

INDIKATOR MUSIK MELALUI EKSPRESI WAJAH

Riyanto Sigit, Achmad Basuki

Politeknik Elektronika Negeri Surabaya

Kampus ITS Keputih Sukolilo Surabaya 60111, Indonesia

Tel: +62-31-5947280 Fax: +62-31-5946114; E-mail: riyanto@eepis-its.edu

ABSTRAKSI

Sistem pengenalan ekspresi wajah yang telah banyak dikembangkan pada umumnya berasal dari sekumpulan data (loadfile) dan tidak bisa mengatasi gerakan perubahan wajah. Pada penelitian ini dilakukan pengenalan ekspresi wajah secara real-time yang diaplikasikan sebagai indikator suatu musik. Prosesnya adalah pengambilan citra wajah secara real-time dengan kombinasi gerakan, bentuk, dan warna, mendeteksi posisi wajah dan mengekstrak fitur dengan mengambil fitur penting dari wajah berdasarkan lokasi alis, mata, dan bentuk mulut, kemudian mengenali ekspresi wajah menggunakan jaringan saraf tiruan. Hasil pengenalan ekspresi wajah secara otomatis akan diindikasikan dengan musik, sehingga musik akan berubah mengikuti perubahan ekspresi wajah.

Kata kunci: real-time, transformasi proyeksi, jaringan saraf tiruan, ekspresi wajah, indikator musik.

1. PENDAHULUAN

Pengenalan ekspresi wajah manusia merupakan salah satu bidang penelitian yang sangat penting karena merupakan perpaduan antara perasaan dan aplikasi komputer seperti interaksi antara manusia dengan komputer, pengkompresan data, animasi wajah dan pencarian citra wajah dari video.

Pada saat ini banyak peneliti membuat sistem komputer yang dapat mengerti dan berkomunikasi dengan manusia melalui citra ekspresi wajah [1,2,3,4]. Tian[1] mencoba menganalisa ekspresi wajah secara real-time melalui pemodelan dari background dan mengambil wajah dengan teknik shot-cut rule. A.Saxena[2] mengembangkan model geometrik menggunakan integral proyeksi dan warna kulit. Fasel[3] mengembangkan ekspresi wajah berdasarkan permasalahan umum dimensi dengan mengambil fitur berdasarkan mulut dan alis untuk mendapatkan wajah. Zhang [4] menyelidiki kegunaan dari dua tipe fitur wajah yaitu posisi geometrik dari 34 titik pada wajah dan sekumpulan multi scale, multi orientation koefisien Gabor wavelet pada titik tersebut untuk mengenali ekspresi wajah.

Hal ini menunjukkan bahwa sangat memungkinkan untuk melakukan pengembangan aplikasi pengenalan ekspresi wajah seseorang yang tampak dan ditampilkan melalui suatu media seperti musik untuk mengetahui kondisi emosinya[5].

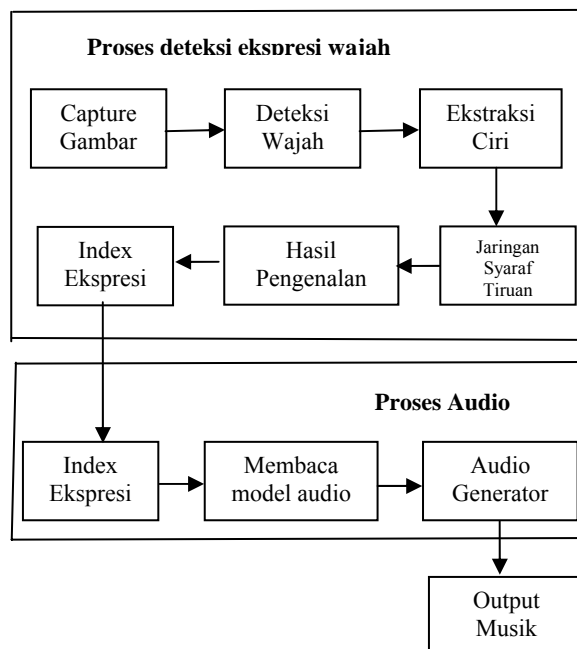
Selain itu komunikasi dengan seseorang akan sangat efisien jika dilakukan pada kondisi yang tepat, misalnya bagaimana berkomunikasi dengan seseorang saat kondisi emosinya sedang dalam keadaan marah. Musik yang dibangkitkan melalui aplikasi ini diharapkan dapat menjadi indikator yang merepresentasikan kondisi emosi seseorang pada saat tertentu. Jadi hal ini akan sangat membantu kita dalam berkomunikasi, selain bisa mengenali dan memahami orang lain juga bisa menempatkan diri sesuai kondisi. Musik diambil

dalam format wav dan dijadikan dalam frame-frame yang digunakan untuk mempermudah dalam pembangkitan lagu.

Aplikasi ini akan mengenali beberapa ekspresi wajah yaitu biasa, tertawa dan terkejut. Ekspresi-ekspresi tersebut merupakan ekspresi umum (universal) yang dialami manusia[6].

2. DISAIN SISTEM

Secara garis besar indikator musik melalui ekspresi wajah seperti pada blok diagram Gambar 1 berikut:



Gambar 1. Blok diagram sistem

2.1 Deteksi Wajah

Pada penelitian ini untuk mendapatkan wajah dengan cara mengkombinasi gerakan tubuh, warna dan bentuk. Gerakan tubuh didapatkan dengan

teknik pengurangan citra saat ini dengan citra sebelumnya. Warna diperoleh dengan memodelkan warna kulit secara YCrCb. Bentuk diperoleh dengan melakukan integral proyeksi.

Gerakan tubuh didapatkan dengan teknik pengurangan citra frame saat ini dengan citra frame sebelumnya menggunakan persamaan:

$$f_i(x,y) = f_i(x,y) - f_{i-1}(x,y) \quad (1)$$

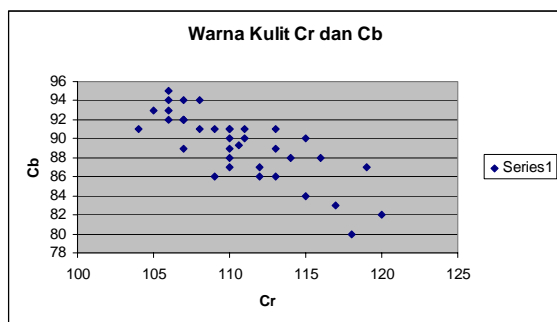
Warna diperoleh dengan memodelkan warna kulit dengan mengkonversi nilai RGB menjadi nilai YCrCb sehingga tidak terpengaruh dengan lingkungan. Dimana nilai :

$$Y = 0.59G + 0.31R + 0.11B, \quad (2)$$

$$Cr = 0.713 * (R-Y), \quad (3)$$

$$Cb = 0.564 * (B-Y). \quad (4)$$

Mula-mula diambil nilai warna kulit Cr dan Cb dari sekumpulan data pada berbagai kondisi pencahayaan kemudian dinormalisasi dengan ditambah 100 didapat nilai rata-rata Cr adalah 111 dan Cb adalah 88 seperti pada Gambar 2.



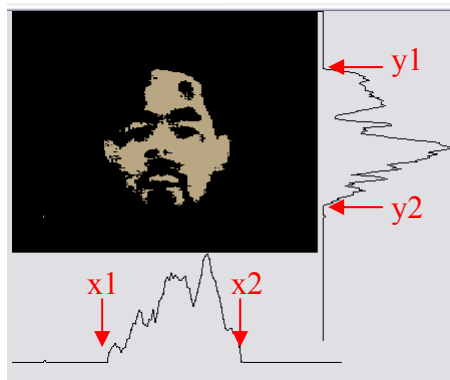
Gambar 2. Data warna kulit Cr dan Cb

Setelah diperoleh nilai rata-rata Cr dan Cb untuk warna kulit maka dilakukan perhitungan jarak pada setiap lokasi yang hendak dikenali. Jika jarak dibawah nilai thresholding maka dianggap sebagai kulit jika tidak dianggap sebagai latar belakang. Dengan menggunakan rumus:

$$\text{Jarak} = \sqrt{(Cr - \bar{Cr})^2 + (Cb - \bar{Cb})^2}, \quad (5)$$

jika $\text{Jarak} < \text{Thresholding}$ adalah kulit

Setelah dilakukan thresholding dari rumus jarak maka dilakukan perhitungan nilai awal wajah $(x1,y1)$ dan akhir wajah $(x2,y2)$ dengan menggunakan integral proyeksi maka diperoleh area wajah meskipun lingkungan berubah nilainya seperti pada Gambar 3.

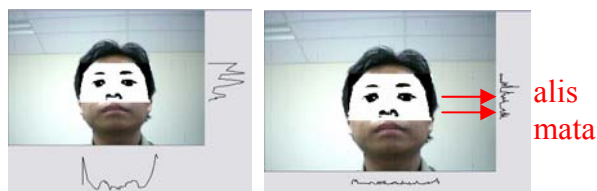


Gambar 3. Hasil pencarian nilai awal dan akhir wajah dengan integral proyeksi

2.2 Ekstraksi Ciri

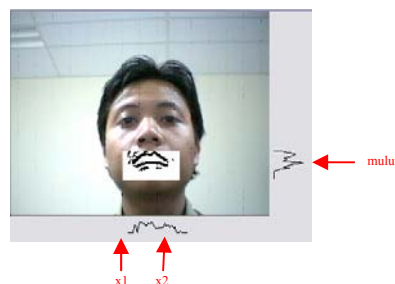
Ekstraksi ciri digunakan untuk mengambil ciri atau fitur yang penting dari ekspresi wajah. Dalam teknik ini berdasarkan lokasi fitur dan bentuk fitur. Ada 6 lokasi fitur penting yaitu 2 lokasi di pusat mata, 2 lokasi di ujung alis dalam dan 2 lokasi di ujung mulut.

Pertama dilakukan pencarian titik pusat mata dan ujung alis dalam dengan metode pencarian berdasarkan titik hitam dari suatu daerah menggunakan iterasi thresholding kemudian dilakukan proses penurunan gambar dengan menggunakan proses newton maju sehingga diperoleh posisi alis dan posisi mata seperti pada Gambar 4.



Gambar 4. Pencarian posisi alis dan mata

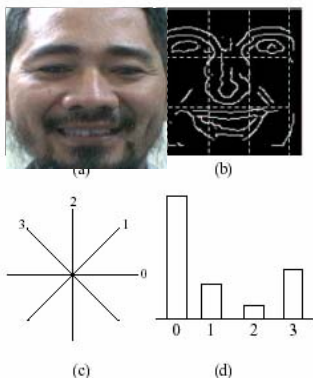
Untuk mendapatkan fitur 2 lokasi di ujung mulut menggunakan teknik deteksi tepi metode sobel dan integral projection sehingga diperoleh posisi mulut dan lebar mulut $(x1,x2)$ seperti pada Gambar 5.



Gambar 5. Pencarian posisi mulut dan lebar mulut

Teknik yang digunakan untuk mendapatkan bentuk fitur adalah dengan mengekstrak bentuk fitur daerah mulut dengan histogram dimana frekwensi menyatakan nilai dari fitur sehingga didapat nilai

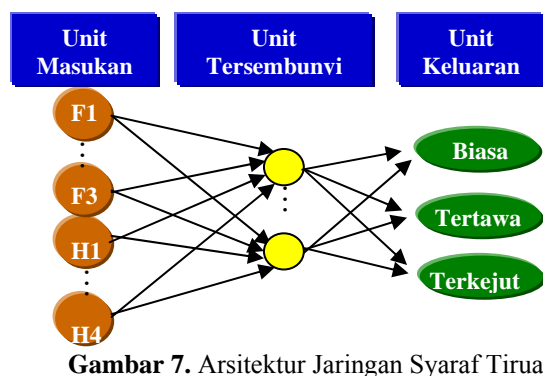
fitur F4 (0), F5 (1), F6 (2) dan F7 (3) seperti pada Gambar 6.



Gambar 6. Pengambilan fitur berdasarkan bentuk

2.3 Jaringan Syaraf Tiruan

Jaringan syaraf tiruan digunakan untuk mengenali ekspresi wajah. Teknik yang digunakan adalah metode propagasi balik dengan arsitektur jaringan lapisan input sebanyak 7 unit, lapisan tersembunyi sebanyak 3 unit dan lapisan output sebanyak 3 unit seperti pada Gambar 7.



Gambar 7. Arsitektur Jaringan Syaraf Tiruan

2.4 Digital Audio Processing

Sinyal adalah besaran fisis yang berubah menurut waktu, ruang, atau variabel-variabel bebas lainnya. Contoh sinyal: sinyal ucapan, ECG, dan EEG. Secara matematis, sinyal adalah fungsi dari satu atau lebih variabel independen[7]. Contoh fungsi matematis dari sinyal:

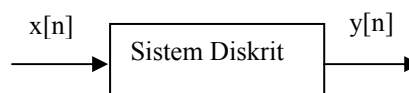
$$s(t) = A(t) \sin(2\pi F(t) * t) \quad (6)$$

2.5 Pembangkitan Musik

Pembangkitan musik dilakukan setelah proses fitur extraction yaitu yang disesuaikan dengan beberapa index ekspresi wajah. Musik-musik instrumen yang digunakan yaitu dalam format wav.

2.5.1 Sistem Waktu Diskrit

Hubungan input/output sistem diskrit di gambarkan dalam persamaan beda. Penggambaran input/output suatu sistem biasa juga disebut sebagai pemodelan sistem. Dibawah ini adalah blok diagram dan persamaan beda linear.



Gambar 8. Blok diagram Sistem Diskrit Linear

2.5.2 Proses pencuplikan (sampling)

Pada beberapa aplikasi, sangat menguntungkan memproses sinyal waktu kontinyu dengan mengubahnya dulu ke sinyal waktu diskrit dan, setelah di proses dalam waktu diskrit, kemudian dikembalikan lagi ke sinyal waktu kontinyu[8].

Kecepatan pengambilan sampel (Frekuensi sampling) dari sinyal analog yang akan dikonversi haruslah memenuhi kriteria Nyquist yaitu:

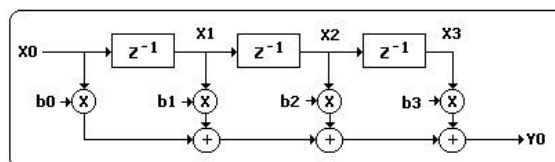
$$F_s > 2 F_{max} \quad (7)$$

dimana frekuensi sampling minimum adalah 2 kali frekuensi sinyal analog yang akan dikonversi. Sehingga apabila terdapat kriteria Nyquist tidak dipenuhi maka akan timbul efek aliasing[7].

2.5.3 Filter Digital

a. Finite Impulse Response (FIR)

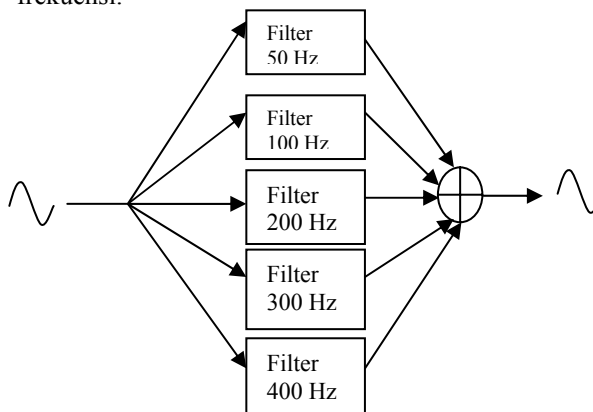
Filter FIR adalah salah satu tipe dari filter digital yang dipakai pada aplikasi Digital Signal Processing (DSP). FIR kepanjangan dari Finite Impulse Response. Disebut respons impulsnya terbatas (finite), karena tidak ada feedback didalam filter, jika memasukkan sebuah impulse (yaitu sebuah sinyal '1' diikuti dengan banyak sinyal '0'), sinyal nol akan keluar setelah sinyal 1 melewati semua delay line dengan koefisiennya[7].



Gambar 9. Flow graph filter FIR orde 4

2.5.4 Mixer

Mixer adalah suatu proses penggabungan antara beberapa filter yang mempunyai perbedaan frekuensi.



Gambar 10. Proses Mixer

3. HASIL DAN PEMBAHASAN

Pada penelitian ini data diambil dilaboratorium PENS-ITS secara online. Data berupa gambar online yang di capture kemudian disimpan. Data terdiri dari 30 orang yang masing-masing memiliki 4 kondisi wajah berbeda yaitu biasa, terkejut, tertawa dan lainnya. Masing-masing data memiliki resolusi yang berbeda yaitu 320*240, 240*180 dan 180*120. Sehingga total data sebanyak 360 yang masing-masing data memiliki karakteristik pencahayaan yang cukup.

3.1 Tahap Pengujian Deteksi Wajah

Pada tahap pengujian deteksi wajah akan diuji berdasarkan resolusi dari wajah. Contoh pengambilan data seperti Gambar 11.



Gambar 11. Hasil deteksi wajah

Secara keseluruhan hasil pendeteksian wajah seperti pada Tabel 1.

Tabel 1. Deteksi Wajah dengan berbagai resolusi dan kondisi

No. Subyek	320*240	240*180	180*120
1(Biasa)	90	90	90
2(Biasa)	90	90	0
3(Biasa)	100	100	90
4(Biasa)	100	100	90
5(Biasa)	70	70	70
6(Biasa)	90	90	80
7(Biasa)	80	80	70
8(Biasa)	70	70	70
9(Biasa)	90	90	90
...
Rata-rata	87.67%	87%	78.33%

3.2 Tahap Pengujian Ekstraksi Ciri

Pada penelitian ini data diekstraksi menjadi beberapa fitur yaitu lokasi alis, lokasi mata, lokasi mulut dan lebar mulut sehingga akan diperoleh fitur 1 (F1) jarak alis terhadap mata, fitur 2 (F2) jarak mata terhadap mulut, fitur 3 (F3) lebar mulut. Cara pengambilan data dengan mengambil contoh sebanyak 12 orang dalam kondisi normal. Nilai tepat menunjukkan bahwa program dapat mengenali

posisi fitur dengan baik. Contoh pengambilan data seperti pada Gambar 12.



Gambar 12. Hasil ekstraksi ciri

Berdasarkan percobaan dengan mengambil data sebanyak 30 orang yang memiliki berbagai kondisi wajah yang berbeda dengan melihat ketepatan posisi alis, posisi mata, posisi mulut dan lebar mulut wajah didapat data seperti pada tabel Tabel 2.

Tabel 2. Fitur Wajah dengan resolusi 320*240

No. Subyek	Alis	Mata	Mulut	Lebar Mulut
1(Biasa)	Tepat	tepat	tepat	tepat
2(Biasa)	Tepat	tepat	tepat	tepat
3(Biasa)	Tepat	tepat	tepat	tepat
4(Biasa)	Tepat	tepat	tepat	tepat
5(Biasa)	Tepat	tepat	tepat	tepat
6(Biasa)	Tepat	tepat	tepat	tepat
7(Biasa)	Tepat	tidak	tidak	tepat
8(Biasa)	Tepat	tepat	tepat	tepat
9(Biasa)	Tepat	tepat	tepat	tepat
10(Biasa)	Tepat	tepat	tepat	tepat
...
Rata-rata	86.67%	80%	90%	93.33%

3.3 Tahap Pengujian Pengenalan Ekspresi Wajah

Pada tahap pengujian pengenalan ekspresi wajah menggunakan jaringan syaraf tiruan metode backpropagasi dengan jumlah unit input sebanyak 7 dan unit output sebanyak 3. Contoh hasil pengenalan seperti Gambar 13.



Gambar 13. Hasil pengenalan ekspresi wajah

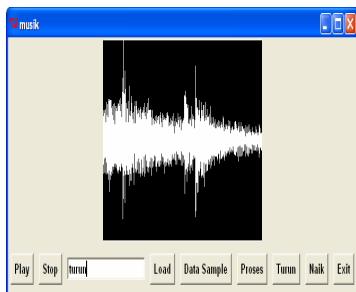
Cara pengujian data campuran baik yang di training maupun yang tidak di Training dengan resolusi 320*240, 240*180 dan 180*120 dengan mengambil contoh sebanyak 90 data dalam kondisi normal, terkejut dan tertawa. Secara keseluruhan hasil pengenalan ekspresi wajah seperti pada Tabel 3.

Tabel 4. Pengenalan Ekspresi Wajah dengan resolusi 320*240, 240*180 dan 180*120

No Subyek	320*240	240*180	180*120
1(Biasa)	dikenali	tidak	tidak
2(Biasa)	dikenali	dikenali	dikenali
3(Biasa)	dikenali	dikenali	dikenali
4(Biasa)	dikenali	dikenali	dikenali
5(Biasa)	dikenali	dikenali	dikenali
6(Biasa)	dikenali	dikenali	dikenali
7(Biasa)	tidak	dikenali	dikenali
8(Biasa)	tidak	tidak	tidak
9(Biasa)	tidak	tidak	tidak
10(Biasa)	dikenali	dikenali	dikenali
...
Rata-rata	70%	66.67%	66.67%

3.4 Tahap Pengujian Button Turun

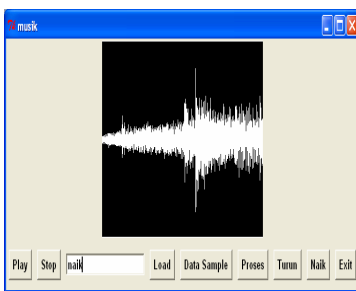
Button Turun di gunakan untuk merubah data amplitudo (dalam .txt) menjadi file .wav, dimana amplitudonya menjadi semakin menurun.



Gambar 14. Penurunan amplitudo

3.5 Tahap Pengujian Button Naik

Button Naik di gunakan untuk merubah data amplitudo (dalam .txt) menjadi file .wav, dimana amplitudonya menjadi semakin meningkat.



Gambar 15. Kenaikan amplitudo

Hasil dari sebuah lagu .wav setelah di proses menghasilkan 2 lagu yang berformat .wav. Lagu pertama mengalami penurunan amplitudo dan lagu kedua mengalami kenaikan amplitudo. Kedua file tersebut akan mengkondisikan keadaan seseorang, untuk menghindari perubahan musik yang sangat drastis dari ekspresi satu ke ekspresi yang lain maka amplitudo dari masing-masing lagu diatur. Tampilan hasil running program pembangkitan lagu tidak ditampilkan, tetapi akan dipanggil dari program pengenalan ekspresi wajah yang dijalankan.

4. KESIMPULAN

Dari hasil uji coba penelitian ini dapat diperoleh beberapa kesimpulan antara lain:

- Kombinasi gerakan, warna dan bentuk dapat digunakan untuk mendeteksi wajah dengan keberhasilan 84.33% dari keseluruhan wajah.

- Lokasi alis, mata dan bentuk mulut dapat digunakan untuk mengekstraksi fitur wajah dengan keberhasilan 85.55%.
- Jaringan syaraf tiruan dapat digunakan untuk mengenali ekspresi wajah dengan keberhasilan 67.78% dan sangat dipengaruhi oleh posisi wajah pada ekstraksi fitur sebelumnya.
- Hasil dari sebuah lagu .wav setelah di proses menghasilkan 2 lagu yang berformat .wav. Lagu pertama mengalami penurunan amplitudo dan lagu kedua mengalami kenaikan amplitudo.
- Keluaran lagu akan mengkondisikan keadaan seseorang, untuk menghindari perubahan musik yang sangat drastis dari ekspresi satu ke ekspresi yang lain maka amplitudo dari masing-masing lagu diatur

DAFTAR PUSTAKA

- Y. Tian and R. Bolle, Real World Real-time Automatic Recognition of Facial Expspressions, *IBM Research Report, PETS 2003*.
- A. Saxena, A. Anand, A. Mukerjee, Robust Facial expression Recognition Using Spatially Localized Geometric Model, *International Conference on Systemics, Cybernetics, February 12-15, 2004*.
- B. Fasel, et al, Automatic Facial Expression Analysis: A Survey, *Pattern Recognition*, 36, 259-275, 2003
- Z. Zhang, Feature based facial expression recognition: Sensitivity analysis and experiments with a multi-layer perceptron, *International Journal of Pattern Recognition and Artificial Intelligence*, 13(6):893-911, 1999.
- Anonymous, *Facial Expression Detection and Recognition System*, pada www.ies.org.sg/journal/past/v44i3/v44i3_2.pdf.
- Anonymous, *Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction*, pada www.mplab.ucsd.edu/projects-home/project1/publications/pdfs/BartlettCVPR2003-6.pdf.
- Miftahul Huda dan Tri Budi S, *Modul Praktikum Digital Signal Processing (DSP)*, PENS-ITS.
- ALAN V. O, ALAN S. W dan S. HAMID NAWAB, *Sinyal dan Sistem edisi kedua*.
- R. L. Hsu dan M. A. Mottaleb, Face Detection in Color Image, *Appear In IEEE Tans. PAMI*, Vol. 24, no.5, pp.696-706, Mei 2002.
- Tian, Ying-li, and Bolle M, Exploratory Computer Vision Group, Automatic neutral face detection using location and shape feature, *IBM Research Report*, 2001.
- Riyanto Sigit, Dadet Pramudianto, Achmad Basuki, *Pengolahan Citra Digital*, Andi Offset, 2005.
- Mauridhi Heri P, *Kecerdasan Buatan*, D4 PENS-ITS, 2002
- User Guide*, Seattle Robotics CMUcam2 Vision System, 2003
- Eru Puspita, Sistem Pendeteksian dan Penjejukan Wajah Secara Real Time, *Seminar Thesis*, ITS, 2004.

