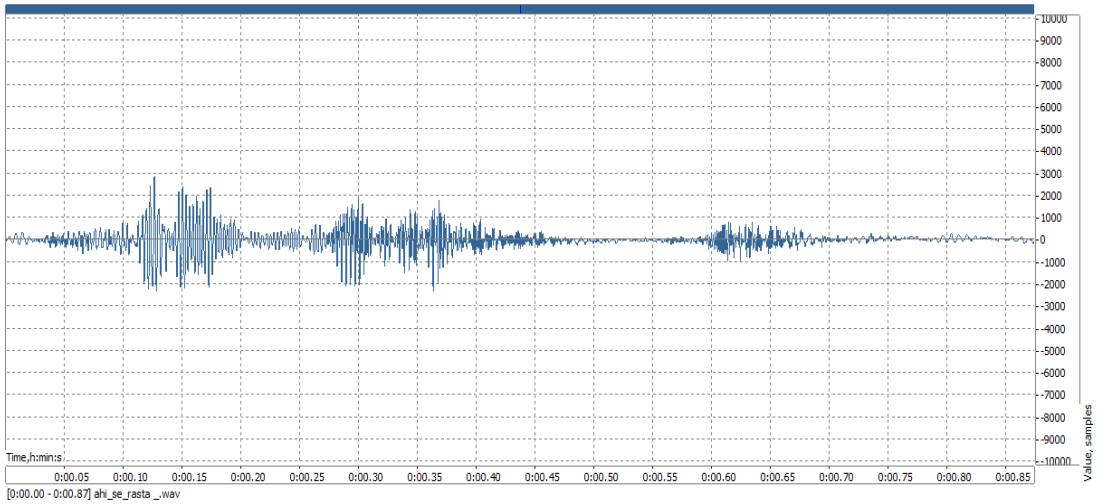


Fig. 9: Plot of amplitude vs time (n db-sec unit) of the extracted audio file after expansion in time-scale, and filter for the part "**ahi se rasta hay**"

AUTHENTICATION:

We were tasked to authenticate the available crime voice sample for different cases. To perform this task, we have several aural parameter analyses, like header file analysis, signal analysis, histogram analysis, levelogram analysis, DC-offset analysis [1-4], etc. Here, we briefly discuss the above analysis along with some graphs of each case. All the graphs are obtained with the help of ACUSTECK LAB+ software.

Header file analysis: It's gives clear meta data by which we easily known about coder.

Sample copy:

File: Evidence.mp4

File size: 2.13 MB (2230730 Byte)

Duration: ~ 1 min 59 s

Format name: QuickTime file format/MPEG-4

Number of streams: 2

Stream №0 video.ID: 1. Format: AVC, 10.11(2.78-25.00) frame/s, VFR, 1280x720, 143 kbit/s, VBR, Block Count: 1204.

- Codec:

- Profile: Main@4.0

- Duration: 1 min 59 s 80 ms

Stream №1 audio.ID: 2. Format: AAC, channels:1, 8000 Hz, 5 Kbit/s, VBR, Block Count: 928.

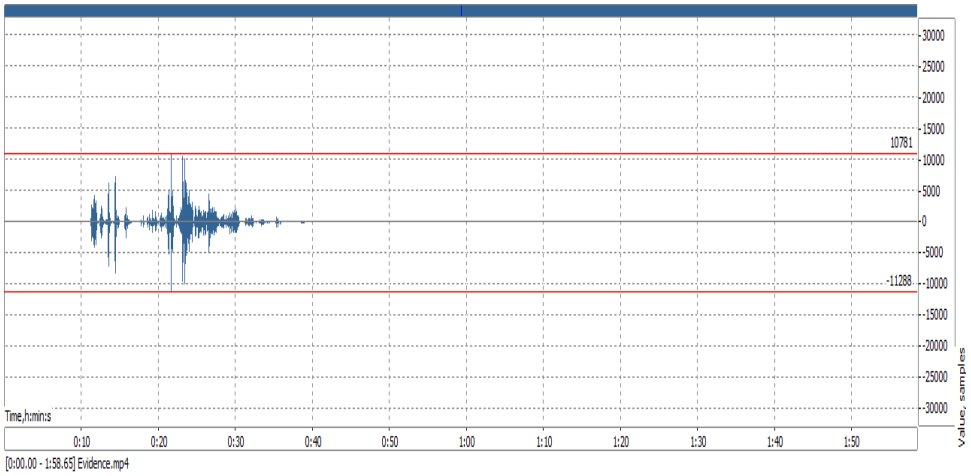
- Codec:

- Duration: 1 min 58 s 656 ms

Creation date and time: No date

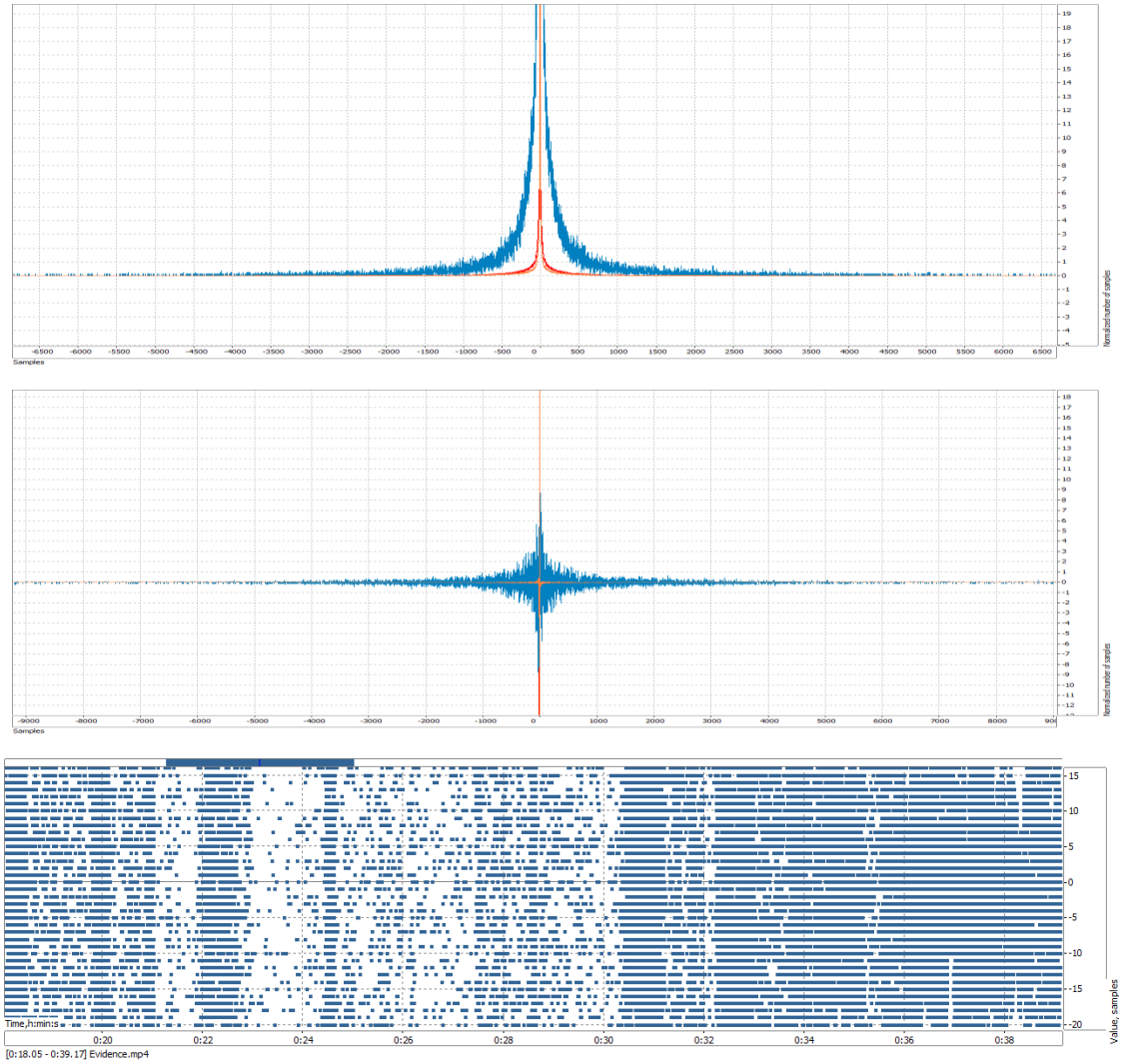
Modification date and time: No date

Number of "broken" blocks: 1



Signal analysis: Signal analysis depicts amplitude range, and continuation of the signal.

Histogram: Histogram of samples reveals signal attenuation due to the rounding algorithm features. Histogram asymmetry and autocorrelation are very important while detecting signal manipulation traces.



Levelogram: Levelogram depicts signal fragments that were amplified, compressed or merged.

DC-offset: DC-offset graph reflects the samples distribution centre time behaviour. After one or two other methods of observation, we reached at a decision that the ".mp4" audio video file was not edited.

Results and Discussion:

Here, our original questioned sample was very noisy. After proper amplification and enhancement using ACCUSTECK LAB+ licensed software, we are able to segregate the voice sample, especially the keywords, which were articulated by the accused persons at the commission of the crime [5, 6]. After that, we checked the authenticity of the questioned audio samples by analysing several pivotal analyses like header file analysis, levelogram analysis [7-9], histogram analysis, and DC offset analysis, and we found that these are unedited in nature. In this ACCUSTECK LAB+ licensed spectrometer, there are so many options for analysing audio samples, especially pitch analysis, pitch table, and energy-frequency (dB-Hz) spectrum analysis [10-13]. In the future, we are interested in speaker identification and clarity enhancement of different crime samples.

AID TO INVESTIGATION AGENCY:

By the techniques mentioned in the above paper, we could separate out the utterances of the accused person with the proper names of the accused, which we have found out often in our analysis. The local investigation agency could trace them out by tower location of their phone and could eventually nab them. This resulted in a breakthrough, whereby the local investigation agency was searching for them for a long time as they were allegedly involved in various nefarious activities in that region.

Authors' contributions

The authors make substantial contributions to conception and design, acquisition of data, analysis, and interpretation of data; the authors participate in drafting the article and revising it critically for important intellectual content; the authors give final approval of the version to be submitted.

Funding

The authors declare that there is no funding for this research. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

No datasets were generated or analysed during the current study

Declarations

Ethics approval and consent to participate

The study was approved by the Egyptian Forensic Science Laboratory, Govt. Of. W.B. India.

Consent for publication

The consent for publishing and presentation of the case report with anonymized data was obtained from the family of the deceased.

Competing interests

The authors declare no competing interests.

REFERENCES:

1. Maher, Robert C. Principles of forensic audio analysis. Vol. 1. Berlin/Heidelberg, Germany: Springer, 2018.
2. Zölzer, Udo. Digital audio signal processing. John Wiley & Sons, 2022.
1. 3.Bessonov, A. A., A. G. Boyarov, and M. V. Stepanov. "The Electric Network Frequency Based Method of Forensic Audio Analysis." Theory and Practice of Forensic Science 14.2 (2019): 43-50.
3. Boyarov, Alexander Grigorievich, and Ivan Sergeevich Siparov. "Forensic investigation of MP3 audio recordings." Теория и практика судебной экспертизы 14.4 (2019): 125-136.
4. Boyarov, A. "Detection and Analysis of MP3, WMA, OGG AND VORBIS Codecs Traces in Audio Signal." Theory and Practice of Forensic Science 1 (41) (2016): 78-86.
2. 6.Korycki, Rafał. "Detection of montage in lossy compressed digital audio recordings." Archives of Acoustics 39.1 (2014): 65-72.
3. 7. Kurniawan, Fajri, et al. "Statistical-based audio forensic on identical microphones." International Journal of Electrical and Computer Engineering 6.5 (2016): 2211.
4. 8. Pedapudi, Srinivasa Murthy, and Nagalakshmi Vadlamani. "Digital forensics approach for handling audio and video files." Measurement: Sensors 29 (2023): 100860.
5. 9. Ikram, Sohaib, and Hafiz Malik. "Digital audio forensics using background noise." 2010 IEEE International Conference on Multimedia and Expo. IEEE, 2010.
6. 10. Bhagtani, Kratika, et al. "An overview of recent work in media forensics: Methods and threats." arXiv preprint arXiv:2204.12067 (2022).
7. 11. Gangwar, D. P., et al. "Authentication of digital audio recording using file's signature and metadata properties." International Journal of Engineering Applied Sciences and Technology 5.3 (2020): 162-165.
8. 12. Fayyad-Kazan, Hasan, et al. "Verifying the audio evidence to assist forensic investigation." Computer and Information Science 14.3 (2021): 1-25.
9. 13. Bijhold, Jurrien, et al. "Forensic audio and visual evidence 2004-2007: A review." 15th INTERPOL forensic science symposium. 2007.