Oral Defense

A generic face processing framework: Technologies, Analyses and Applications

Supervised by: Prof. Michael R. Lyu

Presented by Jang Kim Fung

Presentation Outline

- Introduction
  - Background
  - Face processing framework
- Survey about facial feature
- Face Detection Committee Machine (FDCM)
  - Approach and evaluation
- Facial feature localization
  - Algorithms and evaluations
- Face processing system
  - Architecture
  - Implementation and applications
- Conclusion
- Q&A

Introduction – Background (1)

- Facts
  - Human’s information are valuable.
  - Retrieval information from the video source or camera becomes many researchers' interest, especially for human face.
  - Research area include
    - Face detection
    - Face tracking
    - Extraction of facial features
    - Face recognition
  - Lots of work are done by researchers.

Introduction – Background (2)

- Motivations
  - Face processing system can extract rich contents
  - Wide ranges of usage
    - Security and social issues
    - Multimedia indexing
    - Intelligent vision-based human computer interaction.
  - Organization work is needed for the huge research information

Basic Face Processing Framework

- Three important modules
  - Face detection
  - Face tracking
  - Face recognition

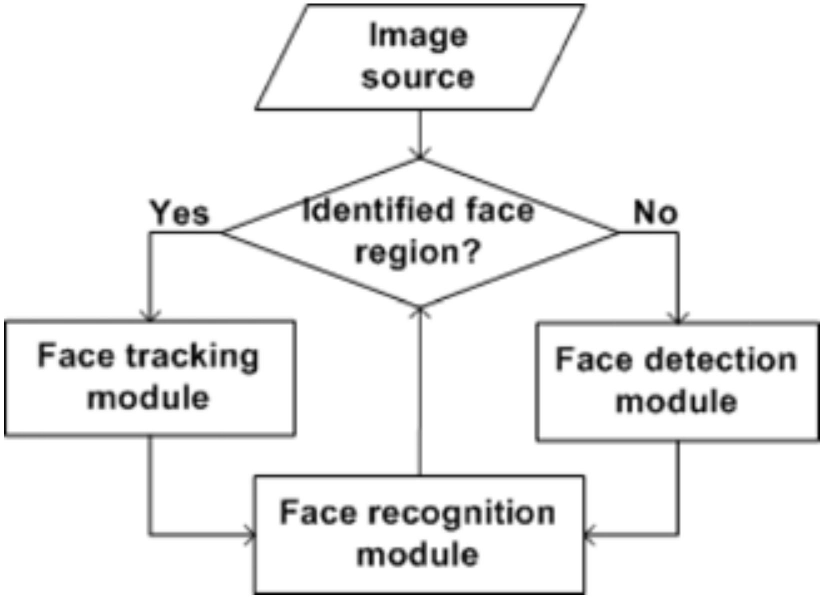

Figure 1: The basic framework for face processing

Facial Feature Analysis

- "Useful" data extracted from the images or image sequences are called ***Facial Features***.
- Facial features are summarized into two main categories
  - Pixel information
    - Apply mathematical operations on the pixel values.
    - Reduce the dimension of data.
    - Not require the pre-defined model or human knowledge.
  - Geometry information
    - Retrieve an intermediate/higher level of facial feature.
    - Employ the characteristic of face to select the features.

Table 1: Category of facial feature - Pixel

		Features
Pixel	Global	Eigenface
		Discrete Cosine Transform (DCT)
		DCT-mod2
		Discriminant Karhunen-Loeve projection (DKL)
		Fisher linear discriminant (FLD)
		Enhanced Fisher linear discriminant (EFC)
		Autocorrelation coefficients
		Direct LDA (D-LDA)
		Direct fractional-step LDA (DF-LDA)
		Log transform
		Gabor + Enhanced FLD (GFC)
		Independent component analysis (ICA)
		Kernel PCA (KPCA)
		Kernel direct discriminant analysis (KDDA)
	Local	Elastic bunch graph
		Eigennose
		Eigeneye
		Eigenmouth
		Eigeneye
		Eigeneyebrow
	Color	PCA in each color channel

Table 2: Category of facial feature - Geometry

		Features
Geometry	Position	Iris center
		Mouth position
		Cheek triangular position
		Nose position
		Vertical position at the eye center
	Shape	Iris radius
		Mouth orientation
		Face elliptical shape
		Eye brow thickness
		Four corner of the mouth
		Eleven radii describing the face chin shape
	Edge	Nasolabial region
		Wrinkle (3 area lateral to eye outer corner)
		Coarse description of brow's arches
		Line Edge Map
	Distance	Distance of two blows
		Mouth width
		Nose width
		Face bigonial breadth
		Face zygomatic breadth
	Motion	Muscle action
		Lips motion

Extraction Technology

Table 3: Coding methods for Face Recognition

Type	Reference	Methods
Linear-based	Turk [122]	Principal Component Analysis (PCA)
	Brunelli [13]	Geometric feature-based matching
	Swets [116]	Discriminant Karhunen-Loeve projection (DKL)
	Belhumeur [7]	Fisher linear discriminant (FLD)
	Yu [132]	Direct LDA (D-LDA)
	Lu [82]	Direct fractional-step LDA (DF-LDA)
	Liu [77]	Enhanced Fisher linear discriminant (EFC)
	Bartlett [5]	Independent component analysis (ICA)
	Goudail [44]	Autocorrelation coefficients
Kernel-based	Eickeler [33]	Discrete Cosine Transform (DCT)
	Sanderson [107]	DCT-mod2
	Kim [63]	Kernel PCA (KPCA)
	Lu [81]	Kernel direct discriminant analysis (KDDA)
	Wiskott [126]	Elastic bunch graph matching (Gabor wavelet)
	Liu [78]	Gabor + Enhanced FLD (GFC)
Edge-based	Takacs [117]	Binary Image Metrics
	de Vel [28]	Random rectilinear line segment
	Gao [41]	Line Edge Map (LEM)
Other	Torres [121]	Skin Color
	Georghiades [42]	Illumination Cone
	Chua [20]	3D Point Signature (PS)

Performance Evaluation

Table 4: Face Recognition results under different database

Method	Database - Recognition rate% (Dimension)					
	ORL	Yale	FERET⋆	UMIST	Bern	AR [87]
Eigenface [122]	93.1(22)	80.6(30)	50(25)	91.5(12)	100(20)	55.4(20)
Fisherface [7]	87.8(22)	99.4(15)	67(25)	94.8(12)	×	×
D-LDA [132]	93.7(22)	×	×	96.5(12)	×	×
DF-LDA [82]	95.7(22)	×	×	97.8(12)	×	×
EFC [77]	×	×	98.5(25)	×	×	×
ICA [5]	×	×	99.8(200)	×	×	×
DCT [33]	100.0	×	×	×	×	×
KPCA [63]	97.5(120)	×	×	86(12)	×	×
KDDA [81]	×	×	×	95(12)	×	×
EBG [126]	×	×	98	×	×	×
GFC [78]	×	×	95(25)	×	×	×
Line segments [28]	99.7∗	×	×	×	99.7∗	×
LEM [41]	×	85.4	×	×	100	96.4

Note: ∗= Combine the ORL and Ben database to form one larger database.

⋆= 600 FERET frontal face images corresponding to 200 subjects.

Evolution and Future Direction

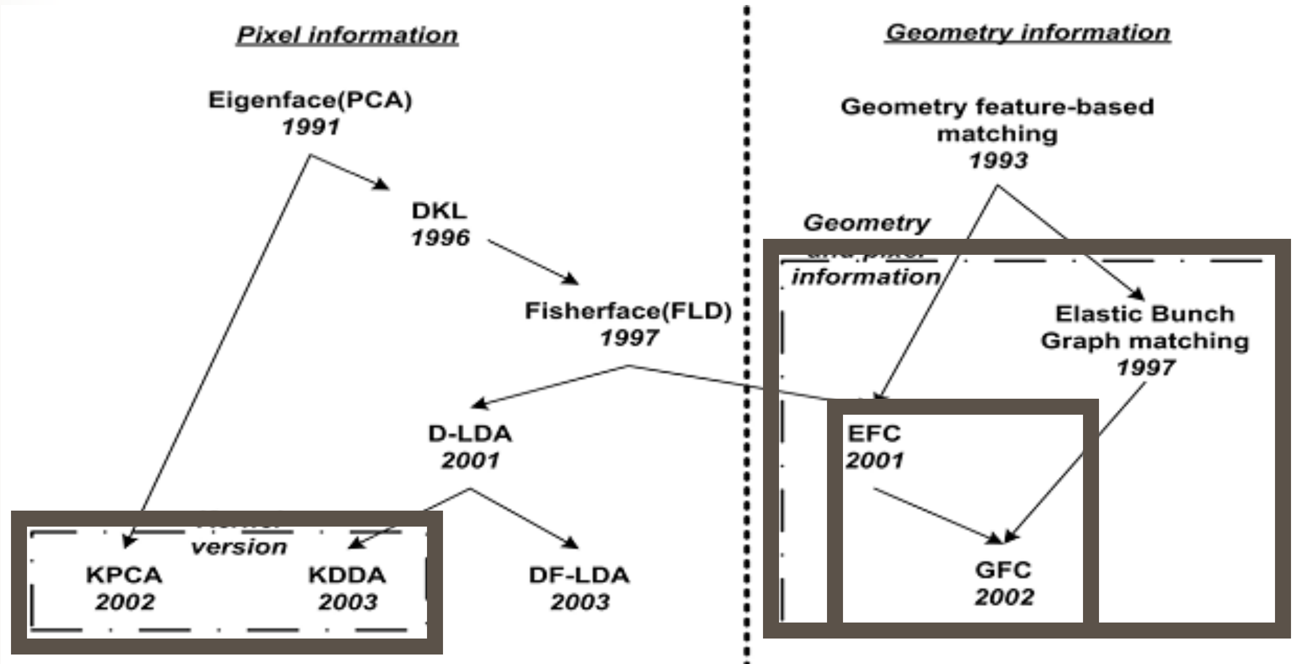

Figure 2: The evolution of the major face recognition approaches

- The trend of the evolution for face recognition
  - Combine the pixel information and geometry information
  - Adopt the deformable model
  - Modify to kernel-based approach

Face Detection Committee Machine (FDCM) - Background

- Face detection approaches
  - Contrasted to the appearance-based methods recently.
  - Rely on statistical analysis and machine learning.
- An ensemble of classifiers has proven to provide a better estimator than the use of a single classifier.
  - Ensemble of neural network
  - Gating network
  - Hierarchical mixtures-of-experts

FDCM – Approach (1)

- We propose the engagement of committee machine with heterogeneous experts:
  - Sparse Network of Winnows (SNoW) algorithm
  - Support Vector Machine (SVM)
  - Neural Networks (NN)

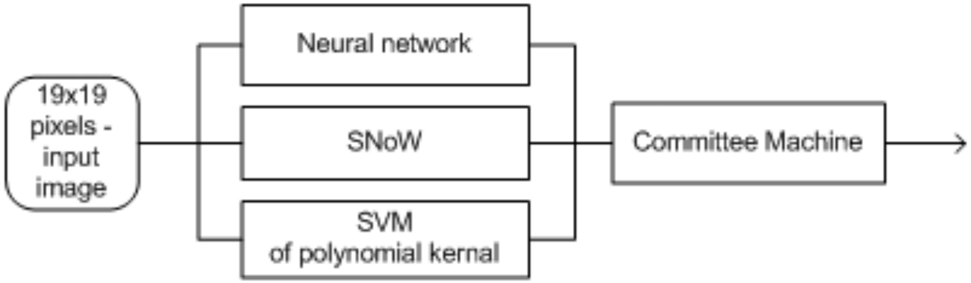

Figure 3: System architecture for FDCM

FDCM – Approach (2)

- Traditional committee machine
  - A combination of homogeneous experts (NNs or RBFs)
  - Trained by different training data sets to arrive at a union decision
- Proposed committee machine
  - A combination of heterogeneous experts
  - Capture more features in the same training data
  - Overcome the inadequacy of each single approach
- Reason for chosen those three approaches
  - Both are statistics approach
  - Operate images without color information
  - No need to use different set of training and testing data

FDCM - Review: SNoW (1)

- Encode the image into a set of active features - Primitive features $= 256 \times (y \times w + x) + I(x, y)$
- The target node *t* is active: $\sum i_j * w_i > \theta_t$

  where w_i is the weight on the edge connecting the *i* th feature to the target node *t*, θ_t is threshold
- Winnow update rule

Prediction	Correct response	update action	update name
1	0	$w_i = \beta * w_i$ **for** $0 < \beta < 1$	demotion step
0	1	$w_i = \alpha * w_i$ **for** $\alpha > 1$	promotion step

FDCM - Review: SVM (2)

- Find a hyperplane that leaves the maximum margin between two classes which will have the smallest generalization error
- The margin is defined as the sum of the distances of the hyperplane from the closest point of the two classes

$$f(x) = sign(\sum_i \lambda_i y_i K(x, x_i) + b)$$

where $K(x,x_i)$ is the kernel function

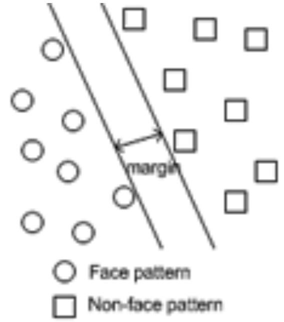

Figure 4: SVM

FDCM - Review: NN (3)

- Back propagation method
  - A number of element simply multiples inputs by a set of weights to calculate the output value
  - Compare the result with a desired response to produce error
  - Adjust the weights to reduce the error

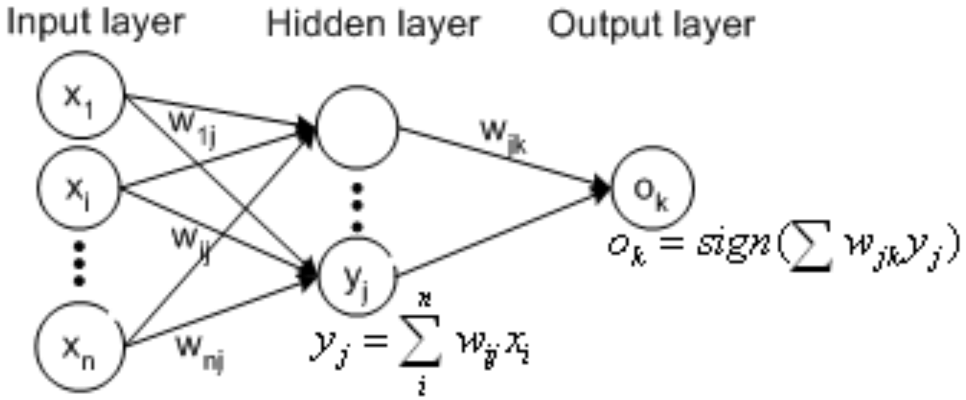

Figure 5: The architecture of the multilayer perceptron

FDCM – Approach (3)

- Based-on the confidence value T_{ij} of each expert j for data i

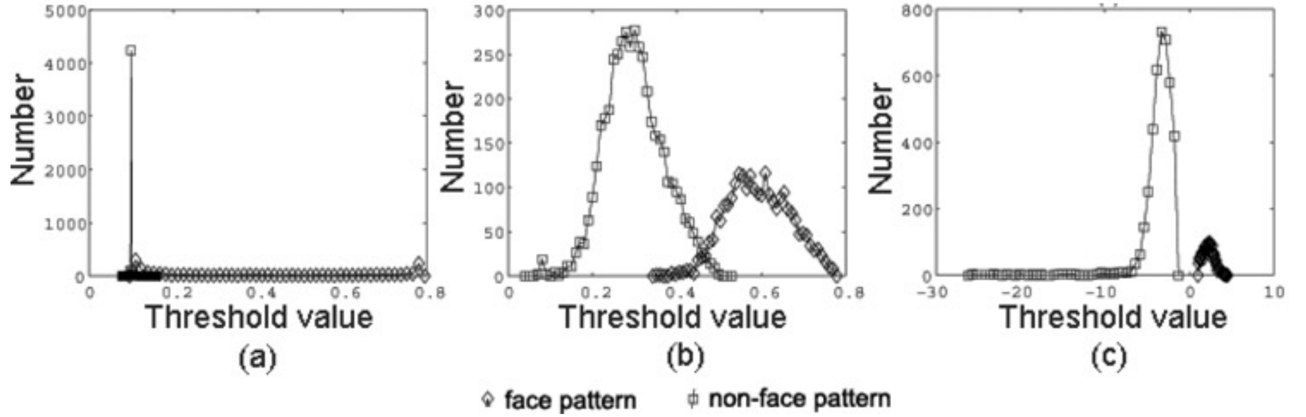

Figure 6: The distribution of confident value of the training data from three different approaches

- The confidence value of each expert are
  - Not uniform function
  - Not fixed output range (e.g. [–1,1] or [0,1])

FDCM – Approach (4)

- Normalization is required using statistics information getting from the training data

$$\alpha_{ij} = (T_{ij} - \mu_j) / \sigma_j$$

where μ_j is the mean value of expert *j* and σ_j is the standard derivation of expert *j*

- The information of the confidence value from experts can be preserved
- The output value of the committee machine can be calculated:

$$\beta_i = \sum_j w_j * (\alpha_{ij} + \sigma_j * \delta_j)$$

where δ_j is the criteria factor for expert *j* and w_j is the weight of the expert *j*

FDCM – Evaluation (1)

- Use the same set of training and testing data to control the condition
- CBCL face database from MIT

Table 5. CBCL face database

	Training Set	Testing Set
Face Pattern	2429	472
Non-face Pattern	4548	23573

Table 6: The best operating point of each approach

	True Positive	False Positive
NN	71.4%	15.2%
SNoW	71.6%	15.1%
SVM	81.2%	13.2%
FDCM	**84.1%**	**11.4%**

FDCM – Evaluation (2)

- Receiver Operating Characteristics (ROC) curves are employed to show the characteristic of each approach

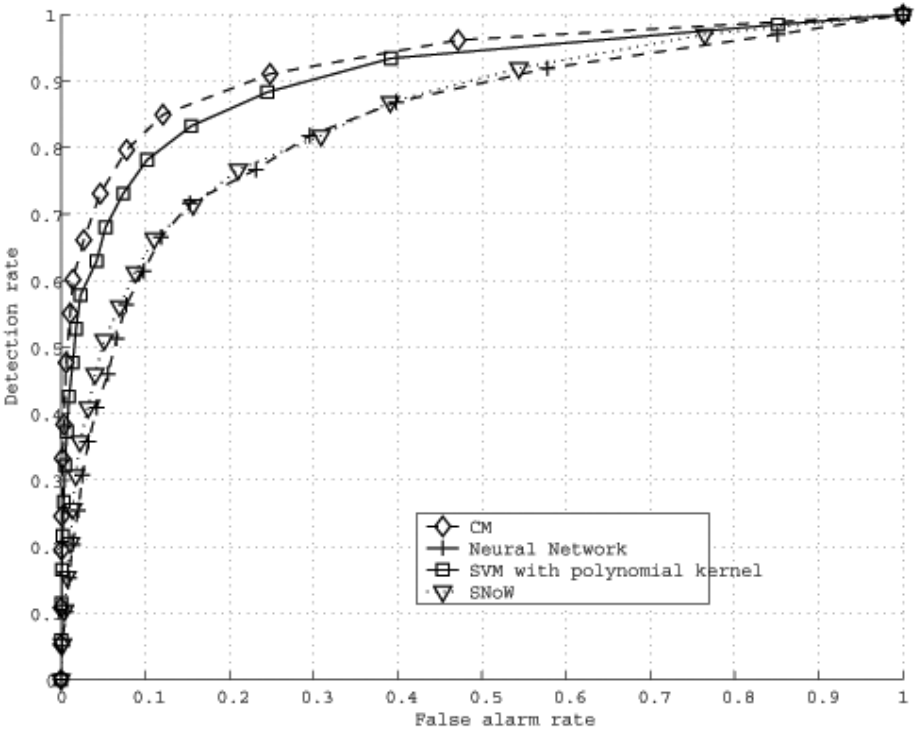

Figure 7: The ROC curves of committee machine and three different approaches

FDCM – Evaluation (3)

Table 7: Experimental results on images from the CBCL

	False Alarm Rate			
Detection Rate	NN	SNoW	SVM	FDCM
10%	0.56%	0.41%	0.05%	0.02%
20%	1.37%	1.09%	0.16%	0.07%
30%	2.54%	1.67%	0.44%	0.14%
40%	4.11%	2.92%	0.83%	0.41%
50%	6.32%	4.91%	1.60%	0.77%
60%	9.47%	8.47%	3.07%	1.41%
70%	13.89%	14.67%	5.98%	3.90%
80%	**26.97%**	**27.62%**	**12.32%**	**7.79%**
90%	48.95%	49.26%	28.60%	22.92%

- FDCM archives lower false alarm rate than the other methods with the same detection rate.
- The false alarm rate of FDCM (***7.79%***) is nearly half of the other approach (***12.32%***).

Facial Feature Localization

- Algorithms
  - Gray-scale image
    - Template matching
    - Separability filter
  - Color image
    - Color information
    - Separability filter

Gray-scale Algorithm (1)

- The cost C of each iris candidates (blob B_i) is defined as

$$C(i) = C_1(i) + C_2(i) + C_3(i) + 1/R(i)$$

$$C_1(i) = \frac{|\eta_{23}(i) - \eta_{24}(i)|}{\eta_{23}(i) + \eta_{24}(i)}, C_2(i) = \frac{|\eta_{25}(i) - \eta_{26}(i)|}{\eta_{25}(i) + \eta_{26}(i)}, C_3(i) = \frac{U(i)}{U_{av}}$$

where η_{xy} is the separability between region R_x and R_y, and R(i) is the value of the normalized cross-correlation result

- The equation of the separability between regions R_1 and R_2 is:

$$\eta = \frac{\delta_b^2}{\delta_T^2}, \delta_T^2 = \sum_{i=1}^{N}\left(P_i - \overline{P_m}\right)^2, \delta_b^2 = n_1\left(\overline{P_1} - \overline{P_m}\right)^2 + n_2\left(\overline{P_2} - \overline{P_m}\right)^2$$

Gray-scale Algorithm (2)

- The position of the iris is selected from the blob with minimum cost value
- Left and right irises are found by using the same method

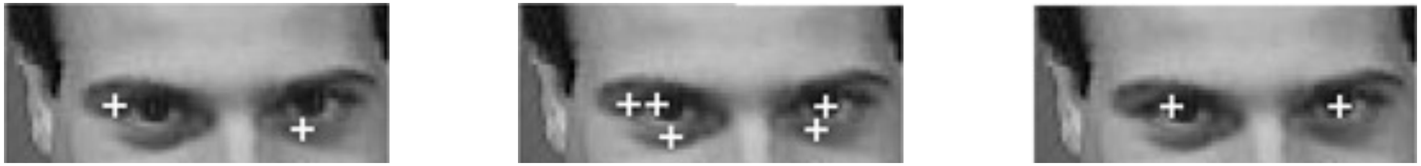

Figure 8: The white cross means (a) the maximum value of normalized cross-correlation, (b) the local maxima of normalized cross-correlation, and (c) the minimum value of the cost function

Color Algorithm (1)

- Selection of eyes candidates
  - Construction of the EYEMAP (EM)

$$EM = ((EC \text{ AND } EC) \text{ AND } EL) \text{ AND } SEP$$

where EC, EL and SEP are the Chrominance eyemap, Luma eyemap and Separability filter map, respectively.

  - Local maxima are selected as eye candidates
- Selection of mouth candidates
  - Construction of the MOUTHMAP (MM)

$$MM = C_r^2 \cdot (C_r^2 - \epsilon \cdot \frac{C_r}{C_b})^2$$

  - Local maxima are selected as mouth candidates

Color Algorithm (2)

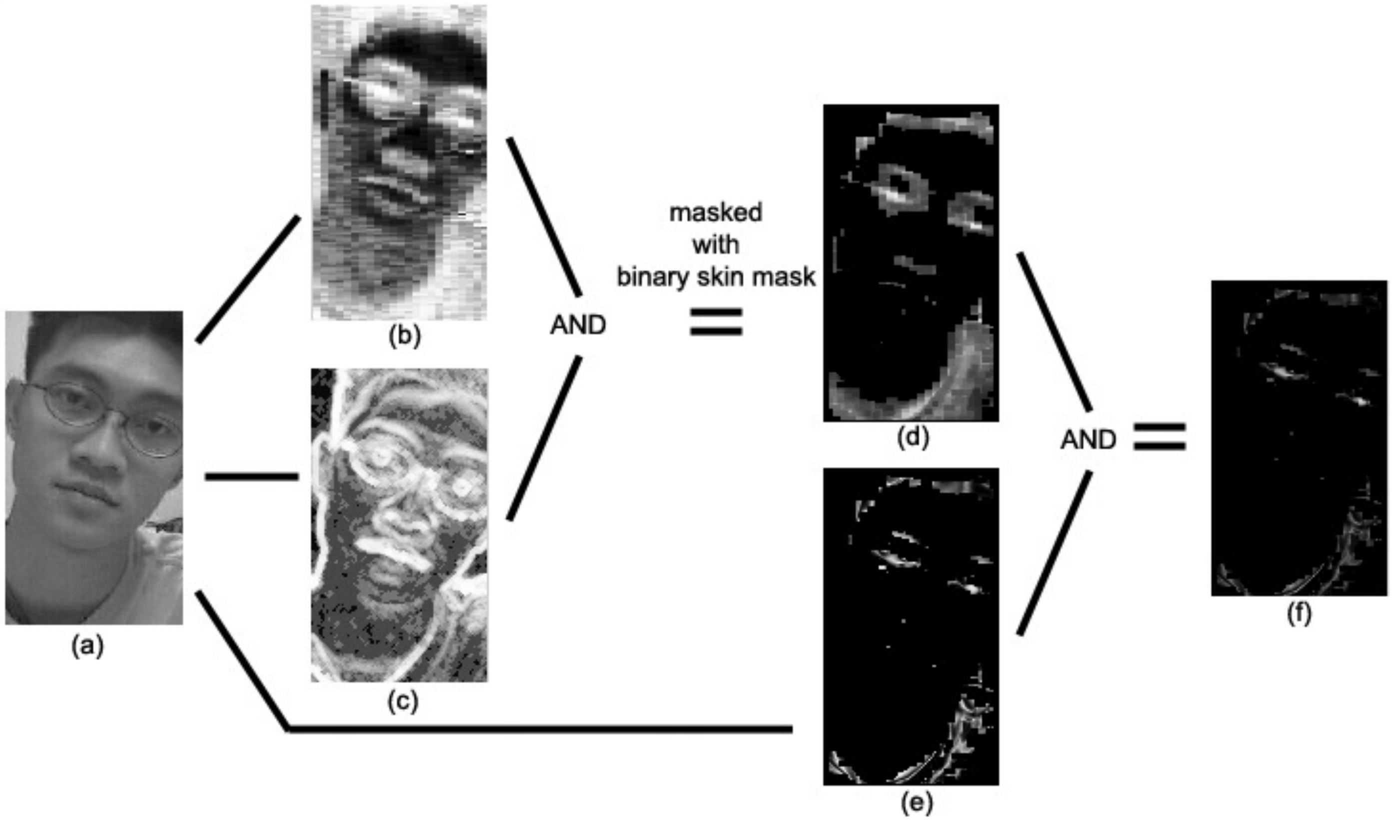

Figure 10: The calculation of EM. (a) The original image, (b) EC after histogram equalization, (c) EL after histogram equalization, (d) the masked intermediate image of ((b) AND (c)), (e) the sepmap SEP, and (f) the resultant eyemap EM.

Color Algorithm (3)

- Selection of face candidates
  - Cost of each eye candidate is calculated by

$$C(i) = EM(x,y) + SEP(x,y) + U(x,y) + V(x,y)$$

where U is the the average intensity of the blob and V is the sum of the value of the blob in EL.

  - Cost function of the face candidate

$$C(i,j,k) = C(i) + C(j) + SVM(i,j) + SYM(i,j)$$

where SVM(i,j) is the confidence value of the SVM classifier and SYM (i,j) is the coefficient of the normalized cross-correlation

Color Algorithm (4)

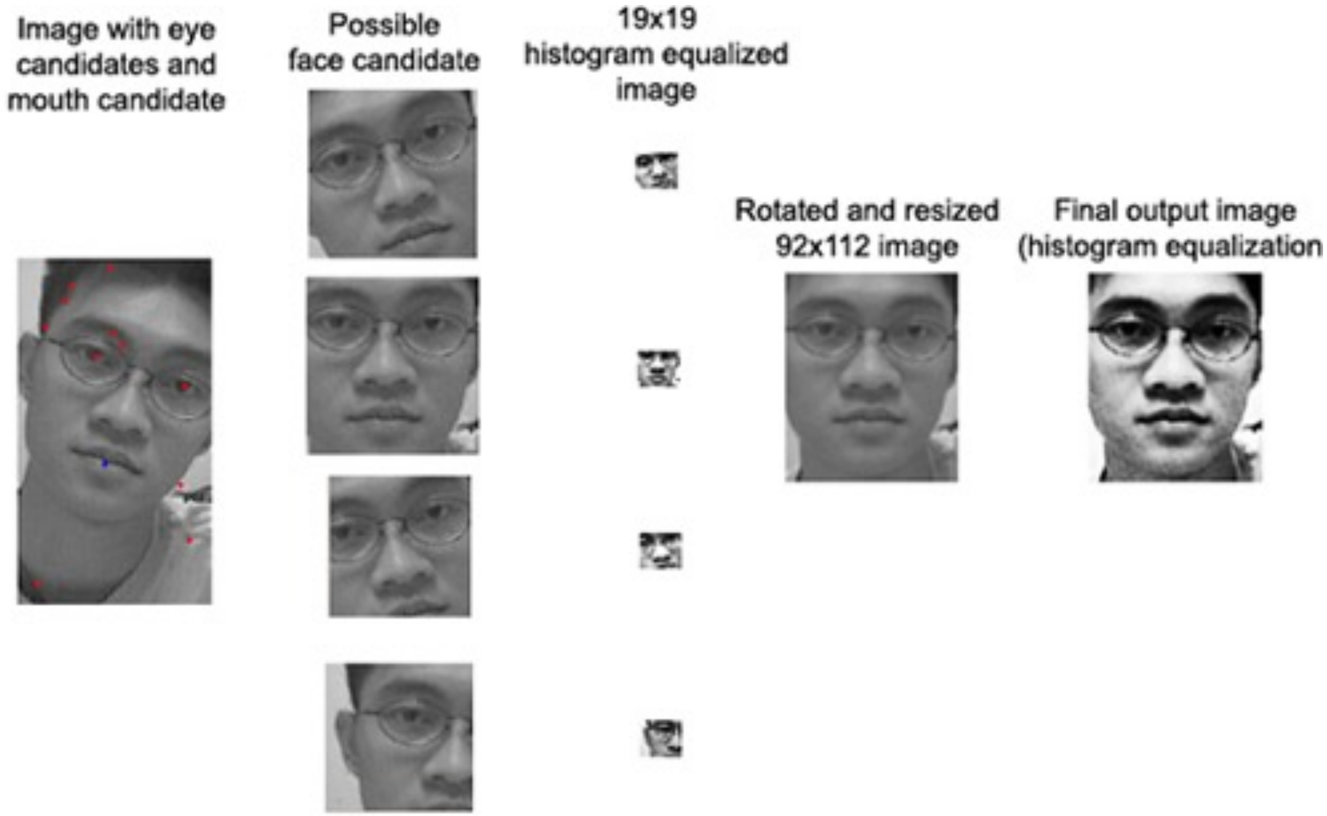

Figure 11: The selection of face candidate and the final output image. (Red cross is eye candidate and blue cross is mouth candidate).

Evaluation – Gray-scale

- ORL face database
  - 40 subjects each with 10 different images

Figure 12: Result of the gray-scale algorithm.

Table 8 : The comparison of the performances of the proposed method, the template matching and the method using Hough transform and separability

Algorithm	Correct rate(%)	CPU time
Proposed method	**88.5**	0.5
Proposed method without separability	84.75	**0.3**
Template matching	63.75	0.6
Hough transform and separability filter method	56.5	19.3

- The separability filter can help to increase the accuracy of finding the position of iris.

Evaluation – Color (1)

- AR face database
  - About 130 subjects that contains two sets of images that are taken in two sessions
  - 4 subsets of the images of each subjects are selected for the experiment
    1. Neutral expression
    2. Smile
    3. Anger
    4. Scream

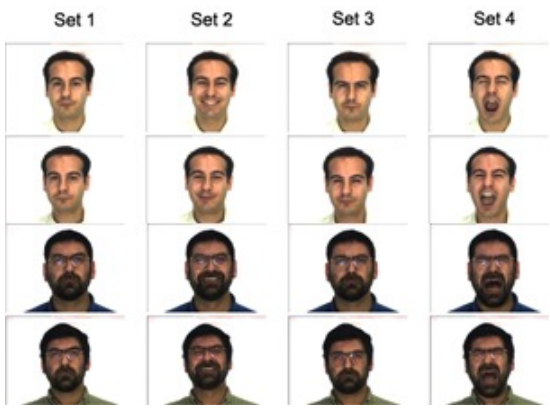

Figure 13: Example of AR database

Figure 14: Example of the normalized image

Evaluation – Color (2)

Table 9 : The results of AR face database

	w and w/o glasses			w/o glasses		
	Set 1-4	Set 1-3	Set 1	Set 1-4	Set 1-3	Set 1
# of images	1,020	765	255	736	552	184
# of errors	220	124	26	97	48	6
Correct rate(%)	78.4	83.8	89.8	86.8	91.3	**96.7**

- Achieves **96.7%** correct rate with neutral expression.
- Reasons for the decreases in accuracy
  - Closed eyes in set 4.
  - Strong reflected light from glasses around the eyes.

Figure 15: Example of the images having strong reflected light

Face Processing System

- Consists of four main components:
  - Pre-processing
  - Face detection
  - Face tracking
  - Face recognition

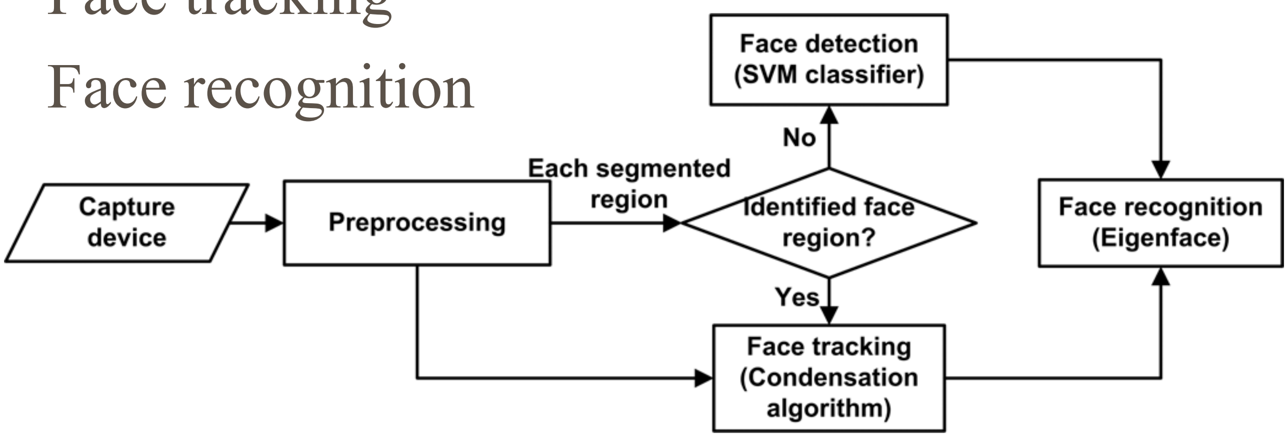

Figure 16: System architecture of the system

Pre-processing Module

1. Employ ellipse color model to locate the flesh color.
2. Perform morphological operation to reduce noise.
3. Apply motion detection to detect the moving person.
4. Apply skin segmentation to find face candidates.

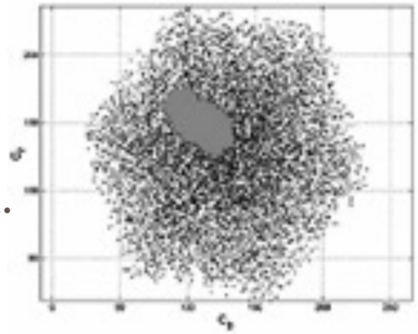

Figure 17: 2D projection in the CrCb subspace

Figure 18: Skin segmentation step (a) original image, (b) binary skin mask, (c) binary skin mask after morphological operation and (d) face candidates

Face Detection Module

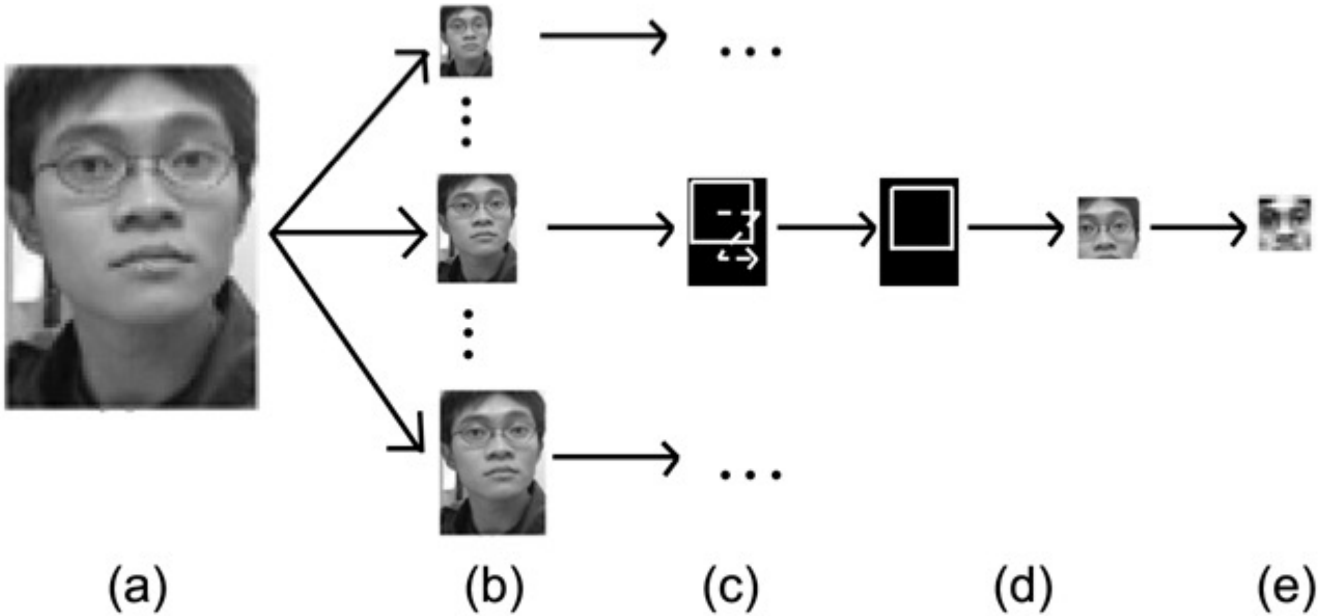

Figure 19: The procedure of face detection: (a) Original face candidate region, (b) resize the region into different scales, (c) the zigzag movement of the 19x19 pixels search window, (d) one of the search window, and (e) histogram equalization on the search window.

Face Tracking Module

- Apply the Condensation algorithm for tracking the detected region.
- Histogram of the Hue channel in HSV color space is defined as the measurement model for the algorithm.

Face recognition module

- Normalization
  - Apply the proposed localization algorithm for color image.
  - Reduce variants in rotation, scaling and pixel value.
  - Return a normalized 92x112 pixels image.
- Recognition
  - Apply the Eigenface method.
  - Extract 30 eigen coefficients from the training image.
  - Employ Euclidian distance for similarity measurement.

Applications

- Recognition system
  - Perform on video conferencing and news report.
  - Record the identity and position of the person on various time.
- Authentication system
  - Perform on door entrance.
  - Verify the identity given by the user.
  - Grant the permission if the result is matched.

Conclusion

- Detailed survey about facial feature.
  - Classification of facial features and the extraction methods.
  - Discussion about the performance, evolution and future direction of face recognition methods.
- FDCM is proposed to enhance the accuracy of classification of face pattern.
  - The false alarm rate of FDCM (7.79%) is nearly half of the best approach (12.32%) among three individual approaches.
- Two localization algorithms for gray-scale and color image are proposed.
  - The accuracy of locating eyes (88.5% and 96.7% for neural expression).
- A face processing system is developed based on the obtained knowledge.

Q&A