

PENGENALAN WAJAH MENGGUNAKAN METODE *PRINCIPAL COMPONENT ANALYSIS* (PCA) DAN *CANBERRA DISTANCE*

Perani Rosyani

Teknik Informatika, Universitas Pamulang

email: dosen00837@unpam.ac.id

ABSTRAK

Wajah merupakan salah satu karakteristik biometrik yang digunakan untuk mengenali seseorang selain karakteristik yang lain seperti ucapan, sidik jari, retina, dll. Wajah adalah struktur multidimesi yang sangat kompleks dan membutuhkan tehnik komputasi yang baik untuk pengenalan. Di dalam penelitian ini, penulis mengambil 10 pose wajah yang berbeda kemudian menggunakan metode PCA untuk pengoptimalan dalam mereduksi dimensi. Setelah citra original di ekstraksi menggunakan PCA maka akan di hitung tingkat kesamaan (*similarity degree*) antara gambar test dengan gambar training menggunakan metode jarak. Hasil dari penelitian ini adalah tingkat kesamaan yang dihasilkan setelah proses reduksi dan ekstraksi menggunakan PCA didapatkan rata-rata nilai untuk *Canberra Distance* adalah 77,59.

Keywords: PCA, Wajah, reduksi dimensi

1. PENDAHULUAN

Memori otak manusia mampu mengenali dan menyimpan kejadian bahkan bentuk yang dilihat. Memori otak akan memberikan sebuah gambaran wajah yang pernah kita kenal sebelumnya. Analogi tersebut bisa diumpamakan sebagai sebuah mesin atau perangkat yang memiliki kemampuan sama seperti manusia untuk mengenal individu melalui citra wajah. [1] Sistem pengenalan wajah banyak dimanfaatkan pada biometrics yang digunakan untuk identifikasi personal pada penggunaan mesin absensi, akses kontrol dan lain-lain. [2]

Secara umum *system* pengenalan wajah tidak secara langsung mengambil image *original* lalu kemudian dapat langsung diolah sebagai *image* yang dikenali didalam *system* pengenalan wajah. *Image* memerlukan ekstraksi *feature* sebelum *image* wajah diolah kedalam *system* pengenalan wajah, hal ini diperlukan untuk menghemat komputasi. Ekstraksi *feature* merupakan tahapan untuk memunculkan ciri dan mereduksi citra dari dimensi tinggi ke dimensi yang lebih rendah. Ekstraksi *feature* yang telah populer untuk mengurangi dimensi antara lain *Principal Component Analysis* (PCA), *Linear Discriminant Analysis* (LDA), *Laplacianfaces*, *Discrete Cosine Transform* (DCT). [3]

Salah satu tehnik yang digunakan didalam penelitian ini adalah *Principal Component Analysis* (PCA) yang merupakan tehnik linear untuk memproyeksikan data vektor yang berdimensi tinggi ke vektor yang mempunyai

dimensi lebih rendah. [4] Dimensi didalam image ini berupa matriks $M \times N$ yang memiliki 3 layer RGB. Oleh karena itu, penulis mengambil 40 pose wajah dengan citra RGB yang kemudian di convert ke citra *grayscale* kemudian matriks tersebut akan di reduksi menggunakan metode PCA. Setelah itu akan di hitung tingkat kesamaan (*similarity degree*) antara citra tes dengan citra uji menggunakan metode jarak.

Metode jarak digunakan untuk menentukan tingkat kesamaan (*similarity degree*) atau ketidaksamaan (*disimilarity degree*) dua vektor fitur. Banyak teknik pengenalan pola dengan menggunakan metode jarak diantaranya yaitu *Euclidean Distance*, *City-Block Distance*, dan *Bray Curtis*. Dari sekian banyak metode jarak yang digunakan dalam pengenalan pola, metode *Euclidean Distance* yang sering dipakai, padahal masih banyak metode lain dalam teknik recognition menggunakan jarak. [5] Dari metode jarak yang disebutkan metode *Euclidean Distance* merupakan metode yang sering dipakai didalam pengenalan pola.

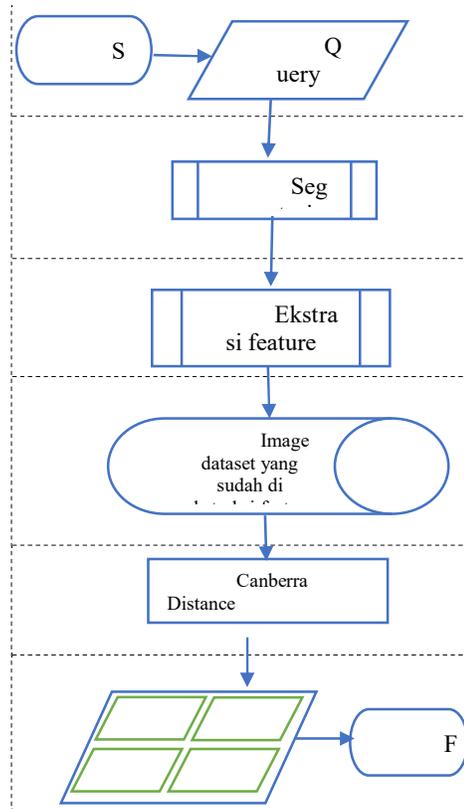
Penelitian Bhaskar Gupta – Anil Kumar Singh [6] membahas tentang pengenalan wajah menggunakan metode PCA dan perbandingan klasifikasi jarak, yang mana klasifikasi jarak yang diambil adalah jarak *Euclidean Distance* dan *Manhattan Distance*. Hasilnya adalah mendapatkan nilai maksimum 94% dari kedua klasifikasi jarak tersebut.

Didalam penelitian ini penulis mencoba salah satu metode klasifikasi jarak untuk

pengenalan pola berupa wajah dengan menggunakan *Canberra Distance*. Yang sebelumnya pola wajah sudah di ekstrak menggunakan PCA. Hal ini dimaksudkan untuk mengetahui tingkat kesamaan dari metode jarak tersebut didalam proses pengenalan wajah. Apakah lebih besar atau lebih kecil dibandingkan dengan penelitian sebelumnya.

2. METODE PENELITIAN

Metode penelitian yang digunakan didalam penelitian ini selain mencari literatur yang berhubungan dengan penggunaan PCA juga membandingkan metode kemiripan antara citra dengan menggunakan Manhattan dan Canberra Distance. Gambar dibawah dapat di lihat adalah diagram alur dari mulai segmentasi, pengekstrakan menggunakan metode PCA hingga membandingkan jarak kemiripan dari citra. Dapat dilihat dari alur proses dibawah ini.



Gambar 2. Flow Diagram Recognition

1. Pre-processing merupakan tahap awal setelah pengumpulan data citra. Ukuran citra yang digunakan sudah disesuaikan semua menjadi 112x92. Setelah ukuran citra disamakan dan dimasukkan kedalam

query citra selanjutnya proses segmentasi. Dimana proses segmentasi didalam penelitian ini sudah diselesaikan oleh database ORL.

2. Processing merupakan pengekstrakan *image* menggunakan PCA. Yaitu membuat matriks gambar yang berbentuk 2dimensi menjadi 1 dimensi yang bertujuan untuk memudahkan dan mempercepat dalam proses perhitungan rata-rata.
3. *Post-processing* merupakan tahap akhir untuk menghitung matriks dengan menggunakan metode jarak yaitu *Canberra Distance*

3. HASIL DAN PEMBAHASAN

Didalam penelitian ini data yang digunakan adalah data citra wajah yang diambil dari database wajah ORL. Berikut adalah tahapan-tahapan proses:

A. Pra-processing

Pra-proses ini bertujuan untuk mempersiapkan citra data training yang diproses oleh sistem agar informasi yang terkandung didalamnya dapat diolah untuk proses selanjutnya. Tahapan pra-proses yang dilakukan adalah tahapan normalisasi sebelum melakukan ekstraksi wajah. Normalisasi ini termasuk normalisasi geometri, koreksi gambar wajah, tingkat grayscale, dll. Dalam penelitian ini kendala-kendala tersebut telah diselesaikan pada ORL database wajah. [7]



Gambar 3. Sample Gambar ORL

B. Processing

Processing : adalah mengekstraksi feature untuk mendapatkan ciri dari kedua citra tersebut, karena didalam paper ini peneliti membahas mengenai wajah. Maka ciri yang diambil menggunakan Principal Component Analysis (PCA) yaitu mereduksi jumlah dimensi yang tinggi ke dimensi yang lebih rendah. Proses ini memiliki beberapa langkah yaitu sebagai berikut:

- a. Pembentukan matriks data citra wajah

Langkah pertama yang dilakukan adalah membentuk matriks data citra dengan mengambil data pixel setiap citranya. Citra berukuran $m \times n$ di mana m adalah jumlah citra training sedangkan n adalah citra tes.

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2n} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & x_{m3} & \dots & x_{mn} \end{bmatrix}$$

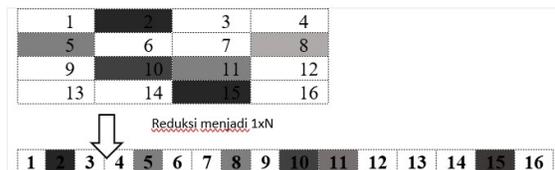
- b. Pencarian rata-rata seluruh citra
Setelah terbentuk matriks data citra wajah, maka proses selanjutnya adalah mencari nilai mean dari seluruh citra. Hal ini bertujuan untuk mengetahui *noise* yang dapat mengurangi tingkat keakuratan didalam perhitungan PCA, yang dapat dihitung dengan menggunakan rumus:

$$\begin{aligned} \tau &= \frac{(x_{12} + x_{21} + x_{31} + \dots + x_{m1})}{m} \\ &= \frac{\sum_{j=1}^m x_{ji}}{m} = [\tau_1, \tau_2, \dots, \dots, \tau_n] \end{aligned}$$

- c. Penggandaan nilai rata-rata.
Penggandaan nilai rata-rata dilakukan untuk menyamakan dimensi nilai rata-rata dengan dimensi data citra sampel sebanyak m , sehingga rata-rata seluruh citra data sampel memiliki dimensi ($m \times n$). Matriks rata-rata citra data sampel yang telah digandakan sebanyak m kali dapat ditulis menggunakan persamaan:

$$\mu = \alpha x \tau$$

α adalah matriks kolom dengan dimensi $m \times 1$ dan α untuk keseluruhan nilai adalah bernilai 1.



Gambar 1. Reduksi Dimensi

Gambar 1 diatas menunjukkan proses reduksi dari citra ukuran $M \times N$ dengan jumlah baris adalah 4 dan kolom adalah

4, menjadi $M \times 1$ dengan jumlah kolom adalah 16 dan baris adalah 1

- d. Perhitungan nilai rata-rata nol
Perhitungan nilai rata-rata nol, berfungsi untuk menghilangkan *noise* yang dapat mengganggu keakuratan pada perhitungan PCA. Perhitungan ini dapat dimodelkan menggunakan persamaan:

$$\phi = x - \mu$$

- e. Pembentukan matriks kovarian
Hasil perhitungan nilai rata-rata nol digunakan untuk mendapatkan nilai matriks kovarian. Berikut ini adalah persamaan matriks kovarian:

$$C = \frac{1}{m-1} \phi_{ji} * \phi^T_{ji}$$

- f. Penentuan matriks *eigenvalue* dan *eigenvector*
Setelah matriks kovarian didapatkan maka langkah selanjutnya adalah menentukan matriks *eigenvalue* dan matriks *eigenvector*.

Matriks *eigen* dapat dicari menggunakan persamaan:

$$C - Z = |C - \lambda I|$$

$$|C - \lambda I| = 0$$

Dimana C adalah matriks kovarian. Z adalah matriks *eigenvalue* dengan λ sebagai *scalar* pembentuknya dan I sebagai matriks identitas.

- g. Pencarian *eigenface* untuk proses pengenalan wajah
Eigenface didapatkan dari hasil perkalian matriks X atau *eigenvector* dengan matriks N yang merupakan matriks hasil perhitungan rata-rata nol pada persamaan yang dibagi oleh akar Z , sebagai *eigenvalue*.

$$Eigenface = \frac{1}{\sqrt{Z}} * X * \phi$$

Setelah proses reduksi di atas semua *query* citra akan di *save* didalam *storage*. Kemudian akan di hitung tingkat kesamaannya (*Similarity*) dengan metode Manhattan Distance, dan Canberra Distance.

- h. Canberra Distance
Untuk setiap nilai 2 vektor yang akan dicocokkan, Canberra Distance

membagi absolute selisih 2 nilai dengan jumlah dari absolute 2 nilai tersebut[3]. Hasil dari dua nilai yang dicocokkan lalu dijumlahkan untuk mendapatkan Canberra Distance. Jika koordinat nol-nol((0,0)) diberikan definisi dengan $0/0=0$. Canberra Distance ini sangat peka terhadap sedikit perubahan dengan kedua koordinat mendekati nol.

$$d_{ij} = \sum_{k=1}^n \frac{|x_{ik} - x_{jk}|}{|x_{ik}| + |x_{jk}|}$$

Keterangan :

d_{ij} = tingkat perbedaan (*dissimilarity degree*)

n = jumlah vector

x_{ik} = vektor citra *input*

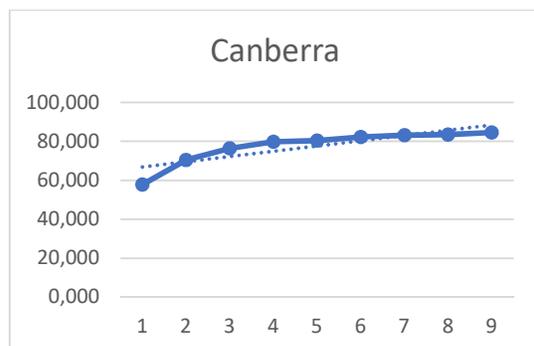
x_{jk} = vektor citra pembandingan */output*

Hasil dari pengambilan gambar training sebanyak 9 secara random. Kemudian dilakukan pengestrakan data menggunakan PCA dan perhitungan jarak *Canberra* didapatkan nilainya sebagai berikut.

tabel 1. Rate PCA

Rate Recognition PCA	
Train Image	Canberra
1	57.806
2	70.469
3	76.464
4	79.833
5	80.350
6	82.188
7	83.167
8	83.500
9	84.500

Dari tabel diatas jika di bentuk didalam grafik akan terlihat seperti gambar di bawah ini.



Gambar 3. Grafik Nilai Recognition PCA

Dari gambar 3 kita dapatkan untuk Canberra Distance dari 9 data training yang digunakan rata-rata ada diposisi 80-90.

4. KESIMPULAN

Mengekstraksi gambar menggunakan PCA cukup rumit bagi peneliti karena harus mentraspose matriks ukuran $m \times n$ menjadi $1 \times n$. Tingkat kesamaan yang dihasilkan setelah proses reduksi dan ekstraksi menggunakan PCA didapatkan rata-rata nilai untuk *Canberra Distance* adalah 77,59.

DAFTAR PUSTAKA

- [1] S. A. Wijaya, "PERBANDINGAN METODE PENGENALAN WAJAH SECARA REAL-TIME PADA PERANGKAT BERGERAK BERBASIS ANDROID".
- [2] D. Parikesit, "FACE RECOGNITION MENGGUNAKAN METODE PCA".
- [3] M. H. Purnomo and A. Muntasa, Konsep Pengolahan Citra Digital dan Ekstraksi Fitur, Surabaya: GRAHA ILMU, 2010.
- [4] M. a. P. Turk, "Face Recognition Using Eigenfaces," in IEEE Conf. of Computer Vision dan Pattern Recognition, 1991.
- [5] S. R. Wurdianto, S. Novianto and U. Rosyidah, "PERBANDINGAN EUCLIDEAN DISTANCE DENGAN CANBERRA DISTANCE PADA FACE RECOGNITION," *Techno.COM*, Vol. 13, No. 1, pp. 31-37, 2014.
- [6] A. K. S. Bhaskar Gupta, "Analyzing Face Recognition Using Pca and Comparison between Different Distance Classifier," *International Journal Of Engineering Sciences & Research Technology*, pp. 683-686, 2013.
- [7] H. D. A Eleyan, "Face recognition system based on PCA and feedforward neural networks Computational Intelligence and Bioinspired Systems," Springer, pp. 935-942, 2005.