Audio Forensic Method: Voice Resemblance Analysis of Mobile Phone's Voice Recorder

Dayu Inten lestari

Universitas Warmadewa

Idaayu240189@gmail.com

Abstract- Audio forensics is one of the sciences that juxtaposes between science and scientific methods in the process of sound recording Audio Forensic Method: Analyze Voice Resemblance Towards Mobile Phone's Voice Recorder, by using Forensic audio method In Audio forensics: Theory and Analysis, are Pitch Statistical Analysis, Formant and Bandwidth Statistical Analysis, Graphical Distribution Analysis and Spectrogram Analysis. However, in this study, researchers only focus on identifying pitch and formants in the data to be analyzed. This study provides an example scenario of applying digital forensic techniques for voice pain recognition. Greetings compare the sound in the evidence and the sound suspect on mobile phone media with a voice case study woman. In this study, PRAAT application was used to help with the audio comparison process from Known Samples and Unknown Samples. The results showed that Obvious results could be seen in words spoken with high intonation. So, it is hoped that the output of this research can become a reference or enrichment material for law enforcement agencies, law, and academics who wish to continue their research related to digital forensics. The words are pronounced with high intonation, while the original voice tends to be low, indicating different formant values. So that it can be easily distinguished from the comparison voice, which is on average high, very high, low, or shallow intonation. In contrast, in words with low and medium intonation, the results of pitch analysis do not show significant differences. Digital voice forensic techniques cannot validate evidence because no standard validation is determined. Digital forensic techniques can only provide sound similarity analysis of good evidence by the suspect's voice.

Keywords: Forensic Linguistic, Forensic audio. Praat Aplication.

I. INTRODUCTION

Audio forensics is one of the sciences that juxtaposes science and scientific methods in sound recording analysis to assist and support the disclosure of a crime required in the trial process. (Subki et al., 2018). The ITE Act No.19 of 2016 states that voice recording is one of the most valid digital instruments and can be used as an indictment. Voice recordings that are digital evidence are straightforward and prone to be manipulated, either intentionally or unintentionally.

The field is known as "forensic linguistics" has been growing in prominence in the past couple of decades. Forensic linguistics is all about taking linguistic insight, method, and knowledge in the context of law, judicial procedures, police investigations, trials, and in short, about studying the language of law and solving crimes. Olsson (2004) defines it as an application of linguistics in the context of crime, court proceedings, or arguments in law. Coulthard and Johnson (2010) mention that forensic linguistics ranges from courtroom discourse and legal language to plagiarism. Briefly, plagiarism is using another person's work for personal advantage without mentioning his/her name. Forensic linguistic experts proficient in plagiarism cases and copyright infringements provide evidence to determine which work is based upon another.

The development of multimedia technology is currently increasingly facilitating human activities in daily life, including how technology can store audio digitally. Digital audio storage is typically used for interviews or education, which is commonly used for storage media is a sound recorder or use similar apps found on a cellphone. Voice recording is often used by someone to immortalize a conversation Directly or by telephone. In practice, voice recordings are used as evidence To strengthen the law enforcement charges during the trial process. The ITE Law No. 19 of 2016 mentions that voice recordings are one of the pieces of evidence, as described in Article 1. It is just that sound recordings cannot be used as evidence without going through a fairly long analysis process, which an expert in audio forensics carries out. AlAzhar Nuh (2011) mentions in his book Audio Forensics: Theory and Analysis that sound recordings can be analyzed through the parameters of tone, formant, and spectrogram. This component can be used to identify the characteristics of a person's voice for speech recognition purposes by using the fragments of the analyzed voice recording. Digital forensic science is, by definition, a combination of the disciplines of law and computer science in collecting and analyzing data from computer systems, networks, wireless communications, and storage devices digital data for later use as evidence in problem-solving in the realm of law. (Binyamin Widi Prasetya, dkk, 2008)

In its application, digital forensic science is often helpful to authorities in uncovering related crime cases the suspect is concerned about through the evidence that has been collected. The science of sound forensics focuses on efforts to analyze of suitability or originality of sound content material with the original content for later testing reliability and validity (Detik.com, 2011). With the increasing development of technology, more and more use is increasing. One of which is the discovery of several cases of legal irregularities accompanied by evidence in the form of sound recording media. Case this indicates a possibility ahead of audio digital will be used more as evidence in legal cases. Legal considerations in using evidence in the form of digital files, Including audio, are the ability of digital evidence to manage the impact associated with risk on the process law. One of the risks in question is using witnesses who have not known with certainty the truth, even though he has sworn to speak the truth. Using digital evidence that has been tested and analyzed will support the discipline of action and accuracy of guesses and help inaccuracy Decision-making. It is important for the parties relevant law enforcement agencies to understand and master digital forensic techniques considering the possibilities of increasing use of sound recordings/good multimedia In terms of variety, quantity, and quality. One of the digital forensic techniques is Voice Recognition, namely digital forensic techniques for detecting records Voice. People who have conversations can identify through audio forensic examination for speech recognition by comparison method, namely, comparing the voices in the recorded evidence (unknown sample) with sound recorded as Comparison (known sample). Suppose the voice recognition result indicates that the sound of the unknown sample is identical to the known sample voice. In that case, the Voice in the conversation in the recording Evidence can be obtained from the owner of the vote Comparison. (Septiyansyah, 2015).

Aligarh (2016) was researched to create the environment as natural as possible, conditions Retrieval, And results of the Forensic method used. In this study, forensic testing of sound evidence is carried out using pitch, formant, and spectrogram values, then comparing the sound of the evidence (unknown samples) with recorded sound as a comparison (known samples).

This study aims to provide an example scenario of applying digital forensic techniques for voice pain recognition. Greetings compare the sound in the evidence and the sound suspect on mobile phone media with a voice case study

woman. So it is hoped that the output of this research can become a reference or enrichment material for law enforcement agencies, law, and academics who wish to continue their research related to digital forensics. In this study, a Praat application was used to help with the audio comparison process from Known Samples and Unknown Samples. *Praat* is a computer program used to sound analysis, synthesis, and manipulation. This app was developed in 1992 by Paul Boersma and David Weenink at the University's Institute of Phoenix Sciences Amsterdam. Several versions are released with customization for some common operating systems used: Mac, Windows, and Linux. Since 2001, it has been 5000 registered users in 99 countries have been using Praat. Septiyansyah (2015) stated that the Praat app could record sound from a microphone or other audio devices. Besides that, this application can also read sound from an input file or disk. With Praat, then the user can see into the audio. This research uses the Forensic audio method. Based on the background above, the formulation of the problem in this research is What are the results of each sample of digital evidence of native speakers' voices and recorded comparisons on the two assessment factors, namely Pitch and Formant, the researcher can distinguish between the original speaker's Voice and the comparison voice?

II. METHODS

In the method, this research uses the Forensic audio method. In Audio forensics: Theory and Analysis, namely Pitch Statistical Analysis, Forman and Bandwidth Statistical Analysis, Graphical Distribution Analysis, and Spectrogram Analysis. However, in this study, researchers only focus on identifying pitch and formants in the data to be analyzed.

III. RESULT
3.1 Process with Praat

This process is the core process of this research. With the Praat application, you can find out the pitch, formant, and spectrogram. The following is the implementation of the pitch, formant, and spectrogram:

3.1.1 Pitch

Pitch analysis can be used to perform voice recognition on a person's voice, that is, through statistical analysis of minimum pitch, maximum pitch, and average pitch. Pitch is the frequency vibrating vocal cords (Jose R.L Batubara, 2010). One of the parameters of the sound signal is the frequency fundamentals. Fundamental frequency in terms of musical

instrument known as pitch or frequency value of a tone type. The faster the vibration of the vocal cords, the higher the pitch, and vice versa. Each person has a distinctive pitch (habitual pitch) strongly influenced by physiological aspects of the human larynx. In normal conversation conditions, habitual pitch levels range from 50-250 Hz for men and 120-500 Hz for women (M. Nuh AL-Azhar, 2012).

To analyze the pitch on the praat, the first thing to do is import sound files that have been noise filtered into the praat app. Open the Praat application, select the Open menu, select Read from a file, and select the sound recording file you want to download analysis.
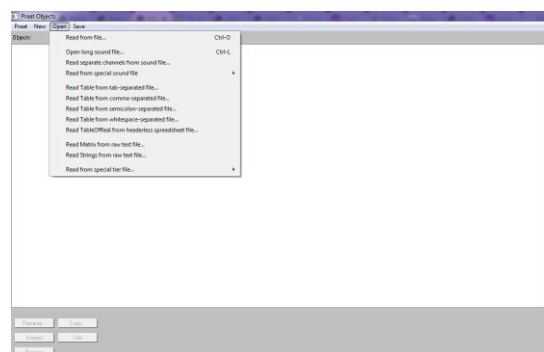
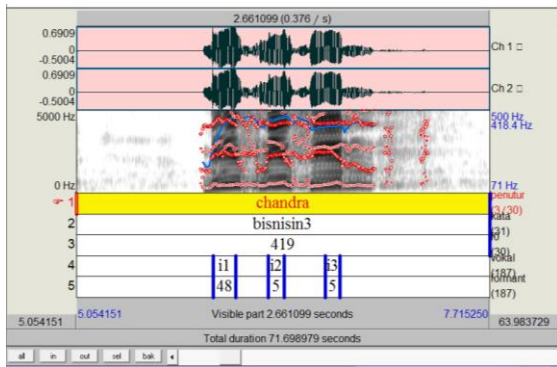**Pitch Analysis Step (1)**
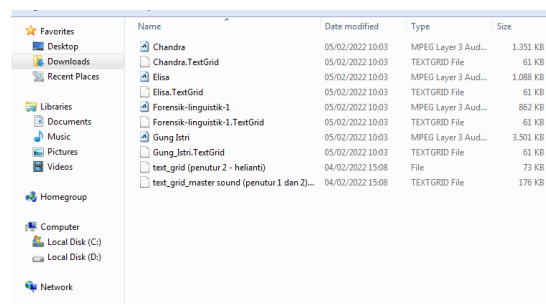
Pitch Analysis Step (2)

After that, select *view and edit*. Then a window like the one below will appear.

View and edit menu window

Because the analysis carried out is the word for word, which includes the vowels of the recorded sound, then listen carefully to the words you want to analyze, then do the blocks on the graph.

To get the right block, click the enter button in the left corner to zoom in. here is a brief explanation button located in the left corner.


Zoom Button

All	All selection (to see the whole chart)
in	Zoom in (to enlarge the graphic view)
out	Zoom Out (to zoom out the graphic)
sel	Selection (to view block graphs only)
back	Back (back to the previous initial view)

After the one-word graphic block, click File – Save the selected sound as a WAV file. In other words, that way, cut or partition each word on each record. If a record contains 30 words, there will be 30 save files as wav. After snipping each word, open the saved file snippet of the word to know its pitch value. Do not forget to turn on the pitch to see pitch value by checking the Show pitch option. The method to find out if the option is checked is to click the pitch menu tab. If you have not already checked, just click Show pitch. Because if not enabled or check show pitch on the menu tab Pitch, then the pitch value will not come out.
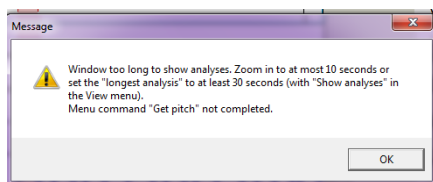
Pitch Value Warning Window

To view the minimum, maximum and mean, select the menu tab Pitch.
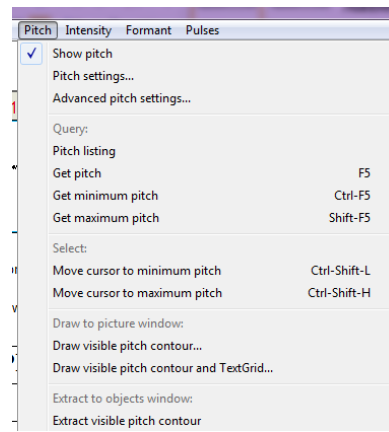
Menu minimum, maximum and mean pitch

3.1.2 Formant

To analyze formant, open praat app, import file audio that has been per-worded by clicking the open menu and select read from file
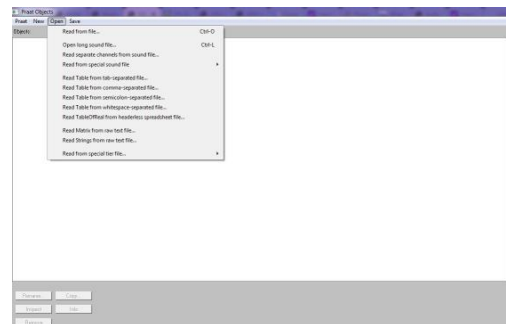
Open file

So a dialog box will appear to select a file as following:
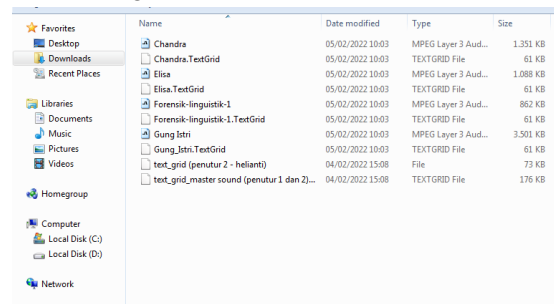
Import file dialog box

After selecting a file, the file will be listed in the left field but still in audio/sound format. Meanwhile, the required data must be in informant format to process formants. To change it, click on the right menu of Analyze Spectrum and select formant.

Pre-PRAAT Window

After that, the file will change its format, as shown In the following image. After the file changes the format, what is done is to find the value of the formant numerically in the form Month so that it can be compared statistically later.

Output Forman

To find the value, click on the file that has been formant, then on the right menu, select tabulate – list.

Selecting Tabulate-List Menu

Forman Values in Numeric

Copy and paste the results into Microsoft Excel, so that the tabular formant results can be processed in Microsoft Excel as shown below:

Ms. Excel

After all the data is summarized, the next step is to compare which suspect voices have a formant value close to that of a native speaker.

3.1.3 Spectogram

Open the praat application, then open the file you want to view the spectrogram (file that has been partitioned word by word from the original recording). Select the View & Edit menu, select the tab menu Spectrum and make sure the Show spectrogram option is checked.

Selecting the Show Spectrogram Menu

Then a window will appear as below:

Then screenshot the spectrogram section to compare with the Spectrogram of the evidence manuscript. With the way, after opening the sound file, select Analyze Spectrum, then select To Spectrogram. Then

click views. Perform this process on all evidence files with suspect voice. Then compare which suspect voices have a Formant value close to that of a native speaker.

Sample data:

Based on the F0 analysis of the voice data with PRAAT and comparing the original voice with ten comparison voices, it can be distinguished between the 11 voices, which are original and which are comparison voices. This is obtained based on the principle that a high pitch value indicates a loud voice with a high intonation, whereas if the pitch or F0 is low, it is a voice with a low intonation. The original F0 tends to be low or very low, while the other comparison f0 sounds are as follows:

PP1: very high
PP2: very high
PP3: low
PP4: high
PP5: low
PP6: very high
PP7: low
PP8: low
PP9: low
PP10: low

The results of the f0 analysis can be seen in the value of the words bisnisin_3, positif_3, akal-akalin_3, tunjangan_3, lima ratus_3, foya-foya_3, and feenya_3. Noticeable results can be seen in words spoken with high intonation, while in words with low and medium intonation, the results of pitch analysis do not show significant differences.

The F1-F4 analysis of the voice data with PRAAT on the original voice with the comparison voice found that the results were quite significant. In a high pitch sound, the formant values of f1, f2, f3, and f4 also show relatively high numbers. This can be seen from the value of formant fi to f4 in the words bisnisin_3, positif_3, akal-akalin_3, tunjangan_3, lima ratus_3, foya-foya_3, and feenya_3. The words are pronounced with high intonation, while the original voice tends to be low, indicating different formant values to be easily distinguished from the comparison voice, which is on average high, very high, low, or very low intonation.

IV. CONCLUSION

Based on the discussion above, the results showed that very clear results could be seen in words spoken with high intonation. In contrast, in words with low and medium intonation, the results of pitch analysis do not show significant differences. The words are pronounced with high intonation, while the original voice tends to be low, indicating different formant values. So that it can be easily distinguished from the comparison voice, which is on average high, very high, low, or very low intonation, it can be concluded that Digital voice forensic techniques cannot validate evidence because there is no standard validation is determined. Digital forensic techniques can only provide results from the sound similarity analysis of good evidence by the suspect's voice.

REFERENCES

Al-Azhar Nuh, M. (2011). *AUDIO FORENSIC :Theory and Analysis*. 1–38.

Aligarh, A., Hidayanto, C., Si, S., & Kom, M. (2016). *Implementasi Metode Forensikdengan Menggunakan Pitch , Formant, danSpectrogram untuk Analisis Kemiripan Suara Melalui Perekam Suara Telepon Genggam Pada Lingkungan yang Bervariasi. 5*(2).

Binyamin Widi Prasetya, Budi Santoso, and J. P. (2008). *Identifikasi Suara Pria Dan Wanita Berdasarkan Frekuensi Suara*. Universitas Kristen Duta Wacana, Yogyakarta.

Coulthard, M., & Johnson, A. (2010). *Handbook of forensic linguistics*. Routledge.

Detik.com. (2011). *Kenapa Suara Laki-Laki Berubah Saat Puber*.

Jose R.L Batubara. (2010). *Adolescent Development (Perkembangan Remaja)*.

M. Nuh AL-Azhar. (2012). *Digital Forensic : Panduan Praktis Investigasi Komputer*. Salemba Infotek.

Olsson, J. (2004). *Forensic linguistics: An introduction to language, crime and the law*. Continuum.

Septiyansyah, H. (2015). *Implementasi Metode Forensik Suara Pria Menggunakan Teknik Voice Recognize Untuk Analisis Kemiripan Suara Pada Media Alat Rekam Telepon Selular*. Universitas Teknologi sepuluh November, Surabaya.

Subki, A., Sugiantoro, B., & Prayudi, Y. (2018). Membandingkan Tingkat Kemiripan Rekaman Voice Changer Menggunakan Analisis Pitch, Formant dan Spectogram. *Jurnal Teknologi Informasi Dan Ilmu Komputer, 5*(1), 17. https://doi.org/10.25126/jtiik.201851500