

penjendelaan_Penjendelaan Gaussian pada Ekstraksi Ciri Perataan Segmen Berbasis Transformasi Sinus Diskrit Pada Pengenalan Nada Alat Musik

by Sumarno Linggo

Submission date: 04-Jul-2022 02:31PM (UTC+0700)

Submission ID: 1866473540

File name: s_Transformasi_Sinus_Diskrit_Pada_Pengenalan_Nada_Alatt_Musik.pdf (2.71M)

Word count: 4262

Character count: 23847

PENJENDELAAN GAUSSIAN PADA EKSTRAKSI CIRI PERATAAN SEGMENT BERBASIS TRANSFORMASI SINUS DISKRIT PADA PENGENALAN NADA ALAT MUSIK

Linggo Sumarno

Program Studi Teknik Elektro; Fakultas Teknik dan Sains
Universitas Sanata Dharma Yogyakarta; Kampus III Paingan Maguwoharjo

e-mail: lingsum@usd.ac.id

Abstrak— Tulisan ini mengusulkan suatu ekstraksi ciri yang dapat digunakan pada pengenalan nada alat musik. Ekstraksi ciri yang diusulkan adalah ekstraksi ciri perataan segmen berbasis Transformasi Sinus Diskrit (*Discrete Sine Transform / DST*), yang menggunakan jendela Gaussian. Kemudian, alat musik yang digunakan adalah yang mempunyai lebih dari satu puncak lokal signifikan pada ranah DST. Berdasarkan eksperimen, jika ekstraksi ciri yang diusulkan digunakan secara optimal untuk nada-nada yang mempunyai banyak atau sedikit puncak lokal signifikan pada ranah DST, akan memerlukan masing-masing 8 atau 4 koefisien ekstraksi ciri. Pada kondisi optimal ini, tingkat pengenalan mencapai 100%. Namun, jika ekstraksi ciri yang diusulkan digunakan secara optimal untuk nada-nada yang mempunyai banyak dan sedikit puncak lokal signifikan pada ranah DST, akan memerlukan 8 koefisien ekstraksi ciri. Pada kondisi optimal ini, tingkat pengenalannya turun menjadi 97,1%.

Kata Kunci— Penjendelaan Gaussian, ekstraksi ciri, perataan segmen, pengenalan nada.

I. PENDAHULUAN

Seiring dengan semakin berkembangnya kemampuan komputer, berkembang pula keinginan manusia untuk membuat komputer yang dapat meniru kemampuan manusia. Salah satu kemampuan manusia yang dicoba untuk ditiru adalah kemampuan manusia untuk mengenali sesuatu. Dalam hal mendengar dan mengenali suatu nada, dewasa ini terus dikembangkan kemampuan komputer yang dapat mampu mendengar dan mengenali suatu nada, seperti halnya yang mampu dilakukan oleh para pemusik profesional [1].

Penelitian-penelitian sebelumnya yang terkait dengan pengenalan nada, pada dasarnya menggunakan pendekatan ranah waktu atau ranah transformasi. Penelitian-penelitian yang menggunakan pendekatan ranah waktu [2] [3], pada dasarnya menggunakan autokorelasi. Pendekatan ranah waktu ini pada dasarnya kurang cocok untuk nada-nada polifonik. Untuk penelitian-penelitian yang menggunakan ranah transformasi, pada dasarnya menggunakan FFT (*Fast Fourier Transform*) [1] [4] [5] [6] [7] [8] [9] [10], atau DCT (*Discrete Cosine Transform*) [1] [12]. Penelitian-penelitian dalam ranah transformasi tersebut dapat dibagi dalam dua kelompok. Kelompok yang pertama berdasarkan frekuensi fundamental [4] [5] [6] [7], dan kelompok yang kedua tidak berdasarkan frekuensi fundamental [1] [8] [9] [10] [11] [12]. Pendekatan ranah transformasi ini pada dasarnya cocok untuk nada-nada polifonik. Selanjutnya, tulisan ini akan membahas pengenalan nada polifonik, yang tidak berdasarkan frekuensi fundamental.

Dalam hal pengenalan nada polifonik, yang tidak berdasarkan frekuensi fundamental, pada dasarnya dapat dilihat adanya masalah pada jumlah koefisien ekstraksi cirinya yang

relatif masih banyak. Sebagai contoh, berdasarkan yang ada pada referensi, jumlah koefisien ekstraksi ciri terkecil yang menghasilkan tingkat pengenalan tertinggi adalah 128 koefisien [1], 64 koefisien [8], 32 koefisien [11], 13 koefisien [9], 12 koefisien [10], dan 8 koefisien [12]. Oleh karena itu, terlihat masih ada peluang, untuk mengurangi lebih lanjut, jumlah koefisien ekstraksi ciri tersebut.

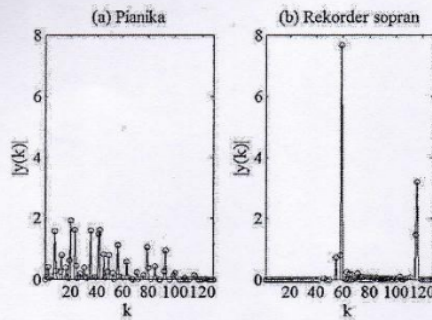
Tulisan ini mengusulkan suatu metode ekstraksi ciri untuk nada-nada yang mempunyai sejumlah puncak lokal signifikan pada ranah transformasi (polifonik), yang tidak berdasarkan pada suatu sinyal fundamental. Metode ekstraksi ciri yang diusulkan tersebut mampu mengurangi lebih lanjut jumlah koefisien ekstraksi ciri yang dideskripsikan di atas. Dalam tulisan ini juga akan dibahas kinerja metode ekstraksi ciri yang diusulkan, jika menghadapi sinyal-sinyal yang mempunyai banyak dan sedikit puncak lokal signifikan pada ranah transformasi. Dalam hal ini, transformasi yang dimaksud adalah yang menggunakan DST.

II. TINJAUAN PUSTAKA

A. Nada alat musik

Berdasarkan persepsi manusia, nada-nada yang berasal dari alat musik mempunyai dua karakteristik aural yaitu skala (*scale*) dan warna suara (*timbre*) [13]. Skala adalah karakteristik yang terkait dengan tinggi rendahnya nada. Dengan demikian, jika skalanya makin tinggi, nadanya juga makin tinggi. Sebaliknya jika skalanya makin rendah, nadanya juga makin rendah.

Warna suara adalah karakteristik yang terkait dengan jenis alat musik, seperti misalnya pianika dan rekorder sopran. Berdasarkan warna suara, dalam ranah DST, nada-nada alat musik dapat dibagi dalam dua kelompok. Kelompok yang pertama adalah yang mempunyai satu puncak lokal signifikan (yang disebut juga monofonik), dan kelompok kedua adalah yang mempunyai lebih dari satu (*multiple*) puncak lokal signifikan (yang disebut juga polifonik). Kelompok kedua ini selanjutnya dapat dibagi lagi dalam dua subkelompok. Subkelompok yang pertama adalah yang mempunyai banyak puncak lokal signifikan, dan subkelompok yang kedua adalah yang mempunyai sedikit puncak lokal signifikan [12]. Gambar 1 memperlihatkan contoh representasi warna suara dalam ranah DST, untuk alat musik pianika dan rekorder sopran. Dalam Gambar 1 tersebut, pianika adalah contoh alat musik dengan banyak puncak lokal signifikan, sedangkan rekorder sopran adalah contoh alat musik dengan sedikit puncak lokal signifikan.



Gambar 1. Representasi warna nada untuk nada 'C' dalam ranah DST $y(k)$ (lihat persamaan 2), untuk pianika dan rekorder sopran, dengan menggunakan DST 128 titik.

B. Jendela Gaussian

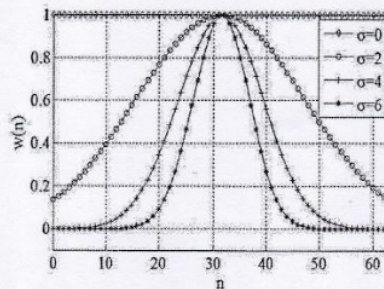
Jendela Gaussian adalah suatu jendela yang dapat digunakan untuk keperluan penjendelaan. Jendela Gaussian $w(n)$ dengan panjang N titik, dirumuskan secara matematis sebagai berikut [15]

$$w(n) = e^{-\frac{1}{2} \left(\frac{n}{N/2} \right)^2} \quad (1)$$

dengan σ adalah deviasi standar, dan $-(N-1)/2 \leq n \leq (N-1)/2$. Jendela Gaussian adalah jendela yang dapat diatur keruncingannya dengan mengatur nilai σ , sebagaimana diperlihatkan pada Gambar 2.

C. DST (Discrete Sine Transform)

DST merupakan suatu jenis transformasi yang diturunkan dari DFT (Discrete Fourier Transform). DST digunakan untuk mentransformasi sinyal dari ranah waktu ke ranah DST. DST dari runtun $u(n)$ dengan panjang N titik, dirumuskan sebagai berikut [16].



Gambar 2. Contoh jendela Gaussian 64 titik, dengan nilai σ yang beragam.

$$y(k) = \sqrt{\frac{2}{n+1}} \sum_{n=0}^{N-1} u(n) \sin \left[\frac{\pi(k+1)(n+1)}{N+1} \right] \quad (2)$$

dengan $0 \leq n \leq N-1$.

D. Perataan segmen

Perataan segmen, yang diinspirasi oleh Setiawan [17], adalah suatu proses untuk mengurangi ukuran sinyal. Pada dasarnya, sinyal yang telah berkurang ukurannya tersebut, masih memperlihatkan bentuk dasar dari pola sinyalnya. Algoritma perataan segmen diperlihatkan di bawah ini.

Algoritma perataan segmen

1. Untuk runtun $y(k) = \{y(0), y(1), \dots, y(N-1)\}$ dengan $N = 2^p$ dan $p \geq 0$.
2. Tentukan panjang segmen L dengan $L = 2^q$ untuk $0 \leq q \leq p$.
3. Bagilah runtun $y(k)$ dengan menjadi segmen-segmen sepanjang L . Dengan demikian akan dihasilkan sejumlah M segment sebagai berikut

$$M = \frac{N}{L}, \quad (3)$$

dan juga runtun $f(r) = \{f(1), f(2), \dots, f(L)\}$ di setiap segmen.

4. Hitung nilai rerata disetiap segmen $z(v)$ sebagai berikut

$$z(v) = \frac{1}{L} \sum_{r=1}^L f_v(r), \quad 1 \leq v \leq M. \quad (4)$$

E. Pencocokan template dan korelasi

Pencocokan template adalah salah satu pendekatan yang paling awal dan juga paling sederhana dalam pengenalan pola [18]. Pada pencocokan template, pola masukan yang akan dikenali dicocokkan dengan sejumlah pola yang disimpan dalam template-template. Suatu pola dalam template yang paling cocok dengan pola masukan, dinyatakan sebagai pola keluarannya. Untuk mengukur tingkat kecocokan antara pola masukan dengan sejumlah pola yang ada dalam template-template, dapat digunakan fungsi kemiripan (similaritas). Korelasi merupakan salah satu pilihan fungsi kemiripan dari sejumlah fungsi kemiripan yang biasanya digunakan dalam suatu sistem belajar (learning system) [19]. Secara matematis, korelasi ini diperlihatkan dibawah ini

$$S(x, y) = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2 \sum_{i=1}^m (y_i - \bar{y})^2}} \quad (5)$$

dengan x dan y adalah dua vektor yang panjangnya sama (dalam hal ini salah satu vektor berasal dari pola masukan dan

vektor yang lain berasal suatu pola dalam *template*), m adalah panjang vektor x atau y , serta $\bar{x}_i$ dan $\bar{y}_i$ masing masing adalah rerata dari elemen-elemen pada vektor x dan y .

III. METODE PENELITIAN

A. Alat dan bahan

Bahan penelitian berupa rekaman nada pianika dan rekorder sopran yang terisolasi, dalam format wav. Bahan penelitian ini diperoleh dengan merekam nada pianika dan rekorder sopran, masing-masing dengan frekuensi pencuplikan 4800 Hz dan 2400 Hz. Besaran frekuensi pencuplikan ini dipilih dengan berdasarkan pada kriteria Nyquist, yaitu frekuensi pencuplikan adalah dua kali atau lebih besar dari frekuensi analog tertingginya. Dalam hal ini, frekuensi analog tertinggi untuk nada 'B' pada pianika dan rekorder sopran adalah masing-masing 2400 Hz dan 995 Hz. Perekaman nada dilakukan selama 2 detik, karena berdasarkan evaluasi secara visual, dengan waktu selama 2 detik tersebut, dapat didapatkan separuh lebih bagian sinyal yang dalam kondisi *steady state*. Setelah perekaman tersebut, bagian *silence* dari sinyal dihilangkan.

Alat penelitian adalah pianika (merek Brother) dan rekorder sopran (merek Yamaha), seperti yang diperlihatkan pada Gambar 3. Untuk perekaman, digunakan mikrofon Genius MIC-01A. Untuk pemrosesan sinyal dan Komputer yang digunakan adalah yang menggunakan Intel Core i3 3220 yang dilengkapi dengan RAM 4 GB.

B. Perancangan sistem pengenalan nada

Sistem pengenalan nada yang dirancang diperlihatkan pada Gambar 4. Masukan sistem adalah *wav file*, dan keluarannya adalah teks yang mengindikasikan nada yang dikenali.

Frame blocking adalah proses pengambilan suatu *frame* sinyal, dari suatu sinyal yang panjang [19]. Tujuan dari *frame blocking* adalah mengurangi jumlah data sinyal yang akan diproses. Adanya pengurangan data ini, akan menyebabkan berkurangnya waktu komputasi. Dalam penelitian ini, suatu *frame* sinyal diambil setelah 100 mili detik dari tepi sinyal.

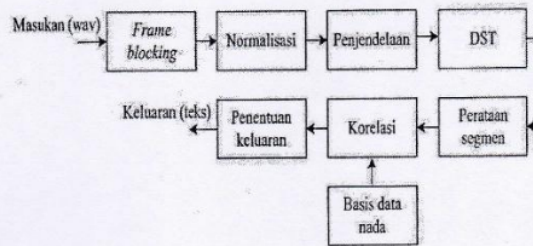


(a) Pianika



(b) Rekorder sopran

Gambar 3. Pianika dan rekorder sopran yang digunakan dalam penelitian [12].



Gambar 4. Diagram blok sistem pengenalan nada.

yang telah dihilangkan bagian *silence*-nya. Berdasar hasil pengamatan secara visual, setelah 100 mili detik tersebut, sinyalnya telah mencapai saat *steady state*. Panjang *frame* sinyal yang digunakan dievaluasi dengan panjang 32, 64, 128, dan 256 titik.

Normalisasi adalah proses pengesetan nilai maksimum menjadi satu. Tujuan dari normalisasi adalah menghilangkan perbedaan nilai maksimum dari sinyal-sinyal hasil perekaman.

Penjendelaan (windowing) adalah proses pengurangan diskontinuitas yang ada pada daerah-daerah tepi sinyal (*signal edges*). Adanya diskontinuitas ini disebabkan oleh proses *frame blocking*, yang merupakan suatu proses sebelum penjendelaan (lihat Gambar 4). Adanya diskontinuitas ini, akan memunculkan adanya sinyal-sinyal harmonik pada sinyal hasil transformasi. Hal ini akan menyebabkan turunnya keakuratan ekstraksi ciri [9]. Untuk mengurangi sinyal-sinyal harmonik, maka diskontinuitas pada daerah-daerah tepi sinyal dapat dikurangi menggunakan penjendelaan [19]. Panjang jendela yang digunakan sama dengan panjang *frame*, yang berasal dari proses *frame blocking*.

DST adalah suatu proses untuk mentransformasi sinyal dari ranah waktu menjadi ranah DST. Panjang DST yang digunakan sama dengan panjang *frame*, yang berasal dari proses *frame blocking*.

Perataan segmen adalah suatu proses untuk mengurangi ukuran sinyal. Pada penelitian ini, perataan segmen digunakan untuk mengurangi ukuran sinyal setelah proses DST. Hasil dari proses perataan sinyal dinamakan ekstraksi ciri dari sinyal masukan.

Korelasi adalah suatu proses untuk menghitung kemiripan, antara ekstraksi ciri sinyal masukan dengan ekstraksi ciri sejumlah sinyal dalam basis data sinyal. Adanya perhitungan kemiripan ini merupakan suatu indikasi bahwa model pengenalan pola pada penelitian ini menggunakan pendekatan pencocokan *template* [17].

Penentuan keluaran adalah suatu proses untuk menentukan nada keluaran yang bersesuaian dengan nada dari sinyal masukan. Penentuan nada ini dilakukan dengan mencari nilai tertinggi dari sejumlah nilai hasil perhitungan korelasi, antara ekstraksi ciri sinyal masukan dengan ekstraksi ciri dari sejumlah sinyal referensi yang ada dalam basis data nada.

Secara teori, suatu nada dalam basis data nada yang mempunyai korelasi tertinggi, maka nada tersebut adalah yang paling mirip dengan sinyal masukan. Oleh karena itu, suatu nada dalam basis data nada yang mempunyai nilai korelasi tertinggi, ditentukan sebagai nada keluaran.

C. Basis data nada

Basis data nada merupakan nada yang diingat dalam sistem pengenalan nada yang dibangun. Untuk membangun suatu basis data nada, untuk setiap alat musik (pianika dan rekorder sopran) diambil 10 sampel dari setiap nada ('C', 'D', 'E', 'F', 'G', 'A', dan 'B'). Dalam hal ini pengambilan 10 sampel untuk setiap nada adalah dengan asumsi, pada peniupan pianika dan rekorder sopran secara normal, semua variasi pola untuk setiap nada telah didapatkan. Karena dari 7 nada, masing-masing nada memiliki 10 sampel, maka terdapat 70 nada untuk membangun basis data nada, untuk setiap alat musik. Berikutnya, untuk membangun basis data nada, untuk setiap alat musik, dilakukan langkah-langkah berikut.

1. Untuk setiap nada ('C', 'D', 'E', 'F', 'G', 'A', dan 'B') yang masing-masing terdiri atas 10 sampel (wav), lakukan proses ekstraksi ciri, seperti yang diperlihatkan pada Gambar 5. Hasil dari proses pada Gambar 5 tersebut dapat dinyatakan secara matematis menggunakan vektor nada $r_{n,i}$, dengan n adalah nada ('C', 'D', 'E', 'F', 'G', 'A', dan 'B') serta nilai i sampel ke 1...10.
2. Lakukan proses rerata untuk setiap vektor nada $r_{n,i}$ sebagai berikut

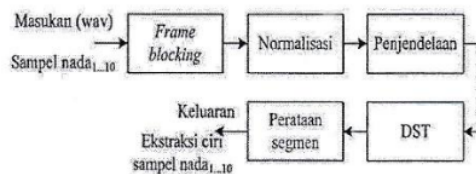
$$x_n = \frac{1}{10} \sum_{i=1}^{10} r_{n,i} \quad (6)$$

dengan x_n adalah vektor nada referensi, dan n adalah nada ('C', 'D', 'E', 'F', 'G', 'A', dan 'B').

3. Selanjutnya, basis data nada akan terdiri atas sejumlah 7 vektor nada referensi x_n .

D. Nada uji

Nada uji adalah nada yang digunakan untuk menguji kinerja sistem pengenalan yang telah dibangun. Dalam penelitian ini, untuk setiap alat musik (pianika dan rekorder sopran), diambil 10 sampel untuk setiap nada ('C', 'D', 'E', 'F', 'G', 'A' and 'B'). Dengan demikian, untuk setiap alat musik terdapat sejumlah 70 nada uji.



Gambar 5. Diagram blok proses ekstraksi ciri untuk basis data nada.

E. Tingkat pengenalan

Tingkat pengenalan adalah salah satu ukuran kinerja dari suatu sistem pengenalan. Tingkat pengenalan ini dihitung menggunakan persamaan berikut

$$\text{Tingkat pengenalan} = \frac{\text{Jumlah nada yang dikenali}}{\text{Jumlah nada uji}} \times 100\% \quad (7)$$

dengan jumlah nada uji adalah 70 nada untuk setiap alat musik (pianika dan rekorder sopran).

IV. HASIL DAN PEMBAHASAN

Hasil pengujian sistem pengenalan nada yang telah dideskripsikan di atas, diperlihatkan pada Tabel 1 (untuk pengujian pianika), dan 2 (untuk pengujian rekorder sopran). Hasil pengujian tersebut memperlihatkan kinerja sistem pengenalan untuk variasi nilai σ jendela Gaussian, panjang DST, dan panjang segmen.

Untuk memahami lebih lanjut kinerja sistem pengenalan pada Tabel 1 dan 2, yang terkait dengan jumlah koefisien ekstraksi ciri, telah dibuat Tabel 3. Tabel tersebut memperlihatkan jumlah koefisien ekstraksi ciri yang dipengaruhi oleh panjang DST dengan panjang segmen. Selanjutnya, dengan berdasarkan Tabel 1, 2, dan 5, telah dibuat Tabel 4 yang memperlihatkan sejumlah kombinasi nilai σ jendela Gauss, panjang DST, dan panjang segmen, yang menghasilkan jumlah koefisien ekstraksi ciri terkecil, pada tingkat pengenalan 100%.

Sebagaimana terlihat pada Tabel 4, jumlah koefisien ekstraksi ciri terkecil, yang menghasilkan tingkat pengenalan 100% pada pianika adalah 8 koefisien, pada penggunaan DST 128 atau 256 titik, panjang segmen 16 atau 32 titik, serta nilai σ jendela Gaussian dari 2 hingga 6. Pada rekorder sopran, jumlah koefisien ekstraksi ciri terkecil, yang menghasilkan tingkat pengenalan 100% adalah 4 koefisien, pada penggunaan DST 64 atau 128 titik, panjang segmen 16 atau 32 titik, serta nilai σ jendela Gaussian sekurang-kurangnya 4.

Berdasarkan Tabel 4, secara umum dapat dikatakan bahwa, metode ekstraksi ciri yang diusulkan dalam penelitian ini lebih efisien untuk digunakan pada nada-nada yang mempunyai jumlah puncak lokal sedikit, seperti misalnya pada rekorder sopran. Hal ini, dapat dijelaskan dengan berdasarkan contoh pada Tabel 5.

Tabel 5 memperlihatkan bahwa pada pianika, nilai rerata korelasi yang relatif tinggi mengindikasikan bahwa pola-pola pada nada referensi 'C' - 'B', pada dasarnya saling mirip. Adanya pola-pola yang pada dasarnya saling mirip ini, akan menyebabkan lebih sulit dibedakan antara pola nada yang satu dengan yang lainnya. Akibatnya adalah tingkat pengenalan yang lebih rendah. Sebaliknya, pada rekorder sopran, nilai rerata korelasinya relatif rendah mengindikasikan bahwa pola-pola pada nada referensi 'C' - 'B', pada dasarnya kurang miripnya. Adanya pola-pola yang pada dasarnya kurang mirip ini, akan menyebabkan lebih mudah dibedakan antara pola nada yang satu dengan yang lainnya. Akibatnya adalah tingkat pengenalan yang lebih tinggi.

Tabel 1.

Hasil pengujian pengenalan nada pianika, pada berbagai nilai σ jendela Gaussian, panjang DST, dan panjang segmen. Hasil yang ditampilkan: Tingkat pengenalan (%).

Nilai σ jendela Gaussian	Panjang g DST (titik)	Panjang segmen (titik)							
		1	2	4	8	16	32	64	128
0	256	100	100	100	100	100	97,1	70,0	1,4
	128	100	100	100	98,6	94,3	74,3	2,9	-
	64	100	100	94,3	91,4	68,6	5,7	-	-
	32	60,0	74,3	68,6	45,7	5,7	-	-	-
2	256	100	100	100	100	100	100	84,3	8,6
	128	100	100	100	100	100	92,9	5,8	-
	64	100	100	100	98,6	72,9	10,0	-	-
	32	75,7	81,4	70,0	47,1	4,3	-	-	-
4	256	100	100	100	100	100	100	87,1	5,7
	128	91,4	100	100	100	100	85,7	4,2	-
	64	91,4	97,1	97,1	77,1	45,7	4,3	-	-
	32	45,7	48,6	47,1	44,3	2,9	-	-	-
6	256	100	100	100	100	100	100	85,7	7,1
	128	78,6	100	100	100	94,3	70,0	5,7	-
	64	75,7	82,8	81,4	62,9	35,7	1,4	-	-
	32	44,3	40,0	38,6	32,9	7,1	-	-	-
8	256	100	100	100	100	100	98,6	80,0	10,0
	128	70,0	94,3	92,9	92,9	75,7	57,1	5,7	-
	64	55,7	54,3	51,4	51,4	34,3	7,1	-	-
	32	25,7	31,4	32,9	25,7	7,1	-	-	-
10	256	97,1	100	100	100	100	98,6	68,6	1,4
	128	60,0	80,0	80,0	80,0	58,6	58,6	1,4	-
	64	40,0	40,0	40,0	38,6	30,0	7,1	-	-
	32	27,1	34,3	34,3	30,0	2,9	-	-	-

Tabel 2.

Hasil pengujian pengenalan nada rekorder sopran, pada berbagai nilai σ jendela Gaussian, panjang DST, dan panjang segmen. Hasil yang ditampilkan: Tingkat pengenalan (%).

Nilai σ jendela Gaussian	Panjang DST (titik)	Panjang segmen (titik)							
		1	2	4	8	16	32	64	128
0	256	100	100	100	100	98,6	98,6	94,3	7,1
	128	100	100	100	98,6	98,6	91,4	11,4	-
	64	100	100	98,6	97,1	92,9	10,0	-	-
	32	95,7	94,3	88,6	91,4	15,7	-	-	-
2	256	100	100	100	100	98,6	97,1	88,6	12,9
	128	100	100	100	98,6	92,9	88,6	14,3	-
	64	100	100	100	95,7	82,8	2,9	-	-
	32	94,3	98,6	97,1	91,4	15,7	-	-	-
4	256	100	100	100	100	100	84,3	81,4	8,5
	128	100	100	100	100	97,1	97,1	18,6	-
	64	100	100	100	100	100	8,6	-	-
	32	77,1	98,6	98,6	98,6	17,1	-	-	-
6	256	100	100	100	100	100	94,3	85,7	12,9
	128	100	100	100	100	100	100	12,9	-
	64	98,6	100	100	100	100	10,0	-	-
	32	68,6	98,6	98,6	97,1	18,6	-	-	-
8	256	100	100	100	100	100	97,1	87,3	15,7
	128	98,6	100	100	100	100	100	7,1	-
	64	90,0	100	100	100	100	12,9	-	-
	32	60,0	98,6	98,6	98,6	17,1	-	-	-
10	256	100	100	100	100	100	98,6	94,2	15,7
	128	97,1	100	100	100	100	100	11,4	-
	64	81,4	100	100	100	100	15,7	-	-
	32	47,1	94,3	95,7	97,1	20,0	-	-	-

Tabel 3.

Jumlah koefisien ekstraksi ciri pada berbagai kombinasi panjang DST dan panjang segmen. Hasil yang diperlihatkan: Jumlah koefisien ekstraksi ciri.

Panjang DST (titik)	Panjang segmen (titik)							
	1	2	4	8	16	32	64	128
256	256	128	64	32	16	8	4	2
128	128	64	32	16	8	4	2	-
64	64	32	16	8	4	2	-	-
32	32	16	8	4	2	-	-	-

Tabel 4.

Berbagai kombinasi nilai σ jendela Gaussian, panjang DST, dan panjang segmen, yang menghasilkan jumlah koefisien ekstraksi ciri terkecil pada pianika dan rekorder sopran, pada tingkat pengenalan 100%.

Hasil yang diperlihatkan: Jumlah koefisien ekstraksi ciri.

Nilai σ jendela Gaussian	Panjang DST (titik)	Panjang segmen (titik)			
		Pianika		Rekorder sopran	
		16	32	16	32
0	256	-	-	-	-
	128	-	-	-	-
	64	-	-	-	-
2	256	-	8	-	-
	128	8	-	-	-
	64	-	-	-	-
4	256	-	8	-	-
	128	8	-	-	4
	64	-	-	4	-
6	256	-	8	-	-
	128	-	-	-	4
	64	-	-	4	-
8	256	-	-	-	-
	128	-	-	-	4
	64	-	-	4	-
10	256	-	-	-	-
	128	-	-	-	4
	64	-	-	4	-

Berdasarkan Tabel 1, 2, dan 3, dapat dilihat bahwa agar supaya metode ekstraksi ciri yang diusulkan dapat digunakan secara optimal, baik untuk alat musik dengan banyak ataupun sedikit puncak lokal pada ranah DST, dapat menggunakan DST 128 titik, panjang segmen 16 titik, serta nilai $\sigma=4$ pada jendela Gaussian. Pada kondisi optimal tersebut, jumlah koefisien ekstraksi cirinya adalah 8 koefisien, dengan tingkat pengenalan 97,1 %.

Tabel 5.

Contoh hasil perhitungan korelasi untuk masukan nada 'C' dengan nada-nada referensi yang ada pada basis data nada. Sebagai catatan, digunakan DST 64 titik, panjang segmen 8, dan nilai $\sigma=6$ pada jendela Gaussian. Hasil yang ditampilkan: Nilai korelasi dan rerata.

Alat musik	Nada referensi						Rerata
	'C'	'D'	'E'	'F'	'G'	'A'	
Pianika	0,945	0,929	0,747	0,767	0,884	0,939	0,738
Rekorder sopran	0,997	0,813	0,274	0,054	0,246	0,326	0,185

Perbandingan dengan metode ekstraksi ciri yang lain

Sebagaimana telah dideskripsikan di atas, metode ekstraksi ciri yang diusulkan sangat efisien untuk nada-nada alat musik yang mempunyai sedikit puncak lokal signifikan pada domain DST. Tabel 6 membandingkan kinerja beberapa metode ekstraksi ciri untuk pengenalan nada yang mempunyai banyak dan sedikit puncak lokal pada ranah transformasi.

Tabel 6.

Perbandingan jumlah koefisien ekstraksi ciri terkecil, yang menghasilkan tingkat pengenalan tertinggi, untuk nada-nada yang mempunyai jumlah puncak lokal banyak dan sedikit pada ranah transformasi. Hasil yang ditampilkan: Jumlah koefisien ekstraksi ciri (alat musik).

Metode ekstraksi ciri	Jumlah puncak lokal pada ranah transformasi	
	Banyak	Sedikit
FFT [1]	128 (pianika)	-
FFT [8]	-	64 (rekorder sopran)
DCT dan penjumlahan koefisien [11]	32 (pianika)	-
MFCC and Codebook [9]	13 (piano pada keyboard)	-
FFT dan penjumlahan koefisien [10]	12 (pianika)	-
DCT and perataan segmen [12]	8 (pianika)	16 (rekorder sopran)
DST, perataan segmen, dan penjumlahan Gaussian (penelitian ini)	8 (pianika)	4 (rekorder sopran)

Sebagaimana terlihat pada Tabel 6, metode ekstraksi ciri yang diusulkan memberikan jumlah koefisien ekstraksi ciri yang secara signifikan lebih sedikit, terutama untuk nada-nada yang mempunyai sedikit puncak lokal pada ranah transformasi. Dengan kata lain, metode ekstraksi ciri dengan penjumlahan yang diusulkan, sangat efisien untuk nada-nada yang mempunyai sedikit puncak lokal pada ranah transformasi.

V. KESIMPULAN

Berdasarkan hal-hal yang telah dideskripsikan di atas, dapat disimpulkan hal-hal berikut.

1. Metode ekstraksi ciri dengan penjumlahan Gaussian yang diusulkan dalam penelitian ini lebih efisien untuk digunakan pada nada-nada yang mempunyai sedikit puncak lokal signifikan, daripada untuk nada-nada dengan banyak puncak lokal signifikan.
2. Jika metode ekstraksi ciri dengan penjumlahan Gaussian digunakan untuk nada-nada dengan banyak puncak lokal signifikan, seperti misalnya pada pianika, jumlah koefisien ekstraksi ciri yang terkecil, yang menghasilkan tingkat pengenalan tertinggi adalah 8 koefisien, pada penggunaan DST 128 atau 256 titik, panjang segmen 16 atau 32 titik, serta nilai σ jendela Gaussian dari 2 hingga 6.
3. Jika metode ekstraksi ciri dengan penjumlahan Gaussian digunakan untuk nada-nada dengan sedikit puncak lokal signifikan, seperti misalnya pada rekorder sopran, jumlah koefisien ekstraksi ciri yang terkecil, yang menghasilkan tingkat pengenalan tertinggi adalah 4 koefisien, pada penggunaan DST 64 atau 128 titik, panjang segmen 16 atau 32 titik, serta pada nilai σ jendela Gaussian sekurang-kurangnya 4.
4. Agar supaya metode ekstraksi ciri dengan penjumlahan Gaussian dapat digunakan secara optimal, baik untuk baik untuk alat musik dengan banyak ataupun sedikit puncak lokal pada ranah DST, dapat menggunakan DST 128 titik,

5. panjang segmen 16 titik, serta nilai $\sigma=4$ pada jendela Gaussian. Pada kondisi optimal tersebut, jumlah koefisien

ekstraksi cirinya adalah 8 koefisien, dengan tingkat pengenalan 97,1 %.

DAFTAR PUSTAKA

- [1] Surya, D. E., dan Sumarno, L., 2012, Pengenalan Nada Pianika Menggunakan Jendela Kaiser, FFT, dan Korelasi, *Prosiding Seminar Nasional Rekayasa Teknologi Industri dan Informasi (RETH) 2012*, STTNAS, Yogyakarta, Hal. 151-157.
- [2] Cheveigné, A.De, dan Kawahara, H., 2002, YIN, A Fundamental Frequency Estimator for Speech and Music, *The Journal of the Acoustical Society of America*, Hal. 111-117.
- [3] McLeod, P., dan Wyvill, G., 2005, A Smarter Way to Find Pitch. *Proceedings of the International Computer Music Conference (ICMC'05)*, Barcelona.
- [4] Noll, A.M., 1970, Pitch Determination of Human Speech by the Harmonic Product Spectrum, the Harmonic Sum Spectrum and a Maximum Likelihood Estimate, *Proceedings of the Symposium on Computer Processing in Communications*, Vol. XIX, Polytechnic Press, Brooklyn, New York, Hal. 779-797.
- [5] Brown, J.C. dan Puckette, M.S., 1993, A High Resolution Fundamental Frequency Determination based on Phase Changes of the Fourier Transform, *The Journal of the Acoustical Society of America*, Vol. 94, Issue 2, Hal. 662-667.
- [6] Mitre, A., Marcelo, Q., dan Régis, F., 2006, Accurate and Efficient Fundamental Frequency Determination from Precise Partial Estimates, *Proceedings of the 4th AES Brazil Conference*, Hal. 113-118.
- [7] Gaffar I., Hidayatno, A., dan Zahra, A.A., 2012, Aplikasi Pengkonversi Nada-nada Instrumen Tunggal Menjadi Chord Menggunakan Metode Pitch Class Profile, *Transient*, Vol. 1, No. 3, Hal. 121-127.
- [8] Hendra, M., dan Sumarno, L., 2012, Pengenalan Nada Suling Rekorder Menggunakan Fungsi Jarak Chebyshev, *Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) 2012*, Institut Sains dan Teknologi AKPRIND, Yogyakarta, Hal. B-82 - B-89.
- [9] Sumarno L., 2013, Pengenalan Nada Pianika Menggunakan Jendela Gaussian, DCT, dan Jarak Kosinus, *Jurnal Penelitian*, Vol. 17, No. 1, Hal. 8-15.
- [10] Fruandta A, dan Buono A., 2011, Identifikasi Campuran Nada pada Suara Piano Menggunakan Codebook, *Prosiding Seminar Nasional Aplikasi Teknologi Informasi (SNATI) 2011*, Yogyakarta, Hal. G8-G13.
- [11] Sumarno, L., 2014, Pengenalan Nada Pianika Menggunakan Jendela Blackman dan Ekstraksi Ciri Transformasi Fourier Cepat, *Media Teknika*, Vol. 9, No. 2, Hal. 84-93.
- [12] Sumarno, L. 2016, On The Performance of Segment Averaging of Discrete Cosine Transform Coefficients on Musical Tone Recognition, *ARPN Journal of Engineering and Applied Sciences*, Vol. 11, No. 9, Hal. 5644-5649, http://www.arpnjournals.com/jeas/volume_09_2016.htm
- [13] Forster, C., 2010, *Musical Mathematics: On the Art of Science and Acoustic Instruments*, Chronicle Books LLC, San Diego, California.
- [14] Harris, F.J., 1978, On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform, *Proceedings of the IEEE*, Vol. 66, No. 1, Hal. 51-83.
- [15] Jain, A. K., 1989, *Fundamentals of Digital Image Processing*, Prentice-Hall International Inc., New Jersey.
- [16] Setiawan, Y.R., 2015, Pengenalan Ucapan Angka Menggunakan Transformasi Fourier Cepat dan Similaritas Kosinus, *Skripsi*, Program Studi Teknik Elektro, Universitas Sanata Dharma, Yogyakarta
- [17] Jain, A.K., Duin, R.P.W., dan Mao, J., 2000, Statistical Pattern Recognition: A Review, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, Hal. 4-37.
- [18] Wilson D.R., dan Martinez, T.R., 1997, Improved Heterogeneous Distance Functions, *Journal of Artificial Intelligence Research*, Vol. 6, Hal. 1-34.
- [19] Meseguer, N.A., 2009, Speech Analysis for Automatic Speech Recognition, *MSc Thesis*, NTNU, Trondheim.

penjendelaan_Penjendelaan Gaussian pada Ekstraksi Ciri Perataan Segmen Berbasis Transformasi Sinus Diskrit Pada Pengenalan Nada Alat Musik

ORIGINALITY REPORT

5%

SIMILARITY INDEX

5%

INTERNET SOURCES

0%

PUBLICATIONS

0%

STUDENT PAPERS

PRIMARY SOURCES

1

repository.usd.ac.id

Internet Source

4%

2

trilasmia.wordpress.com

Internet Source

<1%

3

www.scribd.com

Internet Source

<1%

Exclude quotes On

Exclude matches < 5 words

Exclude bibliography On