

Analisis Pengaruh Sampling Rate Dalam Melakukan Identifikasi Pembicara Pada Rekaman Audio

Roy Rudolf Huizen¹⁾, Ni Ketut Dewi Ari Jayanti²⁾, Dandy Pramana Hostiadi³⁾

STIKOM BALI Jl. Raya Puputan Renon No 86 Denpasar

e-mail: roy@mail.stikom-bali.ac.id¹⁾, daj@stikom-bali.ac.id²⁾, dandy@stikom-bali.ac.id³⁾

Abstrak

Rekaman audio (suara) merupakan metadata berisi informasi diantaranya berupa fitur diperoleh dengan proses ekstraksi. Fitur mempunyai karakteristik khusus digunakan untuk mengetahui identitas individu pemilik rekaman suara. Sampling rate proses perekaman menentukan kualitas rekaman, namun penggunaan sampling rate seringkali tidak diperhatikan, untuk mengetahui pengaruh sampling rate terhadap fitur merupakan fokus penelitian ini. Data yang digunakan berupa rekaman suara langsung dan rekaman suara percakaman ditelepon, perekaman menggunakan varian sampling rate 8 KHz, 16 KHz dan 32 KHz, data diekstraksi menggunakan perangkat lunak praat untuk mendapatkan fitur. Nilai fitur yang diperoleh adalah frekuensi dasar dan formant 1 sampai 4. Hasil pengujian perekaman langsung maupun rekaman melalui telepon didapatkan bahwa karakteristik fitur untuk sampling rate 8 KHz, 16 KHz, dan 32 KHz jika dibandingkan 16 KHz dengan 32 KHz mempunyai fitur yang sama, sedangkan 8 KHz mempunyai fitur yang berbeda. Untuk frekuensi dasar mempunyai nilai fitur yang sama di masing-masing sampling rate yang berbeda.

Kata kunci: Rekaman suara, sampling rate, fitur, Identifikasi

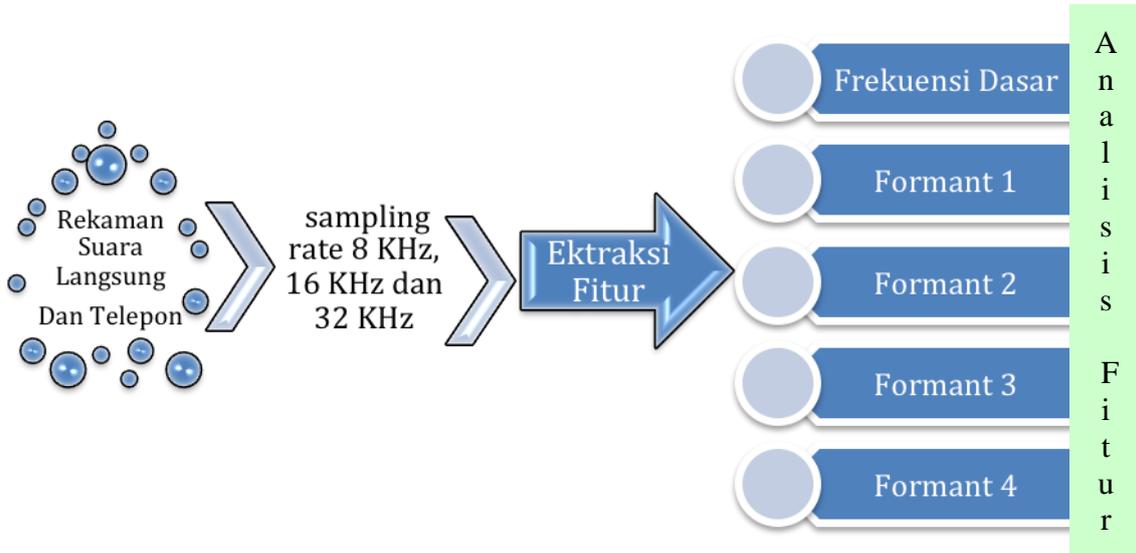
1. Pendahuluan

Rekaman audio [1] merupakan metadata berisi informasi yang dapat digunakan sebagai alat bukti. Untuk melakukan identifikasi di audio forensik terdapat tahapan yang terdiri atas akuisisi, analisis dan evaluasi, hasil dari proses tersebut dapat digunakan sebagai alat bukti dipengadilan ataupun sejenisnya [2]. Audio forensik merupakan suatu proses untuk memunculkan informasi mulai dari fitur, lokasi kejadian, transkrip pembicaraan dan lain sebagainya, dimana informasi tersebut dapat digunakan sebagai alat bukti [3]. Untuk memunculkan bukti direkaman audio digunakan pendekatan ilmu pengetahuan diantaranya adalah bidang pengolahan sinyal [4],[5]. Pada identifikasi identitas individu [6], proses yang dilakukan dengan cara membandingkan fitur antara rekaman bukti dengan rekaman pembanding. Kecocokan fitur dapat diartikan bahwa keduanya berasal dari individu yang sama dan sebaiknya ketidakcocokan berarti bukan berasal dari individu yang sama [7], oleh karena itu fitur merupakan bagian penting pada proses identifikasi [8]. Proses perekaman yang tidak sama nilai sampling rate dapat mengakibatkan proses identifikasi menjadi tidak optimal. Untuk mengetahui pengaruh nilai sampling rate terhadap fitur yang akan dicocokkan merupakan fokus pada penelitian ini. Data yang digunakan pada pengujian ini terdiri atas rekaman secara langsung dan rekaman dari saluran telepon, untuk varian sampling rate yang digunakan mulai dari 8 KHz, 16 KHz dan 32 KHz. Pengujian akan mengamati nilai fitur yang terdiri atas frekuensi dasar, formant 1, formant 2, formant 3 dan formant 4 [9], [10].

2. Metode Penelitian

Untuk mengetahui pengaruh sampling rate pada fitur dilakukan pengujian dengan membandingkan fitur dari individu yang sama dan direkam dengan varian sampling rate dan lafal yang berbeda. Pengujian ini dilakukan menggunakan tiga tahapan yaitu (a) pengambilan sampel suara (b) proses ekstraksi dan (c) analisis fitur. Rekaman suara yang digunakan terdiri atas rekaman secara langsung dan rekaman melalui saluran telepon. Untuk rekaman secara langsung digunakan enam sampel suara dengan lafal pengucapan yang berbeda dan direkam dengan sampling rate 8 KHz, 16 KHz dan 32 KHz, demikian pula untuk rekaman melalui saluran telepon digunakan lafal yang sama. Jumlah data secara keseluruhan adalah 36 data rekam. Data tersebut selanjutnya masing-masing diekstraksi untuk diketahui fiturnya. Fitur yang diamati pada penelitian ini berupa frekuensi dasar, formant 1, formant 2, formant 3 dan formant 4. Nilai masing-masing fitur yang akan digunakan berupa nilai mean. Hasil pengujian dengan melakukan pengamatan masing-masing fitur dengan membandingkan pelafalan yang sama

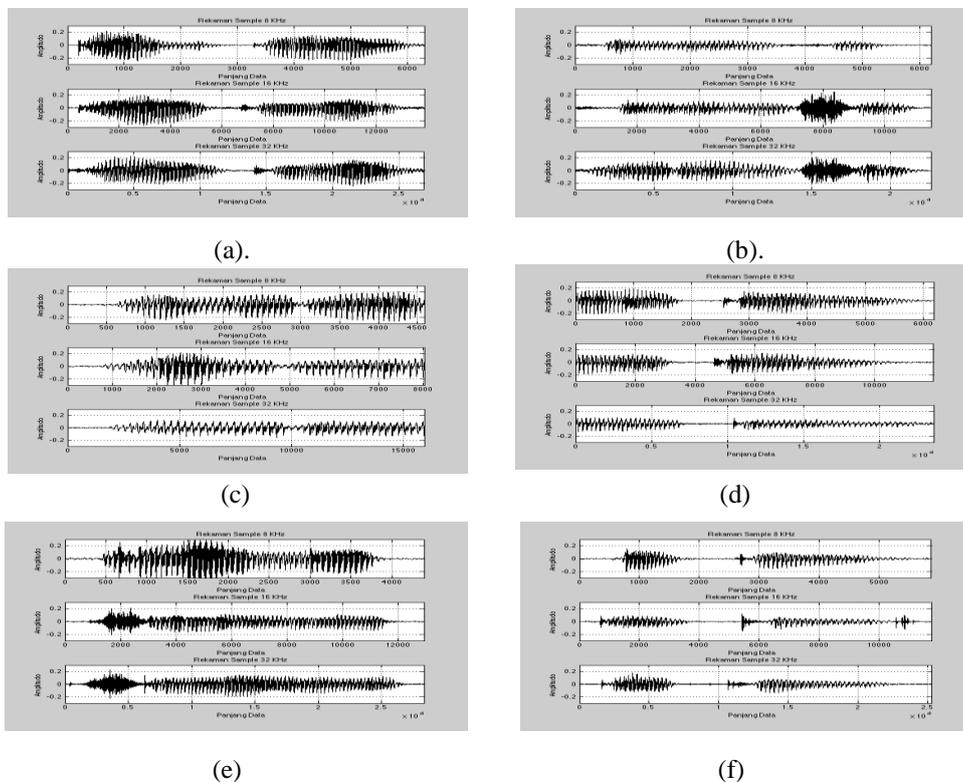
dengan varian sampling rate yang berbeda, hasil pengamatan ini digunakan untuk mengetahui pengaruh sampling rate terhadap fitur. Untuk alur dari proses penelitian ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

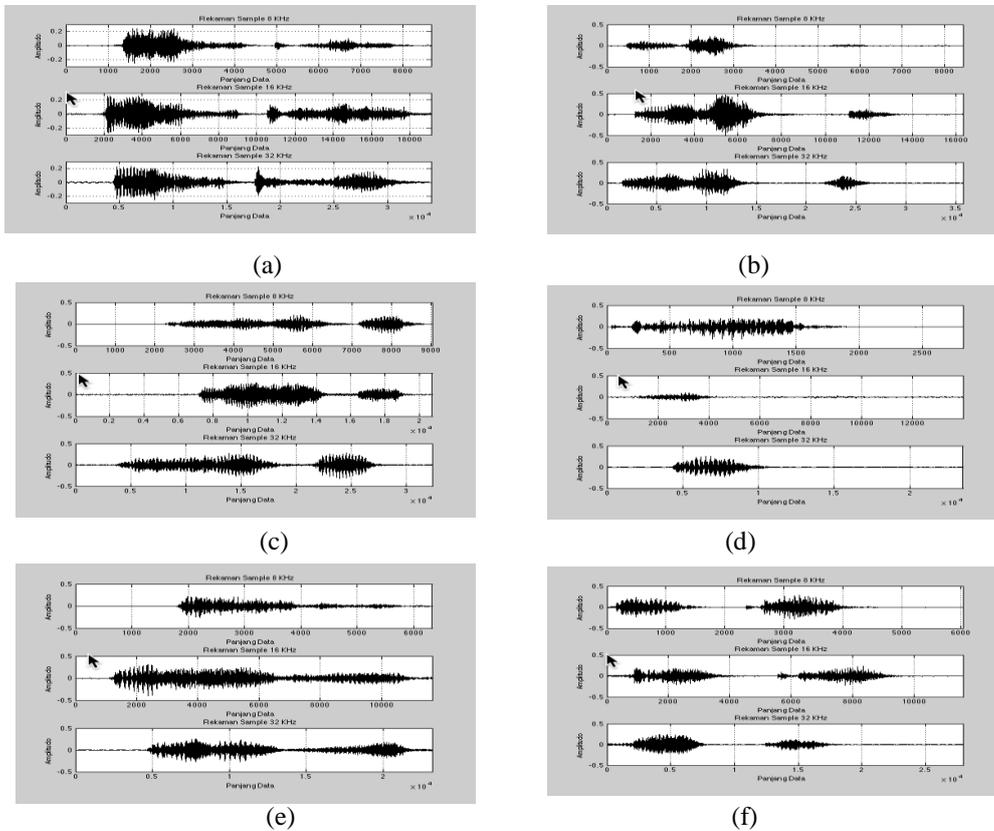
3. Hasil dan Pembahasan

Pengujian menggunakan alur Gambar 1, diawali dari proses pengambilan data rekaman suara. Untuk data rekaman dalam domain waktu ditunjukkan pada Gambar 2, rekaman tersebut diperoleh dari rekaman langsung, dengan berbagai macam ucapan serta varian sampling rate. Untuk data rekaman suara langsung diambil sebanyak 6 sampel kata yaitu Gambar 2(a) sampai 2(f), dengan masing-masing sampel direkam dengan sampling rate yang berbeda 8 KHz, 16 KHz dan 32 KHz.



Gambar 2. Rekaman secara langsung sampel 1 sampai 6 dengan varian sampling rate

Selain itu digunakan pula data dari perekaman suara melalui telepon, kriteria maupun jumlah sampel yang digunakan sama dengan sebelumnya yaitu sebanyak 6 rekaman pada Gambar 3(a) sampai dengan Gambar 3(f), bentuk sinyal rekaman tersebut dalam domain waktu secara keseluruhan ditunjukkan pada Gambar 3.



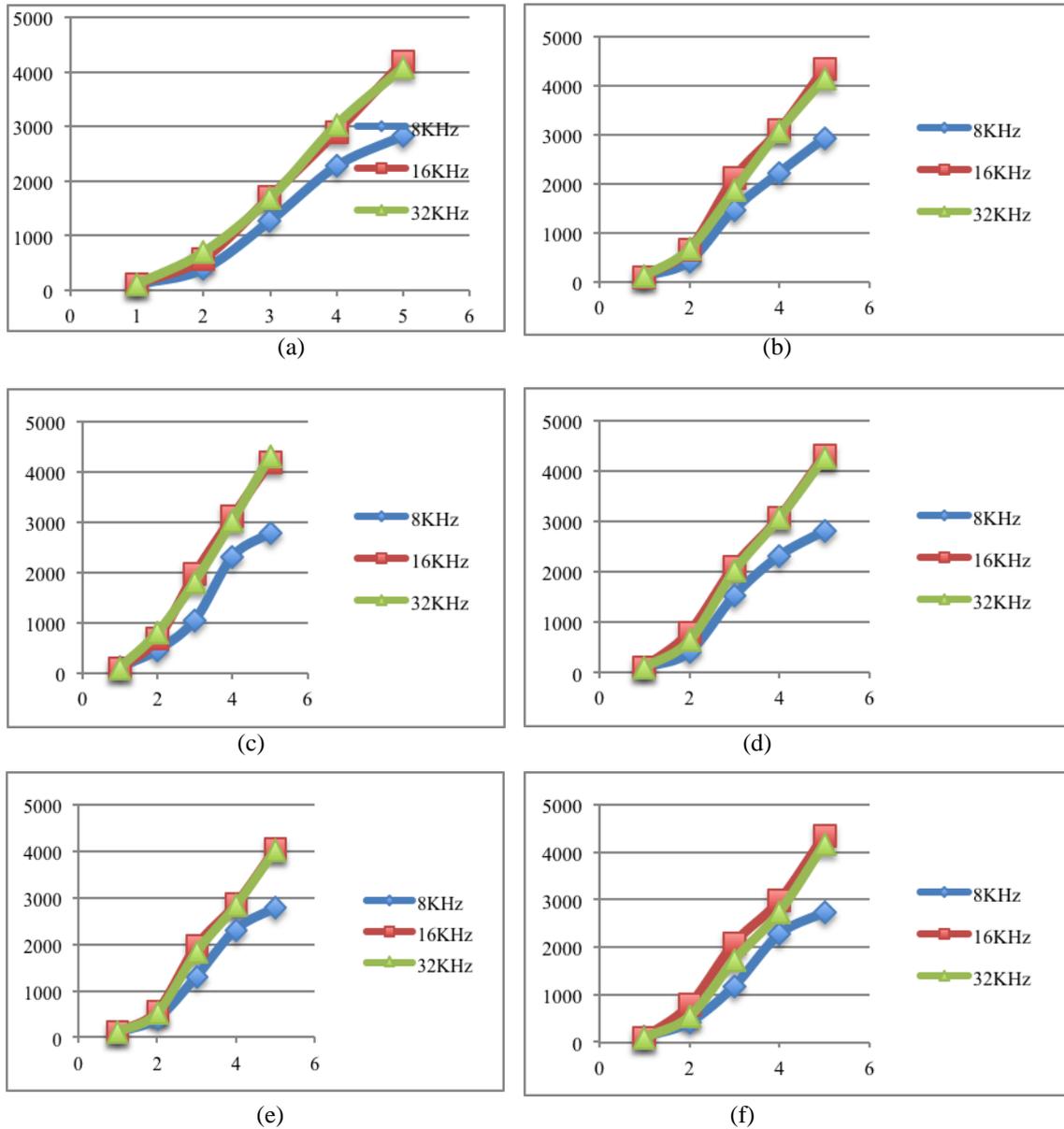
Gambar 3. Rekaman Melalui Telepon sampel 1 sampai 6 dengan varian sampling rate

Jumlah sampel rekaman yang digunakan bersumber dari rekaman langsung maupun melalui jaringan telepon ditunjukkan pada Tabel 1.

Tabel 1. Data uji

No Sampel	Jenis data rekam	Jumlah Data
1	Sampling Rate 8 KHz rekaman langsung	6
2	Sampling Rate 16 KHz rekaman langsung	6
3	Sampling Rate 32 KHz rekaman langsung	6
4	Sampling Rate 8 KHz melalui Telp	6
5	Sampling Rate 16 KHz melalui Telp	6
6	Sampling Rate 32 KHz melalui Telp	6

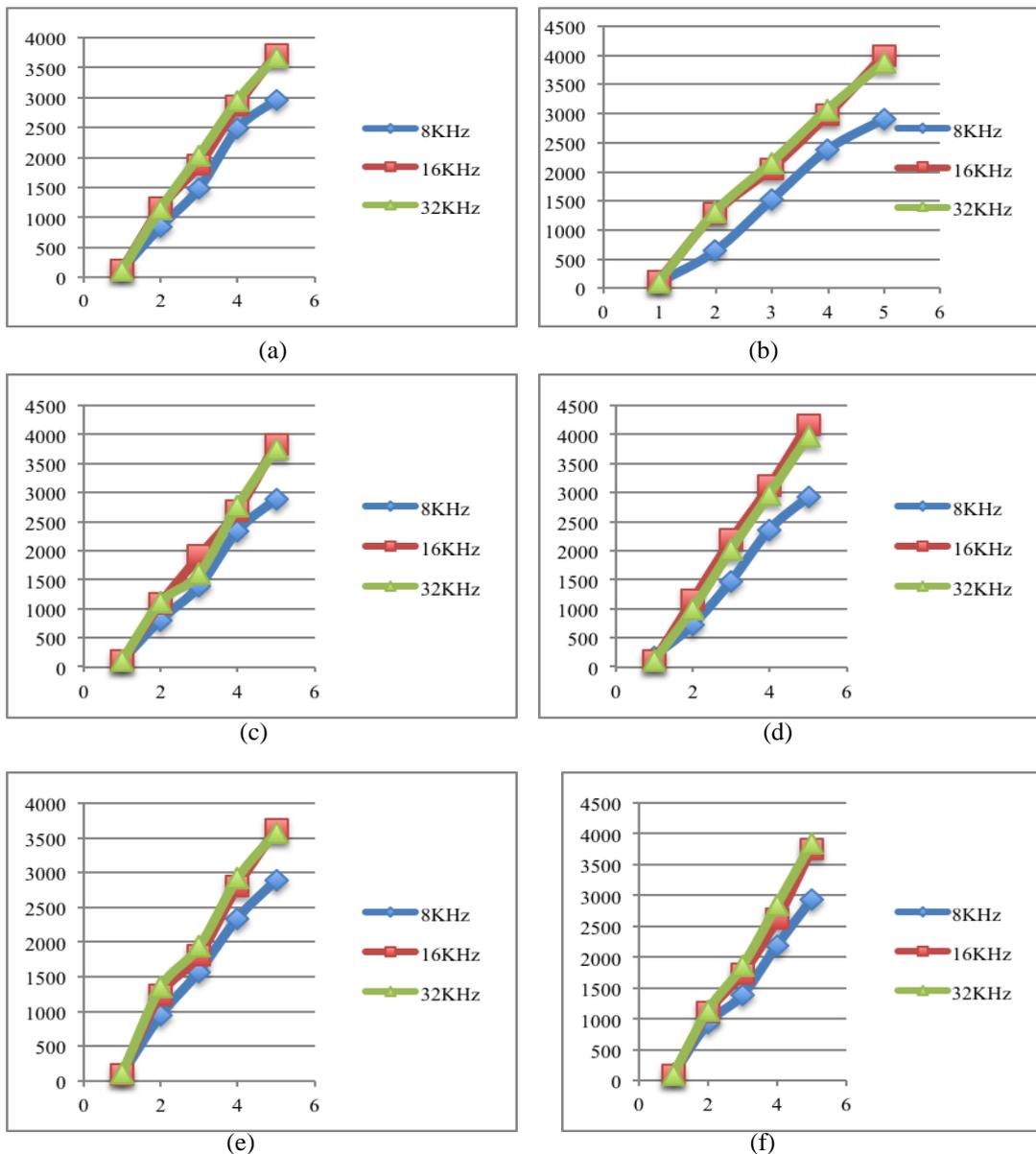
Data rekaman tersebut dilakukan pengujian untuk mengetahui pengaruh sampling rate terhadap nilai fitur yang didapat. Untuk mendapatkan nilai fitur dilakukan dengan mengekstraksi rekaman menggunakan perangkat lunak Praat. Hasil ekstraksi mendapatkan nilai frekuensi dasar, formant 1, formant 2, formant 3 dan formant 4. Masing-masing nilai tersebut diambil nilai meannya. Hasil pengujian untuk masing-masing sampel dengan berbagai varian sampling rate ditunjukkan pada Gambar 4.



Gambar 4. Hasil Perbandingan nilai fitur pada rekaman suara langsung

Pengujian dengan menggunakan suara langsung diberbagai varian, terlihat bahwa untuk frekuensi dasar mempunyai nilai yang sama ditunjukkan pada Gambar 4(a) sampai dengan 4(f). Untuk nilai formant 1 secara keseluruhan pada semua sampel mempunyai nilai yang hampir sama diberbagai varian. Sedangkan untuk nilai fitur dari formant 2 hingga formant 4, dengan sampling rate 8 KHz, nilai fitur tersebut berbeda dibandingkan dengan sampling rate 16 KHz dan 32 KHz. Sedangkan hasil pengujian dengan rekaman menggunakan sampling rate 16 KHz dan 32 KHz fitur yang didapatkan secara rata-rata terlihat sama.

Pengujian berikutnya menggunakan data rekaman melalui saluran telepon, suara yang didapatkan dari hasil rekaman selanjutnya diekstraksi untuk mendapatkan nilai fitur, hasil pengujian tersebut ditunjukkan pada Gambar 5.



Gambar 5. Perbandingan fitur pada rekaman suara dari telepon

Berdasarkan pengujian menggunakan rekaman dari telepon didapatkan bahwa untuk fitur frekuensi dasar dari Gambar 5(a) sampai dengan Gambar 5(f) mempunyai nilai yang sama. Sedangkan untuk nilai formant 1 pada Gambar 5(b) terlihat berbeda pada sampling rate 8 KHz, namun untuk sampel lainnya Gambar 5(a), 5(c), 5(d), 5(e) dan 5(f) mempunyai nilai fitur yang hampir sama. Untuk nilai formant 2 sampai dengan 4 secara umum pada sampling rate 16 HKz dan 32 KHz mempunyai nilai yang identik sedangkan pada sampling rate 8 mempunyai nilai fitur yang cenderung tidak sama mulai di nilai formant 2 sampai dengan formant 4.

4. Simpulan

Berdasarkan pengujian yang telah dilakukan dapat diambil kesimpulan sebagai berikut

- Karakteristik fitur dapat dipengaruhi sampling rate pada proses perekaman, namun tidak dapat menghilangkan ciri dan karakteristik identitas individu disuatu rekaman.
- Untuk sampling rate 16 KHz dan 32 KHz cenderung mempunyai hasil fitur yang sama dibandingkan dengan sampling rate 8 KHz.

-
- Frekuensi dasar mempunyai nilai fitur yang sama di masing-masing rekaman dengan sampling rate yang berbeda.
 - Untuk nilai fitur dengan sampling rate 8 KHz mempunyai nilai formant yang berbeda dibandingkan dengan sampling rate 16 KHz dan 32 KHz. Namun untuk formant 1 mempunyai kecenderungan sama diberbagai sampling rate.

Daftar Pustaka

- [1] G. Wen, J. Tuo, L. Jiang, and J. Wei, "Audio feature extraction for classification using relative transformation," *2012 Int. Conf. Audio, Lang. Image Process.*, pp. 260–265, Jul. 2012.
- [2] R. C. Maher, "Audio Forensic Examination," *Ieee Signal Processing Magazine*, no. March, pp. 84–94, 2009.
- [3] H. Zhao, "Audio Recording Location Identification Using Acoustic Environment Signature," vol. 8, no. 11, pp. 1746–1759, 2013.
- [4] Č. Petar and S. Maravi, "Methodological Frameworks of Digital Forensics," *Int. Symp. Intell. Syst. Informatics*, pp. 343–347, 2011.
- [5] G. Pop, D. Drăghicescu, and D. Burileanu, "On Forensic Speaker Recognition Case Pre-Assessment," *Speech Technol. Hum. - Comput. Dialogue (SpeD), IEEE*, pp. 1–8, 2013.
- [6] Abriyono and A. Harjoko, "Pengenalan Ucapan Suku Kata Bahasa Lisan Menggunakan Ciri LPC, MFCC, dan JST," *IJCCS*, vol. 6, no. 2, pp. 23–34, 2012.
- [7] H. B. Kekre and V. Kulkarni, "Speaker identification using feature vector reduction of row mean of different transforms," *Int. Conf. Commun. Inf. Comput. Technol.*, pp. 1–5, Oct. 2012.
- [8] K. Vijayan, V. Kumar, and K. S. R. Murty, "Feature Extraction from Analytic Phase of Speech Signals for Speaker Verification," no. September, pp. 1658–1662, 2014.
- [9] J. Franco pedroso, F. Espinoza cuadros, and J. Gonzalez rodriguez, "Formant Trajectories in Linguistic Units for Text-Independent Speaker Recognition," *Biometrics (ICB), Int. Conf.*, pp. 1–6, 2013.
- [10] S. Punyayodhin, D. Misra, R. Yadav, and V. Narang, "A Study of F1 Correlation with F0 in a Tone Language: Case of Thai," *2010 Int. Conf. Asian Lang. Process.*, pp. 330–333, Dec. 2010.