

BAB 2

LANDASAN TEORI

2.1 Definisi Citra, Pengolahan Citra, dan Pengenalan Pola

Citra dapat dijelaskan sebagai dua dimensi dari fungsi $f(x,y)$ dimana x dan y tersebut adalah sebuah koordinat pada bidang dan amplitude dari f pada pasangan koordinat adalah intensitas atau sebuah tingkatan keabu-abuan dari suatu citra pada titik tersebut. Jika x , y , dan nilai intensitas dari f tersebut bernilai diskrit, berhingga, citra tersebut dinamakan citra digital (Gonzales and Woods, 2008:1).

Pengolahan citra (*image processing*) merupakan bidang yang berhubungan dengan proses transformasi citra (*image*) yang bertujuan untuk mendapatkan kualitas citra yang lebih baik (Fairhurst, 1988:5).

Pengenalan pola (*pattern recognition*) merupakan bidang ilmu yang melakukan proses analisis gambar yang inputnya adalah gambar ataupun citra digital dan menghasilkan output suatu deskripsi dengan tujuan untuk mendapatkan informasi yang disampaikan oleh gambar atau citra, dengan kata lain meniru kemampuan manusia (otak manusia) dalam mengenali suatu objek atau pola tertentu (Fairhurst, 1988:5).

2.2 Computer Vision

Computer Vision merupakan ilmu yang mempelajari bagaimana computer dapat mengenali obyek yang diamati (Fairhurst, 1988:5). *Computer Vision* adalah kombinasi antara pengolahan citra dan pengenalan pola. *Computer Vision* bersama intelegensia semu (*artificial intelligence*) akan mampu menghasilkan sistem intelegen visual (*visual intelligence system*).

2.3 Face Recognition

Identifikasi (pengenalan) wajah atau *face recognition* adalah sebuah tugas yang dikerjakan oleh manusia secara rutin dan mudah dalam kehidupan sehari-hari. Penelitian dan pengembangan ilmu pengenalan wajah berkembang secara otomatis atas dasar ketersediaan *desktop* kuat dan rendah biaya serta *embedded-system* yang telah menciptakan minat yang sangat besar dalam pengolahan citra digital dan video. Motivasi penelitian dan pengembangan dari pengenalan wajah termasuk dalam lingkup otentikasi *biometric*, pengawasan, interaksi manusia-komputer, dan manajemen multimedia (Li & Jain, 2005:1).

Pengenalan wajah adalah salah satu ilmu yang terdapat di dalam *computer vision*, di mana sebuah komputer dapat menganalisa suatu citra wajah yang terdapat di dalam sebuah gambar dan dapat menemukan identitas atau data diri dari citra wajah tersebut dengan membandingkan terhadap data-data citra wajah yang sudah disimpan sebelumnya di dalam *database*. Pada umumnya *face recognition* dilakukan dari sisi depan dengan pencahayaan yang merata ke

seluruh wajah. Akan tetapi muncul beberapa permasalahan, seperti posisi wajah, skala atau jarak wajah, orientasi, umur, dan ekspresi wajah.

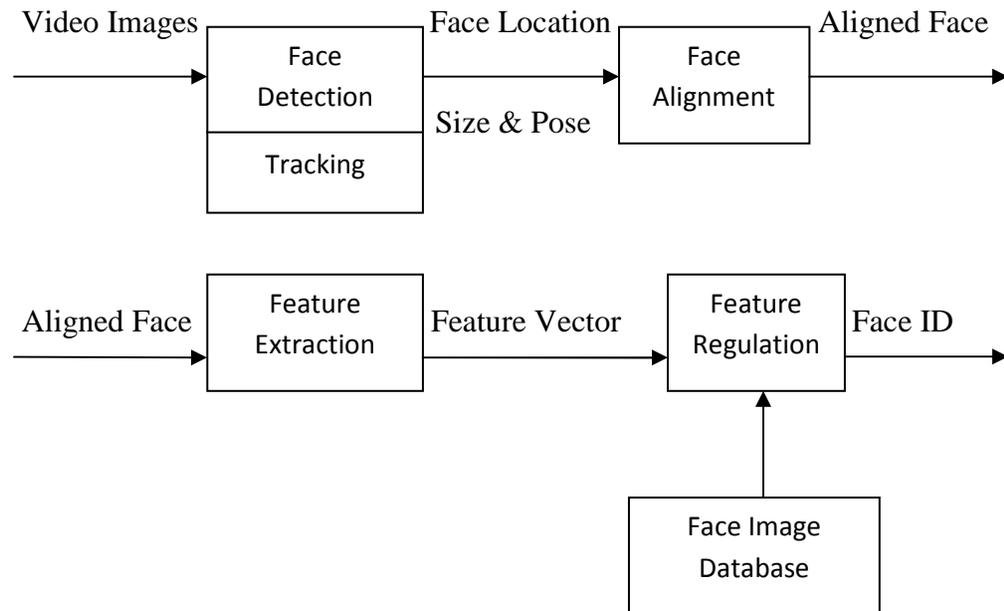
Sistem *face recognition* pada umumnya mencakup empat modul utama (Li & Jain, 2005:2), yaitu: deteksi, *alignment*, ekstraksi fitur dan pencocokan. Proses lokalisasi dan normalisasi (deteksi wajah dan *alignment*) adalah langkah-langkah sebelum proses pengenalan wajah (ekstraksi fitur wajah dan pencocokan) dilakukan.

Deteksi wajah adalah langkah awal untuk melakukan identifikasi wajah atau *face recognition*. Sebuah pendeteksi wajah yang ideal seharusnya mampu mengidentifikasi dan menemukan lokasi dan luas semua wajah yang ada di dalam sebuah gambar tanpa memperhatikan pose, skala, orientasi, umur, dan ekspresi (Li & Jain, 2005:13). Deteksi wajah melakukan segmentasi area citra wajah dengan bagian latar (*background*).

Proses *alignment* bertujuan untuk memperoleh akurasi yang lebih baik dan tinggi untuk lokalisasi dan normalisasi citra wajah sebab deteksi wajah menyediakan batas lokasi dan skala dari setiap citra wajah yang dapat terdeteksi.

Setelah sebuah wajah dilakukan normalisasi, ekstraksi fitur dilakukan untuk mengambil data yang efektif yang berguna untuk memisahkan antara citra-citra wajah dan orang-orang yang berbeda satu sama lain dan cukup stabil untuk bermacam-macam *geometric* dan fotometrik. Pencocokan wajah dilakukan dengan cara melakukan pencocokan fitur yang telah diekstraksi dari citra wajah

masuk dengan kumpulan data latihan dan uji coba citra wajah yang telah tersimpan sebagai *database* wajah.



Gambar 2.1 : Proses Pengenalan Wajah

Citra wajah direpresentasikan sebagai sebuah susunan *pixel* dengan dimensi tinggi. *Face recognition*, dan penelitian *computer vision* secara umum, telah mengamati pertumbuhan minat dalam teknik yang menerapkan aljabar dan fitur statistik untuk melakukan ekstraksi dan analisis kasus jenis ini. Analisa komputer untuk citra wajah dipengaruhi oleh sinyal visual (cahaya yang terpantul pada permukaan wajah) yang disimpan oleh sebuah sensor digital sebagai sebuah susunan dari nilai *pixel*. Nilai *pixel* ini menyimpan warna atau hanya intensitas cahaya. Array *pixel* dari citra wajah berukuran $m \times n$ dapat disimpan dalam bentuk sebuah vektor (contoh: vektor) di dalam sebuah ruang citra

mendimensi dengan menulis nilai-nilai *pixel*-nya dalam urutan tetap. Masalah utama dari data multidimensi adalah dimensionalitasnya, jumlah koordinat yang diperlukan untuk menspesifikasikan sebuah titik data. Jumlah dimensionalitas ruang tersebut, yang dibuat dari representasi *pixel* $m \times n$, adalah jumlah yang sangat tinggi bahkan untuk sebuah ukuran citra wajah yang sangat sederhana. Metode pengenalan wajah untuk mengoperasikan representasi ini mengalami beberapa kesulitan. Menurut Shakhnarovich & Moghaddam, (2005, p. 141), beberapa dari hambatan ini juga disebut *curse* dari dimensionalitas.

- Menangani dimensionalitas yang tinggi, terutama dalam konteks pengenalan berdasarkan kecocokan memerlukan biaya yang mahal secara komputasi.
- Untuk metode parametrik, jumlah parameter yang diperlukan untuk berkembang secara eksponensial berdasarkan dimensionalitasnya. Terkadang jumlahnya lebih banyak dari jumlah citra yang disediakan untuk latihan dan uji coba.
- Untuk metode non-parametrik, kompleksitas dari sampel cukup tinggi.

Menurut Cao (2005:7), *face recognition* merupakan proses penganalisa karakteristik dari bentuk muka yang tidak berubah, seperti:

- Bagian atas dari rongga mata
- Area di sekitar tulang pipi

- Sisi kiri dan kanan dari mulut

Kesulitan dalam pengenalan wajah sering ditemukan pada:

- *Noise* dan *blur* yang disebabkan oleh ketidaksempurnaan kamera
- Skala: Ukuran wajah terhadap citra
- Perubahan bentuk: Posisi wajah, ekspresi, usia
- Intensitas cahaya: Pencahayaan, efek pantulan sinar
- Gangguan: kacamata, janggut, dan kumis

Menurut Turk (2005:3), pengenalan wajah bisa dilihat sebagai suatu cara untuk secara tepat mengenali citra dari sebuah wajah, dengan menggunakan data-data dari wajah yang telah lebih dahulu dikenal. Pengenalan wajah memiliki semua hambatan dalam pengenalan yang berdasarkan pemrosesan citra. Dikarenakan citra yang digunakan dapat berubah-ubah secara drastis disebabkan beberapa faktor yang rumit dan membingungkan, seperti faktor pencahayaan, posisi kamera, pengaturan kamera, dan *noise*.

Jadi hasil dari *face recognition* akan berupa informasi yang dikenal atau tidak sebagai wajah dengan sebelumnya membandingkan dengan informasi dari wajah yang diketahui. Proses *face recognition* ini memiliki permasalahan dari pencahayaan, posisi kamera, parameter kamera dan *noise* yang didapatkan pada sebuah citra.

2.4 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) digunakan untuk menentukan fitur dimensi rendah dari ruang dimensi tinggi yang membantu untuk foto grup dari kelas yang sama dan gambar dari kelas yang berbeda. LDA memilih fitur yang memaksimalkan rasio dari antara kelas dan menyebar ke dalam kelas. Menurut Etemad & Chellappa (1997:4), didefinisikan sebagai :

$$S_b = \sum_{i=1}^C P_i (\mu_i - \mu) (\mu_i - \mu)^T \quad (1)$$

Dimana μ adalah mean dari semua sampel dan μ_i adalah mean gambar dari kelas C_i dengan probabilitas P_i sebelumnya. S_b yang dimaksud adalah *between class scatter matrix*. Dan *within class scatter matrix* S_w didefinisikan sebagai:

$$S_w = \sum_{i=1}^C P_i S_i \quad (2)$$

Dimana

$$S_i = E[(\mu_i - \mu)(\mu_i - \mu)^T | x \in C_i] \quad (3)$$

Jika S_w non-tunggal, proyeksi optimal dipilih dari matriks dengan kolom ortonormal yang memaksimalkan rasio dari determinan *between class scatter matrix* dengan determinan dari *within class scatter matrix*.

$$J(W \text{ opt}) = \arg \max_W \frac{|W^T S_b W|}{|W^T S_w W|} \quad (4)$$

Dimana W_{opt} adalah himpunan vektor-vektor eigen umum dari dan S_w sesuai dengan nilai eigen terbesar λ_i :

$$S_b W_i = \lambda_i S_w W_i \quad (5)$$

Kebanyakan $C-1$ *nonzero* menyimpulkan nilai eigen, dimana C adalah jumlah kelas.

LDA secara luas digunakan untuk menemukan kombinasi linear fitur sambil menjaga kelas keterpisahan. Tidak seperti PCA, LDA mencoba untuk memodelkan perbedaan antara kelas. LDA klasik dirancang untuk memperhitungkan hanya dua kelas. Secara khusus, memerlukan titik data untuk kelas yang berbeda untuk menjadi jauh dari satu sama lain, sementara point dari kelas yang sama dekat. Akibatnya, LDA memperoleh perbedaan vektor proyeksi untuk masing-masing kelas. Algoritma multi-kelas LDA yang dapat mengelola lebih dari dua kelas yang lebih digunakan. Misalkan kita mempunyai m sampel x_1, \dots, x_m milik kelas c ; masing-masing kelas memiliki m_k elemen. Kita mengasumsikan bahwa *mean* telah diekstraksi dari sampel, seperti di PCA. Fungsi objektif LDA dapat didefinisikan sebagai

$$a_{opt} = \operatorname{argmax} \frac{a^T S_b a}{a^T S_t a} \quad (6)$$

$$S_b = \sum_{k=1}^c m_k \mu^{(k)} (\mu^{(k)})^T = \sum_{k=1}^c \left(\frac{1}{m_k} \left(\sum_{i=1}^{m_k} x_i^{(k)} \right) \right) \left(\frac{1}{m_k} \left(\sum_{i=1}^{m_k} x_i^{(k)} \right) \right)^T = X W_{m \times m} X^T \quad (7)$$

$$S_t = \sum_{i=1}^m x_i (x_i)^T = X X^T \quad (8)$$

Dimana $W_{m \times m}$ adalah diagonal matriks yang didefinisikan sebagai

$$W_{maxm} = \begin{bmatrix} W^1 & 0 & \dots & 0 \\ 0 & W^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & W^c \end{bmatrix} \quad (9)$$

Dan W^k adalah matriks $m_k \times m_k$

$$W^k = \begin{bmatrix} \frac{1}{m_k} & \frac{1}{m_k} & \dots & \frac{1}{m_k} \\ \frac{1}{m_k} & \frac{1}{m_k} & \dots & \frac{1}{m_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{m_k} & \frac{1}{m_k} & \dots & \frac{1}{m_k} \end{bmatrix} \quad (10)$$

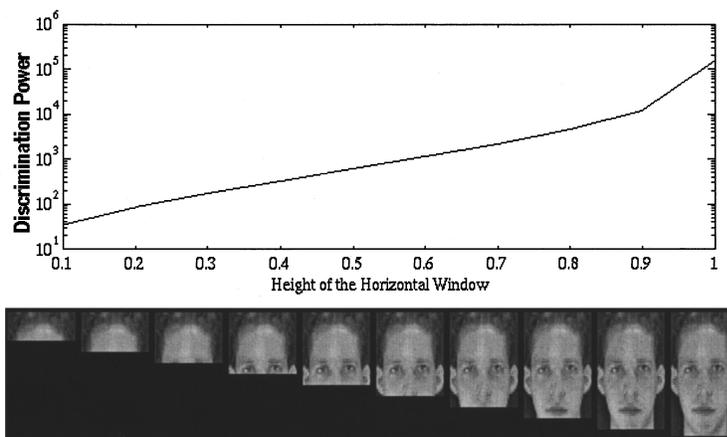
Akhirnya, kita dapat menulis *eigenproblem*:

$$S_b a = \lambda S_t a \rightarrow S_t^{-1} S_b a = \lambda a \rightarrow X W_{\text{int}} X^T (X X^T)^{-1} a = \lambda a \quad (11)$$

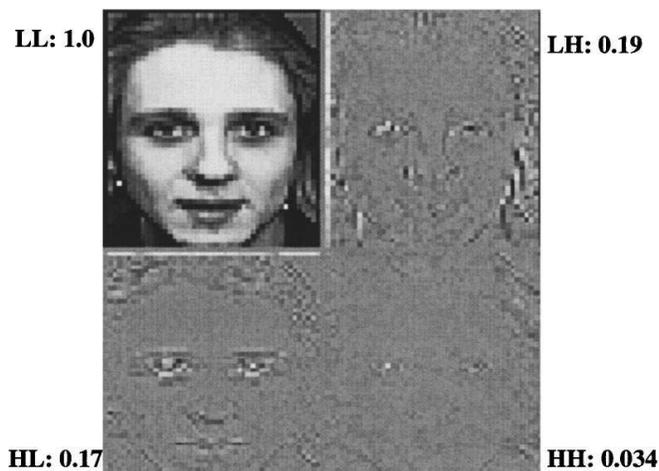
Berdasarkan pembahasan dari para ahli, terdapat beberapa masalah dalam pengenalan wajah. Salah satunya adalah dihadapkan pada kesulitan bahwa *within class scatter matrix* S_w biasanya tunggal, karena jumlah gambar pada set latihan jauh lebih kecil jika dibandingkan dengan jumlah *pixel* pada setiap gambar, yang disebut ukuran sampel masalah yang kecil. Untuk mengatasi masalah ini digunakan algoritma Chen: Memaksimalkan antar kelas dan penyebaran gambar yang diproyeksikan dikenakan batasan bahwa tersebarnya di dalam kelas adalah nol.

Telah diketahui bahwa tujuan dari LDA adalah untuk memaksimalkan *between class scatter matrix* sekaligus meminimalkan tersebarnya di dalam kelas.

Ketika masalah ukuran terjadi, matriks S_w *within class scatter matrix* adalah tunggal, maka teori Aljabar Linear memberitahu kita adanya kemungkinan untuk menemukan beberapa vektor proyeksi w seperti $w^T S_w w = 0$ dan $w^T S_b w \neq 0$. Dalam hal ini kriteria Fisher bisa pasti mencapai nilai maksimum. Jika lebih lanjut memaksimalkan nilai $w^T S_b w$, kita dapat mendapatkan proyeksi vektor dengan *discriminability* kelas maksimal. Oleh karena itu, metode baru ini dimulai dengan menghitung proyeksi vektor yang mengakibatkan $w^T S_w w = 0$, jika tidak ada, berarti *within class scatter matrix* adalah non-tunggal, LDA tradisional dapat dimanfaatkan. Jika tidak, Kriteria Fisher dapat dipenuhi dengan memilih seperti w yang memaksimalkan $w^T S_b w$. Metode ini juga telah disebutkan dalam (2) (Bing, Lianfu & Ping, 2002)



Gambar 2.2 : Variasi kekuatan diskriminasi dari segmen horisontal wajah yang berkembang di ketinggian dari atas ke bawah gambar (Etemad & Chellappa, 1997:3).



Gambar 2.3 : Komponen yang berbeda dari transformasi *wavelet* menangkap bahwa variasi mencolok dari intensitas gambar dalam arah yang berbeda dan memiliki potensi diskriminasi yang berbeda. angka-angka mewakili kekuatan relatif diskriminasi (Etemad & Chellappa, 1997:3).

Analisis diskriminan linear atau metode *fisherface* mengatasi keterbatasan metode *eigenfaces* dengan menerapkan kriteria diskriminan linier Fisher. Kriteria ini mencoba untuk memaksimalkan rasio determinan dari sampel *between class scatter matrix* diproyeksikan sebagai determinan *within class scatter matrix* dari sampel yang diproyeksikan.

Kelompok gambar diskriminan *fisher* dari kelas yang sama dan memisahkan gambar dari kelas yang berbeda. Gambar diproyeksikan dari N^2 -dimensi ruang ke ruang dimensi C (di mana C adalah jumlah kelas gambar). Sebagai contoh, mempertimbangkan dua set poin dalam 2-dimensi ruang yang diproyeksikan ke satu baris. Tergantung pada arah garis, titik-titik dapat dicampur bersama-sama atau terpisah. Diskriminan Fisher menemukan baris

yang paling memisahkan poin. Untuk mengidentifikasi gambar tes masukan, gambar tes proyeksi dibandingkan dengan setiap gambar diproyeksikan pelatihan, dan citra uji diidentifikasi sebagai citra pelatihan terdekat.

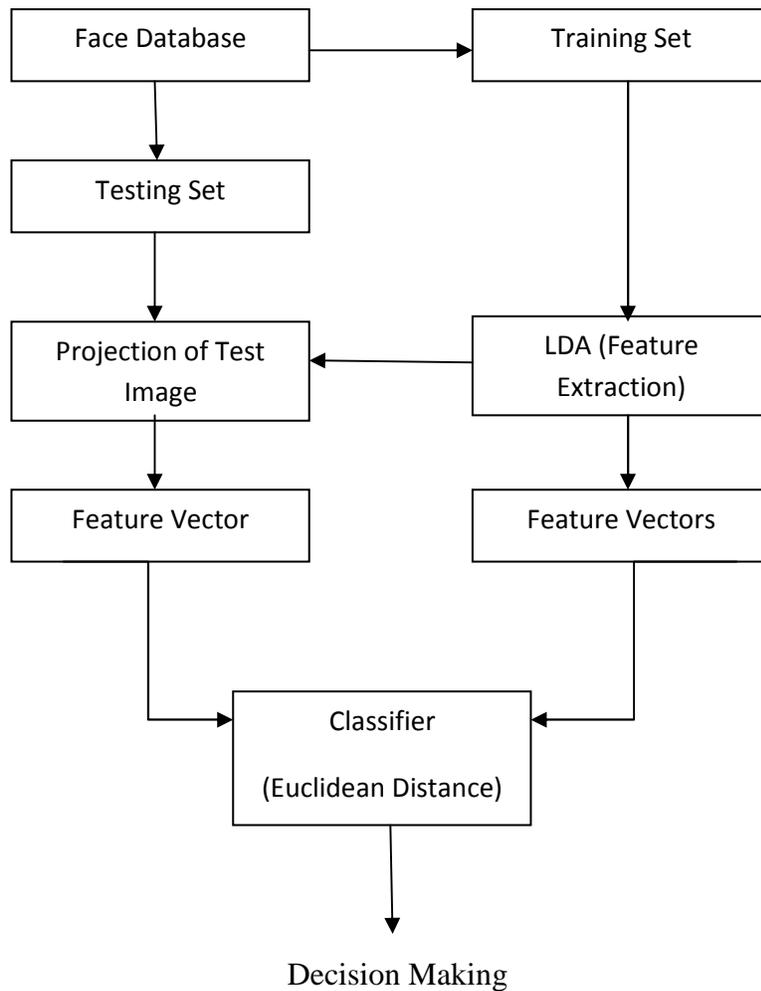
Seperti dengan proyeksi eigen, gambar pelatihan diproyeksikan ke dalam sebuah *subspace*. *Test Images* diproyeksikan ke dalam *subspace* yang sama dan diidentifikasi menggunakan kesamaan ukuran. Yang berbeda adalah bagaimana *subspace* dihitung.

Tidak seperti metode PCA yang mengutip fitur terbaik untuk mewakili gambar wajah, metode LDA mencoba untuk menemukan *subspace* yang paling membedakan kelas wajah yang berbeda seperti yang ditunjukkan. *The within class scatter matrix*, juga disebut intra-personal, mewakili variasi dalam tampilan dari individu yang sama karena pencahayaan yang berbeda dan ekspresi wajah, sedangkan *between class scatter matrix*, juga disebut extra-personal, mewakili variasi dalam tampilan karena perbedaan identitas. Dengan menerapkan metode ini, kita menemukan arah proyeksi yang di satu sisi memaksimalkan jarak antara gambar wajah kelas yang berbeda di sisi lain meminimalkan jarak antara gambar wajah dari kelas yang sama. Dengan kata lain, memaksimalkan *between class scatter matrix* S_b , dan meminimalkan *within class scatter matrix* S_w di *subspace* proyektif.

The within class scatter matrix mewakili bagaimana gambar wajah didistribusikan erat *within classes* dan *between class scatter matrix* menjelaskan bagaimana kelas dipisahkan dari satu sama lain. Ketika citra wajah

diproyeksikan ke diskriminan vektor W , gambar wajah harus didistribusikan erat dalam kelas dan harus dipisahkan antara kelas, sebanyak mungkin. Dengan kata lain, vektor diskriminan ini meminimalkan penyebut dan memaksimalkan pembilang dalam rumus. W sehingga dapat dibangun oleh vektor eigen dari $S_w^{-1} S_b$. Vektor eigen ini juga disebut sebagai fisherface. Ada berbagai metode untuk memecahkan masalah LDA seperti metode *pseudo inverse*, metode *subspace*, atau metode *nullspace*.

Pendekatan LDA mirip dengan metode Eigenface, yang menggunakan proyeksi gambar pelatihan ke sub ruang. *Test Images* diproyeksikan ke dalam *subspace* yang sama dan diidentifikasi menggunakan ukuran kesamaan. Satu-satunya perbedaan adalah metode penghitungan sub ruang mencirikan ruang wajah. Wajah yang memiliki jarak minimum dengan gambar *Test Images* diberi label dengan identitas citra tersebut. Jarak minimum dapat dihitung dengan menggunakan metode jarak Euclidian seperti yang diberikan sebelumnya. Gambar 2.4 menunjukkan fase pengujian pendekatan LDA menurut Eleyan & Demirel (2007:8).



Gambar 2.4 Pendekatan LDA untuk pengenalan wajah

2.5 Euclidean Distance

Euclidean Distance adalah metode yang sering digunakan untuk menghitung jarak antara 2 vektor. *Euclidean Distance* digunakan untuk mengklasifikasi data (*feature vector*). *Euclidean distance* menghitung akar dari kuadrat perbedaan 2 vektor (Teknomo, 2008).

Rumus *euclidean distance* adalah sebagai berikut:

$$D_{ij} = \sqrt{\sum_{k=1}^n (\mu_{ik} - \mu_{jk})^2} \quad (12)$$

μ_{ik} = Koordinat ke k dari vektor i.

μ_{jk} = Koordinat ke k dari vektor j.