

MENENTUKAN EKSPRESI WAJAH DENGAN METODE K-MEANS KLUSTERING

Jasman Pardede¹⁾ Irma Amelia Dewi²⁾ Ade Bambang Kurnia³⁾

Teknik Informatika Institut Teknologi Nasional
Jl. PH. H. Mustofa No.23, Bandung, Jawa Barat, Indonesia
jasman@itenas.ac.id¹⁾, Irma@itenas.ac.id²⁾, adebmbng@gmail.com³⁾

Abstract

To determine human facial expression be required data set as data test. In this research, data set use Japanese Female Facial Expression (JAFFE). To detect the face area of a picture uses GMS and then transformed to be data that need to calculate with K-Means. The data that calculated are probability of the eyes opening, angle of lips, and array pixel by size 32 x 80. By the calculation of Euclidian Distance between all of data set and input result the similarity for each centroid. The precision of research is 87.68%. The precision to women subject is 90,28% and to men subject is 88,61%. The precision using of JAFFE data is 88.2%. The precision the subject who has mustache and beard is only 40%.

Key word: K-Means Clustering, JAFFE, Facial Expression.

1. Pendahuluan

1.1 Latar Belakang

Musik merupakan salah satu hiburan seluruh manusia. Musik dapat memberikan dampak positif ataupun negatif kepada pendengarnya[1]. Oleh karena itu, musik dapat dimanfaatkan sebagai *trigger* untuk memberikan ekspresi atau emosi. Secara tidak sadar, musik dapat menciptakan kondisi nyaman atau tidak nyaman untuk pendengarnya. Penilaian terhadap musik dapat membantu pendengarnya untuk memilih musik yang sesuai.

Ekspresi wajah manusia dapat diklasifikasikan menjadi 6 tipe, diantaranya adalah ekspresi senang, sedih, marah, terkejut, jijik, dan netral. Untuk mengetahui ke-6 ekspresi tersebut, diperlukan 4 data klasifikasi yang harus dipenuhi. Ke-4 data yang dimaksud adalah sudut kemiringan di area mulut, ukuran lebar mata, naik atau turunnya rahang, dan kerutan pada area mata[2].

Penelitian ini mampu mengetahui 3 ekspresi, diantaranya ekspresi senang, sedih dan netral. Data yang diolah untuk mendapatkan ekspresi hanya dipenuhi 2 data, yaitu sudut kemiringan di area mulut, dan ukuran lebar mata. *Japanese Female Facial Expression (JAFFE)*[3] merupakan *dataset* yang digunakan untuk proses klasifikasi. Terdapat 68 data latih yang terdiri dari 3 ekspresi.

Pada penelitian ini menggunakan musik sebagai *trigger* agar pendengarnya memberikan ekspresi. Lalu, ekspresi tersebut diambil gambarnya dengan menggunakan kamera depan *smartphone android* dan akan diklasifikasikan menggunakan K-Means klustering dengan JAFFE.

1.2 Rumusan Masalah

Rumusan masalah pada penilitan ini adalah bagaimana mendapatkan ekspresi wajah manusia dengan menggunakan K-Means klustering.

1.3 Tujuan

Tujuan dari penelitian yang dilakukan adalah mengimplementasikan K-Means klustering dalam menentukan ekspresi wajah terhadap suasana hati menggunakan *smartphone* berbasis *android*.

2. Metodologi

2.1. K-Means Klustering

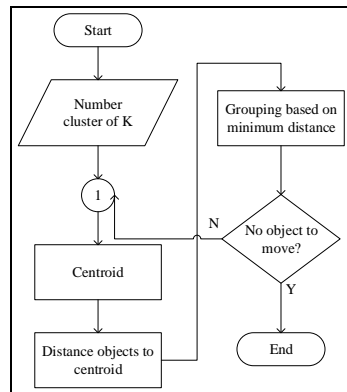
Metode K-Means sudah sangat umum digunakan untuk melakukan pengklasifikasian. Dengan membagi keseluruhan data kedalam beberapa kluster dan memilih secara acak 1 data dari masing-masing kluster untuk dijadikan *centroid*[5].

Seluruh data termasuk data masukan diproses dan dihitung ke masing-masing *centroid* menggunakan persamaan *Euclidian Distance* pada persamaan 1.

$$D(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{in} - x_{jn})^2} \dots\dots(1)$$

Hasil dari persamaan *Euclidian Distance* dibandingkan untuk setiap *centroid*-nya menggunakan metode *Nearest Neighbor* pada persamaan 2 dimana nilai yang mendekati 0 adalah nilai yang diambil. Setelah mendapatkan nilai dari *Nearest Neighbor*, data tersebut dapat diklasifikasikan di kluster yang dituju.

$$Mean = \frac{(x_{i1} + x_{i2} + \dots + x_{in})}{n}, \frac{(y_{i1} + y_{i2} + \dots + y_{in})}{n} \dots(2)$$



Gambar 1. Algoritma K-Means

Proses K-Means ini tidak cukup dilakukan dengan 1 iterasi, diperlukan beberapa iterasi hingga tidak ada data yang berpindah kluster. Pada penelitian ini, proses K-Means dilakukan 2 iterasi. Untuk penentuan kesimpulan ekspresi pada data masukan, dilakukan dengan melihat dimana data masukan tersebut berada di kluster mana pada saat iterasi terakhir, seperti yang dinyatakan pada Gambar 1.

2.2. Google Mobile Vision (GMS)

GMS adalah komponen pendukung yang diberikan oleh Google untuk mempermudah dalam pendeteksian wajah menggunakan kamera. GMS sudah tertanam pada seluruh perangkat android yang memiliki aplikasi Google Play Service. GMS dapat memberikan koordinat wajah dalam hitungan mili detik. GMS dapat mendeteksi lebih dari 1 wajah dalam 1 frame.

Selain mendeteksi wajah, GMS dapat memberikan nilai landmark pada area wajah yang berfungsi sebagai pemberi ciri untuk setiap komponen di area wajah. Terdapat 8 landmark yang disediakan oleh GMS dari mata hingga ujung bibir. GMS juga memberikan nilai probabilitas senyuman, nilai probabilitas bukaan mata kanan dan kiri.

2.3. Perubahan Citra Bibir menjadi Array Piksel

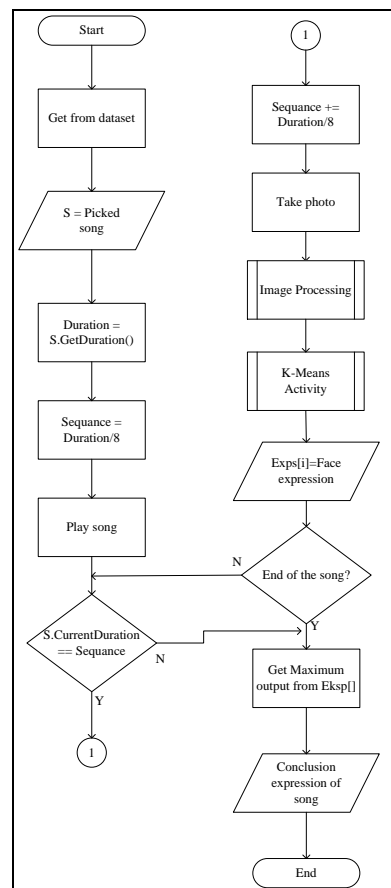
Perubahan citra bibir dilakukan saat data-data dari GMS selesai diolah. Data yang digunakan adalah koordinat landmark pada area bibir. Dilakukan proses scaling antara ukuran citra dengan ukuran citra wajah untuk mendapatkan ukuran bibir. Margin ditambahkan pada citra bibir agar area pendeteksian tidak pada ukuran yang terlalu kecil. Cropping dilakukan pada ukuran 80x32 piksel menyesuaikan dengan ukuran dari citra yang dimiliki oleh dataset. Setelah citra bibir yang berukuran 80x32 didapat, dilakukan proses pengabuan atau grayscaling agar perubahan data dari 3 dimensi array menjadi 1 dimensi array. Hasil akhir yang didapat adalah string yang berisikan array piksel yang telah dilakukan penggabungan data dan dipisahkan menggunakan “;”.

3. Pembahasan

3.1 Diagram Alir Aplikasi

Setelah memutar lagu, ada proses dimana durasi dibagi menjadi 8 bagian. Setiap bagian atau sequence, aplikasi mengambil gambar dari kamera depan handphone untuk dideteksi ekspresi wajahnya. Setelah 8 sequence selesai, aplikasi menghitung jumlah total ekspresi yang terdeteksi dan memberikan kesimpulan ekspresi terhadap lagu yang didengarkan.

Secara bersamaan pada saat handphone mengambil gambar, terdapat 2 kemungkinan untuk melanjutkan prosedur sistem. Pertama, wajah terdeteksi oleh face-detection dari GMS, dimana pada opsi pertama ini menjalankan prosedur selanjutnya yaitu pendeteksian ekspresi menggunakan K-Means klustering. Sedangkan jika tidak ada wajah di dalam frame atau wajah tidak terdeteksi, tidak ada prosedur selanjutnya dan hanya memberikan nilai null pada sequence tersebut seperti yang dinyatakan pada Gambar 2.



Gambar 2. Diagram alir aplikasi

3.2 Penggunaan K-Means Klustering

K-Means klustering yang dilakukan adalah mengambil 9 contoh data dari 68 dataset JAFFE, lalu memberikan 1 data masukan untuk ditentukan ekspresi wajah. Data untuk array piksel bibir pun diminimalisir agar mempermudah perhitungan yang

dilakukan. Tabel 1 menunjukan contoh data sebagai pengganti *dataset* JAFFE, dengan data masukan seperti pada Gambar 3.

Tabel 1. Dataset

Id	Nama Ekspresi	Smiling Probability	Right-Eye Open Probability	Left-Eye Open Probability	Lips Array Pixel
1	Senang	0.9558	0.6170	0.5547	150,170,180,130,100,90,90,80,130,200
2	Senang	0.9677	0.9918	0.9894	210,190,190,170,120,140,150,170,200,190
3	Senang	0.9706	0.9963	0.9940	150,130,130,120,100,140,160,170,200,230
4	Sedih	0.0068	0.9915	0.9969	90,80,100,180,170,190,120,0,110,100,100
5	Sedih	0.0556	0.5715	0.5981	70,70,140,150,150,160,100,100,80,50
6	Sedih	0.1669	0.9915	0.9920	80,100,150,160,200,180,180,120,100,120
7	Netral	0.1722	0.9961	0.9936	210,200,220,230,190,210,200,210,190,220
8	Netral	0.2352	0.9939	0.9956	180,180,190,170,180,180,160,170,190,180
9	Netral	0.1431	0.9965	0.9960	50,50,50,60,70,80,90,50,60,70



Gambar 3. Foto masukan

Tabel 2 adalah data masukan yang didapat dari Gambar 3.

Tabel 2. Data Masukan

Smiling Probability	Right Eye Open Probability	Left Eye Open Probability	Lips Array Pixel
0.9675	0.9873	0.8312	200,180,170,150,150,160,180,190,200,210

Langkah pertama untuk mendapatkan ekspresi wajah yang tepat adalah memilih jumlah kluster dan *centroid* dari *dataset*. Karena pada *dataset* memiliki 3 ekspresi maka, *dataset* tersebut dibagi menjadi 3 kluster dimana *centroid* 1 adalah data dengan id 1, *centroid* 2 adalah data dengan id 4 dan *centroid* 3 adalah data dengan id 7. Persamaan 1 digunakan seluruh data dengan masing-masing *centroid*. Sebagai contoh, *centroid* 1 dibandingkan dengan data masukan. Perhitungan dilakukan sesuai dengan persamaan 1 menghasilkan nilai 0.659253317.

Selanjutnya, data masukan diperhitungkan dengan data *centroid* 2 dan 3. Kemudian dengan menggunakan Persamaan 2 atau *Nearest Neighbor*, data masukan dikelompokkan pada kluster dimana dari ke-3 perhitungan dengan *centroid* memiliki nilai yang mendekati 0. Tabel 3 merupakan pengelompokan data untuk kluster 1 atau ekspresi senang.

Tabel 3. Kluster untuk Ekspresi Senang

Data id	Smiling Probability	Right Eye Open Probability	Left Eye Open Probability	Array Pixel Bibir (Sum mary)
1	0.9558	0.6170	0.5547	1320
2	0.9677	0.9918	0.9894	1730
3	0.9706	0.9963	0.9940	1530
Input	0.9675	0.9873	0.8312	1790
Centroid baru	0.9654	0.8981	0.8423	1592.5

Centroid baru ditentukan dengan mencari rata-rata dari setiap kolomnya untuk digunakan pada iterasi ke-2. Hal yang sama dilakukan pada *centroid* lain untuk mendapatkan nilai *centroid* baru untuk iterasi ke-2. *Centroid* baru pada Tabel 3 menjadi *centroid* 1 untuk iterasi selanjutnya. Perhitungan *Euclidean Distance* dilakukan kembali pada *dataset* terhadap *centroid* baru. Tabel 4 menunjukkan data masukan yang telah dilakukan perhitungan *Euclidean Distance*.

Tabel 4. Hasil Iterasi ke-2 untuk Data Masukan

Data id	Centroid 1	Centroid 2	Centroid 3
Input	0.2170	1.1310	0.7936

Data dari Tabel 4 menunjukkan bahwa nilai terkecil untuk data masukan berada di *centroid* 1 yang berarti data masukan merupakan bagian dari kluster 1 atau ekspresi senang. Disimpulkan, data masukan adalah ekspresi senang pada iterasi ke-2.

3.3 Hasil Pengujian

Terdapat 18 skenario pengujian yang telah diujikan. Setiap skenario pengujian memiliki minimal 15 naracoba. Dan hasil dari ke-18 skenario adalah aplikasi mendapat nilai keberhasilan sebesar 84.78% dengan keberhasilan pada penguji perempuan sebesar 90.28% dan keberhasilan pada penguji pria sebesar 88.61%. Tabel 5 menunjukkan rincian hasil pengujian yang telah dilakukan.

Tabel 5. Hasil Pengujian

No	Pengujian	P	W	Result
1	Pendeteksian ekspresi normal tanpa musik	71.4%	87.5%	86.7%
2	Pendeteksian ekspresi normal dengan musik	85.7%	87.5%	86.7%
3	Tes pendeteksian ekspresi terhadap dataset	-	-	88.2%
4	Tes pendeteksian ekspresi senang	71.4%	87.5%	80%
5	Tes pendeteksian ekspresi sedih	85.7%	75%	80%
6	Tes pendeteksian ekspresi netral	100%	100%	100%
7	Kombinasi ekspresi dalam satu lagu 2 ekspresi	100%	100%	100%
8	Kombinasi ekspresi dalam satu lagu 3 ekspresi	100%	100%	100%
9	Kondisi pada saat tes (normal)	100%	87,5%	93.4%
10	Kondisi pada saat tes (low light)	71,4%	75%	73.4%
11	Kondisi pada saat tes 2 wajah	-	-	93.4%
12	Tidak terdapat wajah	-	-	100%
13	Ekspresi lain	100%	100%	100%

14	Area wajah tidak penuh	100%	100%	100%
15	Pendeteksian wajah yang memiliki kumis dan berjenggot	-	-	40%
16	Pendeteksian ekspresi wajah berdasarkan jarak	-	-	70%
17	Pendeteksian ekspresi wajah kepada pengguna yang menggunakan kacamata	77,8%	83,4%	67%
18	Pendeteksian ekspresi wajah kepada pengguna yang menggunakan lipstik	-	-	73%
Nilai keberhasilan		88,61 %	90,28 %	84,73%

Kendala dalam pengujian yang telah dilakukan adalah pada saat melakukan pengujian pendeteksian wajah yang memiliki kumis dan jenggot dikarenakan *point* pada bibir tidak terdeteksi karena *noise* pada jenggot.

4. Kesimpulan

Berdasarkan penelitian mengenai aplikasi pendeteksian ekspresi wajah dengan menggunakan metode pendekatan *K-Means* diperoleh kesimpulan sebagai berikut:

1. Keberhasilan pendeteksian ekspresi menurut pengujian yang telah dilakukan adalah sebesar 84.73%. Kelemahan aplikasi adalah pada saat mendeteksi wajah yang memiliki kumis dan jenggot. Aplikasi tidak dapat menentukan area bibir yang terhalang oleh kumis dan jenggot tersebut.
2. Tingkat keberhasilan untuk penguji perempuan lebih tinggi dibandingkan dengan tingkat keberhasilan pada penguji pria. Tingkat keberhasilan yang diperoleh dari penguji wanita adalah 90,28% dan penguji pria adalah 88,61%. Ini disebabkan dataset merupakan wajah ekspresi dari wajah perempuan.
3. Tingkat keberhasilan untuk dataset JAFFE adalah sebesar 88.2%. Harapan keberhasilan untuk dataset adalah 100% karena dataset JAFFE digunakan sebagai pengklasifikasian ekspresi wajah.
4. Jarak maksimum untuk pendeteksian ekspresi adalah sejauh 100 cm. Dijarak 100 cm tingkat keakurasian yang didapat adalah 70%. Pada saat subjek yang diuji di jarak 110 cm, ukuran wajah pada citra yang diambil sangatlah kecil dan tidak bisa dilakukan pendeteksian ekspresi.

Daftar Pustaka

- [1] McCraty, Rolling, MA., Barrios-Choplin, Bob, PhD., Atkinson, Mike., dan Tomasino, Dana, BA. (1998). “*The Effects of Different Types of Music on Mood, Tension, and Mental Clarity*”. *Alternative Therapies*. 4 (1), 75-84.
- [2] Calder, A. J., Burton A. M., Miller, P., Young, A. W. & Akamatsu, S. 2001. “*A Principal Component Analysis of Facial Expressions*”. *Vision Research*, vol 41, pp 1179-1208.
- [3] Michael J. Lyons, Shigeru Akamatsu, Miyuki Kamachi & Jiro Gyoba. (1998). “*Coding Facial Expressions with Gabor Wavelets*”. IEEE Computer Society, pp. 200-205.
- [4] Gunadi, I Gede Aris., dan Harjoko, Agus. (2014). “Pengenalan Pola Gerak Mulut Dengan Menggunakan Pendekatan SOM (Self Organizing Feature Map)”. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*. 3 (1), 15-20.
- [5] Utami, Dina Budi. (2012). “Pengenalan Gerak Jari Tangan Menggunakan Hidden Markov Model pada Augmented Interface System”.