

## MUSICAL INSTRUMENT TONE RECOGNITION USING DCT BASED FEATURE EXTRACTION AND GAUSSIAN WINDOWING

**Linggo Sumarno**

Program Studi Teknik Elektro, Universitas Sanata Dharma  
Kampus III, Paingan, Maguwoharjo, Depok, Sleman, Yogyakarta 55282  
Email: lingsum@usd.ac.id

### *Abstract*

*This paper studied a feature extraction subsystem in a musical instrument tone recognition system. The purpose of this study was to obtain a number of feature extraction coefficients that are smaller than those obtained in previous papers. The studied subsystem was a DCT (Discrete Cosine Transform)-based segment averaging and Gaussian windowing. The testing of the musical instrument's tone recognition system was carried out using pianica, tenor recorder, and bellyra musical instruments, each of which represented many, several, and one significant local peaks in the transform domain. The test results showed that the optimal number of feature extraction coefficient was 8 coefficients, which could give a recognition rate of up to 100%. The test results were achieved using a Gaussian window with a alpha value of 2-6, and a 128 points DCT.*

**Keyword:** *Tone recognition, feature extraction, segment averaging, DCT, Gaussian window.*

### **1. Pendahuluan**

Artikel-artikel yang terkait dengan subsistem ekstraksi ciri dalam suatu sistem pengenalan nada alat musik, umumnya terbagi dalam dua pendekatan. Pendekatan pertama adalah pendekatan yang berdasarkan ranah waktu [1] [2]. Pendekatan kedua adalah pendekatan yang berdasarkan ranah transformasi. DCT (*Discrete Cosine Transform*), DST (*Discrete Sine Transform*), dan FFT (*Fast Fourier Transform*), adalah tiga metode transformasi yang umumnya digunakan. Pada pendekatan yang berdasarkan ranah transformasi, ada dua cara yang umumnya dipakai. Cara pertama adalah cara yang menggunakan sinyal-sinyal fundamental [3] - [6]. Cara kedua adalah cara yang tidak menggunakan sinyal-sinyal fundamental [7] -[11].

Pada artikel-artikel sebelumnya di atas, untuk cara yang tidak menggunakan sinyal-sinyal fundamental, kebanyakan artikel hanya untuk alat-alat musik yang mempunyai banyak [7], sedikit [8], banyak dan sedikit [9], serta banyak dan satu [11] puncak lokal signifikan pada ranah transformasi. Masih sangat sedikit artikel untuk alat-alat musik yang mempunyai banyak, sedikit, atau satu puncak lokal signifikan pada ranah transformasi [10]. Pada artikel sebelumnya [10], peneliti mengusulkan suatu subsistem ekstraksi ciri perataan segmen berbasis FFT, untuk digunakan dalam suatu sistem pengenalan nada alat musik. Namun, untuk dapat mengenali suatu nada yang mempunyai banyak, sedikit, atau satu puncak lokal signifikan pada ranah transformasi, sistem pengenalan nada alat musik tersebut masih memerlukan sekurang-kurangnya 16 koefisien ekstraksi ciri. Hal ini merupakan indikasi masih adanya kesempatan untuk mengurangi jumlah koefisien ekstraksi ciri tersebut.

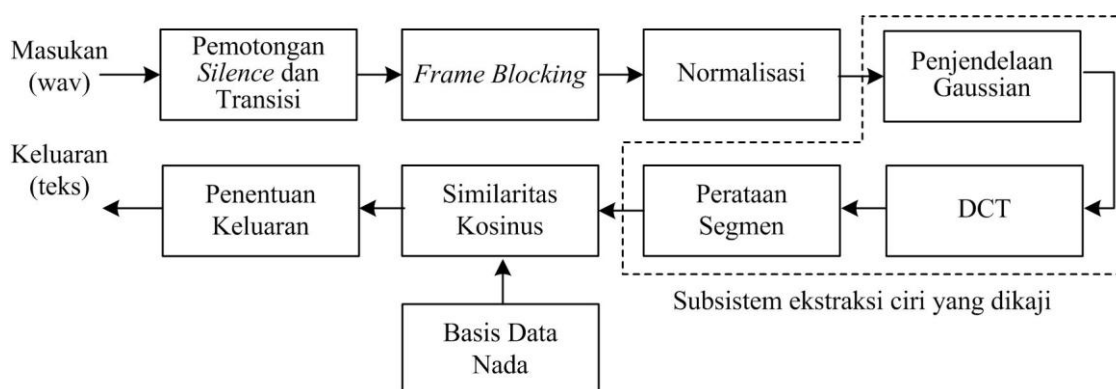
Artikel ini mengkaji suatu subsistem ekstraksi ciri dalam suatu sistem pengenalan nada alat musik, yang tidak berdasarkan sinyal-sinyal fundamental. Subsistem ekstraksi ciri ini adalah perataan segmen berbasis DCT dan penjendelaan Gaussian. Subsistem ekstraksi ciri ini dapat menghasilkan jumlah koefisien ekstraksi ciri yang lebih sedikit dari metode-metode sebelumnya. Alat-alat musik yang digunakan dalam artikel ini adalah

pianika, rekorder tenor, dan belira, yang masing-masing mewakili alat-alat musik yang mempunyai banyak, beberapa, dan satu puncak lokal signifikan pada ranah transformasi.

## 2. Metodologi Penelitian

### 2.1. Pembuatan sistem pengenalan nada dan subsistem ekstraksi ciri

Langkah pertama dalam metodologi penelitian adalah pembuatan sistem pengenalan nada alat musik, yang secara diagram blok diperlihatkan pada Gambar 1. Masukan sistem adalah sinyal nada yang terisolasi dalam format wav. Keluaran sistem adalah teks, yang mengindikasikan sinyal nada yang dikenali. Subsistem ekstraksi ciri yang dikaji, akan dibahas lebih detail pengaruhnya pada bagian pembahasan. Sebagai catatan, implementasi dari sistem pengenalan dibuat menggunakan perangkat lunak Octave. Secara lebih detail, masukan dan fungsi-fungsi dari setiap blok yang ada dalam Gambar 1, dideskripsikan dalam subbagian 2.1.1-2.1.9.



Gambar 1. Sistem pengenalan nada alat musik secara keseluruhan.

#### 2.1.1 Masukan

Masukan berupa sinyal nada dari tiga alat musik yaitu pianika, rekorder sopran, dan belira. Sinyal nada ini merupakan sinyal nada terisolasi, yang direkam dalam format wav. Artikel ini menggunakan delapan sinyal nada yaitu C, D, E, F, G, A, B, dan C', yang merupakan sinyal-sinyal nada satu untuk oktaf. Pengambilan sinyal nada dilakukan dengan merekam sinyal nada tersebut menggunakan frekuensi pencuplikan yang memenuhi teorema pencuplikan Shannon [12]:

$$f_s \geq 2f_{\text{maks}} \quad (1)$$

dengan  $f_s$  adalah frekuensi pencuplikan, dan  $f_{\text{maks}}$  adalah komponen frekuensi tertinggi dari sinyal yang akan dicuplik. Artikel ini menggunakan frekuensi pencuplikan 5000 Hz. Berdasarkan hasil evaluasi spektrum sinyal nada, komponen frekuensi signifikan yang tertinggi dari nada C' untuk alat musik pianika, rekorder sopran, dan belira, masing-masing adalah 2100 Hz, 547 Hz, dan 2109 Hz. Dengan demikian, frekuensi pencuplikan 5000 Hz tersebut sudah memenuhi teorema pencuplikan Shannon. Selanjutnya, perekaman sinyal nada dilakukan selama dua detik. Berdasarkan hasil evaluasi amplitudo sinyal nada, pemilihan dua detik ini sudah mencukupi untuk mendapatkan lebih dari separuh bagian sinyal nada yang berada pada kondisi *steady state*. Sebagai catatan, bagian yang *steady state* ini, terdapat informasi nada yang akurat.

Alat-alat musik yang digunakan dalam artikel ini adalah pianika Yamaha P-37D, rekorder tenor Yamaha YRT-304B II, dan belira Isuzu ZBL-27. Alat-alat musik tersebut diperlihatkan dalam Gambar 2. Untuk menangkap sinyal nada dari alat-alat musik tersebut digunakan mikrofon AKG Perception 120 USB.



Gambar 2. Pianika, rekorder tenor, dan belira yang digunakan dalam artikel ini [10].

### 2.1.2 Pemotongan *silence* dan transisi

Pemotongan *silence* dan transisi adalah proses pemotongan daerah *silence* dan transisi dari sinyal nada. Berdasarkan hasil pengamatan amplitudo, sinyal nada hasil perekaman terdiri atas tiga daerah yaitu *silence*, transisi, dan *steady state*. Pada daerah *silence* tidak ada informasi nada. Pada daerah transisi ada informasi nada, namun belum akurat. Pada daerah *steady state* ada informasi nada yang akurat. Oleh karena itu, untuk mendapatkan daerah *steady state*, perlu dilakukan pemotongan daerah *silence* dan transisi. Berdasarkan juga hasil pengamatan amplitudo, pertama kali pemotongan daerah *silence* dapat dilakukan dengan memotong mulai dari bagian paling kiri dari sinyal nada menggunakan nilai ambang  $|0,5|$  dari nilai amplitudo maksimum. Selanjutnya pemotongan daerah transisi dapat dilakukan dengan memotong lagi mulai dari bagian kiri dari sinyal nada selama 0,2 detik.

### 2.1.3 *Frame blocking*

*Frame blocking* adalah proses pemotongan sebagian kecil sinyal dari suatu sinyal yang panjang [13]. Pada dasarnya, dari sebagian kecil sinyal nada sudah bisa didapatkan informasi nada yang akurat. Dengan kata lain, tidak perlu memproses sinyal nada yang panjang untuk mendapatkan informasi nada. Sebagai catatan, sebagian kecil sinyal nada tersebut berasal dari suatu sinyal nada yang panjang, yang sudah pada kondisi *steady state*. Artikel ini mengevaluasi *frame blocking* yang panjangnya 32, 64, 128, dan 256 titik. Sebagai catatan, panjang *frame blocking* ini mengikuti panjang sinyal yang diperlukan pada masukan dari proses lanjutan yaitu proses perataan segmen. Proses perataan segmen ini memerlukan panjang sinyal dengan rumusan  $2^q$ , dengan  $q$  adalah bilangan bulat.

### 2.1.3 Normalisasi

Normalisasi adalah proses pengaturan nilai maksimum sinyal nada ke nilai 1. Berdasarkan hasil pengamatan amplitudo sinyal, terdapat variasi nilai maksimum absolut dari sinyal-sinyal nada hasil *frame blocking*. Normalisasi bertujuan untuk menyamakan nilai maksimum absolut sinyal-sinyal tersebut ke nilai 1. Proses normalisasi dikerjakan menggunakan persamaan berikut.

$$\mathbf{x}_{\text{out}} = \frac{\mathbf{x}_{\text{in}}}{\max |\mathbf{x}_{\text{in}}|} \quad (2)$$

dengan  $\mathbf{x}_{\text{out}}$  adalah sinyal nada hasil normalisasi, dan  $\mathbf{x}_{\text{in}}$  adalah sinyal nada hasil proses *frame blocking*.

### 2.1.4 Penjendelaan Gaussian

Penjendelaan Gaussian adalah proses penjendelaan menggunakan jendela Gaussian. Proses penjendelaan ini dilaksanakan dengan mengurangi amplitudo di daerah tepi-tepi sinyal. Berdasarkan hasil pengamatan, pengurangan amplitudo ini akan mengurangi munculnya *artifacts*, setelah sinyal nada ditransformasi menggunakan DCT. Sebagai catatan,

*artifacts* yang amplitudonya terlalu besar dapat mengaburkan sinyal nada hasil transformasi. Jendela Gaussian adalah salah satu jenis jendela yang biasa digunakan dalam pengolahan sinyal digital [14]. Secara matematis, jendela Gaussian dirumuskan sebagai berikut.

$$w(n) = e^{-\frac{1}{2}\left(\alpha \frac{n}{N/2}\right)^2} \quad (3)$$

dengan  $\alpha$  adalah faktor kelebaran jendela,  $N$  adalah panjang jendela, dan  $n$  ada dalam rentang  $-(N-1)/2 \leq n \leq (N-1)/2$ . Dalam artikel ini, nilai  $N$  adalah panjang *frame blocking*.

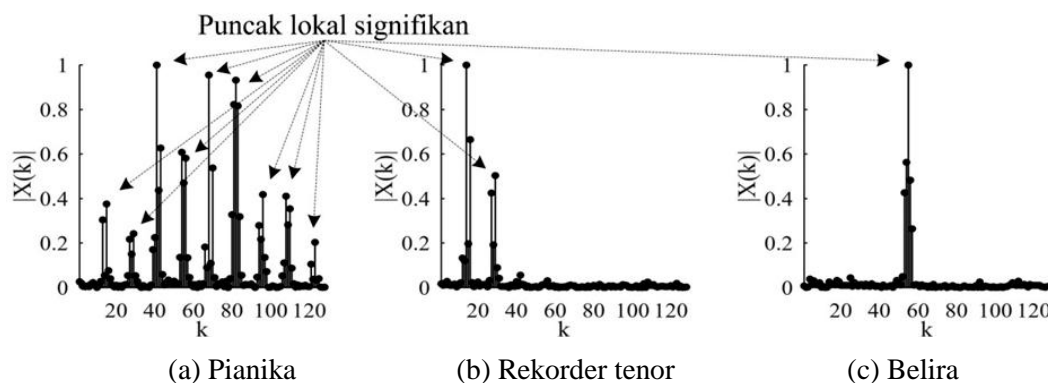
### 2.1.5 DCT

DCT adalah proses transformasi sinyal nada dari ranah waktu ke ranah DCT. Gambar 3 memperlihatkan contoh representasi tiga sinyal nada pada ranah DCT. Secara matematis, untuk sinyal nada  $x(n)$  yang panjangnya  $N$ , DCT dari sinyal tersebut dirumuskan sebagai berikut.

$$X(k) = \beta(k) \sum_{n=0}^{N-1} x(n) \cos\left[\frac{\pi(2n+1)k}{2N}\right], \quad 0 \leq k \leq N-1 \quad (4)$$

dengan

$$\beta(0) \triangleq \sqrt{\frac{1}{N}}, \quad \beta(k) \triangleq \sqrt{\frac{2}{N}} \quad \text{untuk } 0 \leq k \leq N-1. \quad (5)$$



Gambar 3. Representasi sinyal nada C dalam ranah DCT ternormalisir  $X(k)$ , untuk alat musik pianika, rekorder tenor, dan belira. Representasi tersebut diperoleh menggunakan frekuensi pencuplikan 5000 Hz, DCT 128 titik, dan jendela Gaussian dengan nilai  $\alpha=2$ .

Dalam artikel ini panjang DCT yang digunakan sama dengan panjang sinyal yang diperlukan pada masukan proses berikutnya yaitu perataan segmen. Selain itu, dalam artikel ini digunakan nilai hasil DCT yang absolut. Hal ini disebabkan, proses berikutnya yaitu perataan segmen, hanya memproses nilai-nilai yang bertanda positif.

### 2.1.6 Perataan segmen

Perataan segmen adalah proses untuk membuat sinyal nada yang panjang menjadi lebih pendek. Pada dasarnya, sinyal nada yang lebih pendek ini masih memperlihatkan bentuk dasar dari sinyal nada yang panjang. Artikel ini menggunakan jenis perataan segmen yang awalnya diinspirasi dari Setiawan [15]. Algoritma dari jenis perataan segmen ini diperlihatkan sebagai berikut.

1. Pada barisan data bernilai positif  $\{X(k) | 0 \leq k \leq N - 1\}$ , dengan  $N = 2^q$  untuk  $q \geq 0$ .
2. Tentukan panjang segmen  $L$ , dengan  $L = 2^p$  untuk  $0 \leq p \leq q$ .
3. Potong secara seragam sepanjang  $L$  barisan data  $X(k)$ . Pemotongan ini akan menghasilkan jumlah segmen

$$S = N / L \quad (5)$$

dan juga barisan data  $\{D(r) | 1 \leq r \leq L\}$  untuk setiap segmen.

4. Hitung nilai rerata untuk setiap segmen  $R(v)$  sebagai berikut.

$$R(v) = \frac{1}{L} \sum_{r=1}^L D_v(r), \quad 1 \leq v \leq S. \quad (6)$$

Keluaran dari proses perataan segmen adalah  $\{R(v) | 1 \leq v \leq S\}$  yang merupakan hasil ekstraksi ciri dari sinyal nada masukan. Hasil ekstraksi ciri ini mempunyai sejumlah  $S$  koefisien ekstraksi ciri. Dalam artikel ini dievaluasi panjang segmen  $L$  yang panjangnya  $2^{\log_2 N}$ , ..., 4, 2, dan 1 titik, dengan  $N$  adalah panjang sinyal nada pada masukan proses perataan segmen. Berdasarkan persamaan (5), evaluasi panjang segmen  $L$  tersebut akan memberikan jumlah koefisien ekstraksi ciri sebanyak 1, 2, 4, ..., dan  $2^{\log_2 N}$  koefisien.

### 2.1.7 Similaritas kosinus

Similaritas kosinus adalah proses perhitungan nilai similaritas menggunakan similaritas kosinus. Perhitungan nilai similaritas ini dilakukan untuk menghitung nilai similaritas antara ekstraksi ciri sinyal masukan dengan sejumlah ekstraksi ciri sinyal nada (C, D, E, F, G, A, B, dan C') yang tersimpan dalam basis data nada. Dengan demikian, pada keluaran proses similaritas kosinus, ada sejumlah delapan nilai similaritas. Similaritas kosinus merupakan suatu rumusan similaritas yang populer digunakan [16].

### 2.1.8 Penentuan keluaran

Penentuan keluaran adalah proses untuk menentukan teks keluaran (C, D, E, F, G, A, B, atau C') yang mengindikasikan nada yang dikenali. Langkah pertama dari proses ini adalah mencari nilai similaritas terbesar dari sejumlah delapan nilai similaritas, yang merupakan keluaran dari proses similaritas kosinus. Langkah berikutnya adalah penentuan nada yang dikenali. Suatu nada yang berasosiasi dengan nilai similaritas terbesar dinyatakan sebagai nada yang dikenali. Sebagai catatan pertama, nada yang dikenali ini berasosiasi dengan teks keluaran (C, D, E, F, G, A, B, atau C'). Sebagai catatan kedua, adanya penentuan keluaran dengan berdasarkan nilai similaritas kosinus terbesar mengindikasikan bahwa, metode penentuan nada yang dikenali menggunakan metode pencocokan *template* [17].

### 2.1.9 Basis data nada

Basis data nada adalah kumpulan dari sejumlah hasil ekstraksi ciri sinyal nada (C, D, E, F, G, A, B, dan C'). Hasil ekstraksi ciri ini diperoleh dengan melaksanakan proses pemotongan *silence* dan transisi hingga proses perataan segmen, yang diperlihatkan pada Gambar 1. Dalam artikel ini, untuk setiap alat musik (pianika, rekorder tenor, dan belira), diambil sejumlah 10 sampel nada pelatihan untuk setiap nada (C, D, E, F, G, A, B, dan C'). Peneliti mengasumsikan bahwa dengan 10 sampel nada pelatihan, semua variasi dari setiap sinyal nada dari suatu alat musik, telah didapatkan. Untuk setiap nada (C, D, E, F, G, A, B, dan C'), ekstraksi ciri dari 10 sampel nada pelatihan, akan menghasilkan 10 ekstraksi ciri sinyal nada. Selanjutnya untuk setiap alat musik dilaksanakan proses rerata berikut.

$$\mathbf{Y}_T = \frac{1}{10} \sum_{i=1}^{10} \mathbf{R}_i \quad (7)$$

dengan vektor  $\{\mathbf{R}_i | 1 \leq i \leq 10\}$  adalah 10 ekstraksi ciri sinyal nada, dan vektor  $\{\mathbf{Y}_T | T = C, D, E, F, G, A, B, \text{ dan } C'\}$  adalah delapan vektor yang dimasukkan dalam basis data nada suatu alat musik.

## 2.2. Pengambilan Nada Uji

Langkah kedua dalam metodologi penelitian adalah pengambilan nada uji, yang digunakan untuk pengujian sistem pengenalan nada alat musik. Artikel ini menggunakan 160 sampel nada uji yang berasal dari delapan nada (C, D, E, F, G, A, B, dan C), dengan 20 kali perekaman untuk setiap nada.

## 2.3 Pengujian dan Perhitungan Tingkat Pengenalan

Langkah terakhir dalam metodologi penelitian adalah pengujian dan perhitungan tingkat pengenalan. Pengujian dilaksanakan dengan menggunakan 160 sampel nada uji, untuk setiap alat musik, setiap nilai  $\alpha$  jendela Gaussian, setiap panjang DCT, dan setiap jumlah koefisien ekstraksi ciri. Perhitungan tingkat pengenalan adalah perhitungan rasio (yang dinyatakan dalam persen), antara jumlah nada yang dikenali dengan benar, dengan sejumlah 160 sampel nada uji.

## 3. Hasil dan Pembahasan

### 3.1 Hasil Pengujian

Sistem pengenalan nada alat musik yang diperlihatkan dalam Gambar 1 telah diuji untuk setiap alat musik, setiap nilai  $\alpha$  jendela Gaussian, setiap panjang DCT, dan setiap jumlah koefisien ekstraksi ciri. Hasil pengujian tersebut diperlihatkan pada Tabel 1, 2, dan 3. Sebagai catatan, evaluasi jumlah koefisien ekstraksi ciri sudah dijelaskan di subbagian 2.1.6.

Tabel 1. Hasil pengujian untuk alat musik pianika, pada berbagai nilai  $\alpha$  jendela Gaussian, panjang DCT, dan jumlah koefisien ekstraksi ciri. Hasil yang diperlihatkan: Tingkat pengenalan (%).

Nilai $\alpha$ jendela Gaussian	Panjang DCT (titik)	Jumlah koefisien ekstraksi ciri (koefisien)								
		1	2	4	8	16	32	64	128	256
1	32	12,5	58,75	96,88	99,38	98,13	96,25	-	-	-
	64	12,5	66,88	95,63	100	100	100	100	-	-
	128	12,5	67,50	93,75	100	100	100	100	100	-
	256	12,5	68,13	97,50	100	100	100	100	100	100
2	32	12,5	63,75	96,25	99,38	100	96,25	-	-	-
	64	12,5	71,25	97,50	<b>100</b>	100	100	100	-	-
	128	12,5	76,25	98,13	<b>100</b>	100	100	100	100	-
	256	12,5	68,75	99,38	<b>100</b>	100	100	100	100	100
4	32	12,5	40,00	76,88	96,88	97,50	76,25	-	-	-
	64	12,5	66,25	98,75	<b>100</b>	100	100	98,13	-	-
	128	12,5	75,00	97,50	<b>100</b>	100	100	100	100	-
	256	12,5	75,00	98,13	<b>100</b>	100	100	100	100	100
8	32	12,5	39,38	67,50	71,88	72,50	51,25	-	-	-
	64	12,5	48,75	84,38	97,50	98,13	97,50	90,63	-	-
	128	12,5	61,25	96,25	99,38	100	100	100	93,75	-
	256	12,5	72,50	98,13	<b>100</b>	100	100	100	100	100

Tabel 2. Hasil pengujian untuk alat musik recorder tenor, pada berbagai nilai  $\alpha$  jendela Gaussian, panjang DCT, dan jumlah koefisien ekstraksi ciri. Hasil yang diperlihatkan: Tingkat pengenalan (%).

Nilai $\alpha$ jendela Gaussian	Panjang DCT (titik)	Jumlah koefisien ekstraksi ciri (koefisien)								
		1	2	4	8	16	32	64	128	256
0	32	12,5	40,63	71,88	97,50	100	98,75	-	-	-
	64	12,5	52,50	80,00	95,00	100	100	100	-	-
	128	12,5	48,75	79,38	96,88	100	100	100	100	-
	256	12,5	55,63	81,88	96,88	100	100	100	100	100
2	32	12,5	39,38	89,38	<b>100</b>	100	98,75	-	-	-
	64	12,5	39,38	85,00	97,50	100	100	100	-	-
	128	12,5	56,88	83,13	<b>100</b>	100	100	100	100	-
	256	12,5	54,38	88,13	98,13	100	100	100	100	100
4	32	12,5	49,38	94,38	<b>100</b>	100	100	-	-	-
	64	12,5	36,25	90,00	<b>100</b>	100	100	100	-	-
	128	12,5	45,63	88,75	<b>100</b>	100	100	100	100	-
	256	12,5	54,38	88,75	98,75	100	100	100	100	100
8	32	12,5	60,00	88,13	91,25	91,25	66,88	-	-	-
	64	12,5	49,38	95,00	<b>100</b>	100	100	93,13	-	-
	128	12,5	47,50	90,63	<b>100</b>	100	100	100	100	-
	256	12,5	46,88	88,13	98,75	100	100	100	100	100

Tabel 3. Hasil pengujian untuk alat musik belira, pada berbagai nilai  $\alpha$  jendela Gaussian, panjang DCT, dan jumlah koefisien ekstraksi ciri. Hasil yang diperlihatkan: Tingkat pengenalan (%).

Nilai $\alpha$ jendela Gaussian	Panjang DCT (titik)	Jumlah koefisien ekstraksi ciri (koefisien)								
		1	2	4	8	16	32	64	128	256
1	32	12,5	53,75	81,88	95,00	90,63	85,63	-	-	-
	64	12,5	55,00	86,25	96,88	95,00	91,88	88,13	-	-
	128	12,5	56,25	80,00	95,00	95,63	92,50	88,75	87,50	-
	256	12,5	55,63	78,13	94,38	95,63	93,13	90,00	89,38	89,38
2	32	12,5	59,38	88,13	96,25	96,25	86,25	-	-	-
	64	12,5	60,00	88,75	98,13	100	98,13	88,75	-	-
	128	12,5	66,88	89,38	<b>100</b>	100	100	97,50	86,88	-
	256	12,5	63,13	85,00	98,13	96,88	99,38	97,50	95,63	92,50
4	32	12,5	66,25	86,25	90,63	91,88	81,88	-	-	-
	64	12,5	58,75	87,50	96,25	98,75	98,75	86,25	-	-
	128	12,5	67,50	92,50	<b>100</b>	100	100	100	86,88	-
	256	12,5	65,00	86,25	94,38	96,25	99,38	98,13	96,88	88,75
8	32	12,5	51,25	80,63	82,50	82,50	61,88	-	-	-
	64	12,5	60,63	81,88	90,00	91,88	92,50	81,88	-	-
	128	12,5	68,75	91,25	99,38	99,38	100	100	86,88	-
	256	12,5	61,88	87,50	93,13	95,63	96,25	96,88	97,50	88,13

### 3.2 Pembahasan

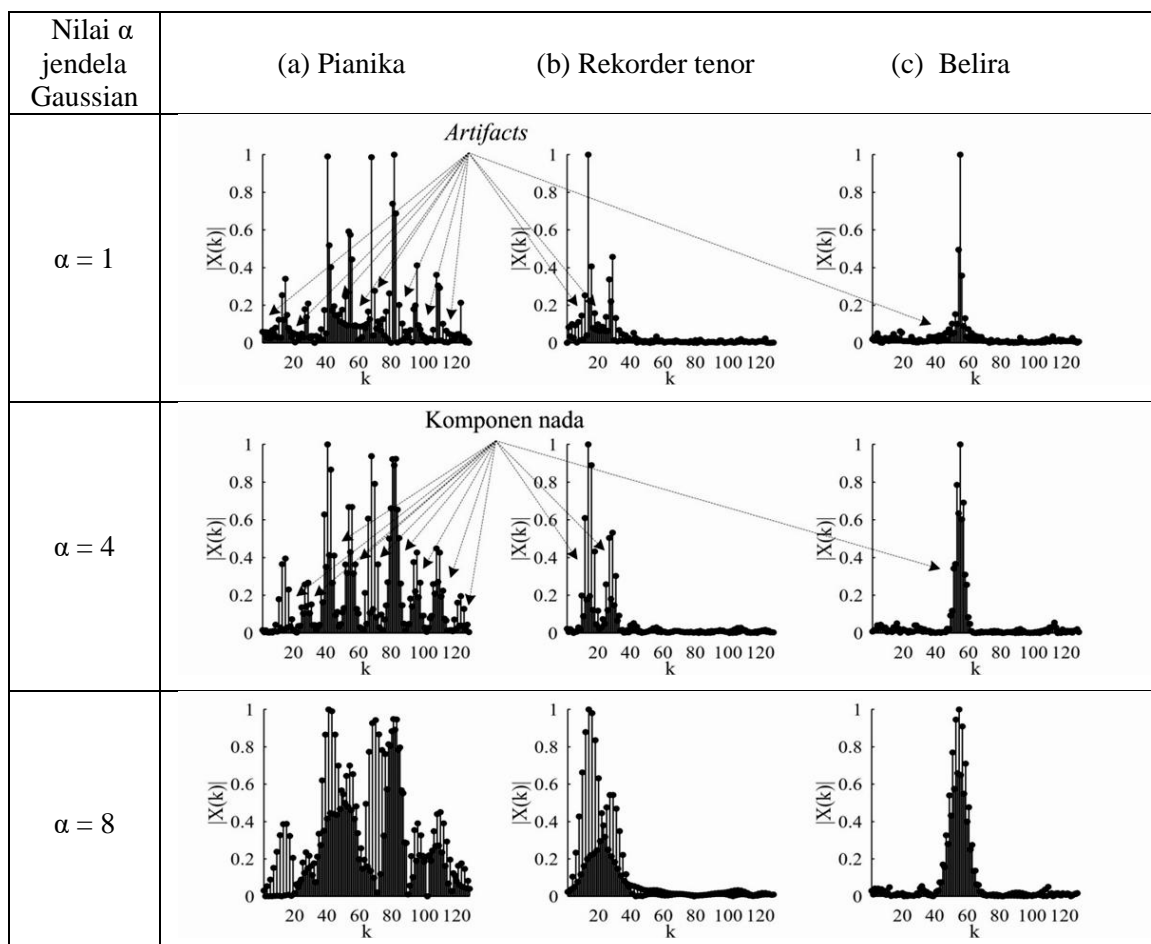
Dari sudut pandang hasil yang optimal, Tabel 1, 2, dan 3 mengindikasikan bahwa, untuk alat musik pianika, recorder tenor, dan belira, penggunaan 8 koefisien ekstraksi ciri dapat memberikan hasil tingkat pengenalan tertinggi hingga 100%. Hasil ini dicapai dengan menggunakan juga jendela Gaussian dengan nilai  $\alpha$  dari 2-6, serta DCT 128 titik. Maka, dapat dikatakan bahwa, untuk alat musik yang mempunyai banyak, beberapa, atau satu puncak lokal signifikan pada ranah transformasi (misalnya pianika, recorder tenor, dan belira), penggunaan 8 koefisien ekstraksi ciri sudah mencukupi untuk merepresentasikan setiap nada. Sebagai catatan, dalam hal ini digunakan jendela Gaussian dengan nilai  $\alpha$  dari 2-6, serta DCT 128 titik.

Dari sudut pandang jumlah koefisien ekstraksi ciri, Tabel 1, 2, dan 3 mengindikasikan bahwa secara umum, jika jumlah koefisien ekstraksi ciri semakin besar, maka tingkat

pengenalan akan semakin meningkat. Hal ini dikarenakan, jika jumlah koefisien ekstraksi ciri semakin besar akan makin banyak dimensi yang digunakan untuk membedakan pola nada satu dengan pola nada yang lain. Dengan makin banyaknya dimensi ini, akan menyebabkan tingkat diskriminasi ekstraksi ciri makin meningkat. Akhirnya, tingkat diskriminasi ekstraksi ciri yang makin meningkat ini akan menyebabkan tingkat pengenalan yang makin meningkat.

Dari sudut pandang nilai  $\alpha$  jendela Gaussian, Tabel 1, 2, dan 3 mengindikasikan bahwa, dari suatu nilai  $\alpha$  yang optimal, jika nilai  $\alpha$  makin kecil atau makin besar, akan menyebabkan makin menurunnya tingkat pengenalan. Berikut ini akan dibahas lebih detail pengaruh dari nilai  $\alpha$  tersebut.

- (a) Jika nilai  $\alpha$  makin kecil, sebagaimana terlihat pada Gambar 4, akan menyebabkan terjadinya dua hal. Hal yang pertama adalah makin kurusnya komponen-komponen nada pada sinyal nada hasil transformasi. Hal yang kedua adalah makin meningkatnya amplitudo *artifacts* yang muncul. Makin meningkatnya amplitudo *artifacts* mengakibatkan makin meningkatnya tingkat *overlap* diantara komponen-komponen nada. Akibat dari meningkatnya tingkat *overlap* ini adalah makin kaburnya sinyal nada hasil transformasi. Selanjutnya, hal ini makin menurunkan tingkat diskriminasi ekstraksi ciri. Akhirnya, tingkat diskriminasi ekstraksi ciri yang makin menurun menyebabkan tingkat pengenalan yang makin menurun.



Gambar 4. Penggambaran sinyal nada C dalam ranah DCT ternormalisir  $X(k)$ , untuk alat musik pianika, rekorder tenor, dan belira. Penggambaran tersebut menggunakan frekuensi pencuplikan 5000 Hz, DCT 128 titik, dan beragam nilai  $\alpha$  pada jendela Gaussian.



- (b) Jika nilai  $\alpha$  makin besar, sebagaimana terlihat pada Gambar 4, akan menyebabkan terjadinya dua hal. Hal yang pertama adalah makin gemuknya komponen-komponen nada pada sinyal nada hasil transformasi. Hal yang kedua adalah makin menurunnya amplitudo *artifacts* yang muncul. Makin gemuknya komponen-komponen nada mengakibatkan makin meningkatnya tingkat *overlap* diantara komponen-komponen nada. Akibat dari meningkatnya tingkat *overlap* adalah makin kaburnya sinyal nada hasil transformasi. Selanjutnya, hal ini makin menurunkan tingkat diskriminasi ekstraksi ciri. Akhirnya, tingkat diskriminasi ekstraksi ciri yang makin menurun menyebabkan tingkat pengenalan yang makin menurun.

### 3.3. Perbandingan dengan Ekstraksi Ciri yang Lain

Perbandingan kinerja pengenalan nada alat musik untuk beberapa metode ekstraksi ciri diperlihatkan pada Tabel 4. Perbandingan kinerja tersebut diperlihatkan untuk nada-nada yang mempunyai banyak, beberapa, dan satu puncak lokal signifikan pada ranah transformasi. Sebagaimana terlihat pada Tabel 4, metode ekstraksi ciri yang diusulkan dalam artikel ini masuk kategori paling efisien. Hal ini disebabkan, metode ekstraksi ciri yang dikaji dapat menghasilkan jumlah koefisien ekstraksi ciri yang paling kecil, untuk nada yang mempunyai banyak, beberapa, dan satu puncak lokal signifikan pada ranah transformasi.

Tabel 4. Perbandingan kinerja beberapa metode ekstraksi ciri.  
Hasil yang diperlihatkan: Jumlah koefisien ekstraksi ciri (koefisien)

Metode ekstraksi ciri	Jumlah puncak lokal signifikan pada ranah transformasi		
	Banyak	Sedikit	Satu
MFCC dan Codebook [7]	13 (piano pada keyboard)	-	-
Spectral Features [8]	-	21 (gamelan)	-
Perataan segmen berbasis DCT [9]	8 (pianika)	16 (rekorder sopran)	-
Perataan segmen berbasis FFT [10]	4 (pianika)	16 (rekorder tenor)	16 (belira)
Perataan segmen berbasis DST [11]	8 (pianika)	-	8 (belira)
Perataan segmen berbasis DCT dan penjendelaan Gaussian (artikel ini)	<b>8 (pianika)</b>	<b>8 (rekorder tenor)</b>	<b>8 (belira)</b>

## 4. Kesimpulan

Artikel ini mengkaji suatu subsistem ekstraksi ciri, yang dapat digunakan dalam suatu sistem pengenalan nada alat musik. Subsistem ekstraksi ciri tersebut adalah perataan segmen berbasis DCT dan penjendelaan Gaussian. Subsistem ekstraksi ciri tersebut diharapkan dapat memberikan jumlah koefisien ekstraksi ciri yang lebih sedikit dari yang pernah didapatkan pada artikel-artikel sebelumnya.

Berdasarkan percobaan, untuk alat musik pianika, rekorder tenor, dan belira, hasil yang optimal diperoleh dengan penggunaan 8 koefisien ekstraksi ciri, karena dapat memberikan hasil tingkat pengenalan tertinggi hingga 100%. Hasil ini dicapai dengan menggunakan juga jendela Gaussian dengan nilai  $\alpha$  dari 2-6, serta DCT 128 titik. Maka, secara umum, untuk alat musik dengan banyak, beberapa, dan satu puncak lokal signifikan pada ranah transformasi (misalnya pianika, rekorder tenor dan belira), diperlukan sekurang-kurangnya 8 koefisien ekstraksi ciri, untuk merepresentasikan setiap nada. Dalam hal ini diperlukan jendela Gaussian dengan  $\alpha$  antara 2-6 dan panjang DCT 128 titik.

Untuk penelitian lebih lanjut, dapat dieksplorasi penggunaan jendela-jendela yang lain selain jendela Gaussian. Selain itu, dapat juga dieksplorasi ekstraksi ciri perataan segmen yang lain selain yang berbasiskan DCT.

### Daftar Pustaka

- [1] Cheveigné, A de, & Kawahara, H. (2002). YIN, A Fundamental Frequency Estimator for Speech and Music. *The Journal of the Acoustical Society of America*. Pp. 111-117.
- [2] McLeod, P., & Wyvill, G. (2005, September). A Smarter Way to Find Pitch. In *International Computer Music Conference (ICMC), Barcelona*. Pp. 138–141.
- [3] Mitre, A., Queiroz, M., & Faria, R. (2006, May). Accurate and Efficient Fundamental Frequency Determination from Precise Partial Estimates. In *4th Audio Engineering Society (AES) Brazil Conference, Sao Paulo*. Pp. 113–118.
- [4] Pertusa, A., & Inesta, J.M. (2008, March-April). Multiple Fundamental Frequency Estimation using Gaussian Smoothness. In *IEEE International Conference on Audio, Speech, and Signal Processing (ICASSP), Las Vegas*. Pp. 105–108.
- [5] Yeh, C., Robel, A., & Rodet, X. (2010). Multiple Fundamental Frequency Estimation and Polyphony Inference of Polyphonic Music Signals. *IEEE Transactions on Audio, Speech, and Language Processing*. Vol. 18(6), pp. 1116–1126.
- [6] Duan, Z., Pardo, B., & Zhang, C. (2010). Multiple Fundamental Frequency Estimation by Modeling Spectral Peaks and Non-peak Regions. *IEEE Transactions on Audio, Speech, and Language Processing*. Vol. 18(8), pp 2121–2133.
- [7] Fruandta, A., & Buono, A. (2011, June). Identifikasi Campuran Nada pada Suara Piano Menggunakan Codebook. In *Seminar Nasional Aplikasi Teknologi Informasi (SNATI), Universitas Islam Indonesia, Yogyakarta*. Pp. G8–G13.
- [8] Tjahyanto, A., Suprpto, Y.K., & Wulandari, D,P. (2013). Spectral-based Features Ranking for Gamelan Instruments Identification using Filter Techniques. *Telkomnika*. Vol. 11(1), pp. 95–106.
- [9] Sumarno, L. (2016). On The Performace of Segment Averaging of Discrete Cosine Transform Coefficients on Musical Instruments Tone Recognition. *ARNP Journal of Engineering and Applied Sciences*. Vol. 11(9), pp. 5644–5649.
- [10] Sumarno, L., & Iswanjono. (2017). Feature Extraction of Musical Instrument Tones using FFT and Segment Averaging. *Telkomnika*. Vol. 15(3), pp. 1280–1289.
- [11] Sumarno, L. (2018). Pengenalan Nada Alat Musik Menggunakan Ekstraksi Ciri Perataan Segmen Berbasis DST dan Pengklasifikasi SVM. *Jurnal Teknologi*. Vol 10(2), pp 23–31.
- [12] Tan, L., & Jiang, J. (2013). *Digital Signal Processing Fundamentals and Applications, Second Edition*. Elsevier Inc. Oxford. Pp. 15–56.
- [13] Meseguer, N.A. (2009). *Speech Analysis for Automatic Speech Recognition*. MSc Thesis. NTNU. Trondheim. Pp. 4–25.
- [14] Roberts, R.A., & Mullis, C.T. (1987). *Digital Signal Processing*. Addison-Wesley. Reading, Massachussets. Pp. 135–136.
- [15] Setiawan, Y.R. (2015). *Pengenalan Ucapan Angka Menggunakan Transformasi Fourier Cepat dan Similaritas Kosinus*. Skripsi. Universitas Sanata Dharma. Yogyakarta. Pp. 66–70.
- [16] Zhu, S., Wu, J., Xiong, H., & Xia, G. (2011). Scaling up top-K similarity search. *Data and Knowledge Engineering*. Vol. 70, pp. 60–83.
- [17] Theodoridis, S., & Koutroumbas, K. (2009). *Pattern Recognition, Fourth Edition*. Elsevier Inc. San Diego, California. Pp. 481–519.