Pengenalan Jenis Kelamin Berdasarkan Citra Wajah Menggunakan Metode Two-Dimensional Linear Discriminant Analysis

Fitri Damayanti
Prodi Manajemen Informatika, Fakultas Teknik, Universitas Trunojoyo Madura
Jl. Raya Telang PO. BOX 2 Kamal, Bangkalan, Madura, 69162
e-mail: fitri2708@yahoo.com

Abstrak

Wajah merupakan salah satu bagian dari tubuh manusia yang mempunyai karakteristik khusus sehingga sering digunakan untuk membedakan identitas individu yang satu dan yang lainnya. Pengenalan wajah menjadi sangat penting untuk dikembangkan sejak aplikasi ini diterapkan dalam sistem keamanan. Pengenalan jenis kelamin merupakan salah satu bagian dari pengenalan wajah. Jenis kelamin memainkan peran penting dalam interaksi kita di masyarakat dan dengan komputer. Klasifikasi jenis kelamin dari gambar wajah dapat diaplikasikan dibidang demografi pengumpulan data, antarmuka manusia-komputer (menyesuaikan perilaku software sehubungan dengan jenis kelamin pengguna) dan lain-lain. Tujuan dari penelitian ini adalah membuat implementasi dari sistem pengenalan jenis kelamin pada citra wajah atau Gender Recognition dengan masukkan berupa citra wajah yang mampu mengenali jenis kelamin seseorang secara cepat dan tepat serta berjalan dengan baik.

Pada penelitian ini digunakan metode Two Dimensional Linear Discriminant Analysis (TDLDA) untuk ekstraksi fitur, yang menilai secara langsung matrik within-class scatter dari matriks citra tanpa transformasi citra ke vektor, dan hal ini mengatasi singular problem dalam matriks within-class scatter. Untuk mendapatkan hasil pengenalan yang optimal metode klasifikasi yang digunakan adalah pengklasifikasian Euclidean Distance. Penelitian ini mengintegrasikan metode TDLDA dan ED untuk pengenalan jenis kelamin berdasarkan citra wajah. Dengan kombinasi kedua metode tersebut terbukti dapat memberikan hasil yang optimal dengan akurasi antara 68% sampai 89% dengan uji coba yang menggunakan database wajah yang diambil dari http://www.advancedsourcecode.com.

Kata kunci: *Euclidean Distance., Jenis Kelamin, Two Dimensional Linear Discriminant Analysis*

1. Pendahuluan

Wajah merupakan salah satu ukuran fisiologis yang paling mudah dan sering digunakan untuk membedakan identitas individu yang satu dengan yang lainnya. Otak manusia mempunyai kemampuan untuk mengenal dan membedakan wajah antara orang yang satu dengan yang lainnya dengan relatif cepat dan mudah. Pengenalan wajah (*face recognition*) manusia merupakan salah satu bidang yang cukup berkembang dewasa ini. Aplikasi dari pengenalan wajah dapat diterapkan dalam bidang keamanan (*security system*) seperti ijin akses masuk ruangan.

Salah satu bagian dari *face recognition* yang telah dikembangkan saat ini adalah pengenalan jenis kelamin (*gender recognition*) Kemiripan antara gender recognition dengan face recognition terletak pada proses ekstraksi fiturnya. Namun, sedikit berbeda pada proses pengklasifikasiannya. Selama ini untuk menghitung berapa jumlah orang yang berjenis kelamin laki-laki atau perempuan yang datang ke sebuah toko atau instansi publik masih secara manual, sehingga dibutuhkan waktu yang lebih lama. Untuk memudahkan iklan apa yang ditampilkan di papan iklan elektronik ditempat umum atau dipinggir jalan bisa disesuaikan dengan jenis kelamin orang yang melewati iklan tersebut. Sehingga untuk memudahkan dan mempercepat waktu pemrosesan dibuat perangkat lunak untuk pengenalan jenis kelamin berdasarkan citra wajah.

Kesuliltan dalam proses *gender recognition* terutama karena kekompleksan dari kondisi wajah, seperti posisi gambar, pencahayaan dan ekspresi gambar yang berbeda-beda yang memiliki dimensi tinggi sehingga harus melalui proses kompresi/ekstraksi terlebih dahulu sebelum diolah datanya dengan metode klasifikasi.

Penelitian sebelumnya yang terkait dengan penelitian ini adalah penelitian yang dilakukan oleh Burhan Ergen dan Serdar Abut yang berjudul "*Gender Recognition Using Facial Images*". Pada

penelitian tersebut dilakukan pengenalan jenis kelamin berbasis citra wajah menggunakan metode *GLCM*. Hasil uji coba mempunyai tingkat akurasi 60% dengan menggunakan *FEI Face database* yang terdiri dari 100 citra wajah perempuan dan 100 citra wajah laki-laki [1].

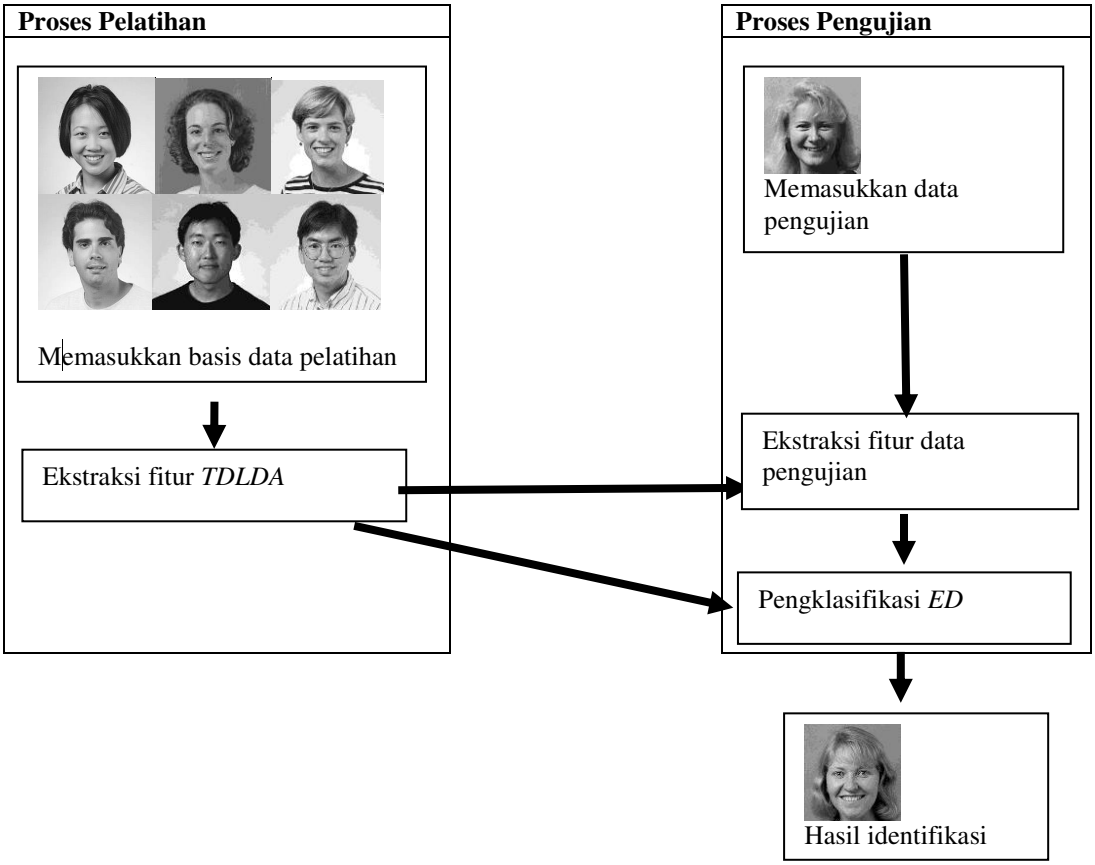
Penelitan yang dilakukan oleh Vladimir Khryashchev, Andrey Priorov, Lev Shmaglit and Maxim Golubev yang berjudul "*Gender Recognition Via Face Area Analysis*". Pada penelitian ini dilakukan pengenalan jenis kelamin berbasis citra wajah menggunakan metode *Adaptive Feature* dan *SVM*. Hasil uji coba mempunyai tingkat akurasi 90,8% dengan menggunakan *FERRET database*[2].

Penelitian yang dilakukan oleh Nafin Fenanda, Arif Muntasa, Fitri Damayanti yang berjudul "Pengenalan Jenis Kelamin Berdasarkan Citra Wajah Menggunakan Metode *Local Binary Pattern* (LBP) dan *Fisherface*". Pada penelitian ini dilakukan pengenalan jenis kelamin berbasis citra wajah menggunakan metode *LBP* dan *Fisherface*. Hasil uji coba mempunyai tingkat akurasi tertinggi sebesar 75% dengan menggunakan database yang diambil dari http://www.advancedsourcecode.com[3].

Penelitian ini mengintegrasikan *TDLDA* dan *ED* untuk pengenalan jenis kelamin berbasis citra wajah. *TDLDA* sebagai metode ekstraksi fitur yang bisa mengatasi *singular problem* dan *ED* sebagai metode klasifikasinya.

2. Metode Penelitian

Secara garis besar sistem ini terdiri dari dua bagian, yaitu proses pelatihan citra dan proses pengujian. Pada Gambar 1 merupakan gambaran garis besar sistem pengenalan jenis kelamin berdasarkan citra wajah. Pada proses pelatihan terdapat proses *TDLDA* yang digunakan untuk mengekstraksi fitur, fitur-fitur yang terpilih pada saat proses pelatihan digunakan dalam proses klasifikasi dan juga digunakan untuk mendapatkan fitur-fitur yang terpilih pada data uji coba. Masing-masing basis data wajah yang digunakan dibagi menjadi dua, sebagian digunakan untuk proses pelatihan (*training*) dan sisanya digunakan untuk proses pengujian (*testing*).



Gambar 1. Sistem Pengenalan Jenis Kelamin Berdasarakan Citra Wajah.

2.1 Ekstraksi Fitur

Ekstraksi fitur pada proses pelatihan dilakukan dengan menggunakan metode *Two-Dimensional Linear Discriminant Analysis*. Tahap ini bertujuan untuk mendapatkan fitur-fitur yang terpilih dari

masukkan data-data pelatihan. Fitur-fitur yang terpilih diperoleh dari semua fitur wajah yang dicari nilai eigen dan vektor eigen terbesar. Fitur-fitur yang terpilih nantinya digunakan untuk proses klasifikasi pelatihan dan digunakan untuk ekstraksi fitur data pengujian.

Ekstraksi fitur pada proses pengujian dilakukan dengan mengambil hasil ekstraksi fitur pada proses pelatihan diterapkan pada data pengujian. Hasil ekstraksi fitur pada data pengujian ini nantinya digunakan sebagai inputan pada proses klasifikasi pengujian.

2.2 Desain Algoritma TDLDA

Berikut ini adalah langkah-langkah dalam proses TDLDA terhadap suatu basis data citra pelatihan [4]:

1. Jika dalam suatu basis data citra wajah terdapat himpunan sebanyak *n* citra pelatihan $A_i = [A_1, A_2, \ldots, A_n]$ $(i = 1,2,\ldots,n)$ dengan dimensi citra (r x c), maka himpunan total matriks dari semua citra tersebut adalah :

$$A_n = \begin{bmatrix} A_{(n)11} & A_{(n)12} & \cdots & A_{(n)1c} \\ A_{(n)21} & A_{(n)22} & \cdots & A_{(n)2c} \\ \cdots & \cdots & \cdots & \cdots \\ A_{(n)r1} & A_{(n)r2} & \cdots & A_{(n)rc} \end{bmatrix} \quad (1)$$

2. Menentukan nilai ℓ_1 (dimensi proyeksi baris) dan ℓ_2 (dimensi proyeksi kolom). Nilai $\ell_1 \leq r$ dan $\ell_2 \leq c$.
3. Tahapan berikutnya adalah perhitungan rata-rata citra pelatihan dari kelas ke *i* :

$$M_i = \frac{1}{n_i} \sum_{X \in \Pi_i} X \quad (2)$$

4. Menghitung rata-rata semua citra pelatihan :

$$M = \frac{1}{n} \sum_{i=1}^{k} \sum_{X \in \Pi_i} X \quad (3)$$

5. Menetapkan matriks transformasi *R* ukuran (c, ℓ_2) yang diperoleh dari gabungan antara matriks identitas ukuran (ℓ_2, ℓ_2) dengan matriks nol ukuran (c-ℓ_2, ℓ_2).
6. Menghitung matriks *between class scatter R* sesuai dengan persamaan (4).

$$S_b^R = \sum_{i=1}^{k} n_i (M_i - M) R R^T (M_i - M)^T \text{ , ukuran matriksnya } (r \text{ x } r). \quad (4)$$

7. Menghitung matriks *within class scatter R* sesuai dengan persamaan (5).

$$S_W^R = \sum_{i=1}^{k} \sum_{x \in \Pi_i} (X - M_i) R R^T (X - M_i)^T, \text{ ukuran matriksnya } (r \text{ x } r). \quad (5)$$

8. Hitung *generalized eigenvalue* (λ_i) dari S_b^R dan S_W^R menggunakan SVD sesuai dengan persamaan (6).

$$J_4(L) = \text{maxtrace}((L^T S_W^R L)^{-1} (L^T S_b^R L)), \text{ ukuran matriksnya } (r \text{ x } r). \quad (6)$$

9. Ambil sebanyak ℓ_1 *eigenvector* dari langkah 8 sebagai matriks transformasi baris (*L*). $L = [\phi_1^L, \ldots, \phi_{\ell_1}^L]$, ukuran matriksnya ($r$ x ℓ_1).
10. Menghitung matriks *between class scatter L* sesuai dengan persamaan (7).

$$S_b^L = \sum_{i=1}^{k} n_i (M_i - M)^T L L^T (M_i - M), \text{ ukuran matriksnya } (c \text{ x } c). \quad (7)$$

11. Menghitung matriks *within class scatter L* sesuai dengan persamaan (8).

$$S_W^L = \sum_{i=1}^{k} \sum_{x \in \Pi_i} (X - M_i)^T L L^T (X - M_i), \text{ukuran matriksnya } (c \text{ x } c). \quad (8)$$

12. Hitung *generalized eigenvalue* (λ_i) dari S_b^L dan S_W^L menggunakan SVD sesuai dengan persamaan (9).

    $J_5(R)=maxtrace((R^T S_W^L R)^{-1}(R^T S_b^L R))$, ukuran matriksnya ($c$ x c). (9)

13. Ambil sebanyak ℓ_2 *eigenvector* dari langkah 12 sebagai matriks transformasi kolom (R). $R = [\phi_1^R$, ..., $\phi_{\ell_2}^R]$, ukuran matriksnya (c x ℓ_2).
14. Hitung matriks fitur ekstraksi adalah $B_i=L^T A_i R$, ukuran matriksnya (ℓ_1 x ℓ_2)
15. Output : matriks fitur ektraksi B_i, matriks transformasi baris L, dan matriks transformasi kolom R.

2.3 Klasifikasi

Klasifikasi pada penelitian ini menggunakan metode *Euclidean Distance*. Metode ini merupakan salah satu pengukuran kemiripan citra dengan konsep menghitung nilai kuadrat dari hasil pengurangan kedua citra [5].

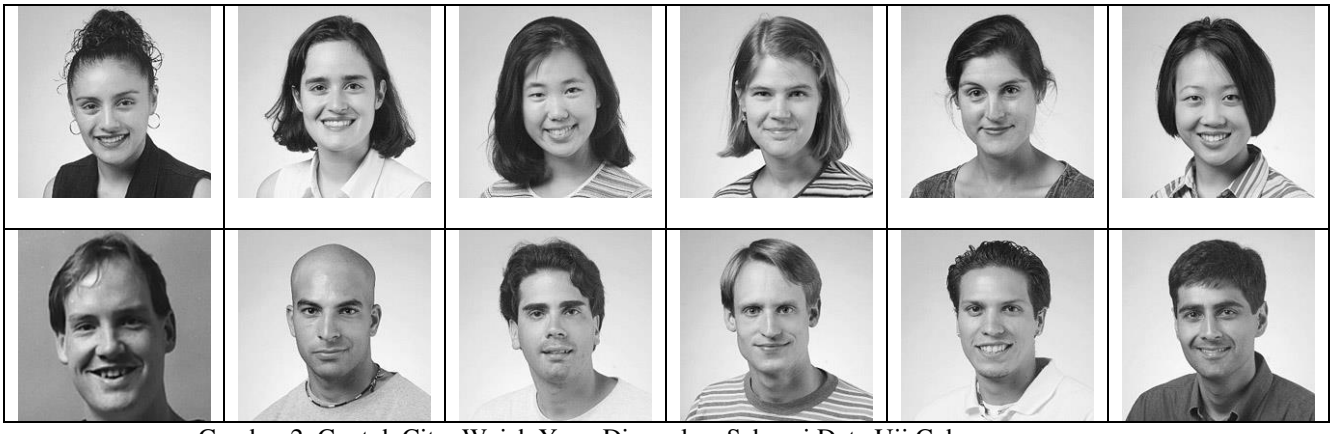
$$d = \sqrt{(\sum_{i=1}^{m}\sum_{j=1}^{n}(X_{i,j} - Y_{i,j})^2)} \quad (10)$$

dengan :

d = tingkat kesamaan objek
X = matrik bobot citra *testing*
Y = matrik bobot citra *training*
i, j = koordinat matrik bobot
m = dimensi panjang matrik bobot
n = dimensi lebar matrik bobot

2.4 Data Yang Digunakan

Pengujian sistem pada penelitian ini menggunakan data uji coba berukuran 200x200 piksel sebanyak 400 citra yang terdiri dari 200 citraa laki-laki dan 200 citra perempuan. Data uji coba ini diambil dari http://www.advancedsourcecode.com dan telah digunakan dalam penelitian sebelumnya. Data uji coba tersebut merupakan data *testing* yang diuji jarak kemiripannya menggunakan *ED* dengan data *training*. Gambar 2 berisi beberapa contoh citra wajah yang digunakan sebagai data *testing* dan data *training*.



Gambar 2. Contoh Citra Wajah Yang Digunakan Sebagai Data Uji Coba

Skenario uji coba yang dilakukan pada penelitian ini terdapat 8 skenario, yaitu skenario 1, skenario 2, skenario 3, skenario 4, skenario 5, skenario 6, skenario 7, skenario 8. Perbedaan dari 8 skenario tersebut terletak pada jumlah data *training* dan data *testing* yang digunakan. Lebih jelasnya dapat dilihat pada Tabel 1.

Tabel 1 Simulasi Skenario Pada Sistem

Skenario	Jumlah Citra Training		Jumlah Citra Testing	
Skenario 1	300 Citra	150 Citra Perempuan 150 Citra Laki-laki	100 Citra	50 Citra Perempuan 50 Citra Laki-laki
Skenario 2	280 Citra	140 Citra Perempuan 140 Citra Laki-laki	120 Citra	60 Citra Perempuan 60 Citra Laki-laki
Skenario 3	260 Citra	130 Citra Perempuan 130 Citra Laki-laki	140 Citra	70 Citra Perempuan 70 Citra Laki-laki
Skenario 4	240 Citra	120 Citra Perempuan 120 Citra Laki-laki	160 Citra	80 Citra Perempuan 80 Citra Laki-laki
Skenario 5	160 Citra	80 Citra Perempuan 80 Citra Laki-laki	240 Citra	120 Citra Perempuan 120 Citra Laki-laki
Skenario 6	140 Citra	70 Citra Perempun 70 Citra Laki-laki	260 Citra	130 Citra Perempuan 130 Citra Laki-laki
Skenario 7	120 Citra	60 Citra Perempuan 60 Citra Laki-laki	280 Citra	140 Citra Perempuan 140 Citra Laki-laki
Skenario 8	100 Citra	50 Citra Perempuan 50 Citra Laki-laki	300 Citra	150 Citra Perempuan 150 Citra Laki-laki

3. Hasil dan Pembahasan

Metode yang digunakan dalam pengujian ini ada dua kelompok. Kelompok pertama menggunakan metode *Local Binary Pattern* (LBP) untuk *preprocessing*, metode *Fisherface* untuk ekstraksi fitur dan metode *Euclidean Distance* sebagai klasifikasi penelitian ini yang dikerjakan oleh Nafin Fenanda, Arif Muntasa, Fitri Damayanti [3]. Kelompok yang kedua menggunakan metode *TDLDA* sebagai ekstraksi fitur dan metode *Euclidean Distance* sebagai klasifikasi. Kelompok yang kedua ini merupakan penelitian yang dikerjakan oleh peneliti. Tabel 2 menunjukkan perbandingan hasil uji coba dari penelitian sebelumnya dan penelitian yang dikerjakan oleh peneliti. Data uji coba dan skenario yang dilakukan pada penelitian ini sudah dijelaskan pada bab 3.

Tabel 2 Perbandingan Hasil Uji Coba

Skenario	Akurasi	
	LBP – Fisherface - ED	*TDLDA - ED*
Skenario 1	75%	89%
Skenario 2	68,33%	88,75%
Skenario 3	67,86%	88%
Skenario 4	73,75%	86%
Skenario 5	60,83%	84%
Skenario 6	49,62%	83%
Skenario 7	45%	70%
Skenario 8	52%	68%

Pada Tabel 2 menunjukkan bahwa prosentase pengenalan *TDLDA-ED* lebih tinggi dibanding dengan metode *LBP-Fisherface-ED*. Keunggulan metode *TDLDA-ED* dibandingkan dengan metode *LBP-Fisherface-ED*, adalah pada *fisherface* prosedur *pre-processing* untuk mereduksi dimensi menggunakan *PCA* dapat menyebabkan kehilangan beberapa informasi diskriminan yang penting untuk algoritma *LDA* yang diterapkan setelah *PCA*. Pada *TDLDA* mengambil keuntungan penuh dari informasi yang diskriminatif dari ruang lingkup wajah (*face space*), dan tidak membuang beberapa subruang (*subspace*) yang mungkin berguna untuk pengenalan.

Terdapat tiga variabel penting yang mempengaruhi tingkat keberhasilan pengenalan, yaitu variasi urutan dari sampel pelatihan per kelas yang digunakan, jumlah sampel pelatihan per kelas yang digunakan, dan jumlah dimensi proyeksi. Pada skenario 1 jumlah data pelatihan lebih banyak dibanding skenario 2, demikian juga pada skenario 2 jumlah data pelatihan lebih banyak dibanding skenario 3, dan seterusnya. Sehingga prosentase pengenalan untuk skenario 1 lebih tinggi hasilnya dibandingkan hasil dari skenario 2.

Dari hasil uji coba yang dilakukan terdapat beberapa pengenalan yang salah, harusnya dikenali perempuan tetapi hasil pengenalannya laki-laki, demikian juga sebaliknya. Beberapa pengenalan yang

salah tersebut disebabkan karena beberapa faktor yaitu miripnya bentuk kepala, bentuk rambut dan ekspresi antara kedua citra perempuan dan laki-laki.

4. Simpulan

Dari uji coba yang sudah dilakukan dapat diambil simpulan sebagai berikut :

1. Metode *TDLDA-ED* mampu menunjukkan akurasi pengenalan yang optimal dibandingkan dengan metode *LBP-Fisherface-ED*. Hal ini dikarenakan *TDLDA* mampu mengatasi *singular problem*, mampu mempertahankan keberadaan informasi diskriminatif, serta mampu memaksimal jarak antar kelas dan meminimalkan jarak inter kelas.
2. Terdapat tiga variabel penting yang mempengaruhi tingkat keberhasilan pengenalan, yaitu variasi urutan dari sampel pelatihan per kelas yang digunakan, jumlah sampel pelatihan per kelas yang digunakan, dan jumlah dimensi proyeksi.
3. Dari hasil uji coba menggunakan metode *TDLDA-ED* didapatkan tingkat akurasi pengenalan antara 68% sampai 89%.
4. Klasifikasi yang salah pada uji coba disebabkan oleh miripnya bentuk kepala, bentuk rambut dan ekspresi antara kedua citra perempuan dan laki-laki.

Daftar Pustaka

[1] Ergen B, Abut S. *Gender Recognition Usiang Facial Images*. International Conference on Agriculture and Biotechnology. 2013. Vol 60 : 112-117.

[2] Khryashchev V, Priorov A, Shmaglit L, and Golubev M. *Gender Recognition Via Face Area Analysis*. Proceedings of The World Congress on Engineering and Compuer Science. 2012. Vol 1 .

[3] Fenanda N, Muntasa Arif, Damayanti F. Pengenalan Jenis Kelamin Berdasarkan Citra Wajah Menggunakan Metode Local Binary Pattern (LBP) dan Fisherface. Bangkalan. Teknik Informatika Universitas Trunojoyo. 2015.

[4] Quan XG, Lei Z, and David Z. Face Recognition Using FLDA With Single Training Image Per Person. *Applied Mathematics and Computation*. 2008. Vol 205 : 726-734.

[5] Ballihi L, Ben Amar B, Daoudi M, Srivastava A, Aboutajdine D. Boosting 3-D-Geometic Feature for Efficient Face Recognition and Gender Classification. *IEEE Transaction on Biometrics Compendium*. 2012. Vol 7 : 1766-1779.