IJFL (International Journal of Forensic Linguistic)

Vol. 2, No. 1 , Page 34-40

P-ISSN: 2723-1542

E-ISSN: 2723-5254

Available Online at https://www.ejournal.warmadewa.ac.id/index.php/ijfl/index

Analyze Voice Resemblance Towards Mobile Phone's Voice Recorder: An Implementation Of The Forensic Method Using Voice Recognition Technique

Stephanus Made Kevin Steven Christian
Universitas Warmadewa
Stphnskevin@gmail.com

Received: DD/MM/YYYY	Revised: DD/MM/YYYY	Published: DD/MM/YYYY

How to cite:

Christian, Stephanus.M.K S. (2022). Analyze voice resemblance towards mobile phone's voice recorder: an implementation of the forensic method using voice recognition technique. *IJFL (International Journal of Forensic Linguistic)*, *Vol*(No), Page. Doi: http://dx.doi.org/xxxxxxjr.xx.xxx.Page

Abstract- Audio forensics is one of the sciences that mnyandingkan between science and scientific methods in the process of sound recording analysis to assist and support the disclosure of a crime required in the trial process. This article is intended to Analyze Voice Resemblance Towards Mobile Phone's Voice Recorder: An Implementation of The Forensic Method Using Voice Recognition Technique, by using Forensic audio method In Audio forensics: Theory and Analysis, namely Pitch Statistical Analysis, Forman and Bandwidth Statistical Analysis, Graphical Distribution Analysis and Spectogram Analysis. However, in this study, researchers only focus on identifying pitch and formants in the data to be analyzed. The results of the data analysis it can be stated that the accuracy of the pitch analysis (F0) through Praat is very low and cannot be used as an indicator to distinguish that a data is correct as the original voice, as evidenced by the results of the F0 value of the original voice data and a very flexible and inconsistent comparison, some are the same, close to and even far from the F0 value of the subject data. This shows that everyone has a different pitch value because the intonation of each person's word pronunciation is different. It is possible that there are pitch values from several subjects that are almost the same.

Keywords: Forensic Linguistic, Forensic audio. Praat Aplication, Pitch, Formant.

I. INTRODUCTION

The field known as “forensic linguistics” is 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 proceeding, 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 other person’s work for personal advantage without mentioning his/her name. Forensic linguistic experts who are proficient in plagiarism cases and copyright infringements provide evidence to find out which work is based upon another.

Audio forensics is one of the sciences that juxtaposes between science and scientific methods in the process of 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 extremely easy and prone to be manipulated, either intentionally or unintentionally. forensic speaker recognition methods, focusing on clustering language-independent utterances to improve voice similarity measurements. It emphasizes using Gaussian mixture modeling and MAP adaptation for effective voice comparison, which can be applied to mobile phone voice recordings, (Singh & Singh, 2020).

The development of multimedia technology is currently increasingly facilitating human activities in daily life, including how technology is able to store audio digitally. Digital audio storage is typically used for the need for interviews or education using 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 That can 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's just that sound recordings cannot be used as evidence without going through a fairly long analysis process, which is carried out by an expert in the field of audio forensics. 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 from the disciplines of law and computer science in collect and analyze 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 authorities in uncovering related crime cases the suspect concerned through the evidence that has been collected. The science of sound forensics focuses on efforts to analysis 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 moreIts 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 that there is 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, is the ability of digital evidence in managing the impact associated with risk on the process law. One of the risks in question is the use of witnesses who are not known with certainty the truth, even though he has sworn to speak the truth. By using digital evidence that has been tested and analyzed will be able to support the discipline of action as well as accuracy of guesses and helps inaccurac Decision-making. For that, it is important for the parties relevant law enforcement agencies to understand and master digital forensic techniques considering the possibilities 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 identity is known 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). If the result of voice recognition indicates that the sound of the unknown sample is identical to known sample voice, then the voice in the conversation in the recording Evidence can be obtained from the owner of the vote Comparison. (Septiyansyah, 2015).

Aligarh (2016) wasconducted research to create environment as natural as possible, conditions Retrieval, And results of the Forensic method that used. In this study,

forensic testing of sound evidence is carried out with using pitch, formant, and spectogram values then compare the sound of the evidence (unknown samples) with recorded sound as a comparison (known samples).

The aims of this study is to provide an example scenario the application of digital forensic techniques for voice pain recognizes Greetings compare the sound in the evidence and the sound suspect on mobile phone media with a voice case study man. 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 doing on In this study, a Praat application was used to help with the audio comparison process from Known Samples and Unknown Sample. Praat is a computer program that is used to sound analysis, synthesis and manipulation. This app developed since 1992 by Paul Boersma and David Weenink at the University's Institute of Phoenix Sciences Amsterdam. There are several versions released with customization for some common operating systems used are Mac, Windows and Linux. Since 2001, it has been 5000 registered users in 99 countries are using Praat. Septiyansyah (2015) stated that Praat app is able to record sound from microphone or other audio devices, besides that this application can also reads sound from an input file or disk. With Praat, then the user is able to see into the audio. This research uses 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 Spectogram Analysis. However, in this study, researchers only focus on identifying pitch and formants in the data to be analyzed.

This research uses software, such as praat application and Microsoft excel. This app praat is used to search information from the comparison between records of native speaker's voice and recorded comparison voice. Microsoft Excel to use to measure the formant, pitch of each word spoken original and comparison.

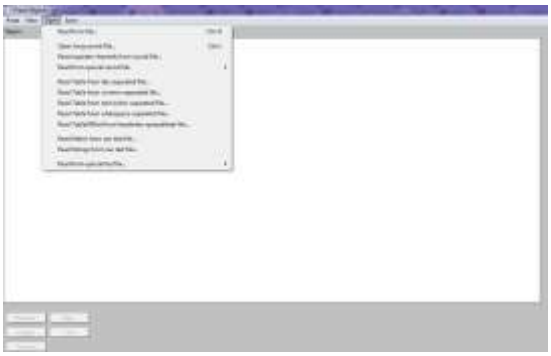
III. RESULT

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 of each sound recording. The following is the implementation of the pitch, formant and spectrogram:

3.1.1 Pitch

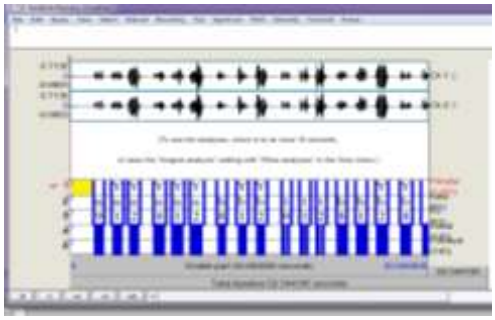
To analyze the pitch on the praat, the first thing to do is import sound files that have been noise filtered into praat app. Open the Praat application, then select the Open menu, select Read from file, then select the sound recording file you want to download analysis.



Langkah Analisis Pitch (1)

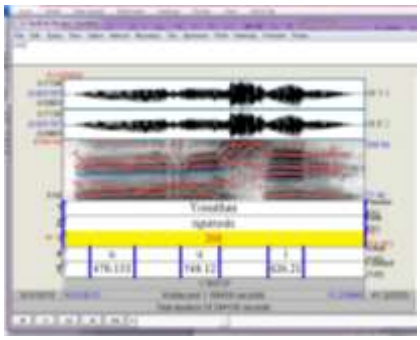
Langkah Analisis Pitch (2)

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

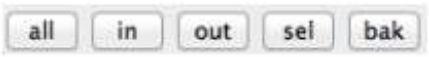


View and edit menu window

Because the analysis carried out is 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 corner left to zoom in. here's a brief explanation button located in the left corner.



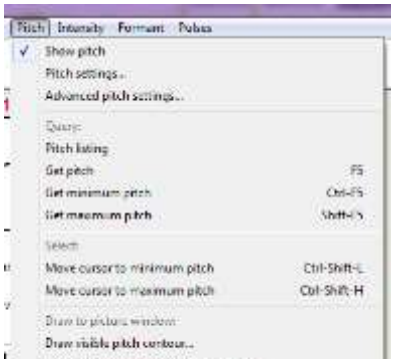
Zoom Button

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 then 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, then there will be 30 save file 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 making sure the Show pitch option is checked. The method to find out if the option is checked or not is to click the pitch menu tab, if you haven't already checked just click Show pitch. Because if not enable 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

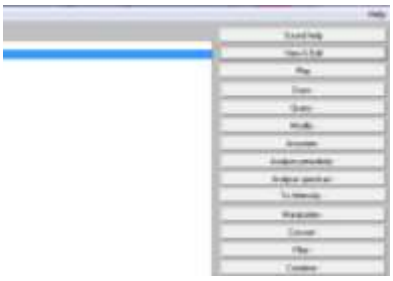
To analyze formant, open praat app, import file audio that has been per-worded by clicking the open menu and select read from file

Open file

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

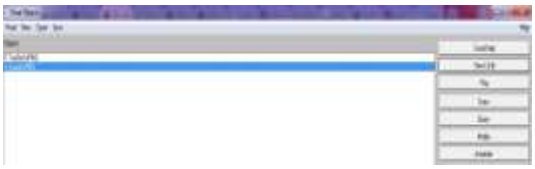
Import file dialog box

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



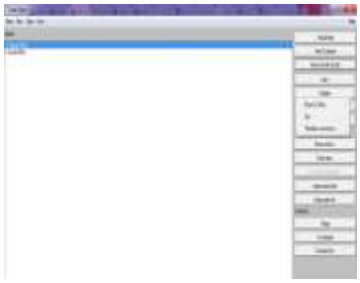
Pre-Praat Window

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.

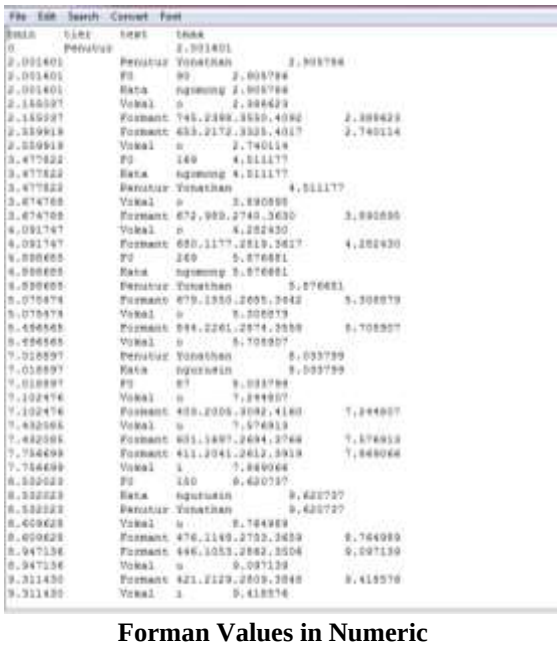


Output Forman

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

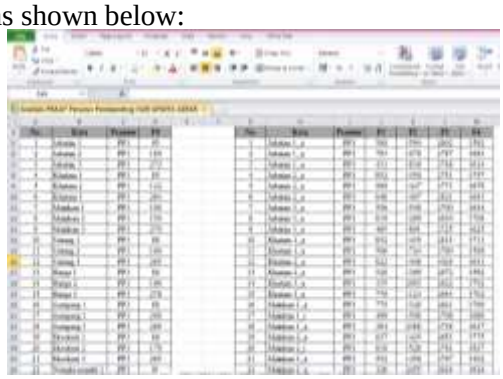


Selecting Tabulate-List Menu



Forman Values in Numeric

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

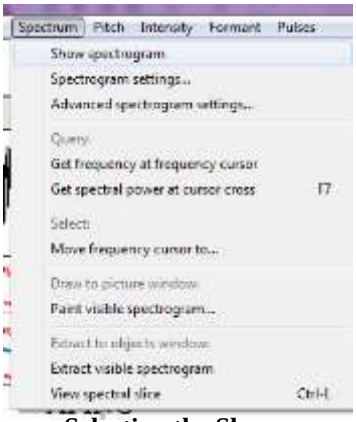


Ms. Excel

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

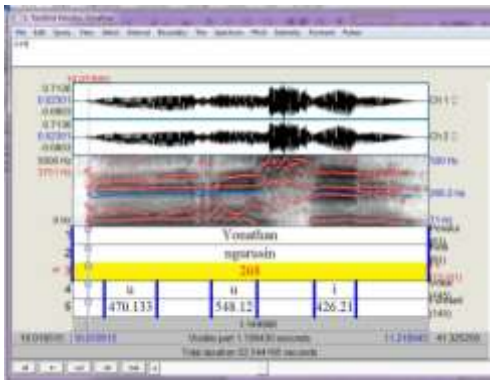
3.1.3 Spectogram

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



Selecting the Show Spectogram Menu

Then a window will appear as below:



Then screenshot the spectrogram section to compare with the spectrogram of the evidence manuscript. Can also with way after opening the sound file, select Analyze Spectrum then select To Spectogram. Then click views. Perform this process on all evidence files with suspect voice. Then compare which suspect voices are has a Formant value close to that of a native speaker.

Sample data:

The pitch character of each voice is compared to the minimum pitch value, maximum pitch value and mean (average)

pitch value. From the difference in the value of this pitch statistic, which can later help assess the level of similarity of the recording. Here is a sample data:

Sample Data
The word “ngomong”

Penutur	**F0**		
Penutur	123	126	128
PP1	90	169	269
PP2	48	61	159
PP3	123	138	205
PP4	125	158	274
PP5	166	206	306
PP6	149	183	288
PP7	144	189	242
PP8	96	136	163
PP9	111	119	181
PP10	107	130	163

The above data was obtained by analyzing the Pitch (F0) value of the respondents (male comparison speakers) three times with variations in tone starting from low, medium, and high notes. Except for the native speaker only at one tone level. Therefore, from the data above, it can be concluded that:

1. The original speaker's voice has a ratio of F0 values in a not too far range in the same pitch,
2. The comparison voice data (PP1-PP10) mostly does not have a pitch value that is identical to the original speaker's F0 voice, this is because the comparison speaker's pitch value does not approach the original speaker's F0 value with a wide range. However, some data also have a value ratio that is almost the same as the voice data of native speakers.
3. Thus, the conclusion is that identical comparison speakers' voice data has a pitch value that is close to the original speaker's voice data pitch value.
4. On the other hand, F0 analysis through praat cannot show the accuracy that a data is true voice data of the original speaker. This is because the data obtained from the comparison speaker's pitch value which has the same value and the range of values is not too far away. As shown by the sample data below;

Sample data:
Word: "ngomong_1" (Low tone voice data)

F0 . speakers	
Spesker	**F0**
Native speaker	**123**
Comparison speaker 3	**123**

Therefore, based on the results of the data analysis above, it can be stated that the accuracy of the pitch analysis (F0) through Praat is very low and cannot be used as an indicator to distinguish that a data is correct as the original voice, as evidenced by the results of the F0 value of the original voice data and a very flexible and inconsistent comparison, some are the same, close to and even far from the F0 value of the subject data. This shows that everyone has a different pitch value because the intonation of each person's word pronunciation is different. It is possible that there are pitch values from several subjects that are almost the same.

From the results of the formant analysis (F1-F4) of the original voice data and the comparison voice through Praat, it can be concluded that the difference between the voice data of the native speaker and the comparison speaker is indicated by the following:

- The native speaker's voice has a formant value (F1-F4) that is not the same as the formant value of the comparison speaker data, as shown by the following sample data:

Data sampel
word “ngomong_1”

Speakers	**F1**	**F2**	**F3**	**F4**
Native speaker	470	817	2611	3433
PP1	745	2398	3550	4092
PP2	548	1260	2680	3843
PP3	512	1196	2663	3504
PP4	541	707	2463	3725
PP5	336	864	2178	2577
PP6	526	2351	3540	4680
PP7	540	1581	3051	3827
PP8	531	1006	2439	3533
PP9	595	2113	2738	3674
PP10	381	1250	2724	4040

The sample above shows the results of the formant value on the word "ngomong_1" from the voice data of a native speaker and 10 voice data from a comparison speaker. from the results of the data analysis above it can be concluded that:

1. From the comparative voice data (PP1-PP10). Having formant value data that is

different from the formant value of the original voice data, however, if the argument is in a range of values, then some of the formant values of the comparison speakers have a value range that is quite close to the original speaker's formant value even though the accuracy is still not met.

2. The formant value of the native speaker's voice data can be clearly distinguished because it has a different value from the formant value of the comparison speaker's data. On the other hand, indicators for the difference in formant values have not been found because the formant values of each data are inconsistent. Starting from a lower formant value, approaching, and even higher than a native speaker's formant value.

IV. CONCLUSION

Based on the discussion above, it can be concluded that Digital voice forensic techniques can not do validation of evidence because there is no standard validation which is determined. Digital forensic techniques can only provide results from the sound similarity analysis of good evidence by the voice of the suspect. Therefore, based on the results of the data analysis above, it can be stated that the accuracy of the pitch analysis (F0) through Praat is very low and cannot be used as an indicator to distinguish that a data is correct as the original voice, as evidenced by the results of the F0 value of the original voice data and a very flexible and inconsistent comparison, some are the same, close to and even far from the F0 value of the subject data. This shows that everyone has a different pitch value because the intonation of each person's word pronunciation is different. It is possible that there are pitch values from several subjects that are almost the same.

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).

Detik.com. 2011. Mei. Kenapa Suara Laki-Laki Berubah Saat Puber.

Binyamin Widi Prasetya, Budi Santoso, and Joko Purwadi, "Identifikasi Suara Pria Dan Wanita Berdasarkan Frekuensi Suara," Informatics Department, Universitas Kristen Duta Wacana, Yogyakarta, 2008.

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

Septiyansyah, Hasbi. 2015. Implementasi Metode Forensik Suara Pria Menggunakan Teknik Voice Recognize Untuk Analisis Kemiripan Suara Pada Media Alat Rekam Telepon Selular. Jurusan Sistem Informasi, Universitas Teknologi sepuluh November, Surabaya.

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

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