

Spectrogram dan Analisis Kemiripan Sinyal Suara dengan Pendekatan *Euclidian Distance*

Sandi Fajar Rodiyansyah

Magister Ilmu Komputer – Universitas Gadjah Mada Yogyakarta

ABSTRAK

Salah satu dari kemampuan komputer adalah melakukan pengolahan sinyal suara. Pengolahan sinyal suara ini dilakukan untuk dapat mengembangkan sistem *index* dan *retrieval* sinyal suara sehingga proses pencarian sinyal suara pada komputer dapat dilakukan dengan prinsip *content based index and retrieval system*. Salah satu proses dari pengembangan *index and retrieval system* adalah perhitungan nilai *similarity*. Sehingga dapat ditentukan nilai *similarity* dari beberapa sinyal suara. Salah satu metode yang dapat digunakan adalah dengan menggunakan metode berbasis *vektor space*, dan *euclidian distance* merupakan teknik untuk menentukan nilai *similarity*-nya.

Keyword : sinyal suara, *index and retrieval*, *content based index and retrieval system*, *similarity*, *vektor space*, *euclidian distance*

I. PENDAHULUAN

Data audio merupakan bagian integral dari komputer modern yang sudah memiliki dukungan multimedia. Data audio direkam dan dibaca dengan menggunakan aplikasi multimedia. Efektifitas penyebaran data audio sangat tergantung pada kemampuan komputer untuk mengklasifikasikan dan mencari data audio yang diinginkan oleh pengguna sesuai dengan apa yang mereka inginkan. Agar dapat memudahkan *indexing* dan *retrieval* dari data audio diperlukan suatu metode komputer yang memungkinkan pengklasifikasian data audio berbasis konten yang efisien dan otomatis dalam mengelola *database* audio.

Salah satu proses pengklasifikasian data audio adalah analisis *spectrogram* sinyal audio untuk perhitungan kemiripan konten data audio. Perhitungan kemiripan konten data audio digunakan untuk menganalisis seberapa mirip antara data suara satu dengan data suara lainnya. Salah satu metode berbasis vektor-space dalam menentukan kemiripan dari sinyal suara adalah metode *euclidian distance*.

II. TINJAUAN PUSTAKA

2.1 Fitur sinyal suara berbasis waktu (*time*)

Representasi data audio dengan domain waktu atau representasi waktu amplitudo adalah teknik representasi paling dasar, di mana sinyal direpresentasikan sebagai amplitudo yang bervariasi dengan waktu. Gambar 1 representasi menunjukkan sinyal audio digital berdasarkan domain waktu.

Dalam gambar 1, *silent* digambarkan sebagai 0. Nilai sinyal dapat menjadi positif atau negatif tergantung pada apakah tekanan suara di atas atau di bawah tekanan keseimbangan atmosfer atau *silent*.

Berikut dijelaskan fitur sinyal suara berbasis waktu (*time*).

2.1.1. Average energy (AE)

Merepresentasikan energi dari sinyal audio. Semakin kencang suaranya energinya pun semakin besar. Adapun persamaan untuk menghitung *average energy* adalah sebagai berikut :

$$AE = \frac{\sum_{n=0}^{N-1} x(n)^2}{N}$$

Dimana AE adalah *average energy*, N adalah jumlah total sample dalam data audio tersebut dan x(n) adalah nilai energi dari sampel n.

2.1.2. Zero Crossing Rate (ZCR)

Merepresentasikan jumlah rata-rata amplitudo mengalami perubahan dari amplitudo positif ke amplitudo negatif dan sebaliknya. Untuk menghitung zero crossing rate digunakan persamaan sebagai berikut :

$$ZCR = \frac{\sum_{n=1}^N |\text{sgn } x(n) - \text{sgn } x(n-1)|}{2N}$$

Dimana ZCR adalah *zero crossing rate*, *sgn* x(n) adalah signal dari x(n), akan bernilai 1 apabila x(n) positif dan -1 apabila x(n) adalah negatif dan N adalah jumlah sampel dalam data audio tersebut.

2.1.3. Silent Ratio

Merupakan proporsi bagian suara yang diam/*silent*. Diam/*silent* didefinisikan sebagai waktu dimana nilai amplitudo dari sejumlah sampel berada dibawah batas ambang tertentu yang ditentukan. Sampel dianggap silent apabila amplitudonya berada dibawah batas ambang.

2.2 Fitur sinyal suara berbasis frekuensi

Representasi data audio dengan domain frekuensi direpresentasikan sebagai amplitudo yang bervariasi dengan frekuensi. Gambar 2 menunjukkan representasi sinyal audio berdasarkan domain frekuensi.

Dalam gambar 2, sumbu y diwakili oleh nilai amplitudo (dB) dan sumbu x diwakili oleh frekuensi.

2.3 Spectrogram

Representasi berbasis waktu dan berbasis spektrum adalah dua representasi sinyal sederhana, sehingga memiliki banyak keterbatasan. Dalam representasi berbasis waktu tidak menunjukkan frekuensi komponen dari sinyal dan dalam representasi berbasis spektrum tidak menunjukkan kapan komponen frekuensi yang berbeda terjadi. Untuk mengatasi masalah ini, sebuah representasi gabungan disebut spektrogram digunakan.

Spektrogram menunjukkan hubungan antara tiga variabel: konten frekuensi, waktu, dan intensitas. Dalam spektrogram, nilai frekuensi ditampilkan sepanjang sumbu vertikal, dan waktu sepanjang sumbu horisontal. Intensitas, atau kekuasaan, komponen frekuensi yang berbeda dari sinyal ditunjukkan dengan skala abu-abu, bagian paling gelap menandai amplitudo terbesar / kekuasaan. Gambar 3 merupakan contoh dari spektrogram.

2.4 Pengukuran Kemiripan Sinyal Suara dengan teknik Euclidean distance dalam Ruang Vektor

Dalam matematika, *euclidean distance* atau adalah jarak antara dua titik yang dapat diukur dan dihasilkan oleh formula pythagoras. Euclidean vector atau sering hanya disebut dengan vector adalah obyek geometri yang memiliki panjang (magnitudo) dan arah (direction). Sedangkan ruang vektor adalah sebuah struktur matematika yang dibentuk oleh sekumpulan vektor. Vektor-vektor tersebut dapat ditambahkan, dikalikan dengan bilangan real dan lain-lain.

Gambar 4 merupakan contoh dari ruang vektor, pada ruang vektor tersebut terdapat 2 vektor yaitu vektor A dan vektor B. Untuk menghitung jarak antara vektor A dan vektor B digunakan persamaan *euclidean distance*.

Berikut merupakan penyelesaian dalam menghitung jarak antara vektor A dan vektor B. Panjang vektor A dan B dapat didefinisikan sebagai berikut:

$$\|\vec{A}\| = \sqrt{X1^2 + Y1^2}$$

$$\|\vec{B}\| = \sqrt{X2^2 + Y2^2}$$

Dengan demikian, untuk menghitung jarak antara kedua vektor tersebut menggunakan persamaan sebagai berikut :

$$d(\vec{A}, \vec{B}) = \sqrt{(X1 - X2)^2 + (Y1 - Y2)^2}$$

Sedangkan untuk n dimensi ruang vektor, jarak *euclidean distance* ditentukan dengan menggunakan persamaan sebagai berikut :

$$d(\vec{u}, \vec{v}) = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \dots + (u_n - v_n)^2}$$

Akan tetapi dalam implementasi pada makalah ini hanya menggunakan perhitungan *euclidean distance* pada ruang vektor 2 dimensi.

Nilai *euclidean distance* merupakan nilai kemiripan sinyal suara. Semakin dekat (mendekati nilai 0) semakin mirip sinyal suara tersebut.

III. METODE

Untuk dapat menyelesaikan permasalahan ekstraksi spektrogram dan perhitungan similarity sinyal digital dikembangkan sebuah model perangkat lunak yang dapat meng-generate spektrogram dan perhitungan nilai similarity. Gambar 5, merupakan flow chart dari perangkat lunak yang dikembangkan.

Proses diawali dengan membacaan dua file sinyal suara *.wav, kemudian dari hasil pembacaan file tersebut data diolah untuk dapat menghasilkan spektrogram dengan menggunakan fungsi spektrogram dan fungsi similar untuk dapat menghasilkan nilai similarity.

IV. IMPLEMENTASI DAN PENGUJIAN

4.1 Implementasi

Pada tahap ini dilakukan pengembangan perangkat lunak untuk menghitung kemiripan sinyal suara. Sinyal suara yang dapat dihitung kemiripan dalam perangkat lunak ini adalah sinyal suara berformat *.wav tanpa kompresi, karena format *.wav tanpa kompresi merupakan format sinyal suara yang paling mudah untuk dianalisis fitur-fiturnya.

Dalam melakukan analisis sinyal suara, penulis menggunakan tools yang banyak dipakai untuk visualisasi data-data yaitu dengan menggunakan Matlab.

```
[filename, pathname] =
uigetfile('*.wav', 'Get Wave');
path=strcat(pathname,filename);
[sound1,fs]=wavread(path,[2000 24000]);
```

Potongan kode program diatas merupakan potongan program yang melakukan pembacaan file *.wav dari folder yang ditunjuk.

```
spectrogram(sound1,fs);
```

Pada potongan kode program ini merupakan proses untuk melakukan perhitungan dan pembangkitan spektrogram dari sinyal suara yang dipakai. Spectrogram merupakan fungsi yang dikembangkan penulis untuk mengekstrak fitur dari sinyal suara yang dipakai.

```
R=similar(sound1,sound2,fs)/10000;
```

Pada potongan kode program ini, merupakan proses perhitungan kemiripan sinyal suara yang kemudian disimpan dalam variabel R. Similar merupakan fungsi yang dikembangkan berdasarkan metode *euclidean distance* yang digunakan untuk menghitung nilai *similarity*.

4.2 Pengujian

Untuk dapat membuktikan keakuratan metode ini, penulis menguji dengan membandingkan dua file

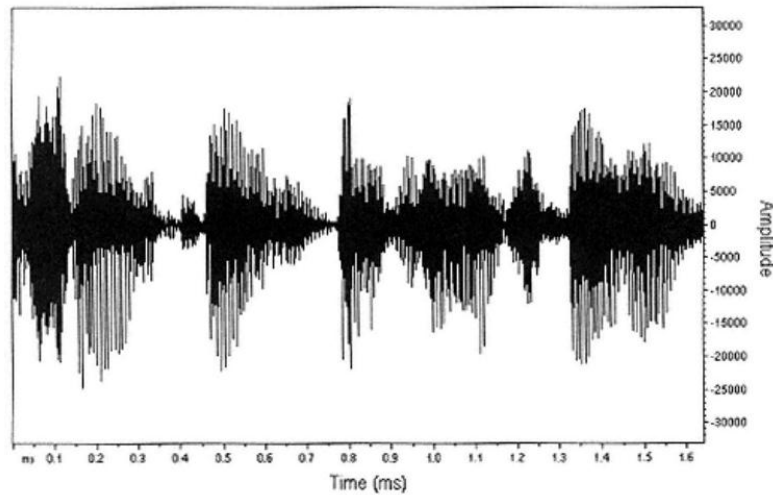
*.wav dan membandingkan satu file *.wav dengan dirinya sendiri.

Setelah dilakukan pengujian hasil dari perbandingan antara dua file *.wav menghasilkan nilai 3,78521 dan hasil perbandingan antara satu file *.wav dengan dirinya sendiri menghasilkan nilai 0. Nilai nol merupakan nilai *similarity* yang dihasilkan apabila satu file *.wav dibandingkan dengan dirinya sendiri (identik). Gambar 6 dan 7 menunjukkan hasil perhitungan nilai similarity.

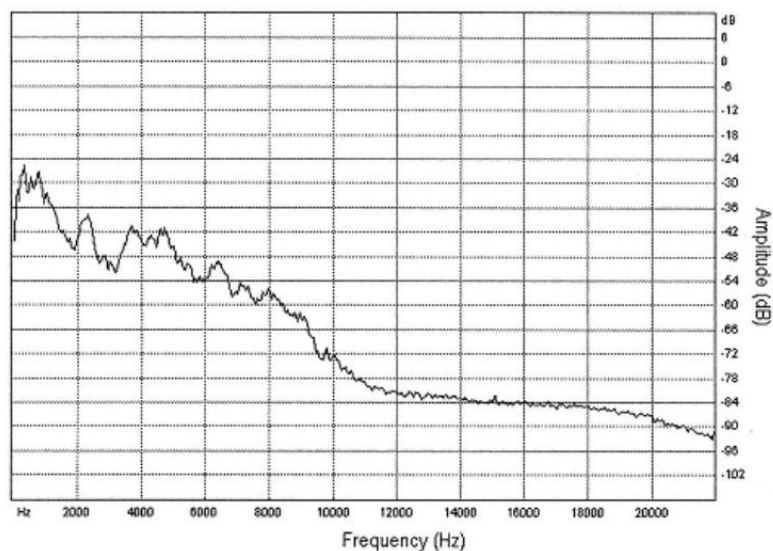
V. KESIMPULAN

Spectrogram merupakan representasi sinyal suara yang merupakan gabungan antara representasi sinyal suara berbasis waktu (*time*) dan representasi sinyal suara berbasis frekuensi.

Euclidean distance merupakan teknik berbasis vektor space yang dapat digunakan untuk menghitung nilai similarity dari dua sinyal digital yang dibandingkan.



Gambar 1 : Representasi sinyal suara berbasis waktu (time)



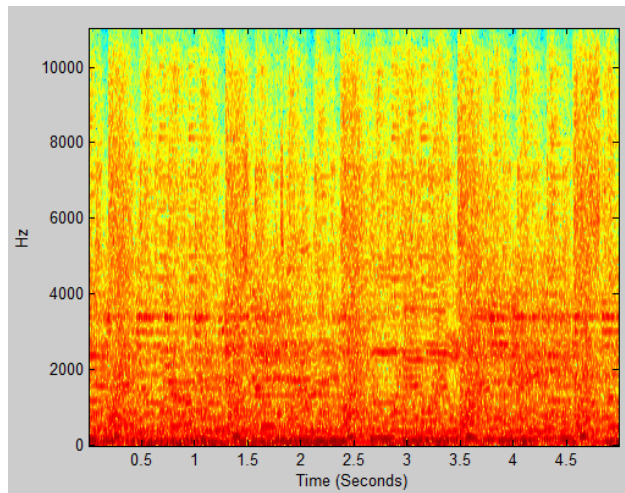
Gambar 2 : Representasi sinyal suara berbasis frekuensi (spectrum)

DAFTAR PUSTAKA

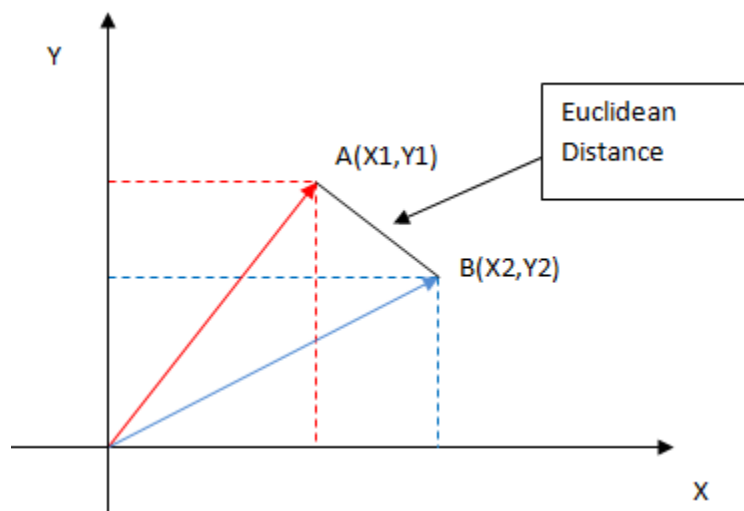
Lu, Guojun. 1963. "Multimedia Database Manajemen Systems. Artech House: Norwood

Julien A, Jean and Pachtet, Francois. 2002. "Music Similarity Measures : What's the Use?"

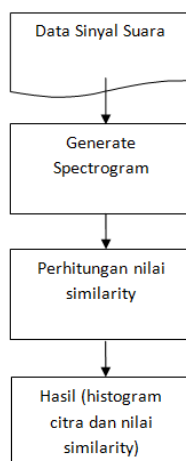
P.W. Ellins, Daniel dan E. Poliner, Graham. 2005. "Identifying 'Cover Songs' With Chroma Features and Dynamic Programming Beat Tracking". Columbia University.



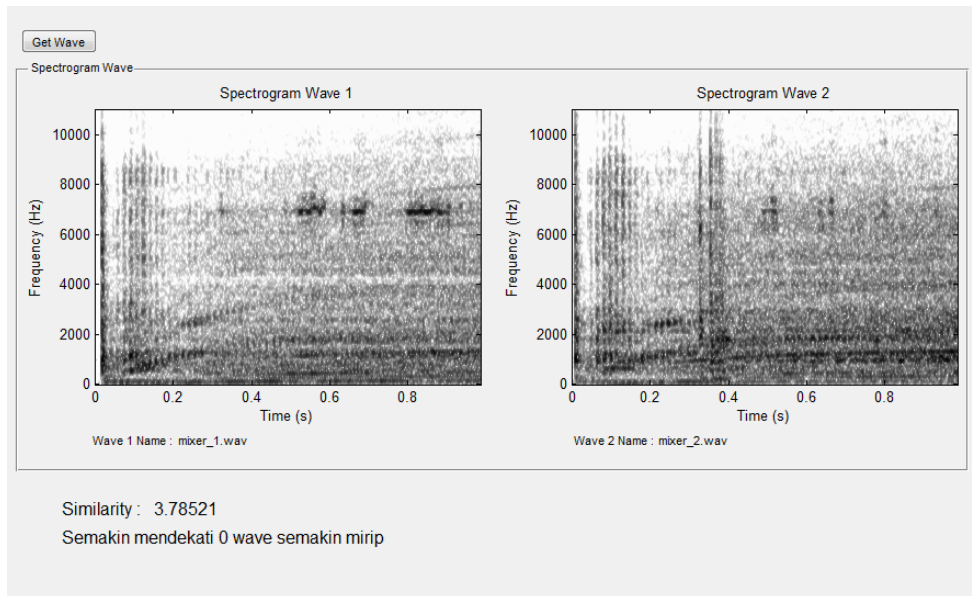
Gambar 3 : Spectrogram dari sinyal suara



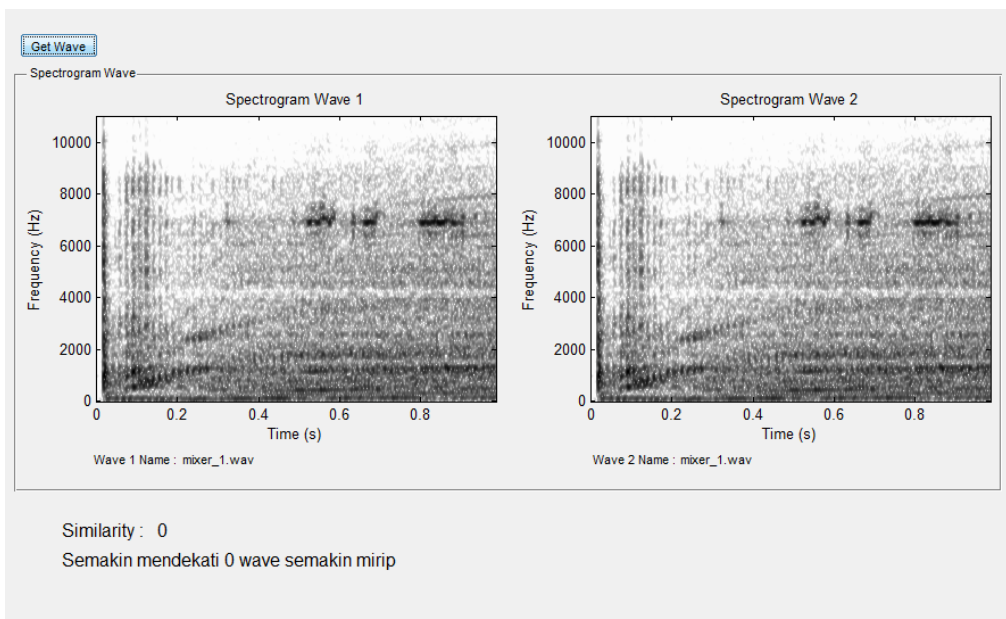
Gambar 4 : Vektor space 2 dimensi



Gambar 5 : Flow chart perangkat lunak



Gambar 6 : Hasil perhitungan similarity pada dua sinyal yang berbeda menghasilkan nilai 3,78521



Gambar 7 : Hasil perhitungan similarity pada dua sinyal yang sama menghasilkan nilai 0