

# MULTIMODAL FACIAL GENDER DAN IDENTIFIKASI ETNIK

Jalaluddin<sup>1</sup>, Fadlisyah<sup>2</sup>

<sup>1</sup>Dosen Teknik Kimia Universitas Malikussaleh

<sup>2</sup>Dosen Teknik Informatika Universitas Malikussaleh

## Abstract

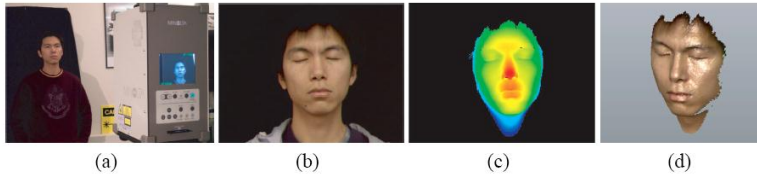
*Human faces provide demographic information, such as gender and ethnicity. Different modalities of human faces, e.g., range and intensity, provide different cues for gender and ethnicity identifications. In this paper we exploit the range information of human faces for ethnicity identification using a support vector machine. An integration scheme is also proposed for ethnicity and gender identifications by combining the registered range and intensity images. The experiments are conducted on a database containing 1240 facial scans of 376 subjects. It is demonstrated that the range modality provides competitive discriminative power on ethnicity and gender identifications to the intensity modality. For both gender and ethnicity identifications, the proposed integration scheme outperforms each individual modality.*

## Pendahuluan

Face manusia memuat berbagai informasi untuk interaksi sosial yang adaptif dengan manusia lainnya. Salah satu informasi yang dimiliki oleh keseluruhan manusia adalah karakteristik demografi seperti gender, etnis, dan bahkan usia. Tidak seperti gender, etnis memiliki suatu definisi yang begitu luas yang berkaitan dengan ras manusia dan variasi alami dari ras tersebut. Di dalam kajian kita di dalam bab ini, permasalahan etnis dibatasi menjadi dua kategori pengklasifikasian, *Asian* dan *non-Asian*. Statistik antropometris telah menunjukkan bahwa etnis memiliki suatu perbedaan *craniofacial*

*morphometric* dan relasi tertutup antara bentuk 3D face manusia dengan sifat keetnisannya.

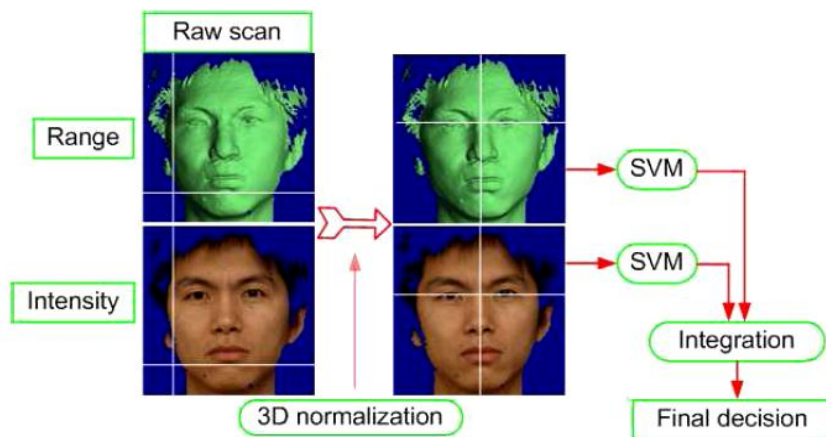
Kita khususkan permasalahan pengidentifikasian gender dan etnis di sini dengan hanya menggunakan dua sifat keadaan face yang berbeda saja, yaitu, range dan intensitas. Dengan kecanggihan teknologi imaging yang ada sekarang, suatu alat sensor 3D yang bersifat komersial tidak hanya mampu menangkap informasi range saja pada citra, tetapi juga dapat memberikan suatu informasi intensitas yang dapat kita manfaatkan untuk riset kita di dalam pengidentifikasian gender dan etnis yang menjadi tema kita di dalam bab ini. Lihat gambar 1, yang merupakan contoh - contoh dari suatu scan face yang diperoleh dari suatu alat sensor Minolta vivid 910.



Gambar 1. Scan face yang diambil oleh Minolta vivid 910, (a) skenario data - capture, (b) citra texture, (c) citra range, titik - titik berwarna merah mewakili domain titik - titik yang dekat terhadap fokus sensor, (d) visualisasi 3D

## Metode

Arsitektur sistem diilustrasikan pada gambar 2. Keseluruhan citra range dinormalisasikan di dalam ruang 3D, dan juga untuk keseluruhan citra intensitas. Sementara, data di dalam suatu region khusus dipangkas dari citra intensitas dan citra range. Dua modul SVM selanjutnya menklasifikasi data intensitas dan data range yang telah dipangkas, dan terakhir hasil dari pengklasifikasian disatukan untuk meraih keputusan final.



Gambar 2. Diagram sistem untuk pengidentifikasian gender dan etnis

### Normalisasi

Untuk mengaplikasikan skema *appearance*, maka raw scan perlu untuk dikombinasikan. Mekanismenya, beberapa raw scan ditranslasikan, diatur skalanya, dan dirotasikan supaya koordinat titik - titik referensinya dapat dikombinasikan.

Hasil scan yang diperoleh dari sensor 3D merupakan suatu himpunan titik-titik  $S = \{(x, y, z)\}$ . Untuk keperluan penormalisasian, kita secara manual menspesifikasikan enam titik jangkar pada citra scan : dalam dan luar sudut mata kiri,  $E_{l,i}$  dan  $E_{l,o}$ , dalam dan luar sudut mata kanan,  $E_{r,i}$  dan  $E_{r,o}$ , ujung hidung  $N$ , dan titik dagu  $C$ . Kita gunakan notasi  $E_{l,i,x}$  dan  $E_{l,i,y}$  untuk merepresentasikan nilai x dan y dari  $E_{l,i}$ , dan  $E_{r,i,x}$  dan  $E_{r,i,y}$  untuk merepresentasikan nilai x dan y dari  $E_{r,i}$ . Setelah dilakukan rotasi, translasi, dan penskalaan, selanjutnya titik - titik tersebut dinormalisasikan supaya pusat dari mata kiri dan mata kanan dilokasikan secara tepat pada koordinat  $(100, 0, 0)$  dan  $(-100, 0, 0)$ , dan bidang yang melalui pusat mata dan titik

dagu, tegak lurus terhadap sumbu - z. Transformasi ini didefinisikan sebagai,

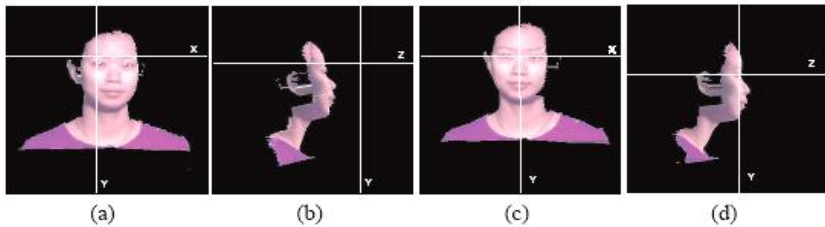
$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = s \cdot R \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$$

dimana

$$\begin{aligned} (t_1 \ t_2 \ t_3) &= -(\vec{E}_{l,i} + \vec{E}_{l,o} + \vec{E}_{r,i} + \vec{E}_{r,o})/4, \quad s = 400/\|\vec{E}_{l,i} + \vec{E}_{l,o} - \vec{E}_{r,i} - \vec{E}_{r,o}\|, \\ (x_0 \ y_0 \ z_0) &= (\vec{E}_{l,i} - \vec{C}) \times (\vec{E}_{r,i} - \vec{C}), \quad R = M_z \cdot M_x \cdot M_y, \end{aligned}$$

$$M_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{pmatrix}, \quad M_y = \begin{pmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{pmatrix}, \quad M_z = \begin{pmatrix} \cos \gamma & \sin \gamma & 0 \\ -\sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\alpha = -\arctan(y_0/\sqrt{x_0^2 + z_0^2}), \quad \beta = \arctan(x_0/z_0), \quad \gamma = \arctan\left(\frac{E_{l,i,y} - E_{r,i,y}}{E_{l,i,x} - E_{r,i,x}}\right),$$



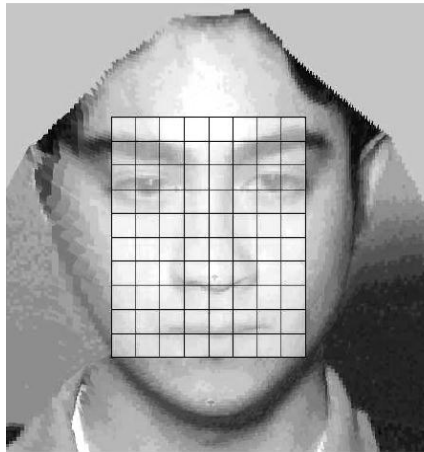
Gambar 3. (a) pandangan frontal sebelum normalisasi, (b) pandangan profil sebelum normalisasi, (c) pandangan frontal sesudah normalisasi, (d) pandangan profil sesudah normalisasi

### Mengkonstruksi vektor fitur

Untuk menghindari pengaruh dari keadaan bentuk atau style rambut dan aksesoris lainnya yang terdapat pada citra face, maka kita dapat mengaplikasikan skema *cropping* scan face tertutup terhadap citra scan. Diberikan suatu data face 3D yang telah dinormalisasikan, dinotasikan oleh  $C$ ,  $x$  dan  $y$  merupakan koordinat-koordinat dari daerah persegi

empat yang telah dipangkas (*cropped*),  $R$ , dan jumlah baris dan kolom dari grid di dalam daerah persegi empat,  $R$ ,  $m$  dan  $n$ . Kita pangkas daerah suatu face dan selanjutnya kita bentuk vektor fitur, dengan tahap – tahap berikut :

- (1) Bangun sebuah grid  $G$  yang merupakan suatu bidang yang paralel dengan bidang  $x - y$  yang memiliki  $m$  baris dan  $n$  kolom. Batas-batas dari  $G$  merupakan himpunan dari daerah persegi empat  $R$ . Lebih jelasnya lihat gambar 4.



Gambar 4. Daerah face yang telah dipangkas untuk konstruksi suatu vektor fitur. Grid  $10 \times 8$  yang melapisi suatu scan face

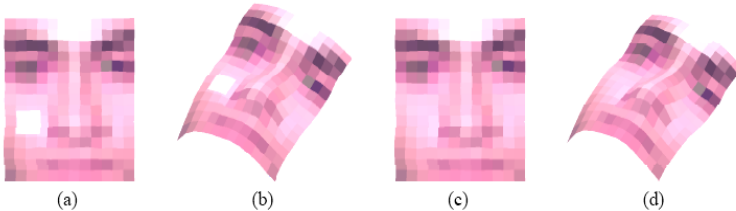
- (2) Bentuk suatu matrik proyeksi  $m \times n$  yang terdiri dari  $XM, YM, ZM$ . Elemen-elemen  $XM(i, j), YM(i, j)$ , dan  $ZM(i, j), i = 1, 2, \dots, n$ , yang berkaitan dengan node - node grid  $G(i, j)$ . Denotasikan himpunan seluruh titik - titik yang berada di dalam  $G(i, j)$  sebagai  $C'$ , di mana

$C' = \{(x, y, z) | (x, y, z) \in C, \text{ dan } x, y \text{ merupakan titik-titik yang berada di dalam } G(i, j)\}$ . Jika  $C'$  bernilai kosong maka elemen terkait diberi label sebagai suatu lubang (lihat gambar 5). sebaliknya, nilai-nilai grid dihitung berdasarkan, di mana  $|C'|$  merupakan jumlah elemen di dalam  $C'$ .

$$XM(i, j) = \frac{1}{|C'|} \sum_{\text{for all } (x, y, z) \in C'} x,$$

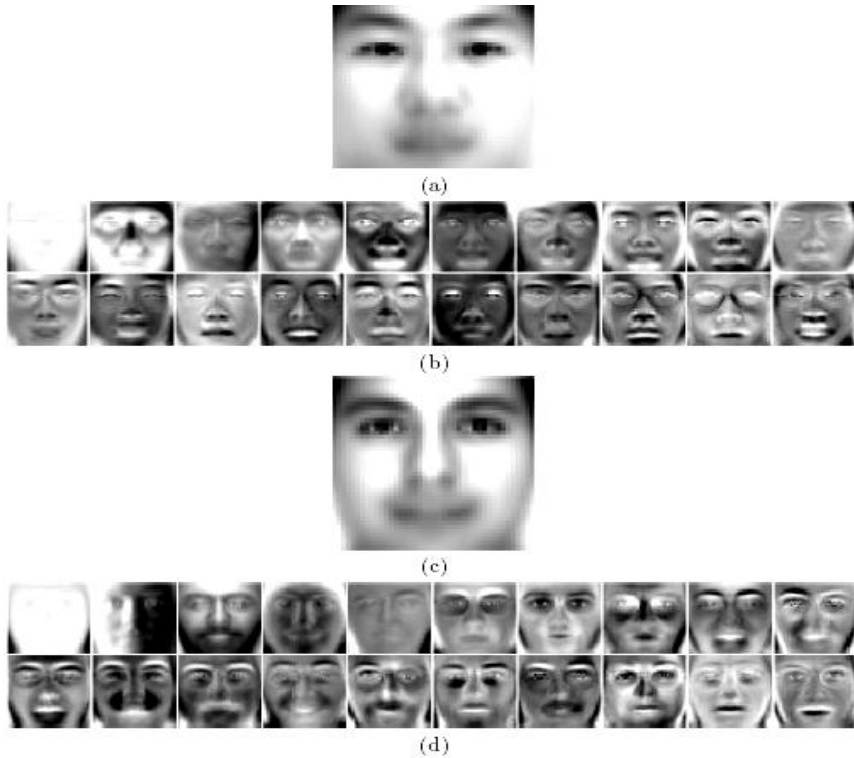
$$YM(i, j) = \frac{1}{|C'|} \sum_{\text{for all } (x, y, z) \in C'} y,$$

$$ZM(i, j) = \frac{1}{|C'|} \sum_{\text{for all } (x, y, z) \in C'} z,$$

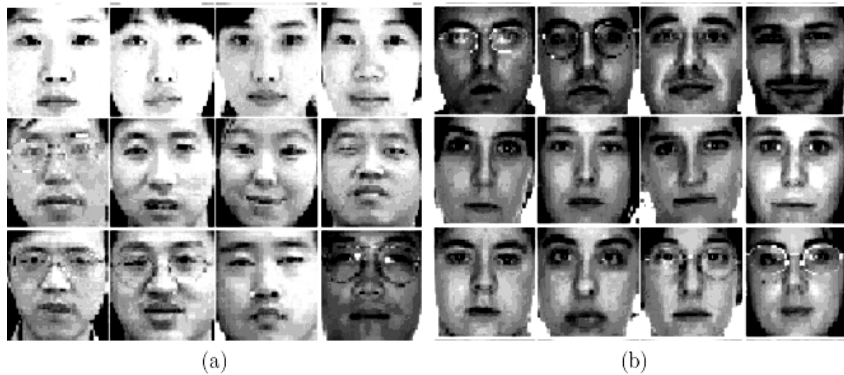


Gambar 5. (a, b) contoh lubang-lubang setelah normalisasi 3D (ditunjukkan sebagai tambalan-tambalan berwarna putih), dan (c, d) lubang-lubang yang telah diisi dengan suatu interpolasi

- (3) Setelah perotasian 3D, maka titik-titik tertentu yang menyebabkan terjadinya lubang di dalam scan yang telah dinormalisasikan dilakukan interpolasi. Lubang-lubang pada  $XM$ ,  $YM$  dan  $ZM$  diisi dengan penginterpolasian titik-titik tetangga terdekat.
- (4) Formasi vektor. Kolom-kolom pada matrik  $ZM$  disatukan untuk menghasilkan suatu vektor  $V$  dengan panjang  $m \times n$  yang selanjutnya digunakan oleh suatu operator pengklasifikasian untuk identifikasi.



Gambar 6. PCA<sup>13,15</sup> pada sekumpulan data asia dan non - asia, (a) "rata - rata face asia, (b) kumpulan data eigenface asia 20 teratas, (c) "rata - rata" face non - asia, (d) kumpulan data eigenface non- asia 20 teratas



Gambar 7. kumpulan face yang representatif, (a) Asia, dan (b) non - Asia

## Hasil

Pengujian dilakukan terhadap sejumlah citra berikut :

	Non-Asian	Asian	Subtotal
Female	106 (255)	33 (110)	139 (465)
Male	176 (563)	61 (212)	237 (775)
Subtotal	282 (918)	94 (322)	376 (1240)

dengan beberapa sampel scan yang dilustrasikan pada gambar 8.



Gambar 8. Sampel scan citra pengujian.

Berdasarkan hasil pengujian yang dilakukan terhadap seluruh citra wajah yang terdapat dalam database, maka unjuk kerja sistem yang dihasilkan adalah sebagai berikut :

	Non-Asian	Asian	Overall
Range	2.7% ± 0.028	6.7% ± 0.052	3.8% ± 0.024
Intensity	2.1% ± 0.027	5.9% ± 0.051	3.2% ± 0.029
Range + Intensity	0.7% ± 0.010	5.5% ± 0.039	2.0% ± 0.016



## Kesimpulan

Pengaruh range informasi yang memuat bentuk wajah 3D, digunakan dalam sistem yang diajukan. Hasil unjuk kerja yang diperoleh membuktikan bahwa range informasi, sangat berperan dalam meningkatkan keakuratan sistem dalam mengidentifikasi gender dan etnik.

## Referensi

- [1]. R. Malpass and J. Kravitz, "Recognition for faces of own and other race," *J. Perc. Soc. Psychol.*, vol. 13, pp. 330-334, 1969.
- [2]. J. Brigham and P. Barkowitz, "Do 'they all look alike?' the effect of race, sex, experience and attitudes on the ability to recognize faces," *J. Appl. Soc. Psychol.*, vol. 8, pp. 306-318, 1978.
- [3]. A. O'Toole, A. Peterson, and K. Deffenbacher, "An other-race effect for classifying faces by sex," *Perception*, vol. 25, pp. 669-676, 1996.
- [4]. A. Golby, J. Gabrieli, J. Chiao, and J. Eberhardt, "Differential responses in the fusiform region to same-race and other-race faces," *Nature Neuroscience*, vol. 4, no. 8, pp. 845-850, 2001.
- [5]. A. Puce, T. Allison, J. Gore, and G. McCarthy, "Face-sensitive regions in human extrastriate cortex studied by functional MRI," *J. Neurophysiol.*, vol. 74, pp. 1192-1199, 1995.
- [6]. A. O'Toole, K. Deffenbacher, D. Valentin, and H. Abdi, "Structural aspects of face recognition and the other race effect," *Memory & Cognition*, vol. 22, pp. 208-224, 1994.
- [7]. A. K. Jain, K. Nandakumar, X. Lu, and U. Park, "Integrating faces, fingerprints, and soft biometric traits for user recognition," in *LNCS 3087*, 2004, pp. 259-269.
- [8]. L.G. Farkas, *Anthropometry of the Head and Face*, Raven Press, 2nd edition, 1994.
- [9]. D. Enlow, *Facial Growth*, W.H. Saunders, 3rd edition, 1990.
- [10]. A. O'Toole, T. Vetter, N. F. Troje, and H. H. Bulthoff, "Sex classification is better with three-dimensional structure than

- with image intensity information," *Perception*, vol. 26, pp. 75–84, 1997.
- [11]. B. Golomb, D. Lawrence, and T. Sejnowski, "Sexnet: A neural network identifies sex from human faces," in *NIPS*, 1990, vol. 3, pp. 572–577.
- [12]. S. Gutta, J. Huang, P. Phillips, and H. Wechsler, "Mixture of experts for classification of gender, ethnic origin, and pose of human faces," *IEEE Trans. Neural Networks*, vol. 11, no. 4, pp. 948–960, Jul. 2000.