

Analisis Ekstraksi Ciri Sinyal Domain Waktu dan Frekuensi pada Suara Senjata Api

Difa Farhani Hakim¹

¹Program Studi Fisika, Universitas Negeri Jakarta

Jl. Rawamangun Muka, Jakarta 13220, Indonesia

Email: difafarhanihakim_1306620040@mhs.unj.ac.id

Abstrak. Ekstraksi ciri sinyal digital termasuk pada pengolahan sinyal audio. Ekstraksi ciri dilakukan untuk mendapatkan informasi yang penting dalam suatu data yang kompleks. Ekstraksi ciri dapat digunakan pada pengolahan citra digital, pengenalan suara, pengenalan bahasa, pengambilan makna penting dari sebuah percakapan, dan lain sebagainya. Pada penelitian ini akan dilakukan ekstraksi ciri sinyal pada domain waktu dan domain frekuensi serta *Principle Component Analysis* (PCA). Ekstraksi ciri yang dilakukan antara lain Energi, *Zero Crossing Rate*, *Energy of Entropy*, MFCC, dan PNCC. Dataset yang digunakan berupa suara sinyal senjata api dengan 9 kelas yang diperoleh dari website kaggle. Hasil PCA dari ekstraksi ciri yang diperoleh kurang baik, sehingga komponen dari PCA tersebut belum dapat menjelaskan dataset secara akurat.

1. Pendahuluan

Pemrosesan sinyal audio melibatkan analisis dan ekstraksi ciri dari sebuah sinyal digital, memprediksi perilakunya, mengenali pola, dan memahami korelasinya dengan sinyal lain yang serupa [1]. Proses dari ekstraksi ciri sangatlah penting untuk mengubah data mentah menjadi representasi yang lebih sederhana dan bermakna [2]. Hal tersebut dilakukan untuk pengambilan atau analisis lebih lanjut. Ekstraksi ciri digunakan dalam berbagai bidang, seperti pengolahan gambar, pengenalan suara, dan pemrosesan sinyal, untuk menemukan informasi yang relevan atau karakteristik penting dari data yang kompleks.

Ada kemungkinan bahwa ciri yang diperoleh bahwa fitur, pola, atau sifat yang dapat digunakan untuk menggambarkan dan membedakan objek atau fenomena yang diamati. Manusia dapat dengan mudah mengklasifikasikan dan membedakan antara berbagai suara tanpa melakukan upaya ekstra, misalnya manusia dapat dengan mudah mengklasifikasikan antara ucapan dan musik, suara mobil dan truk, kualitas ucapan bayi dan orang dewasa, berbagai pembicara, dll. Pada sebuah komputer diharapkan juga mampu mengklasifikasikan berbagai suara seperti manusia [1].

Peningkatan dari kinerja model sistem cerdas, seperti pengenalan wajah, klasifikasi gambar, dan pengenalan suara, bergantung pada keberhasilan ekstraksi ciri. Ekstraksi ciri berpengaruh pada hasil akurasi dan performa dari model sistem cerdas [3]. Dengan kemajuan teknologi, terutama dalam bidang kecerdasan buatan, ekstraksi ciri terus berkembang untuk mengatasi masalah rumit dalam menganalisis dan memahami jumlah data yang semakin besar dan kompleks.

Pada penelitian ini akan dilakukan ekstraksi ciri domain waktu dan frekuensi berupa ekstraksi Energi, *Zero Crossing Rate* (ZCR), dan *Energy of Entropy* (EoE), MFCC, dan PNCC pada dataset senjata api. Lalu, hasil ekstraksi ciri dilakukan *Principle Component Analysis* (PCA).

2. Metode

Ekstraksi ciri sinyal audio adalah proses pengambilan informasi penting atau sebuah karakteristik dari data sinyal audio digital. Ekstraksi ciri dapat digunakan dalam melakukan suatu pekerjaan tertentu pada tahap yang berbeda, seperti pengenalan suara, pengenalan bahasa, pengambilan makna penting dari sebuah percakapan, dan lain sebagainya [4]. Metode ekstraksi ciri terbagi menjadi ekstraksi domain waktu dan ekstraksi domain frekuensi. Ekstraksi domain frekuensi memiliki perbedaan dengan ekstraksi pada domain waktu. Hal tersebut dikarenakan pada ekstraksi domain frekuensi diperlukan adanya transformasi data dari domain waktu ke domain frekuensi dengan menggunakan transformasi Fourier. Beberapa ekstraksi domain waktu antara lain ekstraksi Energi, *Zero Crossing Rate* (ZCR), dan *Energy of Entropy* (EoE). Sedangkan, pada ekstraksi domain frekuensi antara lain *Mel-Frequency Cepstral Coefficient* (MFCC) dan *Power-Normalize Cepstral Coefficient* (PNCC).

2.1. Energi

Energi merupakan temuan dari variasi yang tinggi pada beberapa frame secara berurutan. Energi memiliki makna sebagai ciri kekuatan suara pada audio dan memiliki nama lain sebagai *Power* [5]. Tingkat kebaikan dari sinyal digital dapat dilihat dari tingginya nilai Energi [6]. Energi memiliki persamaan matematis sebagai berikut.

$$E(i) = \sum_{n=1}^{W_L} |x_i(n)|^2 \quad (1)$$

dengan E adalah energi atau *power*, W_L adalah panjang potongan sinyal audio (frame), x_i adalah koefisien nilai audio.

2.2. Zero Crossing Rate (ZCR)

Zero Crossing Rate (ZCR) merupakan nilai perubahan sinyal dalam frame atau perubahan sinyal dari nilai positif ke negatif. Pada sebuah awal percakapan nilai ZCR akan turun, sehingga hal tersebut dapat dijadikan sebagai estimasi awal dan akhir sebuah sinyal suara pada sebuah percakapan [7]. Selain itu, ZCR dapat mengukur *noise* pada sinyal audio [6]. ZCR dapat direpresentasikan dengan selisih dari fungsi *sign*.

$$Z(i) = \frac{1}{2W_L} \sum_{n=1}^{W_L} |\text{sgn}[x_i(n)] - \text{sgn}[x_i(n-1)]| \quad (2)$$

dengan fungsi *sign* adalah sebagai berikut.

$$\text{sgn}[x_i(n)] = \begin{cases} 1, & x_i(n) \geq 0 \\ -1, & x_i(n) < 0 \end{cases} \quad (3)$$

2.3. Energy of Entropy (EoE)

Energy of Entropy (EoE) adalah perubahan signifikan dari sebuah energi pada sinyal audio sebagai urutan probabilitas. EoE mengukur jumlah informasi yang dibawa oleh sinyal digital [8] dan nilai dari EoE harus kecil untuk menunjukkan adanya kestabilan sinyal [6]. Proses EoE dilakukan dengan membagi sebuah frame sinyal audio ke dalam potongan yang lebih kecil lagi. EoE dapat dituliskan sebagai berikut.

$$H(i) = \sum_{j=1}^K e_j \cdot \log_2 e_j \quad (4)$$

dengan H merupakan perubahan energi, j merupakan urutan probabilitas yang memiliki nilai $j = 1, 2, \dots, K$, K merupakan jumlah potongan frame yang dilakukan, e_j adalah perbandingan antara frame utuh $E_{\text{shortFrame}}$ dengan frame yang terpotong menjadi lebih kecil E_{subFrame} . e_j memiliki persamaan sebagai berikut.

$$e_j = \frac{E_{\text{subFrame}_j}}{E_{\text{shortFrame}_i}} \quad (5)$$

$$E_{\text{shortFrame}_i} = \sum_{n=0}^K E_{\text{subFrame}_k} \quad (6)$$

2.4. Mel-Frequency Cepstral Coefficient (MFCC)

MFCC adalah representasi spektrum gelombang dari karakteristik frekuensi dan kecepatan sinyal yang mirip dengan cara pendengaran manusia, sehingga MFCC akan melakukan filter pada frekuensi diatas 1KHz [9]. Proses MFCC dimulai dari melakukan *short-time Fourier Transform* (STFT), melakukan *filter bank triangular*, transformasi logaritmik, transformasi Cosinus Diskrit, normalisasi koefisien, dan diperoleh koefisien PNCC [5]. Pada MFCC, penggunaan filter bank triangular disebut juga sebagai skala Mel dengan rumus persamaan sebagai berikut.

$$Mel_f = 2595 \ln \left(1 + \frac{f}{700} \right) \quad (7)$$

dengan Mel_f merupakan frekuensi dalam skala Mel dan f merupakan frekuensi dalam satuan Hz.

2.5. Power-Normalize Cepstral Coefficient (PNCC)

Metode PNCC memiliki motivasi untuk mendapatkan kualitas yang baik ketika adanya perubahan parameter akustik. Metode ini diharapkan dapat memperoleh informasi tanpa merubah sinyal suara atau adanya penambahan informasi yang hilang. Penggunaan teknik metode PNCC memiliki kompleksitas komputasi yang mirip dengan *Perceptual Linear Prediction* (PLP) [10]. Pada metode PNCC digunakan analisis *meditum-time power* dan penghalusan frekuensi untuk meningkatkan akurasi dalam mendeteksi parameter lingkungan [11]. Pada PNCC, memiliki perbedaan pada penggunaan *filter bank gammatone* dibandingkan *triangular*. *Filter bank* tersebut berupa frekuensi yang berjarak sama secara linier diantara 200 Hz dan 8000 Hz [12].

2.6. Principle Component Analysis (PCA)

Principal Component Analysis (PCA) adalah suatu metode statistik multivariat yang digunakan untuk mengurangi dimensi dari suatu dataset yang kompleks atau berdimensi tinggi [13]. PCA memiliki tujuan untuk mengidentifikasi pola atau struktur yang terkandung dalam data dengan merepresentasikannya dalam bentuk dimensi yang lebih rendah. Hal tersebut disebut sebagai komponen utama atau *principal components* [14]. Proses PCA melibatkan transformasi linear dari variabel asli ke sejumlah variabel baru berupa kombinasi linear dari variabel asli [15]. Komponen utama pertama yang dihasilkan mengandung sebagian besar variabilitas data, diikuti oleh komponen kedua, dan seterusnya. Komponen-komponen ini bersifat ortogonal yang memiliki sifat tidak saling berkorelasi satu sama lain.

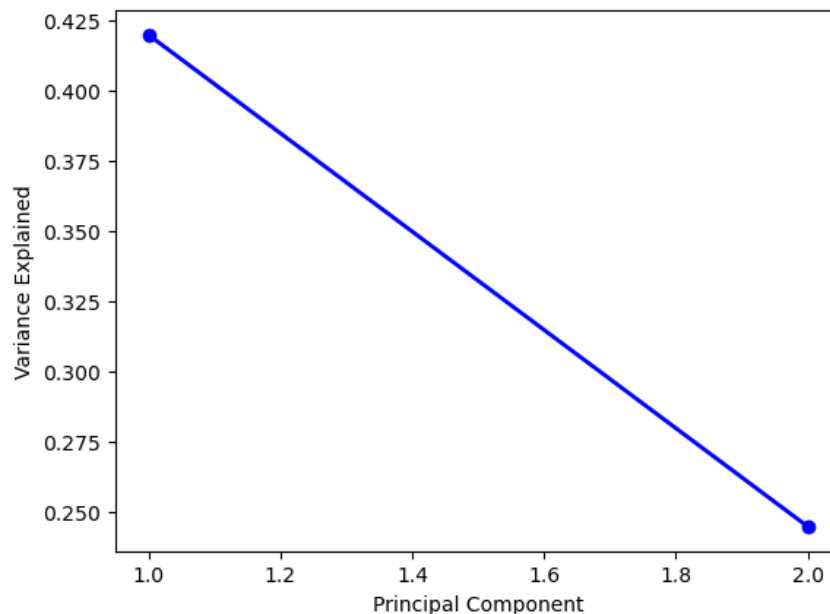
3. Hasil dan Diskusi

Pada penelitian ini digunakan dataset yang diperoleh dari Kaggle. Dataset tersebut memiliki 9 kelas suara senjata api dengan masing-masing kelas berjumlah 10 sampel. Dataset tersebut terdiri dari senjata api bertipe MG-42, M16, MP5, M4, Zastava M92, M249, IMI Desert Eagle, AK-47, dan AK-12. Pada masing-masing sampel, sinyal memiliki waktu suara dari rentang 1 hingga 3 sekon dengan frekuensi sampling sebesar 22050 Hz. Pada penelitian ini sinyal berhasil dilakukan ekstraksi ciri pada domain waktu dan frekuensi yaitu ekstraksi Energi, *Zero Crossing Rate*, *Energy of Entropy*, MFCC, dan PNCC yang terdapat pada TABEL 1. Proses ekstraksi ciri dimulai dengan mengimport modul Python, memasukkan dataset, melakukan ekstraksi ciri pada masing-masing sampel, melakukan *Principle Component Analysis* (PCA), memplot grafik PCA, dan melakukan analisis dari ekstraksi ciri dan hasil PCA yang diperoleh.

TABEL 1. Hasil Ekstraksi Ciri Data

No	Nama File	Energi	ZCR	EoE	MFCC	PNCC
1	AK-12/3 (1).wav	1044.1808	0.1538	2.7757	-36.0553	0.2300
2	AK-12/3 (10).wav	1446.6409	0.1918	3.0950	0.8450	0.2355
3	AK-12/3 (11).wav	225.11473	0.1285	1.0582	-11.7058	0.1697
4	AK-12/3 (12).wav	1088.4072	0.1709	2.739	-37.5438	0.2296
...
87	Zastava M92/9 (15).wav	865.87695	0.1284	3.2347	6.9802	0.3167
88	Zastava M92/9 (16).wav	557.04096	0.1599	3.3179	1.9406	0.3079
89	Zastava M92/9 (17).wav	576.02033	0.1713	3.3131	1.6807	0.3035
90	Zastava M92/9 (18).wav	877.00812	0.1479	3.2371	8.7064	0.3122

Pada penelitian ini diperoleh hasil komponen utama dari PCA terdapat pada GAMBAR 1. Hasil yang diperoleh yaitu variasi sebesar 0.4196 dan 0.2446. Dari hasil tersebut menandakan bahwa PCA1 mewakili sampel data sebesar 41.96%, sedangkan pada PCA2 mewakili sampel data sebesar 24.46%. Dari kedua PCA tersebut diperoleh bahwa 33.58% informasi data telah hilang.

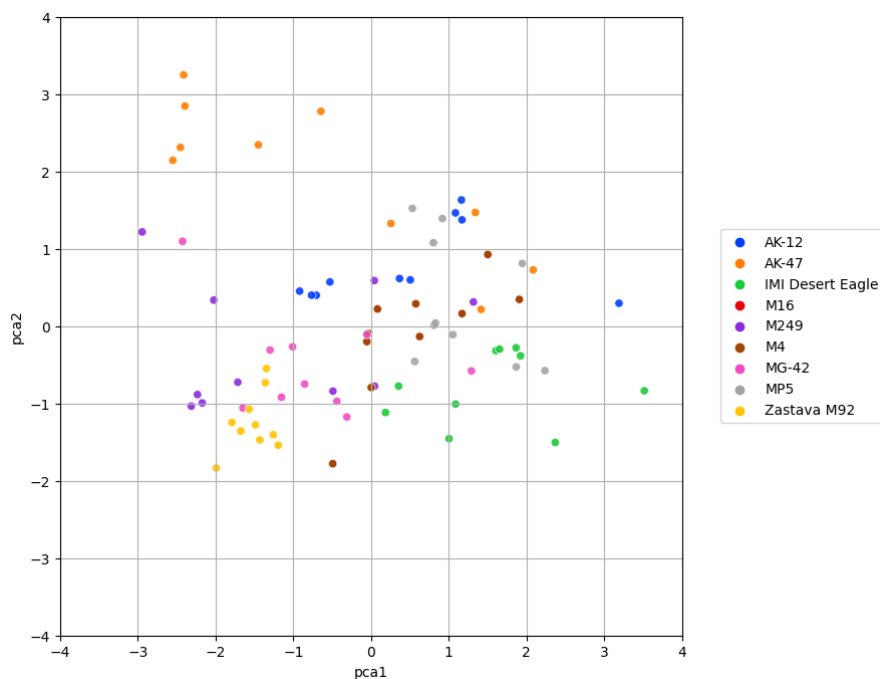
**GAMBAR 1.** *Scree plot* PCA

Kemudian, pada setiap ekstraksi ciri diperoleh hasil PCA terdapat pada TABEL 2. Pada PCA1, variasi data dipengaruhi oleh ciri EoE dan PNCC ditandai dengan nilai korelasi EoE dan PNCC masing-masing sebesar -0.60 dan -0.54. Selain itu, PCA1 memiliki korelasi keterbalikan dengan semua ciri ditandai dengan nilai PCA1 yang negatif pada setiap ciri. Pada PCA2, variasi data dipengaruhi oleh ciri MFCC dan ZCR ditandai dengan nilai korelasi MFCC dan ZCR masing-masing sebesar -0.47 dan 0.68.

TABEL 2. Hasil PCA pada setiap ekstraksi ciri

No	Ciri	PCA1	PCA2
1	Energy	-0.361424	0.385011
2	ZCR	-0.272838	0.6884
3	EoE	-0.600188	0.094492
4	MFCC	-0.371671	-0.471795
5	PNCC	-0.54458	-0.38256

Untuk melihat persebaran data dilakukan visualisasi melalui *scatter plot* pada GAMBAR 2. Dari hasil plot yang diperoleh bahwa senjata api bertipe Zastava M92 memiliki sampel yang sangat berdekatan satu sama lain, kemudian diikuti dengan M4, MG-42, dan IMI Desert Eagle. Pada tipe senjata yang lain, sampel memiliki outlier yang cukup jauh dibandingkan kumpulannya, seperti pada AK-47 yang memiliki outlier terjauh. Namun, secara umum hasil PCA yang diperoleh kurang baik dikarenakan setiap sampel pada tipe senjata belum mengumpul secara baik.



GAMBAR 2. Persebaran data PCA

Selain itu, pada PCA1 nilai terendah ada pada M249 dan AK-47, sedangkan nilai tertinggi ada pada IMI Desert Eagle dan AK-12. Hal tersebut menandakan bahwa variasi pada PCA1 bergantung pada kedua tipe jenis senjata api tersebut. Pada PCA2 nilai terendah ada pada Zastava, M4, dan M16, sedangkan nilai tertinggi ada pada AK-47. Hal tersebut menandakan bahwa variasi pada PCA2 bergantung pada tipe senjata api yang memiliki kesamaan dengan Zastava atau dengan AK-47.

4. Kesimpulan

Pada penelitian ini telah dilakukan PCA pada ekstraksi ciri Energi, *Zero Crossing Rate*, *Entropy of Entropy*, MFCC, dan PNCC. Hasil PCA menandakan bahwa ciri EoE dan PNCC memengaruhi varian pada PCA1, sedangkan ciri PNCC dan ZCR memengaruhi varian pada PCA2. Pada PCA1 dan PCA2 diperoleh hasil variasi yang kurang baik, sehingga kedua komponen tersebut belum mampu secara akurat untuk menjelaskan dataset tersebut. Selain itu, hasil plot PCA yang diperoleh kurang baik dikarenakan ada beberapa kelas pada tipe senjata yang belum mengumpul dengan baik.

Referensi

- [1] Sharma G, Umapathy K and Krishnan S 2020 Trends in audio signal feature extraction methods *Applied Acoustics* **158** 107020
- [2] Kunekar P R, Sadavarte K S, Khambad P R, Lokhande R B and Khara M B 2023 Audio feature extraction: Foreground and Background audio separation using KNN algorithm *In International Journal of Science and Research Archive* **9** 269–76
- [3] Mazumder M A 2019 Feature Extraction Techniques for Speech Processing: A Review. In *International Journal of Advanced Trends in Computer Science and Engineering The World Academy of Research in Science and Engineering* **8** 285–92
- [4] Harsemadi I G 2018 Purwarupa Sistem Ekstraksi Fitur Audio pada Musik Gamelan Tradisional Bali *Seminar Nasional Telekomunikasi dan Informatika*
- [5] Giannakopoulos T and Pikrakis A 2014 *Introduction to Audio Analysis*, ed T Giannakopoulos and A Pikrakis (Oxford: Academic Press)
- [6] Raj V A and Dhas M D K 2022 Analysis of audio signal using various transforms for enhanced audio processing *International journal of health sciences* **6** 12989-3001
- [7] Aggarwal G, Gochhayat S P and Singh L 2021 *Machine Learning and the Internet of Medical Things in Healthcare*, ed K K Singh, *et al.*: Academic Press) pp 209-50
- [8] Djebbar F and Ayad B 2017 Energy and entropy based features for WAV audio steganalysis **8** 168-81
- [9] Shivani J and Brij K 2017 Comparative study of voice print Based acoustic features: MFCC and LPCC *International Journal of Advanced Engineering, Management and Science* **3**
- [10] Shareef S R and Al-Irhayim Y F M 2022 Comparison Between Features Extraction Techniques for Impairments Arabic Speech *Al-Rafidain Engineering Journal (AREJ)* **27** 190-7
- [11] Kim C and Stern R M 2016 Power-Normalized Cepstral Coefficients (PNCC) for Robust Speech Recognition *IEEE/ACM Transactions on Audio, Speech, and Language Processing* **24** 1315-29
- [12] Alasadi A A, Aldhayni T H, Deshmukh R R, Alahmadi A H and Alshebami A S 2020 Efficient Feature Extraction Algorithms to Develop an Arabic Speech Recognition System *Engineering, Technology & Applied Science Research* **10** 5547-53
- [13] Mishra S, Sarkar U, Taraphder S, Datta S, Swain D, Saikhom R, Panda S and Laishram M 2017 Principal Component Analysis *International Journal of Livestock Research* **1**
- [14] Khikmah K N 2021 Penerapan Principal Component Analysis dalam Penentuan Faktor Dominan Cuaca Terhadap Penyebaran Covid-19 di Surabaya *ESTIMASI: Journal of Statistics and Its Application* **2** 11-8
- [15] Tamonob A M, Saefuddin A and Wigena A H 2015 Nonlinear Principal Component Analysis And Principal Component Analysis With Successive Interval In K-Means Cluster Analysis *Forum Statistika dan Komputasi : Indonesian Journal of Statistics* **20** 68-77