ISSN 2347 - 3983
Volume 8. No. 9, September 2020
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter186892020.pdf
https://doi.org/10.30534/ijeter/2020/186892020

Audio enhancement and application of digital filters for speech intelligibility in forensic digital analysis expert work (Computer Forensics)

Lizbardo Orellano-Benancio[1], Ricardo Muñoz-Canales[2], Danny Humpire-Molina[3] Enrique Lee Huamaní[4]
[1] Oficina de Peritajes del Ministerio Público, Perú, orellano01@yahoo.com
[2] Oficina de Peritajes del Ministerio Público, Perú, ramc2240@gmail.com
[3] Oficina de Peritajes del Ministerio Público, Perú, dhumpire@hotmail.com
[4]Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, Lima-Perú, ehuamaniu@uch.edu.pe

ABSTRACT

When clarifying the prosecution of a criminal act, the signs or evidence are found in recordings of digital audio files as questionable evidence, which corresponds to the conversations involved. Therefore, it is necessary to take a sample of the voice of the accused as genuine sample for the "Forensic Expert Work of Voice Homologation" based on voice comparison. Audio authentication, manipulation, edition, transcription or enhancement, among other activities, are part of this forensic expert work. This research focuses on the forensic expert analysis the "Forensic Work of Audio Enhancement" developed in the Forensic Digital Analysis Laboratory of the Bureau of Forensic Expert Analysis of the Prosecution Authority.

Key words: Audio enhancement, digital forensics, filters, forensic computing, frequency, signal processing.

1. INTRODUCTION

The Bureau of Forensic Expert Analysis of the Prosecution Authority receives inquiries for the "Forensic Work of Audio Enhancement", which are carried out by experts in digital forensic analysis; the evidence is received in sealed envelopes containing a chain of custody, and the evidence corresponds to recording equipment such as cell phones, tablets, digital recorders, cameras, among others, and/or digital storage media such as USB, SD, optical discs, etc. Part of the "Forensic Work of Audio Enhancement" is the processing of digital signal that aims to "clean" the audio store without affecting the original source file.

Audio signal quality depends on several factors such as the surreptitious nature of the recording due to the recording equipment, the placement of the microphone which can be affected by clothing rubbing, environmental sounds either open or closed, machinery noise such as air conditioners, fans, distance to the speakers and many others. Therefore, the process of audio enhancement based mainly on the use of digital systems known as frequency-selective filters is [1] for the attenuation of noise signals, which is possible through frequency or spectral analysis by Fourier Transform [2] of the signal that makes up the digital audio file and thus determine which noise signals are sought to be attenuated.

As conversations are significant to the prosecution investigation in order to identify the interlocutors and the context of the dialogue, the "Forensic Work of Audio Enhancement" enables the intelligibility of the conversation.

2. METHODOLOGY

The forensic expert work of audio and video files is associated with the international standards ISO/IEC 27001: 2013 (Management of information security and implementation of security measures), ISO/IEC 27037: 2012 (Information technology - Security techniques - Guidelines for identification, collection, acquisition and preservation of digital evidence) and ISO/IEC 27042: 2015 (Information technology - Security techniques - Guidelines for the analysis and interpretation of digital evidence). In addition, the chain of custody it is the procedure designed to guarantees the individualization, security and preservation of material elements and collected evidences [3]. Al well as guarantees traceability in the processing of evidence, integrity, conservation, inalterability, originality and authenticity, among others, of the evidence at different stages of the clarification process of a criminal act, which is represented in Figure 1.

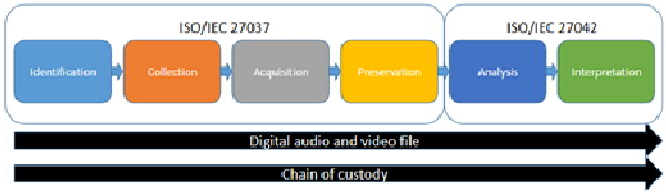

Figure 1: ISO / IEC 27001 and ISO / IEC 27001.

This research is framed within the process of "Analysis" with the **"**Forensic Work of Audio Enhancement" [3], which for this study refers to an unintelligible or noisy digital file recorded in a closed environment with the purpose of enhancing it and allowing a better signal intelligibility, noise attenuation, quality enhancement or ease to listen to the signal of interest, not detailing the recording equipment, the use of write-blocking equipment, audio format conversion, among others, since the audio was recorded in a closed environment and generated for this descriptive explanatory research.

The development of the digital audio signal filtering process for the forensic expert analysis the **"**Forensic Work of Audio Enhancement" is shown in Figure 2.

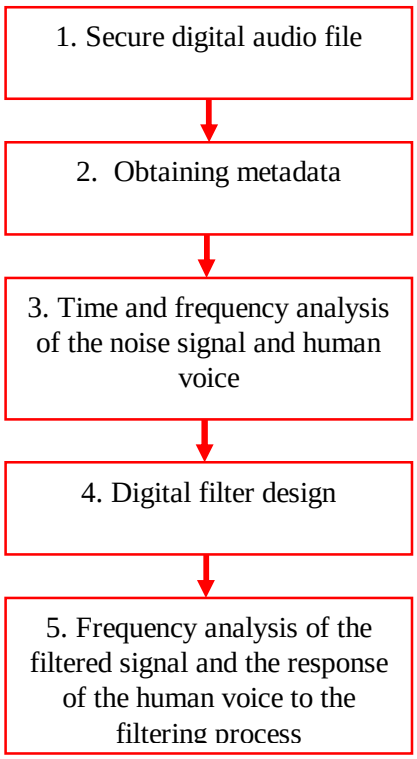

Figure 2: Process design flow of the work performed.

2.1 Ensured copy the digital audio file

This stage consists of securing the digital file received as a sample for the forensic expert analysis the **"**Forensic Work of Audio Enhancement". Such process is performed by obtaining the cryptographic securing code with the hash function (SHA-256, SHA-1, MD5, among others), from the United States National Security Agency (NSA) under the FIPS 180-2 standard [4], which generates a unique code for the binary data that compose the content of the digital audio file, ensuring the integrity of the copies made from the source file. In addition, this process is accomplished through the use of the HashMyFiles software from the company NirSoft, as shown in Figure 3 to the audio file named "audio3.wav".

Figure 3: HASH SHA-256 code obtained from a test audio by using HashMyFiles v2.23.

2.2 Metadata collection

At this stage the technical information recorded in the digital audio file is obtained by the software of the audio recording equipment. The metadata [5] gives us the following information.

- File name.
- Container file format.
- File duration.
- Bit rate.
- Compression code.
- Sample frequency.
- Channels.
- Bit density.
- Date of file creation.
- Date of file modification.
- Other parameters that depend on the hardware and software used for recording the digital audio file.

This information is obtained by using the MediaInfo software from the company MediaArea.net SARL, a metadata display of a test audio called "audio3.wav" is shown as follows in Figure 4.

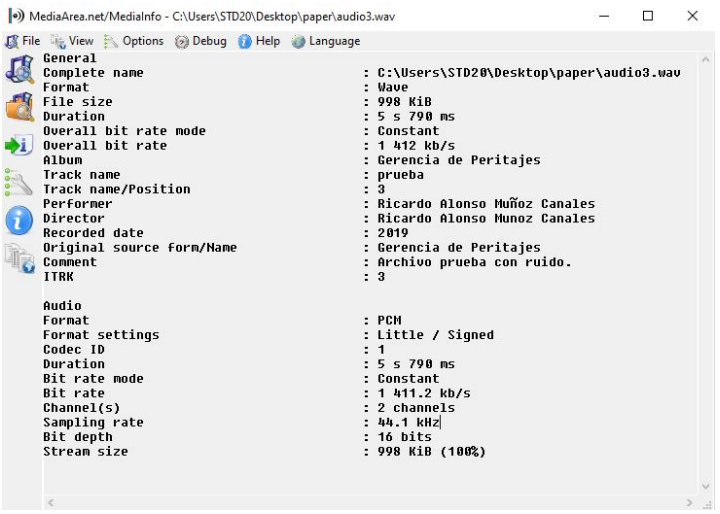

Figure 4: Metadata information collection using Mediainfo v19.09 software.

2.3 Time and frequency analysis of the acoustic signal and human voice

In this stage, the analysis is performed with the Adobe Audition Compilation 13.0.7.38 software, based on a time analysis by which the prosecutor performs the auditory and visual search for noise signals in the recording and thereby documents the time markers where the interference with the voice signal occurs. The two left and right audio channels that make up the digital audio signal are shown in Figure 5.

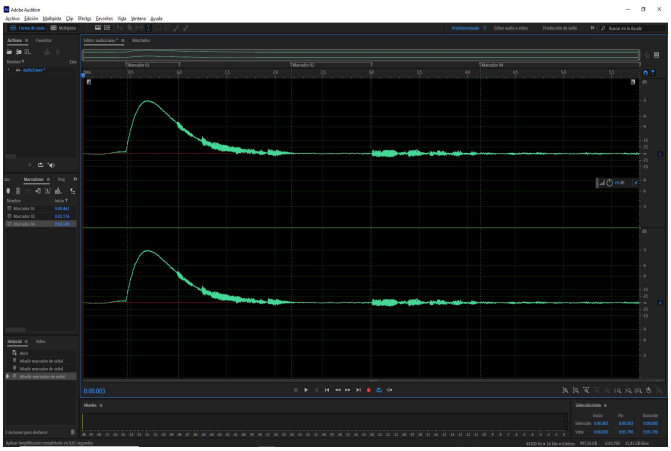

Figure 5: Time display (indicated by the green line) of the audio signal in the Adobe Audition software c13.07.38.

Then, the frequency analysis is performed, as shown in Figure 6, in the Adobe Audition software in order to find the frequency components [2] of the noise signal in each channel.

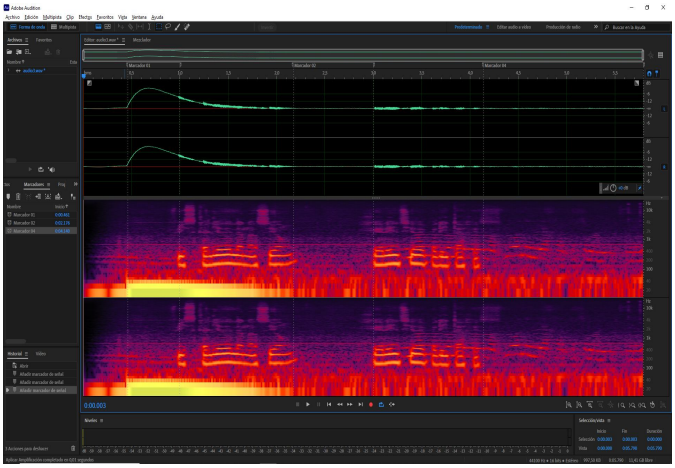

Figure 6: Frequency display (represented by the orange and yellow shades) of the audio signal in the Adobe Audition software c13.07.38.

At this stage, the purpose is to identify the frequency components that affect the voice signal to attenuate them by using digital filters, therefore it is necessary to get to be aware of the frequency components of the voice signal present in the audio file in order to conduct a correct analysis for the audio enhancement. Such analysis will be done using the Praat software version 6116, which is a voice analysis software [6]. In addition, the human voice has physical characteristics as detailed [6]. The spectrum of the audio signal "audio3.wav" is displayed in Figure 7.

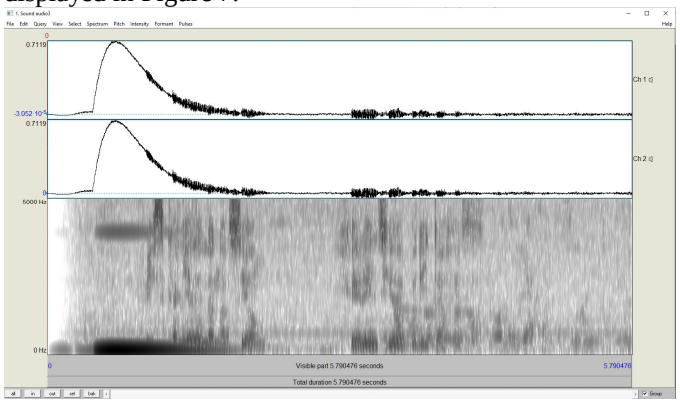

Figure 7: Audio signal display in time and frequency through Praat v6116 software.

In the Praat software, the frequency components that compose the voice signal can be highlighted, as the fundamental frequency or tone [6] shown in Figure 8, while the frequency components of the voice signal called formants [6] are shown in Figure 9.

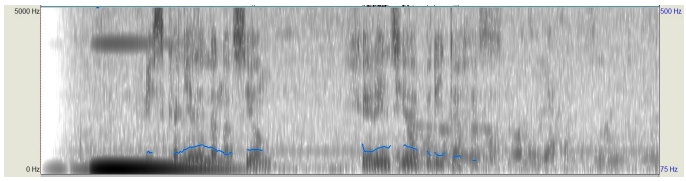

Figure 8: Fundamental frequency visualization (represented by blue).

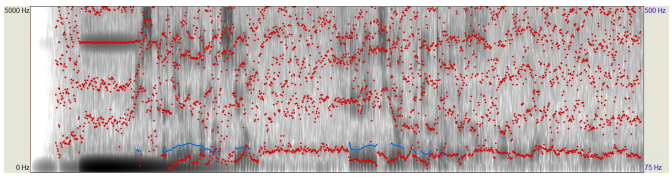

Figure 9: Visualization of the frequencies that form the voice signal (represented by red).

As stated in [6], the voice signals are composed of frequency components called "Pitch" and "Formants".

However, the frequency components "Pitch" and "Formants" differ depending on each individual, as well as on other parameters associated with linguistic analysis which are not covered in the present paper; consequently, those frequency components of the voice signal in the digital audio file should be considered during the design of the digital filters to ensure that no attenuation is caused by them.

The frequency analysis displays noise signals in Figure 10. In the time marker, two frequency components of the noise signal are detected, the first one at 4000 Hz and the second one at 650 Hz, which interfere with the frequency components called voice signal formants. At this point, the frequency selective filtering process must be applied to attenuate the mentioned signals; in the present research, the filtering process will be performed on the digital audio signal exclusively in the noise component at 650 Hz as this signal has higher energy than the one at 4000 Hz and the filtering effects will be more noticeable.

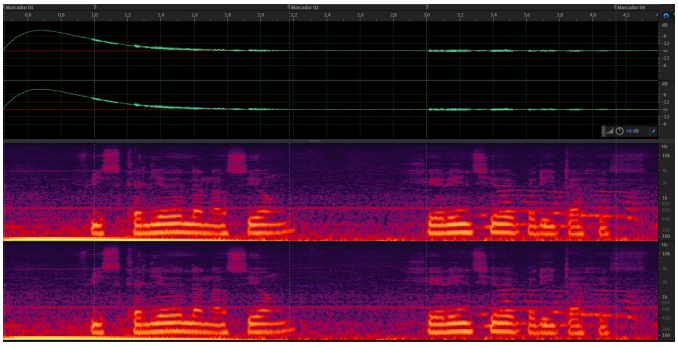

Figure 10: Visualization of the forming frequencies of the voice signal (represented by red).

2.4 Digital Filter Design

There are many software-based filters [5]., the design of a digital filter type IIR was proposed for this design, according to [6]. This is defined in the equation given as a Z-domain transfer function as described in (1), which is based on the division of the sums of the exponential Z elements and the coefficients of the numerator "bi" and the denominator "aj".

$$H(z) = \frac{\sum_{i=0}^{P} b_i z^{-i}}{1 + \sum_{j=1}^{P} a_j z^{-j}} \quad (1)$$

The expression represented in the time domain can be seen as a delayed version of the output and (n), as shown in (2).

$$y(n) = \sum_{i=0}^{n} b_{(i)} x(n-i) + \sum_{i=1}^{m} a_{(i)} y(m-i) \quad (2)$$

Where a (i) and b (i) stand for the IIR filter coefficients. This expression is shown in relation to the processes of multiplying, adding and subtracting.

For the design stage, with MATLAB 2019b scientific software, it is possible to establish the design criteria of the digital filters for the attenuation of the components in frequency of 4000 Hz and 650 Hz. Also, a Butterworth filter based on scripts and verifying the frequency responses is used as shown in figure 11, those frequency responses are standardized with the sampling frequency, in addition the filter obtained is of order 3 for attenuation in the 3dB pass band and in 40 dB rejection, as it was done in [7].

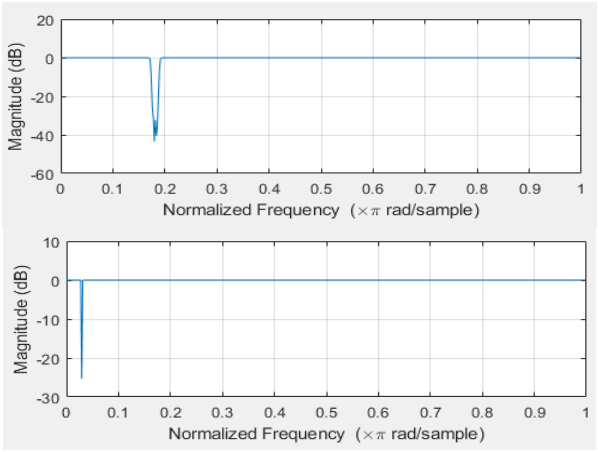

Figure 11: Top: Filter frequency response at 650Hz. Bottom: Filter frequency response at 4000Hz.

However, in the analysis stage, the digital forensic process for audio enhancement requires it to be immediate and in line with the auditory analysis of the digital audio signal, in addition to the use of a quick-access user interface, therefore, the filter application process is done with Adobe Audition software.

2.5 Frequency analysis of the filtered signal and the response of the human voice within the filtering process

When the frequency response of the filter and its stability have been checked, the application of the digital filter in the Adobe Audition software is assessed using the effect called Scientific Filter, by applying the filter to the noise signal at 650 Hz as shown. Refer to Figure 12.

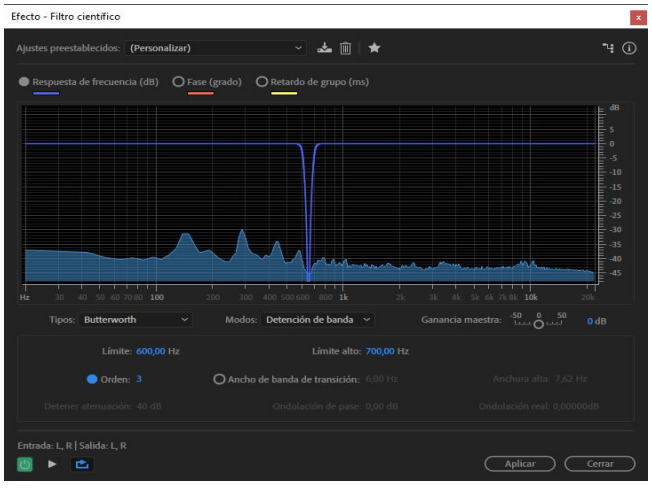

Figure 12: Butterworth filter rejection to 650 Hz frequency band.

The digital filter is implemented in Adobe Audition software as shown in figure 13 within the white line boxes.

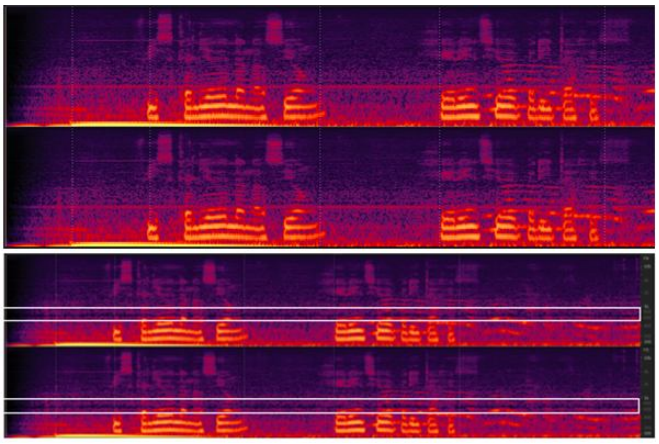

Figure 13: Top: Audio signal of the maximum unfiltered energy along the frequency of 650 Hz.
Bottom: Attenuation of filtered audio signal along the 650Hz frequency.

The frequency response of the filtering process in the attenuated noise signal and the voice signal, allows the intelligibility of the characteristic components of the human voice signal constituted by the tone and its formants, which can be seen in Figure 14.

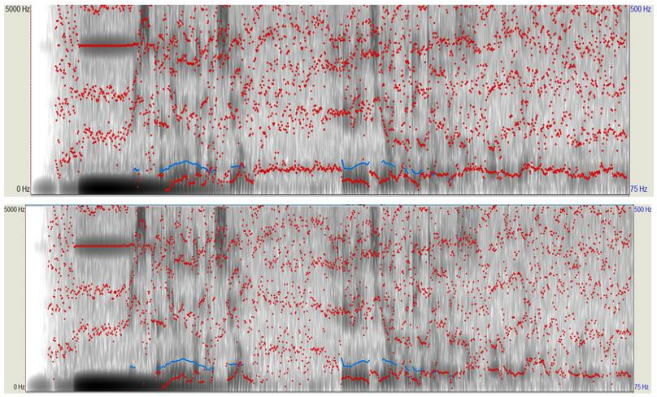

Figure 14: Top: Audio signal frequency response without filtering at 650 Hz including the tone (represented in blue) and formants (represented in red).
Bottom: Audio signal frequency response filtered at 650Hz including the tone (represented by blue) and formants (represented by red).

The enhancement process is performed in an interactive way through the use of a copy of the recording secured by HASH code in order to keep the original evidence unaltered. In this way, the filtered audio signal is obtained and exported with the name "audio3_01.wav" thus preserving the technical characteristics of the original audio signal without any alteration to the original source file.

3. RESULTS

Critical listening of the enhanced audio is performed and a greater intelligibility in the voice is verified, which makes it possible in longer conversations to reduce the fatigue of the listener for the transcription of the speech or to reveal signs in the prosecution, among other activities. Likewise, a comparison of the filtered audio signal called "audio3_01.wav" is made with the original source audio signal with noise called "audio3.wav", which is the sample involved in the illegal act, and the differences through cross-correlation of the two digital signals are shown graphically.

The cross-correlation calculation [8] is carried out to compare the difference between the original signal and the enhanced signal; by using the MATLAB software, the correlation coefficient vector is obtained and the graphs of the resulting vector are displayed. A graphic comparison of the coefficients of these two signals can be seen in Figure 15.

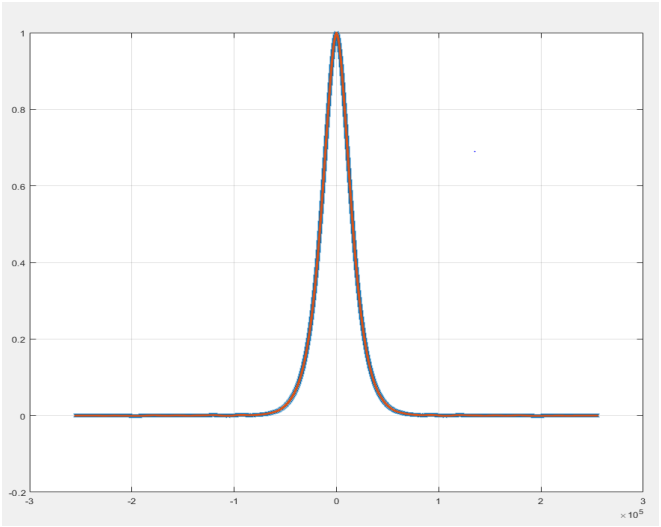

Figure. 15: Cross-correlation graph between unfiltered noise audio signal and filtered audio signal.

The cross-correlation coefficient matrix between the unfiltered signal and the filtered signal is obtained with the "corrcoef" function as illustrated in Figure 16.

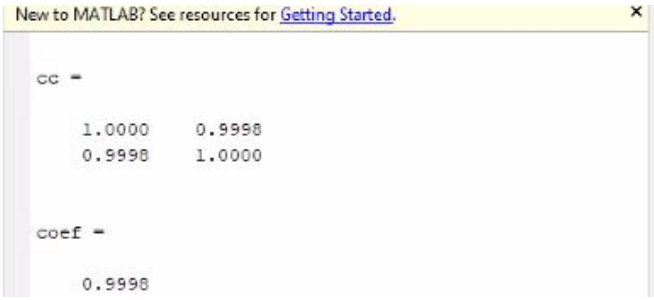

Figure. 16: Cross correlation coefficient matrix results.

4. DISCUSSIONS AND CONCLUSIONS

The "Forensic Work of Audio Enhancement" detailed in red box in part of Figure 17, as well as "Forensic Expert Work of Extraction and Securing" [9], is part the forensic expert work associated with recordings that may be of interest to the investigation process, either to identify an individual, to understand the context of the conversation or to seek homologation when there is a suspect. Some of the forensic expert work in the areas of digital forensic and acoustic analysis, and forensic phonetics as shown in Figure 17.

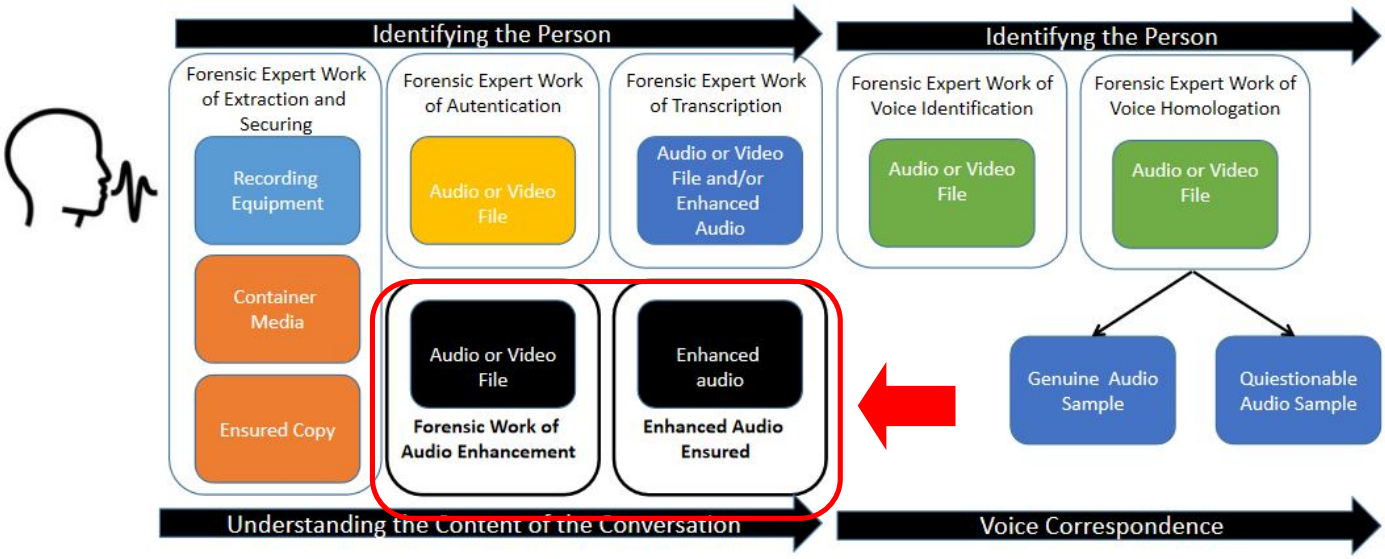

Figure 17: Part of the work of forensic experts in phonetics, acoustics, and digital forensic analysis

From Figure 15, it is highly important to emphasize that audio enhancement expertise does not exclude nor sequence different prosecutorial requirements, nor is it essential to other prosecutorial work, as the prosecutor responsible for the investigation evaluates the case strategies and prosecutorial requirements.

The Butterworth filter model used is one of many types of digital filters employed in digital signal processing and it has been demonstrated that its design and application is specific to the selected noise frequencies to be attenuated and that the impact of the application of the filters on the audio under study increases its intelligibility, permitting less fatigue when listening to the audio transcription of intelligible words or phrases, while not altering the original audio and enabling enhanced and secure copies with the HASH code.

Finally, future work is expected to provide a broader analysis of the filtering process on the voice components (i.e., "Pitch", "Formants") of the audio signal and its impact on the content of the recorded digital audio signal of a conversation of criminal interest, accepted as evidence.

REFERENCES

[1] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, Señales y Sistemas. 1998.

[2] A. Kuzmenko, P. Iezhov, and J.-T. Kim, "Weighting Iterative Fourier Transform Algorithm for Kinoform Implemented with Liquid-Crystal SLM," Fourier Transform. - Approach to Sci. Princ., no. March 2011, 2011.

[3] Ministerio de Justicia y Derechos Humanos. "Protocolo de autenticación interinstitucional para la aplicación de la incautación, comiso, hallazgo y cadena de custodia". Lima - Perú: 2018.

[4] Q. H. Dang, "Secure Hash Standard," FIBS 180-4 Publ., vol. 4, no. August, p. 36, 2015.

[5] Bruce E. Koenig, Douglas S. Lacey and Steven A. Killion Mathworks, "Forensic Enhancement of Digital Audio Recordings" vol. 55, pp. 355, 2007.

[6] F. I. Response, I. I. Response, T. Iir, and T. Mathworks, "Structures in Xilinx FPGAs," vol. 330, pp. 1–19, 2009.

[7] B. S. Fpga, R. Alonso, and M. Canales, "Diseño y simulación de filtro iir butterworth sobre fpga."

[8] P. Dimitris, "Digital Signal Processing: principles, devices and applications," Digit. Signal Process. Princ. devices Appl., 1990.

[9] A. Lazarte, L. Orellano, W. Gutiérrez, J. Soto, H Rodríguez "Documento Interno Ministerio Público- Guía Análisis Digital Forense" Pag. 14, 2020.