

# Hasil Ekstraksi Algoritma *Principal Component Analysis* (PCA) untuk Pengenalan Wajah dengan Bahasa Pemrograman Java Eclipse IDE

Fiqih Ismawan

Dosen Program Studi Teknik Informatika, FMIPA Universitas Indraprasta PGRI

Email : vQ.ismaone@gmail.com

**Abstrak**— Algoritma *Principal Component Analysis* (PCA) merupakan salah satu metode yang dapat digunakan untuk mengolah citra wajah seseorang sehingga secara otomatis sistem akan mengenali wajah seseorang melalui ciri-ciri utamanya seperti mata, hidung, bibir, alis sebagai identitas. Identitas dari citra wajah seseorang tersebut oleh sistem akan dikenali melalui berbagai pelatihan (*training*) yang disimpan di database. Fase pelatihan (*training*) merupakan hasil ekstraksi dari kumpulan berbagai wajah yang berbeda kemudian dikumpulkan dan disimpan di sebuah database. Hasil citra wajah yang telah diekstraksi menggunakan algoritma PCA tersebut nantinya akan dibandingkan dengan citra wajah baru sebagai citra wajah yang akan dites apakah mempunyai kemiripan atau hampir mirip untuk dikenali oleh sistem.

**Kata kunci**— *Face Recognition*, PCA, Java Eclipse IDE, Eigenface.

## I. PENDAHULUAN

Wajah merupakan salah satu komponen dari tubuh manusia yang mempunyai ciri-ciri utama seperti mata, hidung, bibir, alis sebagai identitas. Ciri-ciri utama tersebut yang dapat membedakan antara manusia satu dengan manusia lainnya. Image processing merupakan salah satu teknik untuk menyelesaikan masalah dalam pemrosesan citra. Dalam *image processing* citra yang sudah ada dapat diolah sedemikian rupa sehingga citra tersebut lebih mudah untuk diproses lebih lanjut untuk tujuan-tujuan tertentu.

Algoritma *Principal Component Analysis* (PCA) merupakan salah satu metode yang dapat digunakan untuk mengolah citra wajah seseorang sehingga secara otomatis sistem akan mengenali wajah seseorang melalui ciri-ciri utamanya seperti mata, hidung, bibir, alis sebagai identitas. Identitas dari citra wajah seseorang tersebut oleh sistem akan dikenali melalui berbagai pelatihan (*training*) yang disimpan di database. Fase pelatihan (*training*) merupakan hasil ekstraksi dari kumpulan berbagai wajah yang berbeda kemudian dikumpulkan dan disimpan di sebuah database. Hasil citra wajah yang telah diekstraksi menggunakan algoritma PCA tersebut nantinya akan dibandingkan dengan citra wajah baru sebagai citra wajah yang akan dites apakah mempunyai kemiripan atau hampir mirip untuk dikenali oleh sistem.

Berdasarkan hal tersebut di atas, maka dalam penelitian ini akan mengkaji bagaimana sistem dapat mengenal pola wajah seseorang melalui Hasil Ekstraksi Algoritma *Principal Component Analysis* (PCA).

## II. LANDASAN TEORI

Dalam statistika, analisis komponen utama (*Principal Component Analysis*/PCA) adalah teknik yang digunakan untuk menyederhanakan suatu data dengan cara mentransformasi linier sehingga terbentuk sistem koordinat baru dengan varians maksimum. PCA dapat digunakan untuk mereduksi dimensi suatu data tanpa mengurangi karakteristik data tersebut secara signifikan (Smith, 2002).

Menurut Santosa (2007), *Principal Component Analysis* (PCA) adalah suatu teknik handal untuk mengekstraksi struktur dari suatu set data dengan dimensi yang cukup banyak. Problem dalam PCA adalah menemukan eigenvalue dan eigenvectors. PCA adalah transformasi orthogonal (tegak lurus) dari sistem koordinat dimana data dideskripsikan. Koordinat baru dimana data dideskripsikan dinamakan *principal component* atau PC. Koordinat tersebut dipilih dimana variansi dari data mencapai maksimum.

Prinsip dasar algoritma PCA adalah memproyeksikan citra ke dalam bidang ruang eigen-nya dengan cara mencari eigenvector yang dimiliki setiap citra dan memproyeksikannya ke dalam ruang eigen yang didapat tersebut. Besar ruang eigen tergantung dari jumlah citra training yang dimiliki (Gökberk, 2005).

### A. *Principal Component Analysis* (PCA)

Seringkali dengan beberapa *Principal Component* (PC) sudah cukup untuk menjelaskan struktur data asli. Jika data dalam dimensi asli sulit untuk direpresentasikan melalui grafik, maka dengan dua PC atau satu PC bisa dicitrakan melalui grafik (Santosa 2007). Misalkan terdapat data dengan susunan sebagai berikut:

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ X_{m1} & X_{m2} & \cdots & X_{mn} \end{bmatrix}$$

Dimana n adalah jumlah variabel/atribut dan m adalah banyaknya observasi. Kemudian data dapat ditransformasi ke dalam suatu kolom, misalkan dengan cara centering yakni mengurangi setiap data dengan rata-rata dari setiap atribut yang ada dengan rumus sebagai berikut:

$$\hat{X} = X - \bar{X}$$

$\hat{X}$  adalah vektor hasil setelah centering,  $X$  adalah vektor kolom dan  $\bar{X}$  adalah rata-rata dari kolom yang bersangkutan. Proses tersebut dilakukan untuk semua kolom dari  $i = 1$  sampai  $i = n$ . Untuk mendapatkan besaran baru digunakan matriks *Covariance* ( $C$ ):

$$C = \frac{\hat{X}^T \hat{X}}{m - 1}$$

Dalam *pattern recognition*, fitur-fitur suatu citra menggunakan PCA direpresentasikan sebagai *eigenvectors*. *Eigenvectors* didefinisikan sebagai kumpulan hubungan karakteristik-karakteristik dari suatu citra untuk mengenali citra tersebut secara spesifik.

Setelah matrik *covariance* dihitung, langkah berikutnya adalah mencari *eigenvalue* dan *eigenvector*. *Eigenvalue* yang didapat diurutkan mulai yang terbesar sampai yang terkecil, dan *eigenvector* yang bersesuaian dengan *eigenvalue* tersebut juga diurutkan.

### III. METODOLOGI PENELITIAN

#### A. Hardware dan Software

Alat yang digunakan untuk penelitian ini terdiri dari perangkat keras dan perangkat lunak sebagai berikut:

- 1) Sistem Operasi Windows XP (rekomendasi)
- 2) Processor Intel Pentium 4 2.0 GHz
- 3) Harddisk 40 GB
- 4) RAM 256 MB
- 5) Keyboard dan Mouse
- 6) VGA Card 64 MB
- 7) Monitor Samsung 15"
- 8) Piranti bantu Eclipse IDE for Java Developers Build id: 20090619-0625

#### B. Metode Penelitian

Metodologi yang dilakukan dalam penelitian model deteksi wajah adalah sebagai berikut:

1. Melakukan pengumpulan data-data citra wajah yang dibutuhkan dari hasil pemotretan menggunakan alat sensor kamera atau mengumpulkan citra wajah melalui browsing dari internet. Mempelajari buku-buku tentang Pengolahan Citra Digital (*Digital Image Processing*), Pengenalan Pola (*Pattern Recognition*) serta literatur – literatur dari internet yang sesuai dengan topik penelitian ini.
2. Melakukan perancangan sistem, diantaranya adalah:
  - a. Perancangan sistem untuk fase pelatihan (*training*) berupa normalisasi dan mengekstrak fitur-fitur utama pada beberapa sampel wajah menggunakan algoritma *Principal Component Analysis* (PCA).
  - b. Perancangan sistem untuk fase pengujian atau deteksi (*detection*) berupa normalisasi

dan menganalisa citra wajah baru sebagai data input dengan membandingkan fitur-fitur utama citra wajah yang telah terbentuk dari hasil fase training.

- c. Perancangan sistem secara *Graphical User Interface* (GUI) agar memudahkan pengguna dalam berdialog atau berinteraksi dengan sistem.
3. Melakukan pengkodean/implementasi sistem menggunakan software Eclipse IDE for Java Developers.
4. Melakukan pengujian (*testing*) sistem terhadap data sampel citra wajah hasil training dengan data citra wajah baru sebagai data input untuk dideteksi atau pencocokkan pola.

### IV. HASIL DAN PEMBAHASAN

#### A. Proses Training Citra Wajah

Proses Training citra wajah dilakukan sebelum dilakukan proses pencocokkan pola wajah. Dalam proses training ini, dilakukan beberapa tahapan proses diantaranya proses normalisasi, proses mengekstrak citra wajah menggunakan PCA (*Principal Component Analysis*) sampai diperoleh fitur-fitur utama wajah. Fitur-fitur utama tersebut nantinya sebagai data pembanding dengan citra wajah baru sebagai citra wajah yang dikenali. Pada Gambar 1 merupakan form hasil proses normalisasi citra wajah.



Gambar 1. Citra Wajah Hasil Proses Normalisasi

Hasil proses normalisasi merupakan hasil konversi citra wajah berwarna 24-bit menjadi citra grayscale dan dari dimensi resolusi awal menjadi citra wajah berukuran 125 x 150 piksel. Sebagai batasan masalah pada penelitian ini, posisi wajah yang digunakan adalah citra wajah yang diambil dengan arah ke depan dan sedikit menyerong sehingga fitur-fitur utama dari wajah dapat dikenali dengan algoritma PCA ini. Hal tersebut sudah cukup mewakili data citra untuk dilakukan proses pencocokkan pola.

Proses Normalisasi ini dimaksudkan untuk mendapatkan data yang berukuran kecil sehingga dapat membantu mempercepat pembacaan data citra, selanjutnya untuk diproses training.

Vektor baris dihitung dengan mencari rata-rata vektor citra wajah untuk mengekstrak ciri utama wajah dari n vektor baris.

$$\psi = \frac{\sum_{i=1}^n \vec{x}_i}{n}$$

Suatu matrik  $\Phi$  dikonstruksi dengan mengurangi rata-rata vektor citra wajah  $\Psi$  dari setiap sampel citra wajah  $X$ .

$$\Phi_i = \vec{x}_i - \psi$$

Matrik covariance merupakan matrik yang digunakan untuk memperoleh ruang wajah (eigenspace) dari ciri utama masing-masing sampel wajah.

Adapun perhitungan matrik covariance yakni dengan mengalikan matrik baru  $\Phi$  dengan transpose matriknya  $\Phi^T$ .

$$C = \Phi \times \Phi^T$$

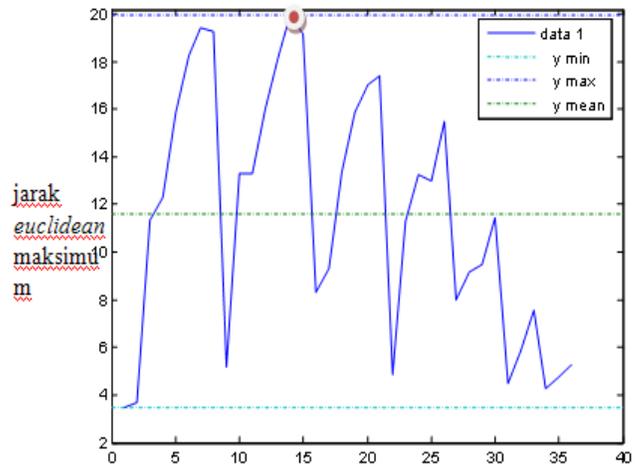
Setelah matrik covariance diperoleh, berikutnya adalah menghitung eigenvector dan eigenvalue. Proses menghitung eigenvector dan eigenvalue dimaksudkan untuk memperoleh kumpulan hubungan karakteristik-karakteristik dari suatu citra, dalam hal ini citra wajah. Nilai eigenvector yang bersesuaian dengan eigenvalue kemudian diurutkan mulai dari nilai yang terbesar sampai yang terkecil, untuk mendapatkan fitur utama citra wajah. Fitur utama citra wajah yang terbentuk tersebut disebut eigenfaces.

Menghitung matrik covariance dimaksudkan untuk mendapatkan nilai eigenvector dan eigenvalue sehingga didapatkan eigenfaces yang merupakan fitur-fitur utama dari citra wajah.

Setelah beberapa eigenvector didapat, kemudian dilakukan proses pengurutan eigenvector yang berkorelasi dengan eigenvalue. Proses pengurutan ini dilakukan untuk mempermudah memperoleh matriks eigenfaces. Eigenfaces yang terbentuk tersebut merupakan nilai-nilai piksel yang merepresentasikan fitur-fitur utama wajah.

Nilai matrik eigenfaces yang diperoleh tersebut dapat digunakan untuk mencari nilai threshold, dengan cara menghitung nilai maksimum jarak.

Hasil nilai maksimum yang diperoleh dalam penelitian ini yaitu 19,864 yang diambil dari nilai eigenvector yang berkorespondensi dengan eigenvalue tertinggi, seperti ditunjukkan dalam bentuk grafik pada Gambar 2. Nilai 19,864 ini akan digunakan sebagai nilai threshold atau batas ambang pada citra wajah yang akan dicocokkan.

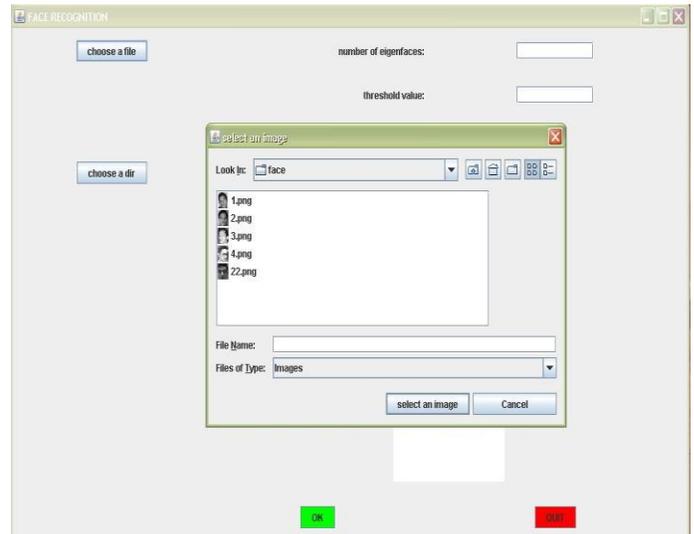


Gambar 2. Perolehan Nilai Maksimum Dari Jarak Euclidean Citra Hasil Training

B. Proses Pencocokan Pola

Proses pencocokan dilakukan setelah data-data citra wajah yakin telah di-training dan telah tersimpan dalam basis data, sehingga data-data tersebut dapat digunakan sebagai data pembandingan antara citra wajah baru dengan citra wajah hasil training untuk mendapatkan output citra wajah yang diharapkan. Gambar 3 merupakan gambaran menginputkan citra wajah yang diuji dari database untuk dicocokkan.

Citra wajah yang akan diuji diperoleh melalui kumpulan file wajah yang tersimpan dalam format PNG. Citra wajah tersebut akan dipilih sebagai citra pembandingan dengan citra sampel hasil training, apakah pola diantaranya sama atau mendekati sama sehingga menghasilkan output wajah yang dikenali sistem.



Gambar 3. Form Proses Pengujian Ketika Menginputkan Data Citra Wajah Untuk Dicocokkan

Kotak dialog yang ditampilkan saat proses pencocokkan seperti pada Gambar 3, berisikan data-data wajah yang telah di training dan data-data wajah untuk diuji atau dicocokkan yang tersimpan dalam satu folder, hal ini dimaksudkan hanya untuk mempermudah pengambilan data saja.

C. Mengekstrak Citra Wajah yang Dideteksi

Citra wajah yang akan dicocokkan, terlebih dahulu diekstrak untuk mendapatkan fitur-fitur utamanya atau komponen-komponen utamanya. Kemudian diproyeksikan ke ruang citra wajah (eigenfaces), yang nantinya akan dicari jarak antar pikselnya.

$$\omega_i = UT \times (Z - \Psi)$$

Berikut kode untuk mengekstrak citra wajah yang akan dicocokkan:

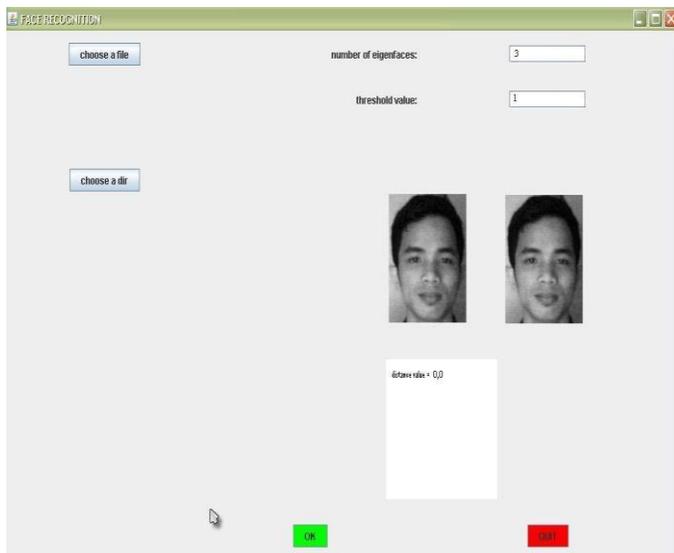
```
MeanInputImage=double(Imagevector)-m1;
ProjectInputImage=Eigenfaces'*MeanInputImage;
```

Ekstrak vektor citra wajah yang akan diuji ( $\omega_i$ ) atau dengan nama variabel ProjectInputImage dihasilkan 7 citra wajah dengan ukuran 7 x 1 dan disimpan dalam vektor  $\Omega$ . Hasil nilai ekstrak vektor citra wajah yang akan diuji ( $\omega_i$ ) adalah seperti pada Tabel 1.

Tabel 1 Nilai Citra Wajah Yang Akan Diuji Setelah Diekstrak

No	Nilai Citra Wajah x 1.0e+008
1	0.877787877700
2	0.0
3	4,5589585

Pada Gambar 4 menampilkan hasil pencocokkan ketika wajah yang diuji atau dicocokkan sesuai dengan wajah hasil training.

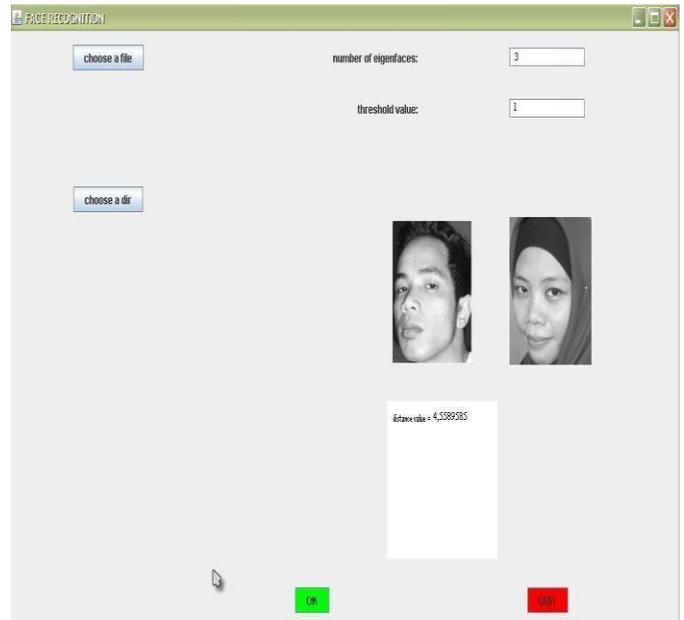


Gambar 4. Form Hasil Pencocokkan Pola

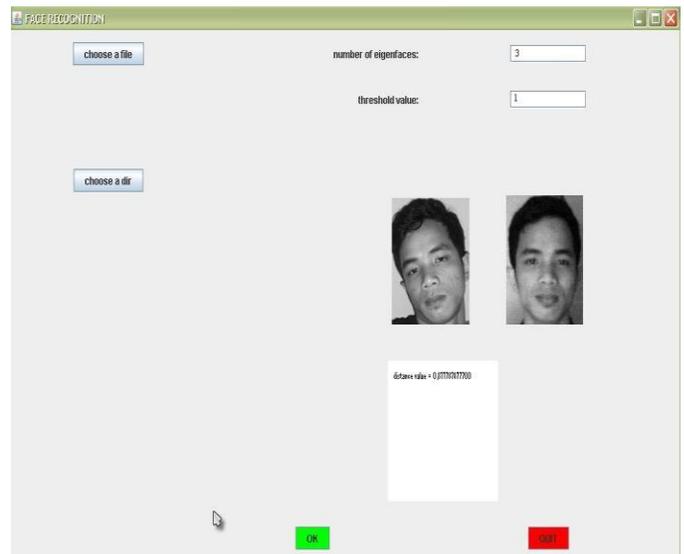
Gambar 5 menampilkan hasil pencocokkan ketika wajah yang diuji atau dideteksi tidak cocok atau bukan wajah yang

diharapkan.

Nilai threshold merupakan sebagai perbandingan antara citra wajah yang diuji dengan citra wajah hasil training ditetapkan nilai threshold yang diperoleh adalah 4,5589585.



Gambar 5. Form Hasil Deteksi Ketika Wajah yang Diuji Tidak Diterima



Gambar 6. Form Hasil Deteksi Ketika Wajah yang Diuji Hampir Mendekati

D. Hasil Pengujian

Pada penelitian ini dilakukan pengujian sebanyak 5 citra wajah yang terdiri dari 5 buah citra wajah. Dari sejumlah 5 buah citra wajah tersebut terdapat 5 buah citra wajah yang akan diujikan yaitu berupa citra wajah yang belum dikenal oleh sistem maupun citra wajah yang mendekati ataupun citra wajah yang blur dimungkinkan terjadi kesalahan saat pengambilan citra. Setelah dilakukan pengujian, diperoleh hasil deteksi pada Tabel 2.

Tabel 2. Hasil Deteksi Dengan 5 Citra Wajah Yang Diuji

		Wajah Uji	
		Wajah dikenal	Wajah tidak dikenal
Hasil Uji	Wajah dikenal	3	1
	Jumlah	3	1

Pada hasil pengujian Tabel 2 diperoleh:

Sebanyak 5 buah citra wajah dicocokkan dengan wajah yang telah ditraining, terdapat 3 buah citra wajah yang sesuai dengan pola wajah yang diharapkan. Diperoleh 1 buah citra wajah yang tidak dikenal atau tidak sesuai dengan pola wajah yang ada pada database. Hasil pengukuran akurasi dalam satuan persen (%) tertera pada Tabel 3.

Tabel 3. Hasil Pengukuran Akurasi (%) Pengenalan Wajah

		Wajah Uji	
		Wajah dikenal	Wajah tidak dikenal
Hasil Uji	Wajah dikenal	86	14
	Wajah tidak dikenal	14	-

Berdasarkan hasil pengukuran akurasi Tabel 3, diperoleh rata-rata akurasi sebesar 86.5 %.

### V. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan dan uji coba aplikasi, maka dapat ditarik kesimpulan sebagai berikut:

1. *Principal Component Analysis* (PCA) relatif mudah menangani sejumlah data yang cukup besar serta kemampuannya menangani data-data dimensi yang kompleks, seperti fitur-fitur pada citra wajah yang diekstrak berupa dimensi eigenvektor dan eigenvalues sehingga diperoleh ruang wajah (eigenspace) atau eigenfaces.
2. Eigenfaces merupakan sejumlah eigenvektor yang berkorelasi dengan eigenvalue tertinggi dan selanjutnya digunakan pada tahap pencocokkan pola wajah.
3. Pencocokkan wajah yang dilakukan diantaranya menghitung jarak minimum antara citra wajah hasil training dengan wajah yang diuji menggunakan euclidean distance.
4. Nilai *threshold* sangat berpengaruh dalam penentuan pengenalan pola suatu citra khususnya pada citra wajah ini, karena jika citra yang diuji mempunyai nilai fitur utama lebih besar dari nilai *threshold*, maka citra yang diuji kemungkinan bukan suatu citra yang diharapkan. Sebaliknya jika citra yang diuji

mempunyai nilai fitur utama sama atau dibawah nilai *threshold*, maka citra tersebut akan teridentifikasi atau dikenali sebagai citra yang sama dengan citra yang sudah di training.

5. Sebuah citra wajah yang rusak berpengaruh pada hasil pencocokkan pola wajah walaupun bentuk citra wajah tersebut sama.
6. Uji coba aplikasi pencocokkan wajah menggunakan PCA, dihasilkan keakuratan 86,5 %.

### DAFTAR PUSTAKA

- [1] Anton H., *Elementary Linear Algebra 2000*, Publisher John Wiley & Sons Inc. 2000.
- [2] Bajwa I.S., and Hyder S.I, *PCA Based Classification Of Single-Layered Cloud Types*, Market Forces, Vol. 1, No.2., 2005.
- [3] Duda, R.O., Stork, D.G., and Hart, P.E., *Pattern Classification* . 2nd edition John Wiley & Sons, Inc., New York, 2000.
- [4] Gerad B., and Maurice C., *Digital Signal and Image Processing using Matlab*, ISTE Ltd., Newport Beach, USA, 2006.
- [5] Mudrova M, et.al., *Principal Component Analysis (PCA) in Image Processing*, Institute of Chemical Technology, Prague Department of Computing and Control Engineering, 2002.
- [6] Santosa B., *Data Mining (Teori dan Aplikasi)*, Graha Ilmu, Yogyakarta, 2007.
- [7] Smith I, *A Tutorial on Principal Components Analysis*, Publisher John Wiley & Sons Inc., 2002.
- [8] Wijaya, dkk, *Pengolahan Citra Digital Menggunakan Matlab*, Bandung: Informatika, 2007.
- [9] Krueger J, *Thresholds for Eigenface Recognition*, Connexions module: m12533.