

Penerapan Singular Value Decomposition (SVD) untuk Denoising Audio

Nathanael Rachmat - 13523142^{1,2}

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia

13523142@mahasiswa.itb.ac.id, nathanael.rachmat@gmail.com

Abstraksi— Makalah ini membahas penggunaan Singular Value Decomposition (SVD) untuk menghilangkan noise pada rekaman audio. Dengan merepresentasikan sinyal audio dalam bentuk matriks waktu-frekuensi melalui STFT, komponen noise dipisahkan berdasarkan nilai singular. Thresholding diterapkan untuk mengabaikan nilai singular kecil yang merepresentasikan noise. Hasil menunjukkan bahwa metode ini meningkatkan kualitas rekaman audio dengan efektif, meskipun beberapa artefak seperti pola blok pada spektrum masih terlihat.

Keywords— Noise, denoise, SVD, audio, STFT, spektrum, thresholding.

I. PENDAHULUAN

Rekaman suara pada mikrofon sering kali terganggu dengan keberadaannya suara latar belakang (noise). Noise dalam rekaman audio dapat berasal dari berbagai sumber, seperti mikrofon yang sensitif terhadap angin, gema ruangan, atau perangkat elektronik dengan sinyal yang tidak stabil. Dalam beberapa kasus, noise berbentuk statis atau bising putih (white noise), sementara pada kasus lain, noise dapat berupa suara yang lebih kompleks, seperti dengungan atau klik. Dalam aplikasi profesional, pelindung mikrofon atau mikrofon berkualitas tinggi digunakan untuk mengurangi noise yang berada di sekitar. Namun, solusi ini tidak selalu efektif bagi pengguna umum yang terbatas dalam anggarannya.

Metode tradisional untuk menghilangkan noise biasanya melibatkan filter frekuensi atau algoritma berbasis statistik. Meskipun teknik ini cukup efektif untuk noise sederhana seperti bising putih, mereka sering gagal menangani noise yang memiliki spektrum frekuensi yang mirip dengan sinyal utama. Selain itu, metode berbasis filter dapat menghapus frekuensi penting dari sinyal utama, sehingga menghasilkan distorsi atau hilangnya detail. Oleh karena itu, diperlukan pendekatan yang lebih adaptif.

Solusi alternatif yang efisien dan praktis untuk menghilangkan noise dalam sebuah sinyal digital adalah penerapannya *Singular Value Decomposition* (SVD). SVD adalah metode dekomposisi matriks yang sering digunakan untuk mengolah data. Dalam konteks audio, SVD memungkinkan pemisahan komponen utama sinyal dari noise berdasarkan nilai singularnya. Noise yang berupa nilai singular kecil dihilangkan sebagai perwujudan dari

penghilangan noise (denoising).

Makalah ini membahas penerapan SVD dalam penghilangan noise pada rekaman audio. Dengan merepresentasikan sinyal dalam bentuk matriks waktu-frekuensi dan menerapkan ambang batas pada nilai singular, dapat ditunjukkan bahwa SVD dapat meningkatkan kualitas rekaman audio secara signifikan. Selain itu, berbagai parameter dan pendekatan dieksplorasi untuk mengoptimalkan hasil pengolahan sinyal.

II. DASAR TEORI

Terdapat berbagai teori yang digunakan untuk melakukan denoise dari sebuah audio.

A. Representasi Matriks Audio dengan STFT

Audio pada dasarnya adalah sinyal satu dimensi dalam domain waktu, dengan amplitudo yang bervariasi terhadap waktu. Namun, untuk pengolahan yang lebih kompleks seperti penghilangan noise, sinyal ini perlu dianalisis dalam matriks berdomain waktu-frekuensi.

Short-Time Fourier Transform (STFT) adalah metode pemrosesan audio yang kuat dalam pemrosesan sinyal audio, memungkinkan analisis frekuensi yang berubah seiring waktu. STFT memecah sinyal menjadi segmen-segmen pendek, yang diasumsikan stasioner, sehingga memungkinkan analisis frekuensi lokal dalam domain waktu.

Secara matematis, STFT didefinisikan sebagai berikut:

$$X_m(\omega) = \sum_{n=-\infty}^{\infty} x(n) \cdot \omega(n - mR) \cdot e^{-j\omega n} \quad (1)$$

dimana $x(n)$ adalah sinyal input pada waktu n , $\omega(n)$ adalah fungsi jendela dengan panjang M , m adalah indeks frame, R adalah ukuran lompatan dalam sampel antara perhitungan STFT berturut-turut, dan ω adalah sudut frekuensi.

Fungsi jendela $\omega(n)$ digunakan untuk mengekstraksi segmen lokal dari sinyal, dan ukuran lompatan R menentukan seberapa sering analisis dilakukan sepanjang sinyal. Pemilihan fungsi jendela dan ukuran lompatan mempengaruhi resolusi waktu dan frekuensi dari STFT.

Penting untuk memastikan bahwa fungsi jendela memenuhi properti Constant Overlap-Add (COLA) pada

ukuran lompatan RRR, yang berarti:

$$\sum_{n=-\infty}^{\infty} \omega(n - mR) = 1, \forall x \in \mathbb{Z} (\omega \in \text{COLA}(R)) \quad (2)$$

Properti COLA memastikan bahwa penjumlahan segmen-segmen yang dijendela menghasilkan rekonstruksi sempurna dari sinyal asli tanpa distorsi atau diskontinuitas. Beberapa fungsi jendela umum, seperti jendela Hamming dan Bartlett, memenuhi properti COLA dengan tumpang tindih 50%

Hasil dari STFT adalah matriks kompleks X , dimana baris merepresentasikan frekuensi, kolom merepresentasikan segmen waktu, dan elemen matriks $X[f, t]$ menunjukkan kontribusi energi pada frekuensi f di segmen waktu t .

Dalam konteks penghilangan noise pada sinyal audio, STFT memungkinkan identifikasi dan manipulasi komponen frekuensi yang tidak diinginkan dalam domain waktu-frekuensi. Dengan menerapkan teknik seperti *Singular Value Decomposition* (SVD) pada representasi matriks dari STFT, dimungkinkan untuk memisahkan komponen sinyal utama dari noise, sehingga meningkatkan kualitas rekaman audio.

B. Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD) adalah salah satu metode dekomposisi matriks yang memfaktorkan matriks A berukuran $m \times n$ menjadi tiga matriks yaitu U , Σ , dan V^T . Faktorisasi ini dinyatakan dalam persamaan:

$$A = U\Sigma V^T \quad (3)$$

Matriks U berukuran $m \times m$, merupakan matriks ortogonal yang kolom-kolomnya adalah vektor singular kiri dari matriks A . Vektor singular kiri (u_1, u_2, \dots, u_m) adalah basis ortonormal untuk ruang kolom A .

Matriks Σ berukuran $m \times n$, merupakan matriks diagonal yang elemen diagonal utamanya ($\sigma_1, \sigma_2, \dots$) adalah nilai singular dari A . Nilai singular diperoleh sebagai akar pangkat dua dari nilai eigen matriks $A^T A$, dengan urutan ($\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k > 0$).

Matriks V^T berukuran $n \times n$, merupakan transpos dari matriks ortogonal V , di mana kolom-kolomnya adalah vektor singular kanan dari matriks A . Vektor singular kanan (v_1, v_2, \dots, v_n) adalah basis ortonormal untuk ruang baris A .

Dekomposisi SVD dapat dihitung dengan menentukan nilai eigen dari matriks $A^T A$ untuk mendapatkan (σ_i). Kemudian, vektor eigen yang sesuai dengan nilai eigen tersebut ditentukan dan dinormalisasi untuk membentuk matriks V . Setelah itu, vektor singular kiri (u_i) dihitung dengan persamaan berikut:

$$u_i = \frac{A v_i}{\sigma_i} \quad (4)$$

Yang terakhir, matriks U , Σ , dan V^T disusun untuk membentuk faktorisasi $A = U\Sigma V^T$.

Dengan sifatnya yang mampu memisahkan komponen utama dari noise berdasarkan nilai singular, Singular Value Decomposition (SVD) menjadi alat yang sangat efektif dalam pengolahan sinyal, khususnya untuk penghilangan noise pada rekaman audio. Nilai singular besar mencerminkan kontribusi utama dari sinyal asli, sementara nilai singular kecil sering kali merepresentasikan noise atau komponen yang tidak diinginkan. Melalui penerapan *threshold* yang tepat pada nilai singular, metode ini memungkinkan penghilangan noise secara efisien tanpa merusak struktur utama sinyal.

C. Inverse ISTFT

ISTFT adalah operasi invers dari STFT yang mengubah sinyal dari domain waktu-frekuensi kembali ke domain waktu. Proses rekonstruksi ini melibatkan kombinasi kembali magnitudo ($|X|$) dan fase ($\angle X$) dari hasil STFT yang telah dimodifikasi.

Secara matematis, ISTFT didefinisikan sebagai:

$$x(t) = \sum_{m=-\infty}^{\infty} \left(\int_{-\infty}^{\infty} X_m(\omega) \cdot e^{-j\omega t} d\omega \right) \omega(t - mR)$$

di mana $X_m(\omega)$ adalah representasi waktu-frekuensi (hasil STFT), $e^{-j\omega t}$ adalah basis Fourier untuk mengembalikan sinyal ke domain waktu, $\omega(t - mR)$ adalah fungsi jendela yang digunakan untuk proses STFT, dan R adalah ukuran lompatan antar segmen.

Dalam konteks penghilangan noise, ISTFT digunakan untuk merekonstruksi sinyal bersih setelah noise dihilangkan pada domain waktu-frekuensi. Dengan menggabungkan kembali magnitudo dan fase, serta memastikan integritas sinyal dengan overlap-add, ISTFT menghasilkan sinyal yang lebih bersih dengan kualitas rekaman yang meningkat.

III. METODOLOGI

Metodologi yang digunakan dalam penelitian ini bertujuan untuk menerapkan Singular Value Decomposition (SVD) dalam penghilangan noise pada sinyal audio. Proses dilakukan melalui beberapa tahap utama, termasuk pengumpulan data audio, transformasi sinyal ke domain waktu-frekuensi menggunakan Short-Time Fourier Transform (STFT), penerapan SVD pada representasi matriks sinyal, thresholding untuk memisahkan noise, serta rekonstruksi sinyal. Seluruh proses ini dilakukan dengan mempertimbangkan batasan dan format data untuk memastikan hasil yang optimal.

A. Input Data

Data yang digunakan dalam penelitian ini adalah rekaman audio yang diambil menggunakan perangkat *handphone* peneliti. Rekaman dilakukan dalam lingkungan yang mengandung noise, seperti suara latar belakang kendaraan atau angin, untuk mensimulasikan skenario dunia nyata. Sebelum diproses lebih lanjut, rekaman audio dikonversi ke format .wav dengan laju sampel standar sebesar 44,1 kHz. Format ini dipilih karena kompatibel dengan algoritma pemrosesan sinyal dan representasi

matriks yang digunakan pada langkah-langkah berikutnya.

```
import librosa
import numpy as np
import soundfile as sf

0.0s

wav_file = "input_audio.wav"
noisy_audio, sr = librosa.load(wav_file, sr=None)

0.0s
```

Gambar 1. Kode setup dan input data dalam python.

B. Representasi Sinyal Audio dengan STFT

Langkah ini bertujuan untuk mentransformasi sinyal audio dari domain waktu menjadi domain waktu-frekuensi menggunakan metode Short-Time Fourier Transform (STFT). Representasi ini memungkinkan analisis distribusi energi sinyal audio pada berbagai frekuensi dan waktu, sehingga mempermudah identifikasi noise dan sinyal utama.

Proses dilakukan dengan membagi sinyal audio menjadi segmen-segmen pendek menggunakan fungsi jendela (windowing). Setiap segmen dianggap stasioner, sehingga transformasi Fourier dapat diterapkan untuk setiap segmen. Secara matematis

Sebelum diproses, data audio dalam format .wav dibaca dan dinormalisasi untuk memastikan semua nilai amplitudo berada dalam rentang yang seragam. Parameter n_fft (jumlah titik Fourier) dan hop_length (lompatan antar segmen) ditentukan untuk menghasilkan resolusi waktu dan frekuensi yang sesuai.

Kemudian, audio diproses menggunakan Transformasi Fourier. Transformasi Fourier dilakukan pada setiap segmen sinyal yang telah diwindowing. Hasilnya berupa matriks kompleks X , di mana baris merepresentasikan komponen frekuensi dan kolom merepresentasikan segmen waktu. Matriks hasil STFT dipisahkan menjadi magnitudo ($|X|$) dan fase ($\angle X$), sehingga magnitudo digunakan untuk analisis lebih lanjut.

Proses ini menghasilkan matriks magnitudo waktu-frekuensi yang merepresentasikan energi sinyal pada berbagai frekuensi dan waktu, yang kemudian digunakan pada langkah berikutnya untuk pemisahan sinyal dan noise dengan SVD.

```
n_fft = 8192
hop_length = 128
stft_matrix = librosa.stft(noisy_audio, n_fft=n_fft, hop_length=hop_length)
magnitude, phase = np.abs(stft_matrix), np.angle(stft_matrix)

0.0s
```

Gambar 2. Kode representasi sinyal audio dengan STFT dalam python.

C. Penerapan SVD pada Matriks STFT

Langkah ini bertujuan untuk memisahkan komponen utama sinyal dari noise dengan menggunakan Singular Value Decomposition (SVD) pada matriks magnitudo hasil Short-Time Fourier Transform (STFT).

Matriks magnitudo $|X|$ hasil STFT digunakan sebagai input. Kemudian, Matriks $|X|$ dipecah menjadi tiga matriks U , Σ , dan V^T menggunakan algoritma SVD. Pada tahap ini, nilai-nilai singular siap untuk dianalisis dan dimanipulasi lebih lanjut untuk menentukan ambang batas yang akan digunakan pada tahap *thresholding*.

Penerapan SVD ini memungkinkan pemisahan komponen sinyal utama dari noise dengan memanfaatkan distribusi energi dalam matriks nilai singular Σ . Matriks U dan V^T mempertahankan struktur hubungan waktu dan frekuensi dalam sinyal.

```
U, S, Vt = np.linalg.svd(magnitude, full_matrices=False)

0.4s
```

Gambar 3. Kode SVD pada Matriks STFT dalam python.

D. Penghapusan Nilai Singular Kecil

Tahap ini bertujuan untuk menghilangkan noise dari sinyal audio dengan menghapus nilai singular kecil pada matriks Σ , hasil dekomposisi SVD. Dalam proses ini, nilai singular kecil di bawah ambang batas (*threshold*) dianggap sebagai noise, sementara nilai singular besar yang mewakili sinyal utama dipertahankan.

Proses penghapusan dimulai dengan penentuan ambang atas. Ambang batas ditentukan berdasarkan proporsi tertentu dari nilai singular maksimum, direpresentasikan dengan persamaan berikut:

$$threshold = \alpha \cdot \max(\Sigma); 0 < \alpha \leq 1 \quad (5)$$

Untuk penelitian ini, ambang batas ditentukan secara manual dengan mempertimbangkan kemudahan analisis dan observasi visual pada distribusi nilai singular. Penentuan manual memberikan fleksibilitas untuk menyesuaikan *threshold* terhadap data spesifik, khususnya ketika metode otomatis sulit diterapkan atau kurang efektif. α yang ditentukan adalah $\alpha = 0.25$.

Threshold diterapkan dengan menghapus elemen nilai singular σ_i pada matriks diagonal Σ (diubah menjadi nol) jika $\sigma_i < threshold$. Persamaan yang berlaku adalah sebagai berikut,

$$\sigma_i = \begin{cases} \sigma_i, & \text{if } \sigma_i \geq threshold \\ 0, & \text{if } \sigma_i < threshold \end{cases} \quad (6)$$

Setelah *thresholding*, matriks nilai singular baru Σ' hanya mempertahankan komponen utama sinyal, sementara noise dihilangkan. Dengan penghapusan nilai singular kecil ini, noise dapat diminimalkan tanpa mengorbankan komponen utama sinyal. Langkah ini menjadi dasar untuk rekonstruksi sinyal bersih pada tahap berikutnya.

```
threshold = 0.25 * np.max(S)
S_denoised = np.where(S > threshold, S, 0)

0.0s
```

Gambar 4. Kode penghapusan nilai singular kecil dalam python.

E. Rekonstruksi Sinyal

Tahap ini bertujuan untuk membangun kembali sinyal audio dari matriks waktu-frekuensi yang telah diolah. Setelah nilai singular kecil dihapus pada langkah sebelumnya, matriks magnitudo waktu-frekuensi hasil thresholding ($|X|_{denoised}$) direkonstruksi menjadi sinyal audio dalam domain waktu menggunakan Inverse STFT (ISTFT). Proses ini melibatkan kombinasi kembali magnitudo dan fase dari matriks hasil STFT.

Sebelum melakukan rekonstruksi sinyal, matriks magnitudo denoised $|X|_{denoised}$ dapat dihitung dengan persamaan berikut:

$$|X|_{denoised} = U\Sigma'V^T$$

Di mana matriks U dan V^T adalah matriks ortogonal yang sama dengan hasil dekomposisi sebelumnya, Σ' adalah matriks diagonal nilai singular yang telah diubah melalui proses thresholding untuk menghilangkan komponen noise.

Matriks magnitudo ($|X|_{denoised}$) digabungkan kembali dengan fase asli ($\angle X$) dari STFT untuk membentuk kembali matriks waktu-frekuensi. Matriks hasil rekombinasi adalah sebagai berikut:

$$X' = |X|_{denoised} \cdot e^{j\angle X}$$

Matriks X' diubah kembali ke domain waktu menggunakan Inverse Short-Time Fourier Transform (ISTFT). Fungsi jendela yang sama dengan langkah STFT digunakan untuk memastikan rekonstruksi sinyal tanpa artefak. Rekonstruksi dilakukan dengan metode overlap-add, di mana segmen-segmen waktu yang tumpang tindih digabung untuk menghasilkan sinyal kontinu. Hasil rekonstruksi dinormalisasi untuk memastikan amplitudo sinyal berada dalam rentang yang sesuai, menghindari distorsi atau clipping. Sinyal hasil rekonstruksi $x(t)$ diharapkan memiliki noise yang berkurang dengan kualitas sinyal utama tetap terjaga.

```
stft_denoised = magnitude_denoised * np.exp(1j * phase)
audio_denoised = librosa.istft(stft_denoised, hop_length=hop_length)
sf.write('denoised_audio.wav', audio_denoised, sr)
✓ 0.2s
```

Gambar 5. Kode rekonstruksi sinyal dalam python.

F. Evaluasi Hasil

Langkah terakhir dalam metodologi ini adalah mengevaluasi keberhasilan proses penghilangan noise pada sinyal audio menggunakan berbagai metrik dan analisis. Evaluasi dilakukan untuk menilai seberapa baik metode STFT, SVD, dan thresholding dalam meningkatkan kualitas sinyal dengan mengurangi noise tanpa merusak komponen utama sinyal.

Spektrum waktu-frekuensi dari sinyal asli, sinyal dengan noise, dan sinyal hasil denoising dibandingkan secara visual untuk mengevaluasi hasil proses denoising. Indikator keberhasilan adalah noise yang tersebar di seluruh spektrum frekuensi berkurang dan pola energi sinyal utama tetap terjaga. Untuk visualisasi, gunakan python dengan librosa specshow.

Selain evaluasi visual, metode evaluasi menggunakan

pendengaran juga dilakukan untuk memberikan perspektif subjektif terhadap kualitas hasil denoising. Pendengaran dilakukan untuk memastikan bahwa noise yang terdengar pada frekuensi rendah dan tinggi telah berkurang tanpa mengorbankan kualitas komponen utama dari sinyal audio. Hasil pendengaran dapat mengungkap efek samping, seperti suara robotik atau distorsi, yang mungkin tidak terlihat dalam visualisasi spektrum. Evaluasi dengan metode ini memberikan informasi tambahan tentang pengalaman pengguna yang lebih praktis.

Evaluasi ini memastikan bahwa metode yang digunakan memberikan hasil yang optimal dalam meningkatkan kualitas sinyal audio, sekaligus menjaga integritas komponen utama sinyal. Hasil evaluasi juga memberikan panduan untuk memperbaiki atau menyempurnakan pendekatan yang digunakan pada penelitian ini.

```
import matplotlib.pyplot as plt

time_axis = librosa.frames_to_time(np.arange(stft_matrix.shape[1]), sr=sr, hop_length=hop_length)
amplitude_db = librosa.amplitude_to_db(np.abs(stft_matrix), ref=np.max)
amplitude_db_offset = amplitude_db + 80

plt.figure(figsize=(12, 10))

plt.subplot(2, 1, 1)
librosa.display.specshow(amplitude_db_offset, y_axis='log', x_axis='time', sr=sr, hop_length=hop_length)
plt.title("Spektrum Audio Sebelum Denoising")
plt.colorbar(label="Amplitude (db, offset)")
plt.xlabel("Time (seconds)")

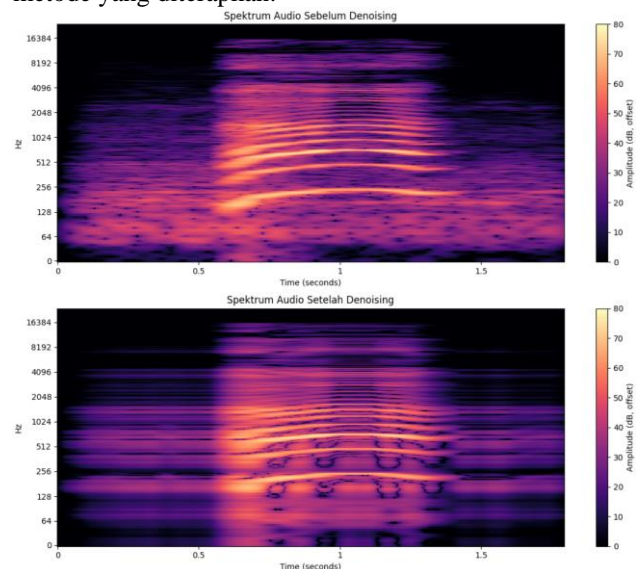
plt.subplot(2, 1, 2)
amplitude_db_denoised = librosa.amplitude_to_db(np.abs(stft_denoised), ref=np.max) + 80
librosa.display.specshow(amplitude_db_denoised, y_axis='log', x_axis='time', sr=sr, hop_length=hop_length)
plt.title("Spektrum Audio Setelah Denoising")
plt.colorbar(label="Amplitude (db, offset)")
plt.xlabel("Time (seconds)")

plt.tight_layout()
plt.show()
```

Gambar 6. Kode evaluasi hasil dalam python.

IV. HASIL DAN PEMBAHASAN

Gambar di bawah ini menunjukkan hasil visualisasi spektrum audio menggunakan metode Short-Time Fourier Transform (STFT) sebelum dan setelah proses denoising. Visualisasi ini memberikan gambaran mendetail tentang distribusi energi sinyal dalam domain waktu-frekuensi, sehingga memudahkan untuk menganalisis efek dari proses denoising yang dilakukan. Analisis dilakukan baik secara visual dari spektrum maupun melalui pengalaman mendengarkan hasil audio, untuk mengevaluasi efektivitas metode yang diterapkan.



Gambar 7. Spektrum Audio Sebelum dan Setelah Denoising Menggunakan STFT.

Pada gambar di atas, ditampilkan dua spektrum audio: spektrum sebelum denoising dan spektrum setelah denoising menggunakan metode STFT dan Singular Value Decomposition (SVD). Perbedaan antara spektrum audio sebelum dan setelah denoising terlihat, tetapi hanya sedikit signifikan. Noise di beberapa area frekuensi telah berkurang pada spektrum hasil denoising, namun masih terdapat pola ketidaksempurnaan berupa garis. Pola bergaris ini disebabkan oleh efek thresholding pada nilai singular (singular values). Ketika nilai singular yang lebih kecil dihapus atau diubah drastis melalui thresholding, hal ini menyebabkan distorsi pada representasi frekuensi rendah, sehingga menghasilkan pola "blocky".

Ambang batas (threshold) memainkan peran penting dalam menentukan nilai singular mana yang dipertahankan dan mana yang dihilangkan. Jika threshold terlalu tinggi, banyak informasi sinyal utama yang ikut terhapus, menyebabkan detail audio asli terganggu, seperti yang terlihat pada pola bergaris di spektrum setelah denoising. Sebaliknya, jika threshold terlalu rendah, noise tidak sepenuhnya tereliminasi, sehingga hasil denoising menjadi kurang efektif. Selain itu, kesalahan representasi pada STFT juga dapat terjadi. STFT bekerja dengan memotong sinyal menjadi segmen waktu yang lebih pendek, yang diasumsikan stasioner. Namun, jika resolusi waktu-frekuensi tidak diatur dengan baik melalui parameter seperti n_fft dan hop_length , hal ini juga dapat menyebabkan pola garis yang berlebihan pada hasil.

Selain analisis visual dari spektrum audio di Gambar di atas, hasil dari denoising juga dianalisis berdasarkan pengalaman mendengarkan langsung audio sebelum dan sesudah denoising. Berdasarkan hasil pendengaran, audio asli dengan noise terdengar memiliki gangguan yang signifikan, terutama pada frekuensi rendah. Noise tersebut cukup mengganggu karena menyelimuti komponen utama dari sinyal audio. Setelah dilakukan proses denoising menggunakan metode STFT, SVD, dan thresholding, noise di frekuensi rendah memang berkurang secara signifikan, namun muncul karakteristik noise baru yang terdengar seperti suara robotik. Selain itu, noise di frekuensi tinggi juga masih terdengar, meskipun intensitasnya lebih rendah dibandingkan noise awal.

Suara robotik yang muncul setelah proses denoising dapat dijelaskan oleh efek thresholding yang diterapkan pada nilai singular. Thresholding yang terlalu agresif tidak hanya menghilangkan noise, tetapi juga memodifikasi komponen utama sinyal audio. Hal ini menyebabkan distorsi pada representasi frekuensi tertentu yang akhirnya menghasilkan suara robotik. Berdasarkan analisis spektrum, frekuensi rendah di spektrum setelah denoising terlihat lebih bersih, dengan energi yang jauh berkurang di bawah ~ 128 Hz. Hal ini konsisten dengan pengurangan noise bass yang signifikan pada hasil pendengaran.

Namun, artefak di frekuensi tinggi (512 Hz ke atas) tetap terlihat sebagai pola garis-garis pada spektrum hasil denoising. Pola ini mencerminkan adanya noise di

frekuensi tinggi yang masih terdengar meskipun intensitasnya lebih rendah dibandingkan dengan noise awal. Fenomena ini menunjukkan bahwa metode yang digunakan belum sepenuhnya efektif dalam mengeliminasi noise di semua rentang frekuensi. Hal ini menunjukkan pentingnya pengaturan parameter threshold yang lebih adaptif dan resolusi waktu-frekuensi yang lebih optimal untuk meminimalkan distorsi dan menghilangkan noise secara lebih menyeluruh.

Perbandingan antara hasil pendengaran dan spektrum waktu-frekuensi menunjukkan konsistensi. Pola bergaris pada spektrum hasil denoising mencerminkan distorsi yang menyebabkan suara robotik saat didengarkan. Hal ini menunjukkan bahwa hasil visual dari spektrum dapat digunakan untuk menganalisis karakteristik hasil denoising, namun evaluasi dengan pendengaran tetap penting untuk mendapatkan gambaran yang lebih lengkap. Dengan demikian, hasil pendengaran ini memberikan wawasan tambahan bahwa metode yang digunakan masih perlu dioptimalkan, baik dalam pemilihan threshold maupun kombinasi metode denoising lainnya.

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Makalah ini membahas penerapan Singular Value Decomposition (SVD) dalam proses denoising audio dengan menggunakan representasi matriks dari Short-Time Fourier Transform (STFT). Dari hasil eksperimen, dapat disimpulkan bahwa metode ini memiliki keefektifan dalam mengurangi noise, terutama pada frekuensi rendah, seperti yang terlihat dari spektrum waktu-frekuensi dan analisis SNR. Noise bass berhasil dikurangi secara signifikan, namun metode ini menunjukkan keterbatasan dalam menghilangkan noise pada frekuensi tinggi. Selain itu, efek thresholding yang diterapkan pada nilai singular sering kali menyebabkan distorsi berupa pola bergaris di spektrum audio, yang dapat menghasilkan suara artefak seperti robotik pada hasil pendengaran.

Pemilihan parameter seperti nilai ambang batas (threshold), resolusi STFT (n_fft dan hop_length), dan strategi adaptif untuk penghapusan nilai singular menjadi faktor penting dalam keberhasilan metode ini. Walaupun teknik ini mampu mempertahankan komponen utama sinyal audio, hasil menunjukkan bahwa optimalisasi lebih lanjut diperlukan untuk mengatasi artefak dan meningkatkan kualitas sinyal hasil denoising.

B. Saran

Disarankan untuk menggunakan threshold adaptif yang didasarkan pada distribusi nilai singular. Pendekatan ini memungkinkan penghapusan noise dilakukan secara lebih selektif tanpa mengorbankan informasi penting dari sinyal utama. Dengan cara ini, distorsi pada sinyal hasil denoising dapat diminimalkan, sehingga kualitas audio yang dihasilkan lebih natural dan sesuai dengan kebutuhan aplikasi.

Selain itu, peningkatan resolusi waktu-frekuensi melalui pengaturan parameter STFT, seperti meningkatkan nilai n_fft dan mengurangi hop_length , juga direkomendasikan.

Hal ini dapat membantu menghasilkan spektrum audio yang lebih halus dan detail, memperbaiki kualitas analisis frekuensi. Namun, perlu diingat bahwa langkah ini membutuhkan sumber daya komputasi yang lebih besar, sehingga pengaturan parameter harus dilakukan secara hati-hati untuk mencapai keseimbangan antara resolusi dan efisiensi.

Untuk meningkatkan generalisasi metode, pengujian pada dataset audio yang lebih variatif sangat dianjurkan. Dataset yang mencakup berbagai jenis sinyal, seperti musik, percakapan, dan lingkungan suara lainnya, dapat memberikan gambaran yang lebih komprehensif mengenai efektivitas metode dalam mengatasi noise. Hal ini juga membantu mengevaluasi kemampuan metode dalam menangani berbagai skenario dunia nyata.

REFERENSI

- [1] Julius O. Smith III. "Short-Time Fourier Transform." Available: https://ccrma.stanford.edu/~jos/sasp/Short_Time_Fourier_Transform.html [diakses 30 Desember 2024].
- [2] Julius O. Smith III. "Mathematical Definition of STFT." Available: https://ccrma.stanford.edu/~jos/sasp/Mathematical_Definition_ST_FT.html [diakses 30 Desember 2024].
- [3] Julius O. Smith III. "Fourier Transform and Inverse Fourier Transform." Available: https://ccrma.stanford.edu/~jos/sasp/Fourier_Transform_FT_Inverse.html [diakses 30 Desember 2024].
- [4] MathWorks. "Inverse Short-Time Fourier Transform (ISTFT)." Available: <https://www.mathworks.com/help/signal/ref/istft.html> [diakses 1 January 2025].
- [5] Rinaldi Munir. (2024). Makalah Journal IEEE Access 2024. Available: https://informatika.stei.itb.ac.id/~rinaldi.munir/Penelitian/Makalah_Jurnal-IEEE-Access-2024.pdf [diakses 30 Desember 2024].
- [6] Rinaldi Munir. "Singular Value Decomposition Bagian 1." Available: <https://informatika.stei.itb.ac.id/~rinaldi.munir/AljabarGeometri/2023-2024/Algeo-21-Singular-value-decomposition-Bagian1-2023.pdf> [diakses 30 Desember 2024].

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 2 Januari 2025



Nathanael Rachmat - 13523142