

Abstrak

Perbandingan Metode *Itakura-Saito Distance* dan Manual Statistik (*Pitch, Formant, Spectrogram*) untuk Akurasi Identifikasi Suara pada Audio Forensik

Teknologi sering digunakan dalam kasus kriminal, hal ini dibuktikan dengan ditemukannya bukti elektronik dalam kasus kriminal seperti korupsi, penipuan, pencurian, dan lainnya. Bukti yang biasa ditemukan dalam bukti digital biasanya berupa file dalam bentuk dokumen, gambar, audio, video, dan lainnya. Untuk mendapatkan bukti-bukti ini, perlu memiliki proses analisis yang berbeda di setiap karakteristik bukti digital yang ditemukan, misalnya penemuan bukti digital dalam bentuk audio harus melalui prosedur yang sesuai dengan *Standart Operational Procedure* (SOP) untuk penanganan audio forensik, karena rekaman suara seseorang memiliki karakter dengan pengucapan kosakata yang berbeda - beda, pengucapan yang tidak jelas dan memiliki banyak *noise*, perlu adanya penanganan khusus untuk mendapatkan akurasi analisa yang maksimal. Proses secara umum untuk menganalisis audio forensik, diperlukan sejumlah pendekatan analisis statistik yang digunakan, diantaranya analisis statistik *Pitch, formant bandwidth, likelihood Ratio* (LR), Analisis Statistik *Anova*, Analisis *Graphical Distribution*, dan analisis *Spectrogram*, atau sering disebut dengan metode konvensional. Selain metode tersebut, beberapa saksi ahli juga menggunakan metode lain untuk proses audio forensik yaitu metode *Itakura-Saito Distance*. Metode ini menggunakan model *speech processing* yang berfungsi untuk mengukur perbedaan antara *spectrum* suara asli dan *spectrum* perkiraan suara yang diduga memiliki kemiripan. Dengan adanya beberapa saksi ahli yang menggunakan metode yang berbeda, maka dilakukan perbandingan metode dengan tujuan untuk meningkatkan akurasi pada proses audio forensik. Beberapa metode tersebut memiliki perbedaan diantaranya metode konvensional memiliki beberapa analisa diantaranya adalah *pitch, formant, likelihood, graphical distribution, dan spectrogram*, sedangkan metode *speech processing* hanya menganalisa nilai *formant* dan memvisualisasi kedalam bentuk grafik. Analisa tersebut dapat ditarik kesimpulan yaitu bukti digital yang ditemukan proses investigasi terutama barang bukti digital dengan bentuk rekaman suara, memiliki karakter yang berbeda – beda, belum tentu memiliki kualitas yang baik, selain itu metode konvensional dengan metode *speech processing* memiliki perbedaan dalam pengolahan terhadap rekaman barang bukti (*unknown*) dengan rekaman suara pembanding (*known*), diantanya adalah metode konvensional dengan analisa manual statistik memiliki beberapa analisa diantaranya adalah *pitch, formant, likelihood, graphical distribution, dan spectrogram*, sedangkan metode *speech processing* hanya menganalisa nilai *formant* dan memvisualisasi kedalam bentuk grafik. Hal ini mengakibatkan analisa menggunakan metode konvensional membutuhkan waktu yang lebih lama dibandingkan metode *Itakura-Saito Distance*, akan tetapi metode konvensional mampu menganalisa rekaman suara dengan lebih detail dibandingkan dengan metode *Itakura-Saito Distance*.

Kata kunci

Pitch, Formant, Spectrogram, Audio, Itakura-saito, Distance, likelihood, Graphical, Distribution.

Abstract

Comparison of Itakura-Saito Distance Method and Statistical Manual (Pitch, Formant, Spectrogram) for Accuracy of Voice Identification in Forensic Audio

Technology is often used in criminal cases, this is evidenced by the discovery of electronic evidence in criminal cases such as corruption, fraud, theft, and others. Evidence commonly found in digital evidence is usually in the form of files in the form of documents, images, audio, video, and others. To obtain this evidence, it is necessary to have a different analysis process in each of the characteristics of digital evidence found, for example the discovery of digital evidence in audio form must go through procedures in accordance with the Standard Operational Procedure (SOP) for handling audio forensics, because someone's recorded voice has characters with different vocabulary pronunciation, unclear pronunciation and have a lot of noise, special handling is needed to get maximum analysis accuracy. The general process for analyzing audio forensics, many statistical analysis approaches are used, including Pitch statistical analysis, formant bandwidth, likelihood Ratio (LR), Anova Statistical Analysis, Graphical Distribution Analysis, and Spectrogram analysis, or often called conventional methods. Besides this method, several expert witnesses also used another method for the audio forensic process, namely the Itakura-Saito Distance method. This method uses a speech processing model that functions to measure the difference between the original sound spectrum and the estimated sound spectrum that is thought to have a similarity. With the presence of several expert witnesses using different methods, a comparison of methods was conducted to improve accuracy in the audio forensic process. Some of these methods have differences between conventional methods that have some analysis including pitch, formant, likelihood, graphical distribution, and spectrogram, while the speech processing method only analyzes formant values and visualizes them into graphical form. The analysis can be concluded that digital evidence found in the investigation process, especially digital evidence in the form of sound recordings, has a different character, not necessarily of good quality, besides the conventional method with the speech processing method has a difference in the processing of the recording of goods evidence (unknown) with comparative sound recordings (known), among them is the conventional method with statistical manual analysis having several analyzes including pitch, formant, likelihood, graphical distribution, and spectrogram, while the speech processing method only analyzes formant values and visualizes them into graphical form . This results in the analysis using conventional methods requiring a longer time than the Itakura-Saito Distance method, but the conventional method can analyze sound recordings in more detail compared to the Itakura-Saito Distance method.

Keywords

Pitch, Formant, Spectrogram, Audio, Itakura-saito, Distance, likelihood, Graphical, Distribution.