

## IDENTIFIKASI EMOSI MANUSIA BERDASARKAN UCAPAN MENGGUNAKAN METODE EKSTRAKSI CIRI LPC DAN METODE EUCLIDEAN DISTANCE

Siti Helmiyah<sup>\*1</sup>, Imam Riadi<sup>2</sup>, Rusydi Umar<sup>3</sup>, Abdullah Hanif<sup>4</sup>, Anton Yudhana<sup>5</sup>, Abdul Fadlil<sup>6</sup>

<sup>1,3,4</sup>Program Studi Teknik Informatika, Universitas Ahmad Dahlan

<sup>2</sup>Program Studi Sistem Informasi, Universitas Ahmad Dahlan

<sup>5,6</sup>Program Studi Teknik Elektro, Fakultas Teknologi Industri, Universitas Ahmad Dahlan, Yogyakarta

Email: <sup>1</sup>siti1708048022@webmail.uad.ac.id, <sup>2</sup>imam\_riadi@is.uad.ac.id, <sup>3</sup>rusydi@mti.uad.ac.id,

<sup>4</sup>abdullah1708048026@webmail.uad.ac.id, <sup>5</sup>eyudhana@ee.uad.ac.id, <sup>6</sup>fadlil@mti.uad.ac.id

\*Penulis Korespondensi

(Naskah masuk: 09 November 2019, diterima untuk diterbitkan: 26 November 2020)

### Abstrak

Ucapan merupakan sinyal yang memiliki kompleksitas tinggi terdiri dari berbagai informasi. Informasi yang dapat ditangkap dari ucapan dapat berupa pesan terhadap lawan bicara, pembicara, bahasa, bahkan emosi pembicara itu sendiri tanpa disadari oleh si pembicara. Speech Processing adalah cabang dari pemrosesan sinyal digital yang bertujuan untuk terwujudnya interaksi yang natural antar manusia dan mesin. Karakteristik emosional adalah fitur yang terdapat dalam ucapan yang membawa ciri-ciri dari emosi pembicara. Linear Predictive Coding (LPC) adalah sebuah metode untuk mengekstraksi ciri dalam pemrosesan sinyal. Penelitian ini, menggunakan LPC sebagai ekstraksi ciri dan Metode Euclidean Distance untuk identifikasi emosi berdasarkan ciri yang didapatkan dari LPC. Penelitian ini menggunakan data emosi marah, sedih, bahagia, netral dan bosan. Data yang digunakan diambil dari Berlin Emo DB, dengan menggunakan tiga kalimat berbeda dan aktor yang berbeda juga. Penelitian ini menghasilkan akurasi pada emosi sedih 58,33%, emosi netral 50%, emosi marah 41,67%, emosi bahagia 8,33% dan untuk emosi bosan tidak dapat dikenali. Penggunaan Metode LPC sebagai ekstraksi ciri memberikan hasil yang kurang baik pada penelitian ini karena akurasi rata-rata hanya sebesar 31,67% untuk identifikasi semua emosi. Data suara yang digunakan dengan kalimat, aktor, umur dan aksen yang berbeda dapat mempengaruhi dalam pengenalan emosi, maka dari itu ekstraksi ciri dalam pengenalan pola ucapan emosi manusia sangat penting. Hasil akurasi pada penelitian ini masih sangat kecil dan dapat ditingkatkan dengan menggunakan ekstraksi ciri yang lain seperti prosidis, spektral, dan kualitas suara, penggunaan parameter *max*, *min*, *mean*, *median*, *kurtosis* dan *skewenes*. Selain itu penggunaan metode klasifikasi juga dapat mempengaruhi hasil pengenalan emosi.

**Kata kunci:** emosi, ucapan, Linear Prediction Coding (LPC), Euclidean Distance, Berlin Emo DB

## IDENTIFICATION EMOTION RECOGNITION USING LPC EXTRACTION FEATURE METHOD AND EUCLIDEAN DISTANCE METHOD

### Abstract

*Speech is a signal that has a high complexity consisting of various information. Information that can be captured from speech can be in the form of messages to interlocutor, the speaker, the language, even the speaker's emotions themselves without the speaker realizing it. Speech Processing is a branch of digital signal processing aimed at the realization of natural interactions between humans and machines. Emotional characteristics are features contained in the speech that carry the characteristics of the speaker's emotions. Linear Predictive Coding (LPC) is a method for extracting features in signal processing. This research uses LPC as a feature extraction and Euclidean Distance Method to identify emotions based on features obtained from LPC. This study uses data on emotions of anger, sadness, happiness, neutrality, and boredom. The data used was taken from Berlin Emo DB, using three different sentences and different actors. This research resulted in inaccuracy in sad emotions 58.33%, neutral emotions 50%, angry emotions 41.67%, happy emotions 8.33% and bored emotions could not be recognized. The use of the LPC method as feature extraction gave unfavorable results in this study because the average accuracy was only 31.67% for the identification of all emotions. Voice data used with different sentences, actors, ages, and accents can influence the recognition of emotions, therefore the extraction of features in the recognition of speech patterns of human emotions is very important. Accuracy results in this study are still very small and can be improved by using other feature extractions such as provides, spectral, and sound quality, using*

parameters max, min, mean, median, kurtosis, and skewness. Besides the use of classification methods can also affect the results of emotional recognition.

**Keywords:** emotion, speech, Linear Prediction Coding (LPC), Euclidean Distance, Berlin Emo DB

## 1. PENDAHULUAN

Ucapan merupakan cara manusia menyampaikan informasi terhadap sesamanya. Karena hal tersebut, ucapan seringkali disebut sebagai sinyal. Ucapan merupakan sinyal kompleks yang berisi informasi (Likitha *et al.*, 2017). Selain menyampaikan pesan dan maksud si pembicara, seringkali lawan bicara mampu mengenali pembicara tanpa harus mengenali wajahnya, juga memahami *mood* dan kondisi emosional pembicara tersebut.

Emosi adalah bagian dari manusia yang diekspresikan secara nyata (Gumelar *et al.*, 2019). Emosi adalah keadaan psikologis seseorang sebagai respon terhadap terjadinya suatu keadaan yang dipengaruhi kondisi sekitarnya baik dari kondisi dalam orang tersebut maupun sekitarnya. Paul Ekman pada tahun 1972, psikolog dari Amerika (Aswari and Diana, 2016) mengklasifikasikan bahwa emosi manusia berasal dari enam emosi dasar, yaitu: marah, jijik, takut, bahagia, sedih dan terkejut. Dalam ucapan, kondisi emosional kerap terekspresikan tanpa disadari pembicara. Sehingga akan terdapat fitur-fitur tertentu yang secara alami membawa informasi terkait emosi (Kumar and Mahajan, 2019). Hal ini dapat dimanfaatkan dalam bidang interaksi antara manusia dan mesin sehingga dapat terjadi interaksi natural yang mirip dengan interaksi antar manusia.

Bidang pemrosesan sinyal digital merupakan bidang yang memproses sinyal agar informasi yang ada di dalam sinyal dapat dimanfaatkan. Pemrosesan ucapan merupakan cabang penelitian dari bidang pemrosesan sinyal digital yang menarik perhatian ketika bertemakan emosi. Pengenalan ucapan dapat dimanfaatkan dalam bidang interaksi antara manusia dan komputer seperti identifikasi dan perintah suara. Pengenalan emosi dalam ucapan dapat memberikan keunggulan dalam interaksi antara manusia dan mesin. Pengenalan emosi berdasarkan ucapan sangat berguna dalam interaksi manusia dan mesin (Wang *et al.*, 2015).

Suara manusia adalah sinyal analog yang memiliki informasi. Beberapa parameter statistik dapat digunakan sebagai ekstraksi ciri untuk melakukan pengenalan emosi manusia melalui ucapan diantaranya adalah parameter *pitch*, energi, waktu, dan spektral seperti LPC, LPCC, dan MFCC (Prasetyo, Kurniawan & Ichsan, 2017).

Penelitian ini melakukan identifikasi ucapan untuk pengenalan pola emosi manusia dengan menggunakan metode Linear Predictive Coding (LPC) sebagai ekstraksi ciri ucapan karena sederhana. LPC adalah metode yang sering digunakan untuk pemrosesan sinyal audio dan pemrosesan ucapan.

LPC adalah salah satu metode yang berguna untuk memodelkan ucapan yang berkualitas pada *bit rate* rendah dan memberikan perkiraan parameter ucapan yang akurat (Chamoli, Semwal and Saikia, 2017). Metode Euclidean Distance digunakan sebagai identifikasi atau pencocokan ucapan. Metode Euclidean Distance merupakan metode yang sangat sederhana dan simpel karena cara yang dilakukan adalah menghitung kemiripan nilai ekstraksi ciri pada data pengujian dengan data pelatihan menggunakan rumus Pythagoras (Dewi, Zulkarnain and Lestari, 2018). Penggunaan metode Euclidean Distance pada penelitian ini untuk mengetahui ekstraksi ciri yang digunakan cocok atau tidak pada *database* dan kasus yang diteliti.

Tujuan dari penelitian adalah sebagai penelitian awalan yang akan digunakan sebagai dasar dan fondasi awal dari sebuah penelitian mengenai emosi. Penelitian ini masih menggunakan data emosi yang sedikit, untuk melihat hasil dari ekstraksi cirinya. Metode klasifikasi yang digunakan juga sangat sederhana untuk. Penelitian awalan ini akan dilanjutkan dengan data yang lebih banyak, metode ekstraksi ciri yang lain, dan metode klasifikasi yang lebih baik lagi.

## 2. METODE PENELITIAN

Penelitian ini dilakukan untuk mengetahui identifikasi emosi manusia berdasarkan ucapan dengan menggunakan metode ekstraksi ciri Linear Coding Prediction (LPC) dan identifikasi menggunakan metode Euclidean Distance. Tahapan pada penelitian ini dapat dilihat pada blok diagram yang ditunjukkan Gambar 1.



Gambar 1. Blok Diagram

Gambar 1 adalah tahapan-tahapan pada penelitian ini. Tahap pertama menunjukkan data yang digunakan adalah Berlin Emo DB. Tahap kedua pemrosesan awal yaitu proses peng-*input*-an data dan normalisasi. Tahap ketiga yaitu proses ekstraksi ciri menggunakan LPC. Tahap keempat adalah klasifikasi emosi menggunakan metode Euclidean Distance. Tahap kelima adalah hasil identifikasi emosi yang didapatkan dari proses klasifikasi.

## 2.1. Berlin EMO DB

Emotion Speech Berlin (Emo DB) adalah *database* yang sering digunakan dalam bidang pemrosesan ucapan (Lanjewar, Mathurkar and Patel, 2015). Emotion Speech Berlin adalah *database* emosi ucapan yang direkam oleh Technical University of Berlin di Jerman pada tahun 1997. Emo DB ini adalah proyek penelitian yang dipimpin oleh Prof. Dr. W. Sendlmeier dengan dana yang diperoleh dari German Research Community. Alat rekaman yang digunakan pada Emo DB adalah mikrophone Sennheiser pada frekuensi 16 kHz. Rekaman ini disuarakan oleh sepuluh aktor yang dipilih dari proses seleksi. Sepuluh aktor ini terdiri dari lima orang laki-laki dan lima orang perempuan. Kisaran umur aktor adalah berkisar antara 21 tahun sampai 35 tahun. Data aktor yang digunakan pada penelitian ini adalah semua aktor yang ada di Berlin Emo DB, data aktor dapat dilihat pada Tabel 1.

Tabel 1. Aktor Berlin Emo DB

Kode	Jenis Kelamin, Umur
03	Laki-laki, 31 tahun
08	Perempuan, 34 tahun
09	Perempuan, 21 tahun
10	Laki-laki, 32 tahun
11	Laki-laki, 26 tahun
12	Laki-laki, 30 tahun
13	Perempuan, 32 tahun
14	Perempuan, 35 tahun
15	Laki-laki, 25 tahun
16	Perempuan, 31 tahun

Kalimat yang digunakan pada Berlin Emo DB adalah kalimat netral yang diucapkan dalam bahasa Berlin. Kalimat tersebut diucapkan dalam percakapan sehari-hari dan dapat diucapkan dalam semua emosi. Berlin Emo DB memiliki tujuh emosi yaitu marah, bosan, bahagia, netral, sedih, jijik, dan takut. Penelitian ini menggunakan menggunakan lima emosi dari Berlin Emo DB yaitu emosi marah, sedih, bahagia, netral dan bosan. Kalimat yang digunakan dalam penelitian ini dapat dilihat pada Tabel 2

Tabel 2. Kalimat Berlin Emo DB

Kode	Kalimat Berlin Emo DB
a01	Der lappen liegt auf dem Eisschrank.
a02	Das will sie am Mittwoch abgeben.
a04	Heute abend könnte ich es ihm sagen.

Data yang diambil dibagi menjadi dua kategori yaitu data pelatihan dan data pengujian. Data pelatihan digunakan untuk mendapatkan pola masing-masing emosi. Sedangkan data pengujian digunakan pada proses pengujian untuk mengetahui hasil kerja sistem dalam melakukan identifikasi (Surya, Fadlil and Yudhana, 2017). Data ucapan yang digunakan sebagai data pelatihan tidak digunakan dalam pengujian. Data yang diambil adalah tiga ucapan dari setiap emosi dari pengucapan aktor yang berbeda dan kalimat yang berbeda juga. Vektor ciri yang didapat dari data pelatihan dicocokkan dengan data pengujian, satu per satu dan diukur berdasarkan nilai jarak (*distance*) yang paling minimum (Yudhana *et al.*, 2017).

## 2.2. Pemrosesan Awal

Data yang digunakan dalam penelitian ini diambil dari *dataset* Berlin EMO-DB. Rekaman ucapan berupa *file* dalam bentuk \*.wav. Maka pada pemrosesan awal ini dilakukan untuk mengubah atau membaca *file* audio menjadi nilai-nilai dalam bentuk matriks, dimana panjang matriks ditentukan dari panjangnya rekaman audio dan digunakan sebagai data *input*.

Pada pemrosesan awal ini juga data terlebih dahulu dinormalisaikan agar data rekaman ucapan menjadi sama rentang nilai amplitudonya. Normalisasi dilakukan dengan cara membagi nilai data masukan dibagi dengan nilai maksimal dari data masukan tersebut. Rumusnya dapat dilihat pada persamaan (1).

$$X_{normalisasi} = \frac{x_{data}}{\max(abs(x_{data}))} \quad (1)$$

Keterangan

$X_{normalisasi}$  = hasil data normalisasi

$X_{data}$  = data masukan suara ucapan

Normalisasi ini digunakan pada awal sebelum ekstraksi ciri dan sesudah ekstraksi ciri. Setelah itu dilakukan juga normalisasi data agar menjadi rentang nilai 0 sampai 1.

## 2.3. Ekstraksi Ciri

Ekstraksi ciri yang digunakan dalam penelitian ini menggunakan Metode Linear Predictive Coding (LPC). Tahapan umum dalam LPC (Chaudhari and Alex, 2016) ditunjukkan pada Gambar 2. Penjelasan dari tahapan-tahapan LPC sebagai berikut:

### a. Pre Emphasis

Tahapan awal dalam LPC ini bertujuan untuk mengurangi *noise* suara yang tidak dibutuhkan. Ketika sebuah sinyal memiliki perbedaan yang signifikan dengan nilai sinyal di sekitarnya, sinyal ini dianggap sebagai *noise*. Pada tahapan ini, diterapkan pada persamaan (2) pre-emphasis :

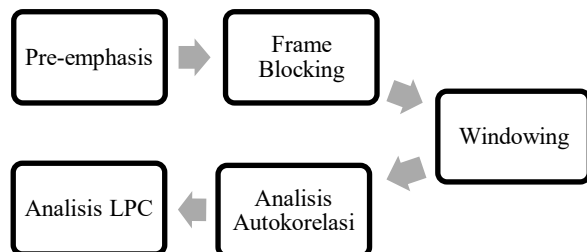
$$y(n) = s(n) - \alpha s(n-1) \quad (2)$$

Keterangan

$y$  = Data setelah pre-emphasis

$s$  = Data sebelum pre-emphasis

$\alpha$  = Nilai dasar sebagai tolak ukur signifikansi



Gambar 2. Tahapan LPC

Nilai default  $\alpha$  yang digunakan dalam LPC adalah 0.95 (Irmawan *et al.*, 2014). Nilai dari  $y(n)$  ditambahkan dengan nilai  $s(n)$  sehingga didapatkan nilai baru yang berkurang *noisenya*.

#### b. Frame Blocking

Dalam tahap ini, sinyal ucapan yang sudah melewati proses pre-emphasis dipotong-potong per-*frame* sepanjang 20ms di setiap *timelapse*.

#### c. Windowing

Dalam tahap Frame Blocking, terdapat kemungkinan adanya data sampel yang terputus. Efek tersebut dapat dikurangi dengan melakukan *windowing*. Hamming Window adalah model *window* yang sering digunakan untuk model LPC yang ditampilkan pada persamaan (3).

$$w(n) = 0,54 - 0,46\cos(2\pi n/N - 1) \quad (3)$$

Dimana  $N$  adalah estimasi jumlah *frame* yang ada di masing-masing sampel.

#### d. Analisis Autokorelasi

Pada tahapan ini sinyal yang telah melalui tahap *windowing* dikorelasikan antar *frame*. Bagian sinyal yang telah melalui tahap *Windowing*, dibentuk autokorelasinya dengan persamaan (4).

$$r(m) = \sum_{n=0}^{N-1-m} \tilde{x}(n) \cdot \tilde{x}(n+m) \quad (4)$$

$$m = 0,1,2,\dots,p$$

Nilai autokorelasi tertinggi pada  $m=p$  adalah orde dari analisa LPC dimana  $m$  adalah hasil matriks. Orde LPC biasanya menggunakan angka 8 sampai 16 (Thiang, 2005). Nilai ke 0 merupakan nilai energi dari *frame* data masukan.

#### e. Analisis LPC

Data masukan yang sudah diautokorelasi selanjutnya dianalisa LPC nya. Tahap analisis LPC ini mengubah  $p+1$  hasil dari autokorelasi setiap *frame* menjadi satu set parameter LPC yaitu *LPC coefficient*, *reflection coefficient*, dan *log area ratio coefficient* (Rahmawanthi, Raharjo and Rusdinar, 2019). Penelitian ini akan menggunakan orde LPC 16.

### 2.4. Klasifikasi menggunakan metode Euclidean Distance

Hasil dari proses ekstraksi ciri LPC digunakan sebagai *input* pelatihan dalam sistem pengenalan emosi. Tahap pelatihan ini dijadikan sebagai referensi ciri yang digunakan pada proses pengujian data. Data pengujian yang digunakan dalam penelitian ini sebanyak 12 data untuk emosi marah, sedih, bahagia, netral, dan bosan dari tiga kalimat berbeda dan aktor yang berbeda.

Hasil dari tahap pengujian ini kemudian dihitung jaraknya menggunakan metode Euclidean Distance. Metode Euclidean Distance adalah salah satu metode jarak yang sering digunakan dalam penelitian untuk pengenalan pola. Metode ini digunakan untuk mengukur kemiripan nilai ekstraksi ciri pada data pengujian ke seluruh data pelatihan yang lain. Proses ini dapat diukur dengan menggunakan rumus Pythagoras yang dapat dilihat pada persamaan (5).

$$d = \sqrt{\sum_i^n (x_i - y_i)^2} \quad (5)$$

Keterangan

$d$  = jarak Euclidian

$W_i$  = vektor bobot ke- $i$

$X_i$  = vektor *input* ke  $X_i$

## 3. HASIL DAN PEMBAHASAN

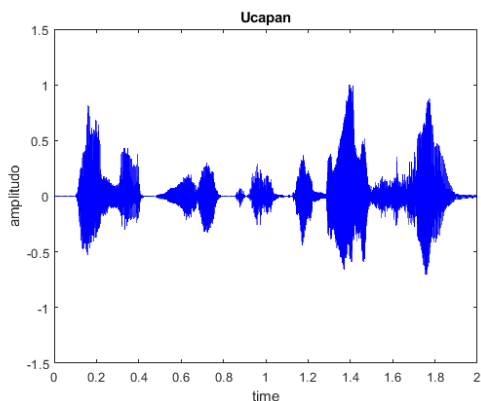
### 3.1. Tahap Pemrosesan Awal

Tahap pemrosesan awal yaitu merubah *file* rekaman ucapan menjadi bentuk matriks. Rekaman ucapan Berlin EMO-DB, menggunakan tiga data untuk pelatihan, yang terdiri dari tiga kalimat berbeda dan aktor yang berbeda baik laki-laki maupun perempuan. *File* rekaman ucapan dari bentuk *file* \*.wav diubah menggunakan fungsi di Matlab yaitu  $[x] = \text{audioread}(\text{wav\_file})$ ; fungsi tersebut menghasilkan nilai-nilai matriks yang diperoleh dari waktu ( $x$ ) panjang nya rekaman. Proses data rekaman menjadi matriks dilakukan pada data pelatihan dan data pengujian. Hasil dari fungsi ini kemudian yang digunakan untuk ekstraksi ciri. *Source code* yang digunakan untuk mengubah data rekaman ucapan suara emosi menjadi matriks dapat dilihat pada Source Code 1.

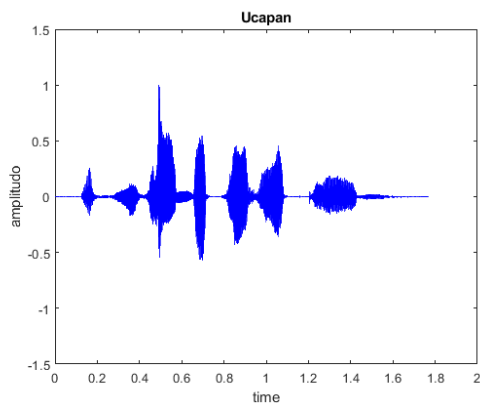
```
for i = 1:3
    data{i} = audioread(['marah\'
num2str(i)'.wav]);
end;
```

Source Code 1 Mengubah Data Rekaman Ucapan Suara Emosi menjadi Matriks

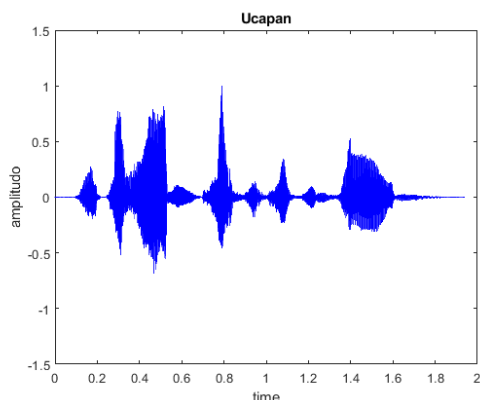
Source Code 1 mengubah data rekaman ucapan emosi diambil dari tiga data untuk lima emosi. Hasil data rekaman ucapan emosi dapat dilihat Gambar 3 sampai Gambar 5 Gambar yang disajikan dimana sumbu  $x$  adalah waktu dan sumbu  $y$  adalah amplitudo. Data yang ditampilkan adalah data rekaman dari emosi marah.



Gambar 3. Sinyal Rekaman Emosi Marah data marah1



Gambar 4. Sinyal Rekaman Emosi Marah data marah2



Gambar 5. Sinyal Rekaman Emosi Marah data marah3

Tiga data dari setiap emosi dinormalisasi menggunakan persamaan (1).

### 3.2. Hasil Ekstraksi Ciri Emosi

Proses dari ekstraksi ciri untuk data pelatihan menggunakan tiga data rekaman ucapan emosi. Sedangkan data pengujian menggunakan 12 data rekaman ucapan emosi. Proses ekstraksi ciri menggunakan metode LPC. Penelitian ini menggunakan fungsi bawaan dari Matlab yaitu  $lpc()$ . Umumnya, fungsi ini ditulis  $a = lpc(x, p)$ , dimana  $a$  adalah variabel untuk menyimpan nilai dari koefisien LPC,  $x$  adalah data rekaman yang sudah dinormalisasi dan  $p$  adalah orde LPC. Orde LPC yang digunakan yaitu orde 16 yang akan menghasilkan  $m = 16+1$ , seperti penjelasan pada tahapan LPC di bagian nilai autokorelasi. *Source code* ekstraksi ciri yang digunakan untuk data latih dapat dilihat pada Source Code 2.

```
for i = 1:3
    x1{i} = lpc(data1{i}, 16);
end;
```

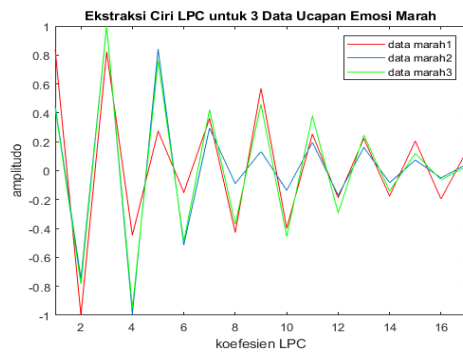
Source Code 2 Ekstraksi Ciri

Source Code 2 digunakan untuk data pelatihan sebanyak tiga data, dimana  $x1$  adalah variabel untuk menyimpan koefisien LPC,  $data1$  adalah data matriks dari rekaman ucapan dan 16 adalah orde LPC. Tiga data pelatihan tersebut menghasilkan nilai koefisien dalam bentuk matriks berukuran  $1 \times 17$ . Tabel 3 adalah salah satu contoh hasil dari ekstraksi ciri LPC untuk pada emosi marah.

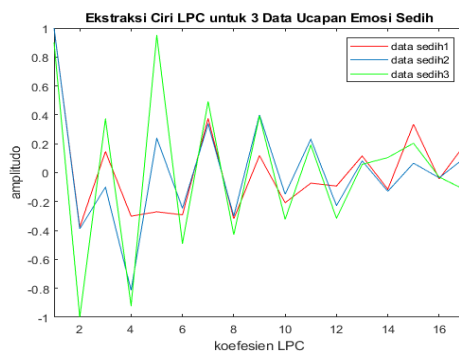
Tabel 3. Koefesien LPC Emosi Marah

LPC: marah1	LPC: marah2	LPC: marah3
0.7073	0.4799	0.3931
-1	-0.8563	-0.8165
0.7166	1	1
-0.4662	-0.9339	-0.9085
0.3969	0.7416	0.7068
-0.3839	-0.5867	-0.4160
0.4481	0.5217	0.2796
-0.4094	-0.4378	-0.2605
0.3838	0.4236	0.3207
-0.3045	-0.3310	-0.4180
0.2398	0.2656	0.4078
-0.2669	-0.2453	-0.3134
0.2652	0.2247	0.2132
-0.1718	-0.1928	-0.1116
0.1347	0.1442	0.0245
-0.0338	-0.0693	-0.0232
0.0214	0.0316	0.0181

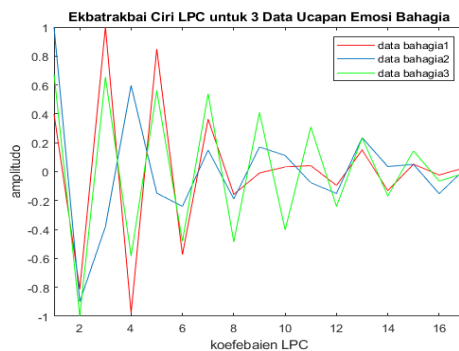
Hasil ekstraksi ciri dari data pelatihan masing-masing emosi, nilai koefisien LPC nya dibuat plot untuk mengetahui ciri masing-masing emosi. Hasil pola ekstraksi ciri untuk data pelatihan dapat dilihat pada Gambar 6 sampai Gambar 10.



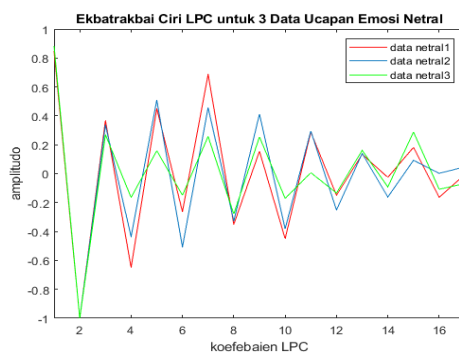
Gambar 6. Pola Ciri Emosi Marah



Gambar 7. Pola Ciri Emosi Sedih



Gambar 8. Pola Ciri Emosi Bahagia

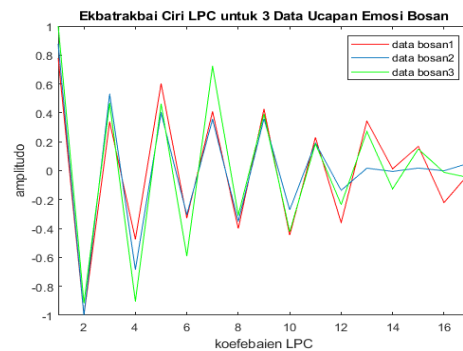


Gambar 9. Pola Ciri Emosi Netral

Gambar 6 sampai Gambar 10 adalah ekstraksi ciri metode LPC dimana sumbu  $x$  adalah koefisien LPC dan sumbu  $y$  adalah amplitudo.

Tiga data pelatihan dari masing-masing emosi tersebut dinormalisasi menggunakan persamaan (1) dan dinormalisasikan lagi menjadi rentang nilai antara 0 sampai 1. Hasil dari normalisasi kemudian

diambil nilai *mean* nya yang digunakan sebagai parameter identifikasi (Prasetio, Kurniawan and Ichsan, 2017).



Gambar 10. Pola Ciri Emosi Bosan

Setelah nilai *mean* ditemukan dari tiga data latih masing-masing emosi, langkah selanjutnya adalah mencari nilai jarak antara data pelatihan dengan data pengujian dengan menggunakan metode Euclidean Distance. Fungsi dari persamaan (5) diimplementasikan ke Matlab dengan Source Code 3.

```
function d = euclidean (x,y)
    d = sqrt(sum((x-y).^2));
end;
```

Source Code 3 Fungsi Jarak

Semakin kecil jarak yang diperoleh ( $d$ ), maka semakin besar kemiripan antara dua objek. Sebaliknya, jika semakin besar jarak yang diperoleh ( $d$ ), maka semakin kecil juga kemiripan antara dua objek (Deza & Deza, 2009).

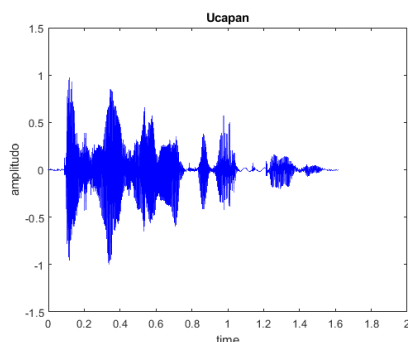
### 3.3. Hasil Identifikasi

Hasil identifikasi adalah proses untuk menentukan rekaman ucapan yang di-input-kan, masuk ke dalam kelas emosi marah, sedih, bahagia, netral, atau bosan. Penentuan kelas ini berdasarkan *template* atau nilai yang dihasilkan dari data pelatihan. Proses identifikasi data pengujian dilakukan dengan cara yang sama dengan data pelatihan yaitu dilakukan pengubahan rekaman menjadi menjadi nilai matriks, normalisasi sinyal, dan ekstraksi ciri LPC.

Salah satu contoh proses identifikasi dilakukan pada data sedih4. Sinyal rekaman data sedih4 dapat dilihat pada Gambar 11, dimana sumbu  $y$  adalah amplitudo dan sumbu  $x$  adalah *time* (waktu).

Data sedih4 dinormalisasi dengan persamaan (1). Selanjutnya proses ekstraksi ciri menggunakan LPC yang menghasilkan nilai koefisien LPC yang dapat dilihat pada Tabel 4.

Tabel 4 menampilkan hasil koefisien LPC dalam matriks 1x17 yang merupakan hasil dari ekstraksi ciri LPC. Nilai koefisien LPC pada Tabel 4 diuji kemiripannya dengan mencari jarak *euclidean* minimum dari kelima kelas.



Gambar 11. Sinyal digital Emosi Sedih data sedih4

Tabel 4. Koefisien LPC data sedih4

LPC data sedih4
1
-0.5935
-0.0372
-0.7118
0.3096
-0.1479
0.1150
-0.1941
0.3366
-0.0287
0.1436
-0.3302
0.0234
0.0760
0.2107
-0.1186
0.0103

Nilai *mean* dari emosi data pelatihan dijadikan sebagai *template* untuk menentukan kelas emosi menggunakan metode Euclidean Distance yang telah ditentukan. *Source code* untuk mencari jarak antara data pelatihan dengan data pengujian dilihat Source Code 4.

```
d1 = euclidean(px,p1);
d2 = euclidean(px,p2);
d3 = euclidean(px,p3);
d4 = euclidean(px,p4);
d5 = euclidean(px,p5);
```

Source Code 4 Mencari Jarak Data

Source Code 4 digunakan untuk menghitung jarak ( $d1-d5$ ) nilai *mean* pada masing-masing emosi, dimana  $px$  adalah nilai *mean* data pengujian dan  $p1-p5$  adalah nilai *mean* untuk masing-masing emosi. Langkah selanjutnya adalah identifikasi untuk pengambilan keputusan hasil pengenalan emosi. Identifikasi yang dilakukan menggunakan Source Code 5.

Variabel *min* digunakan untuk mencari nilai minimum di semua kelas emosi. Nilai minimum ini digunakan untuk mengidentifikasi data pengujian sebagai salah satu emosi dari lima kelas emosi. Hasil identifikasi dari data emosi sedih4 dapat dilihat pada Tabel 5.

```
%pengambilan keputusan hasil
pengenalan
min = min ([d1 d2 d3 d4 d5])
if (d1==min)
    class = 'marah'
elseif (d2==min)
    class = 'sedih'
elseif (d3==min)
    class = 'bahagia'
elseif (d4==min)
    class = 'netral'
elseif (d5==min)
    class = 'bosan'
```

Source Code 5 Identifikasi Emosi

Tabel 5. Identifikasi Emosi untuk data sedih4

	Nilai Mean	Nilai Min	Emosi
p1	0.5527	d1 0.9036	Marah
p2	<b>0.4460</b>	d2 <b>0.8587</b>	<b>Sedih</b>
p3	0.5536	d3 0.9049	Bahagia
p4	0.5658	d4 0.9235	Netral
p5	0.5547	d5 0.9065	Bosan

Tabel 5 menampilkan hasil identifikasi emosi untuk data sedih4, dimana kolom pertama menampilkan hasil nilai *mean* sebagai parameter ciri, kolom kedua menampilkan nilai *min* untuk mencari nilai kemiripan emosi, dan kolom ketiga menampilkan hasil identifikasi emosi. Hasil dari pengujian data sedih4, teridentifikasi sebagai emosi sedih. Karena nilai jarak minimum terdekat dari parameter nilai *mean* adalah  $d2$ , dimana  $d2$  adalah kelas emosi sedih.

Data pengujian diidentifikasi seperti langkah diatas. Hasil data pengujian yang dilakukan terhadap 60 data emosi, dapat dilihat pada Tabel 6.

Tabel 6. Hasil Pengujian

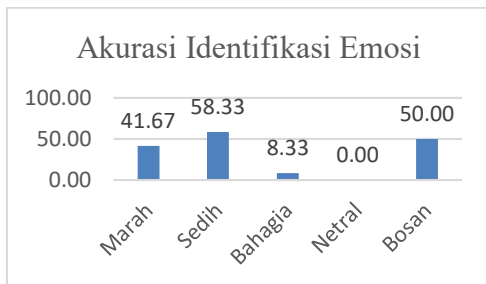
Emosi	M	S	B	N	Bo
Marah (M)	5	3	0	0	4
Sedih (S)	4	7	0	0	1
Bahagia (B)	1	0	1	1	9
Netral (N)	2	0	8	0	2
Bosan (Bo)	2	2	2	0	6
<b>Akurasi</b>	$\frac{19 \times 100}{60} = 31,67\%$				

Pada Tabel 6 emosi sedih dikenali sebanyak 7 data dengan akurasi 58,33%, emosi bosan 6 data dengan akurasi 50%, emosi marah 5 data dengan akurasi 41,67%, emosi bahagia 1 data dengan akurasi 8,33% dan emosi bosan tidak dapat dikenali sebagai kelas emosinya sendiri. Grafik dari akurasi emosi dapat dilihat pada Gambar 12.

#### 4. UCAPAN TERIMA KASIH

Riset ini didukung dan didanai oleh Kemenristek Dikti dengan Nomer Kontrak : PTM-026/SKPP.TT/LPPM UAD/111/2019.





Gambar 12. Grafik Akurasi Emosi

## 5. KESIMPULAN

Pengenalan emosi menggunakan LPC dan metode Euclidean Distance hanya mampu mendeteksi emosi marah, sedih dan netral. Sedangkan emosi bahagia dan bosan tidak dikenali atau terdeteksi sama sekali. Sistem pengenalan emosi pada penelitian ini menggunakan nilai *mean* sebagai parameter pengujian sistem. Sistem ini menghasilkan akurasi tertinggi pada emosi sedih 58,33%, emosi netral 50%, emosi marah 41,67% dan emosi bahagia 8,33%. Sedangkan untuk emosi bosan tidak dapat terdeteksi.

Penggunaan Metode LPC menggunakan parameter *mean* dan klasifikasi menggunakan metode Euclidean Distance memberikan hasil yang kurang baik pada penelitian ini karena akurasi rata-rata hanya sebesar 31,67% dari hasil pengujian untuk identifikasi semua emosi. Metode LPC dan Euclidean Distance menghasilkan ciri berupa jarak. Ciri yang digunakan pada pengenalan ucapan memberikan korelasi yang kecil pada metode konvensional (metode jarak Euclidean).

Data suara yang digunakan dengan kalimat, aktor, umur dan aksen yang berbeda kemungkinan dapat mempengaruhi dalam pengenalan emosi, maka dari itu ekstraksi ciri dalam pengenalan pola ucapan emosi manusia sangat penting. Hasil akurasi pada penelitian ini masih sangat kecil dan dapat ditingkatkan dengan menggunakan ekstraksi ciri yang lain seperti prosidis, spektral, dan kualitas suara, penggunaan parameter *max*, *min*, *mean*, *median*, *kurtosis* dan *skewenes*. Selain itu penggunaan metode klasifikasi juga dapat mempengaruhi hasil pengenalan emosi.

## DAFTAR PUSTAKA

- ASWARI, P. and DIANA, N. E. (2016) 'Identifikasi Emosi Berdasarkan Action Unit Menggunakan Metode Bézier Curve', *Sinergi*. Mercu Buana University, 20(1), pp. 74–84. doi: 10.22441/sinergi.2016.1.010.
- CHAMOLI, A., SEMWAL, A. and SAIKIA, N. (2017) 'Detection Of Emotion In Analysis Of Speech Using Linear Predictive Coding Techniques (L.P.C)', in *Proceedings of the International Conference on Inventive Systems and Control, ICISC 2017*, pp. 1–4. doi: 10.1109/ICISC.2017.8068642.
- CHAUDHARI, P. R. and ALEX, J. S. R. (2016) 'Selection of Features for Emotion Recognition from Speech', *Indian Journal of Science and Technology*, 9(39), pp. 1–5. doi: 10.17485/ijst/2016/v9i39/95585.
- DEWI, I. A., ZULKARNAIN, A. and LESTARI, A. A. (2018) 'Identifikasi Suara Tangisan Bayi menggunakan Metode LPC dan Euclidean Distance', *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, 6(1), p. 153. doi: 10.26760/elkomika.v6i1.153.
- DEZA, M. M. and DEZA, E. (2009) 'Encyclopedia of distances', in *Encyclopedia of distances*. Springer, pp. 1–583.
- GUMELAR, A. B. *et al.* (2019) 'Human Voice Emotion Identification Using Prosodic and Spectral Feature Extraction Based on Deep Neural Networks', in *2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH)*. IEEE, pp. 1–8.
- IRMAWAN *et al.* (2014) 'Pengenalan Kata dengan Metode Linear Predictive Coding dan Jaringan Syaraf Tiruan Pada Mobile Robot', in *Conference on Information Technology and Electrical Engineering*, pp. 139–144.
- KUMAR, Y. and MAHAJAN, M. (2019) 'Machine Learning Based Speech Emotions Recognition System', *International Journal of Scientific and Technology Research*, 8(7), pp. 722–729.
- LANJEWAR, R. B., MATHURKAR, S. and PATEL, N. (2015) 'Implementation and Comparison of Speech Emotion Recognition System Using Gaussian Mixture Model (GMM) and K-Nearest Neighbor (K-NN) Techniques', *Procedia Computer Science*. Elsevier, 49, pp. 50–57. doi: 10.1016/J.PROCS.2015.04.226.
- LIKITHA, M. S. *et al.* (2017) 'Speech based human emotion recognition using MFCC', in *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pp. 2257–2260. doi: 10.1109/WiSPNET.2017.8300161.
- PRASETIO, B. H., KURNIAWAN, W. and ICHSAN, M. H. H. (2017) 'Pengenalan Emosi Berdasarkan Suara Menggunakan Algoritma HMM', 4(3), pp. 168–172. doi: 10.25126/jtiik.201743339.
- RAHMAWANTHI, I., RAHARJO, J. and RUSDINAR, A. (2019) 'Deteksi Suara Manusia dalam Keadaan Emosi dengan menggunakan Linear Predictive Coding (LPC) dengan Klasifikasi Coarse To Fine Search (CFS) Berbasis Pengolahan Data', *eProceedings of Engineering*, 6(1).
- SURYA, R. A., FADLIL, A. and YUDHANA, A. (2017) 'Ekstraksi Ciri Metode Gray Level Co-Occurrence Matrix (GLCM) dan Filter Gabor untuk Klasifikasi citra Batik Pekalongan',



*JURNAL INFORMATIKA : Jurnal Pengembangan IT*, 2(2), pp. 23–26. Available at:

<http://ejournal.poltektegal.ac.id/index.php/informatika/article/view/520>.

- THIANG, H. S. (2005) ‘Sistem Pengenalan Kata dengan Menggunakan Linear Predictive Coding dan Nearest Neighbor Classifier’, *Jurnal Teknik Elektro*, 5(1), pp. 19–24.
- WANG, K. *et al.* (2015) ‘Speech Emotion Recognition Using Fourier Parameters’, *IEEE Transactions on Affective Computing*, 6(1), pp. 69–75. doi: 10.1109/TAFFC.2015.2392101.
- YUDHANA, A. *et al.* (2017) ‘Pengolahan Sinyal Fleks Sensor pada Sarung Tangan Pintar Penerjemah Bahasa Isyarat’, in *Annual Research Seminar (ARS)*, pp. 296–299.

*Halaman ini sengaja dikosongkan*