

# SINTESA EKSPRESI WAJAH DENGAN MENGGUNAKAN RADIAL BASIS FUNCTION NETWORK

**Wiwik Anggraeni, Handayani Tjandra**

Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember  
Kampus ITS, Jl. Raya ITS, Sukolilo – Surabaya 60111, Tel. + 62 31 5939214, Fax. + 62 31 5913804  
Email: [wiwik@its-sby.edu](mailto:wiwik@its-sby.edu), [handatj@rad.net.id](mailto:handatj@rad.net.id)

## ABSTRAK

Pada penelitian yang sebelumnya [4] telah dilakukan penelitian tentang letak-letak (koordinat) facial characteristic points (FCP) yang digunakan sebagai dasar untuk mengenali ekspresi-ekspresi wajah manusia. Diantaranya ada enam ekspresi wajah, yaitu gembira, sedih, marah, takut, terkejut, dan jijik yang digunakan dalam penelitian tersebut.

Dengan adanya dasar tersebut, maka dalam penelitian ini penulis berusaha mensintesa ekspresi wajah yang dikategorikan menjadikan enam ekspresi dengan menggunakan facial characteristic points tersebut. Prinsip dasar dari mensintesa ekspresi wajah adalah mencari pemindahan spasial relatif facial characteristic points pada setiap ekspresi.

Permasalahan utamanya adalah bagaimana menghasilkan wajah dengan ekspresi tertentu dari sebuah citra input wajah tanpa ekspresi. Dengan menggunakan 30 pasang perpindahan titik FCP, dilakukan training terhadap Radial Basis Function Network (RBFN)  $6 \times n \times 60$  (6 input yang merupakan kadar dari keenam ekspresi,  $n$  hidden unit, dan 60 output yang merupakan 30 pasang perpindahan FCP, dimana  $n$  adalah variabel). RBFN yang telah ditraining dapat menghasilkan perpindahan FCP sesuai dengan ekspresi yang diinputkan. Informasi pemindahan FCP ini kemudian dimasukkan ke dalam algoritma Image Warping bersama-sama dengan citra input wajah tanpa ekspresi untuk menghasilkan citra wajah berekspresi tertentu.

**Kata kunci :** Facial Characteristic Points, Radial Basis Function Network, Sintesa Ekspresi Wajah

## 1. PENDAHULUAN

### 1.1. LATAR BELAKANG MASALAH

Dalam perkembangan teknologi yang semakin canggih ternyata simulasi ekspresi wajah mempunyai peranan yang sangat besar, antara lain:

1. Komunikasi: Sintesa Ekspresi Wajah dapat digunakan untuk mengurangi bandwidth transmisi dengan pengkodean ekspresi wajah saat mentransmisikan wajah pada aplikasi telekonferensi dan video-phone.
2. Pembuatan film: Banyak film-film yang menggunakan animasi wajah, seperti *Terminator2* dan *Robocop2*.
3. Riset Medis: Sintesa Ekspresi Wajah dapat memperluas wawasan tentang kontribusi informasi visual gerakan wajah dalam hubungannya dengan *speech perception* sehingga analisa proses persepsi pendengaran dapat lebih sistematis.
4. Bantuan Pengajaran dan Berbicara: Mengajari cacat pendengaran untuk membaca gerak bibir atau membuat yang bisu dapat berbicara memudahkan dengan sistem Sintesa Ekspresi Wajah. Dengan membuat proses belajar seperti permainan, perhatian murid dapat lebih terpusat sehingga dapat belajar lebih banyak.

### 1.2. RUMUSAN PERMASALAHAN

Dalam penelitian ini akan disimulasikan ekspresi wajah normal tanpa ekspresi menjadi wajah yang berekspresi tertentu dengan bantuan komputer. Permasalahan yang terjadi adalah bagaimana cara mendapatkan FCP yang ada di dalam wajah, kemudian bagaimana cara kerja RBFN dalam mempelajari dan menghasilkan pemindahan FCP untuk ekspresi tertentu, dan bagaimana cara kerja image warping untuk memetakan citra wajah tanpa ekspresi menjadi wajah dengan ekspresi tertentu sesuai yang diinginkan.

### 1.3. BATASAN MASALAH

Batasan masalah yang diberikan adalah sebagai berikut:

1. Input, berupa citra gray-scale 256 level, 512x512 pixel dengan persyaratan:
  - Wajah frontal tanpa kacamata
  - Tidak ada rotasi pada aksis vertikal.
  - Wajah berada di tengah citra dengan lebar wajah kurang lebih setengah lebar citra
  - Iluminasi yang baik / tidak buram
  - Daerah wajah yang penting, seperti alis dan mata, tidak tertutup.
2. Ekspresi wajah yang dikategorikan dalam enam kelompok dasar, yaitu: Gembira, Sedih, Marah, Takut, Kaget, dan Jijik.

- Input untuk training adalah citra wajah tanpa ekspresi dan citra wajah berekspresi tertentu, keduanya dengan persyaratan yang sama dengan citra input untuk sintesa.

## 2. BLOK DIAGRAM PROSES

Pada penelitian kali ini, secara garis besar proses sintesa diawali dengan proses utama yang terdiri dari bagian sampling, bagian training, dan bagian testing.

Proses utama dibagi menjadi 3 bagian utama yaitu:

### 1. Sampling

Kegiatannya adalah mengumpulkan data *FCP displacement* dari foto yang diinputkan.

Input : wajah tanpa ekspresi dan wajah dengan ekspresi tertentu

Output : *FCP displacement*

### 2. Training

Kegiatan yang dilakukan adalah melatih *radial basis function network* untuk mengenali *FCP displacement* berbagai ekspresi

Input : *FCP displacement*

Output : *Miu* dan *weight*

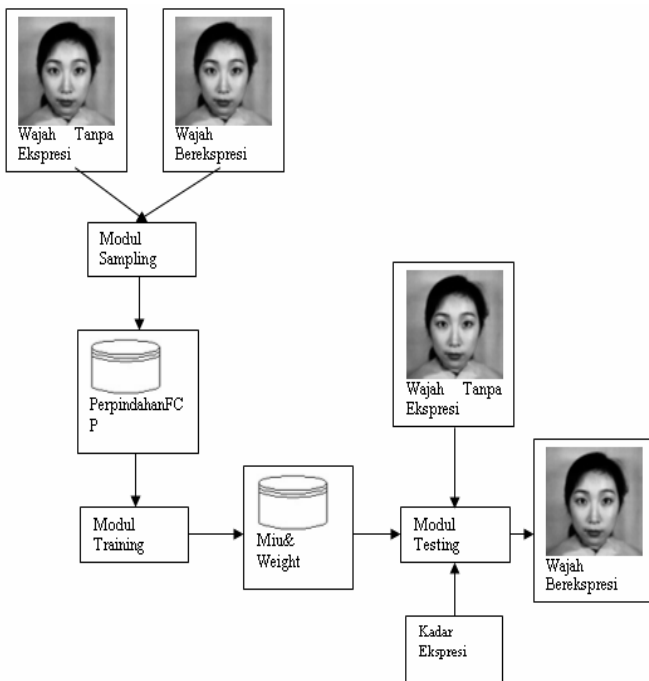
### 3. Testing

Didalam bagian testing pengetahuan yang dimiliki *radial basis function network* akan digunakan untuk mensintesa ekspresi wajah sesuai dengan ekspresi yang diinginkan.

Input : *Miu* dan *Weight* serta wajah tanpa ekspresi

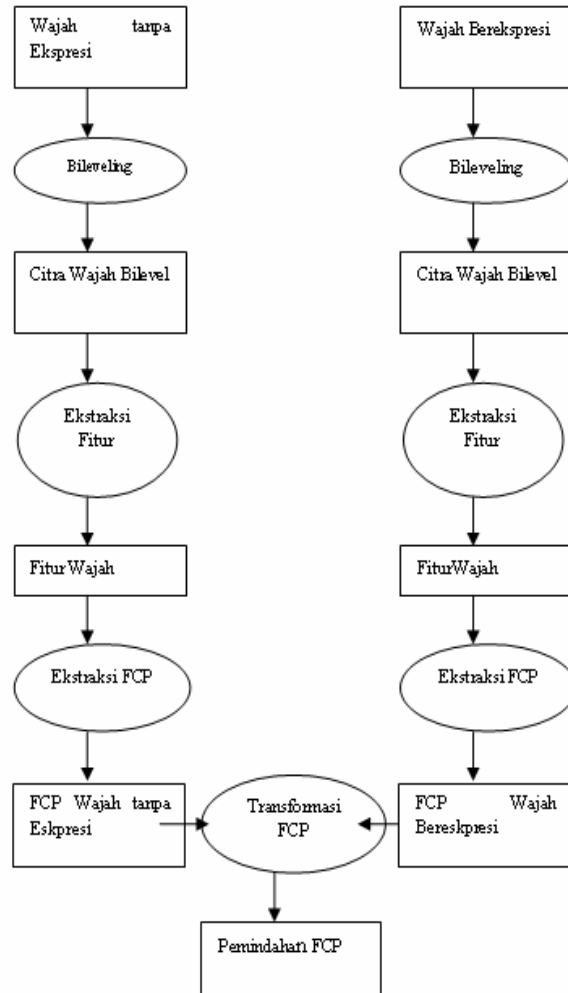
Output : wajah dengan ekspresi tertentu

Blok diagram dari proses utama dapat dilihat pada gambar berikut :



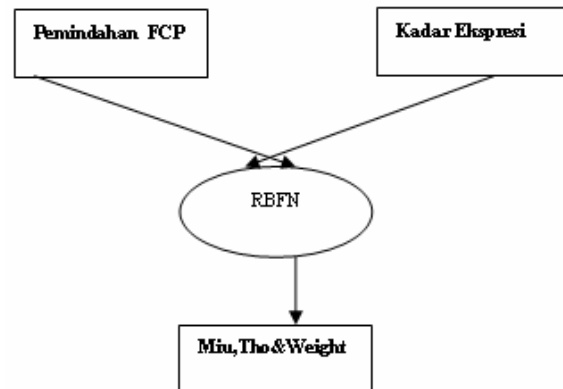
Gambar 1. Blok diagram proses utama

Blok diagram dari modul sampling digambarkan sebagai berikut :



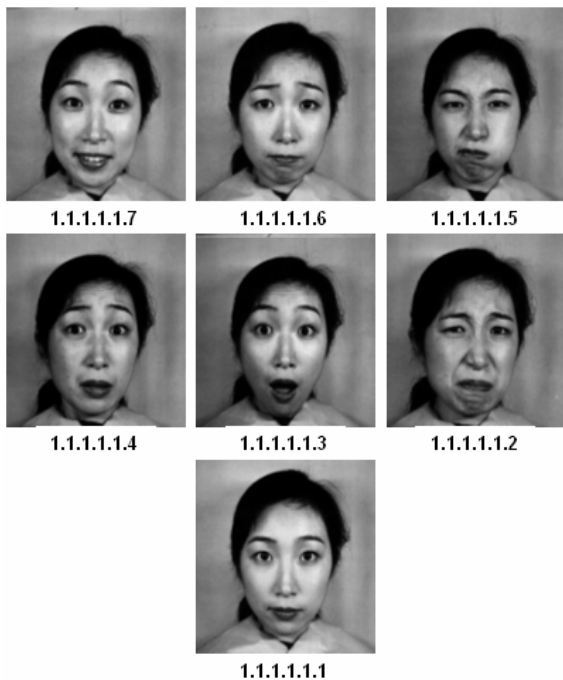
Gambar 2. Blok diagram modul sampling

Sedangkan untuk modul training, blok diagramnya digambarkan sebagai berikut :



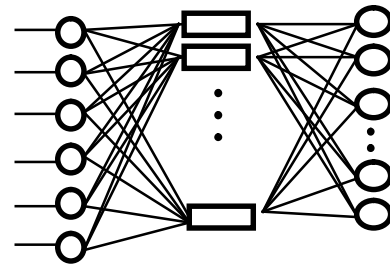
Gambar 3. Blok diagram modul training

Di dalam bagian training, foto-foto yang digunakan untuk training adalah foto berbagai ekspresi yang diambil dari database JAFFE (Japanese Female Facial Expression). Database tersebut memiliki 213 citra dari 7 ekspresi wajah (6 ekspresi wajah dasar + 1 netral) yang diambil dari model 10 wanita Jepang. Bersama dengan database tersebut terdapat file yang berisi data rate semantik dari eksperimen psikologi menggunakan citra-citra tersebut, karena setiap ekspresi tidak pernah murni dari satu emosi, tapi selalu merupakan gabungan dari berbagai emosi. Ukuran kadar ekspresi yang digunakan berskala 5 (1-rendah dan 5-tinggi). Contoh foto-foto dari database JAFFE dengan kadar ekspresinya dapat dilihat pada gambar berikut :



**Gambar 4. Contoh citra pada database JAFFE**

Arsitektur Radial Basis Function yang digunakan dalam penelitian ini dapat dilihat pada gambar dibawah. Jumlah Input adalah enam unit yang merupakan inputan kadar dari enam ekspresi wajah. Outputnya terdiri dari 60 unit yang merupakan 30 pasang FCP displacement. Dan hidden unitnya adalah variabel yang nilai maksimalnya adalah 100.



**Gambar 5. Arsitektur radial basis function**

Fungsi Interpolasinya adalah:

$$F_k(d_i) = \sum_{j=1}^c w_{jk} g(\|d - \mu_j\|)$$

G(.) adalah fungsi aktivasi Gauss:

$$g(\|d - \mu\|) = \exp\left[\frac{-(d - \mu)}{2\sigma^2}\right] \quad (1.2)$$

Perbaikan Weight:

$$\omega[j,k] = \omega[j,k] + \alpha * (\text{Output}[k] - \text{SampleOutput}[k]) * g[j] \quad (1.3)$$

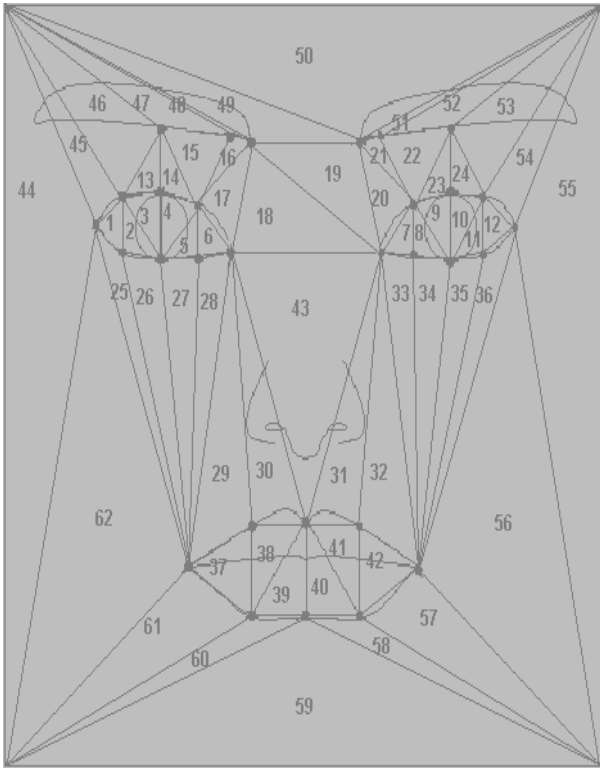
Inputnya berupa vektor yang terdiri dari 6 elemen. Sample-sample ini dikelompokkan dengan algoritma kMeans menjadi sejumlah (tergantung jumlah hidden unit) dengan pusat Miu. Setiap cluster selain memiliki pusat cluster juga memiliki lebar yaitu rata-rata jarak sample dari pusat, Tho

Untuk setiap unit hidden dihitung nilai Gaussnya Nilai Gauss akan diteruskan ke unit output dengan rumus interpolasi. Kemudian hasilnya akan dibandingkan dengan nilai sample output. Perbaikan nilai Weight akan dilakukan dengan rumus di atas yang mana perbaikan akan terus dilakukan selama belum kesalahan lebih kecil daripada nilai toleransi error atau perbaikan sudah dilakukan sebanyak looping maksimal.

Bagian yang berikutnya adalah bagian testing. Berdasarkan pengetahuan tentang Miu dan weight yang telah diperoleh pada bagian training, berikutnya dapat diperoleh facial characteristic points displacement yang kemudian akan ditransformasikan dan diwarping menjadi citra wajah dengan ekspresi tertentu berdasarkan citra wajah tanpa ekspresi yang diinputkan.

Pada proses warping, metode yang digunakan adalah Triangle Image Warping, dengan input koordinat FCP hasil RBFN yang telah ditransformasi dan koordinat FCP dari wajah tanpa ekspresi. Sedangkan outputnya adalah wajah hasil sintesa dengan ekspresi tertentu. Berikut ini adalah gambar segitiga-segitiga yang dibentuk dari hubungan titik-titik FCP yang digunakan untuk keperluan image warping. Setiap segitiga diberi nomor agar dapat

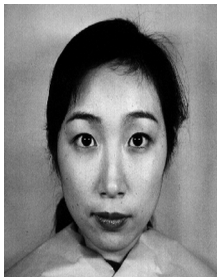
diketahui segitiga asal dari setiap titik pada citra tujuan.



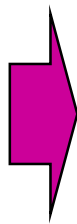
**Gambar 6. Nomor segitiga warping**

**3. HASIL TESTING**

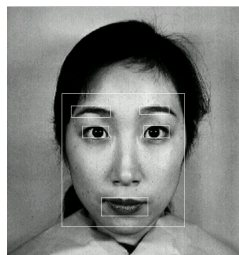
Berikut ini adalah contoh hasil dari proses yang telah dijelaskan sebelumnya.



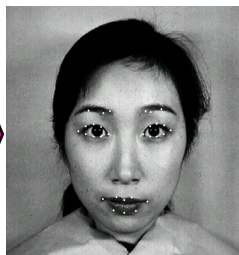
**Gambar 7.1**



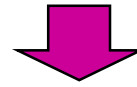
**Gambar 7.2**



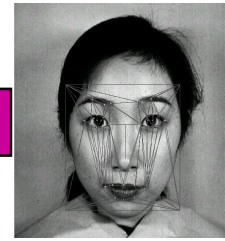
**Gambar 7.3**



**Gambar 7.4**



**Gambar 7.6**



**Gambar 7.5**

Berikut penjelasan dari masing-masing gambar pada gambar 7

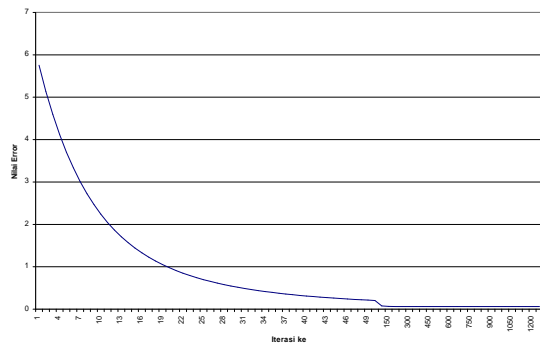
- Gambar 7.1 : Gambar wajah tanpa ekspresi / normal – inputan
- Gambar 7.2 : Hasil proses bileveling yang dilakukan pada bagian sampling
- Gambar 7.3 : Hasil ekstraksi fitur wajah yang dilakukan dengan cara template matching. Proses ini dilakukan pula pada bagian sampling
- Gambar 7.4 : Hasil ekstraksi facial characteristic points
- Gambar 7.5 : Hasil proses segitiga warping. Dari gambar dapat dilihat bahwa segitiga yang terbentuk hanya berada dalam kotak wajah saja.
- Gambar 7.6 : Hasil terakhir dari image warping atau contoh inputan dari ekspresi campuran

**4. PERBANDINGAN ANTARA HASIL TESTING DENGAN GAMBAR ASLI**

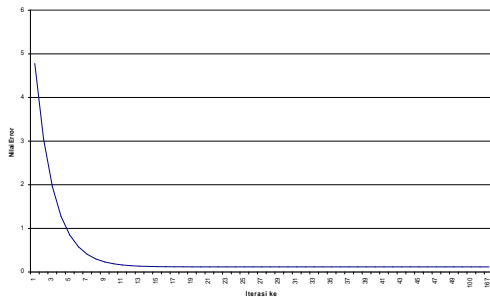
Setelah uji coba perangkat lunak dilakukan, proses berikutnya adalah membandingkan hasil antara hasil uji coba / testing dengan gambar aslinya. Hal ini dilakukan untuk menguji seberapa jauh terjadi penyimpangan dari ekspresi yang dihasilkan. Tolok ukur yang digunakan adalah error hasil *training set* dan *Mean Square Error* (MSE). Uji coba juga dilakukan pada jumlah sampel dan jumlah neuron yang berbeda. Berikut adalah hasil perbandingannya.

**4.1. ERROR TRAINING SET**

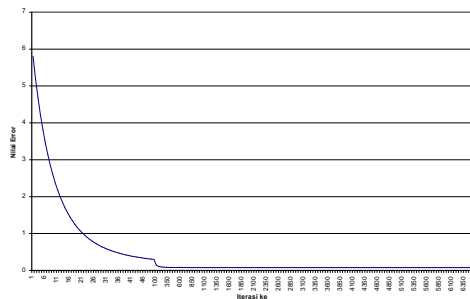
Yang dimaksud error disini adalah error yang terjadi antara output hasil setiap proses training dengan image yang digunakan sebagai supervised training. Nilai error yang terjadi dapat dilihat pada gambar 8 berikut :



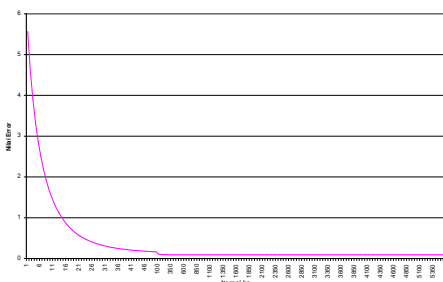
**Gambar 8.1**  
Grafik error untuk sampel 24 wajah hidden layer 25



**Gambar 8.2**  
Grafik error untuk sampel 24 wajah hidden layer 50



**Gambar 8.3**  
Grafik error untuk sampel 160 wajah hidden layer 25



**Gambar 8.4**  
Grafik error untuk sampel 160 wajah hidden layer 50

Dari hasil diatas dapat diketahui bahwa semakin banyak jumlah sampel yang digunakan mak atraining errornya semakin besar. Hal ini disebabkan karena error adalah jumlah kumulatif dari semua yang ada, sehingga apabila jumlahnya semakin besar maka errornya semakin besar pula. Selain itu pencapaian konvergensinya juga lama.

**4.2. MSE**

MSE adalah merupakan tolak ukur dari tingkat keberhasilan hasil sintesa. Hasil uji cobanya dapat dilihat pada tabel dibawah. Dalam hal ini pada input normal tanpa ekspresi dilakukan sintesa dengan memberi input kadar untuk gembira 4.39, sedih 1.35, marah 1.16, takut 1.26, terkejut 2.29, dan jijik 1.23. Dengan menggunakan *training set* dengan jumlah sampel masing-masing didapatkan nilai MSE seperti pada tabel 1 berikut :

**Tabel 1. Nilai MSE uji coba 1**

No	Jumlah Sampel Input (Wajah)	MSE
1	24	0.00499
2	160	0.00025

Kemudian pada input yang sama dilakukan sintesa dengan memberi input kadar untuk gembira 1.42, sedih 4.0, marah 2.39, takut 3.03, terkejut 1.55, dan jijik 3.26. Dengan menggunakan *training set* dengan jumlah sampel masing-masing didapatkan nilai MSE MSE seperti pada tabel 2 berikut :

**Tabel 2. Nilai MSE uji coba 2**

No	Jumlah Sampel Input (Wajah)	MSE
1	24	0.00133
2	160	0.00018

Dari hasil uji coba di atas kelihatan bahwa untuk jumlah wajah sample yang berbeda ternyata menghasilkan MSE yang berbeda pula. Untuk *training set* yang mempunyai jumlah sample lebih banyak ternyata menghasilkan MSE yang lebih kecil bila dibanding *training set* yang mempunyai jumlah sample lebih sedikit. Hal ini disebabkan karena dengan semakin besar jumlah sample kemungkinan displacement FCP nya akan semakin banyak sehingga pada waktu kadar input dimasukkan kemungkinan mencari displacement FCP dengan kadar yang hampir sama dengan input semakin bertambah.

**5. KESIMPULAN**

Dari penelitian ini dapat diambil kesimpulan bahwa :

1. Konvergensi error pada saat training ditentukan oleh banyaknya sampel wajah yang digunakan pada saat training
  2. MSE suatu ekspresi hasil sintesa akan lebih kecil jika jumlah sampel yang digunakan pada saat training semakin besar
  3. Semakin banyak dan bervariasi sampel yang digunakan untuk training, semakin baik kemampuan RBFN untuk mensintesa berbagai macam ekspresi wajah
  4. Lamanya proses perbaikan *cluster* tergantung dari jumlah *cluster* dan jumlah sampel yang digunakan dalam proses k-means *clustering*.
  5. Lebar *cluster* akan semakin berkurang apabila jumlah *cluster* dan jumlah sampel yang digunakan dalam proses k-means *clustering* lebih banyak.
- 6. DAFTAR PUSTAKA**
1. Jain, L.C., U. Halici, I. Hayashi, S.B. Lee, dan S. Tsutsui. "Intelligent Biometric Techniques in Fingerprint and Face Recognition." USA: CRC Press. 1999.
  2. Michael J. Lyons, Shigeru Akamatsu, Miyuki Kamachi dan Jiro Gyoba. "Coding Facial Expressions with Gabor Wavelets". Japan: Third IEEE International Conference on Automatic Face and Gesture Recognition, IEEE Computer Society. 1998.
  3. Nalwan, Agustinus. "Seri Pemrograman Aplikasi: Movie and Special Effects". Jakarta: Elex Media Komputindo. 2000.
  4. Neural Networks. <http://scitec.uwichill.edu.bb/cmp/p21h>. Oktober 2001.
  5. NeuroComputing. <http://www.brainstorm.co.uk/NCTT/tech/tb.htm>. Oktober 2001.
  6. Ong Hee Seng. "Recent Advances in Radial Basis Functions Network". <http://www.comp.nus.edu.sg/~inns/OngHeeSeng/rbfNetwork.htm>. Oktober 2001.
  7. Radial Basis Functions (RBFs). <http://www-ra.informatik.uni-tuebingen.de/SNNS/UserManual/node182.html>. Oktober 2001.