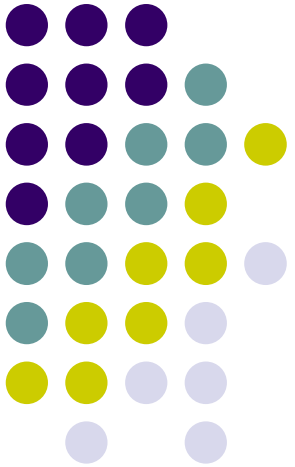
Face Alive Icons

Xin Li
Dept. of Computer Science
University of Pittsburgh

Graduate Research Competition
Feb. 17, 2005



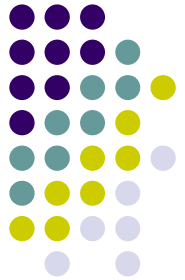
Expressions are Expressive

- Facial expression is one of primary means by which we communicate with each other.
- Nowadays, more and more people are “talking” without seeing each other, so…

  :-) , :-(, :-8, ^_^ , *_^,

- The popularity of these representations has proven the great needs for facial expressions in numerous applications.
- No representations are more nature and friendly than the facial images of the real human.

Main Challenges



- Acquisition: It costs user too much effort to take photos for every expression.
- Transference: Images are too big to transfer.
- Display: Screen is too small, and Images are too large.

Our Goal

- A solution to synthesize realistic facial expression images from photographs for the portable devices such as cell phones and PDAs, which have
  - Limited processing power
  - Limited network bandwidth
  - Limited display area

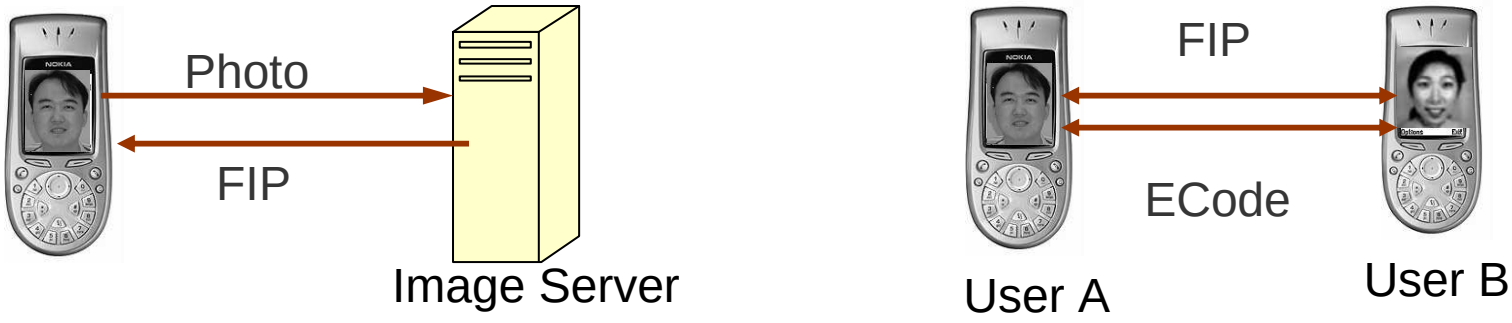
  Referred as----

  the *“LLL”* environment

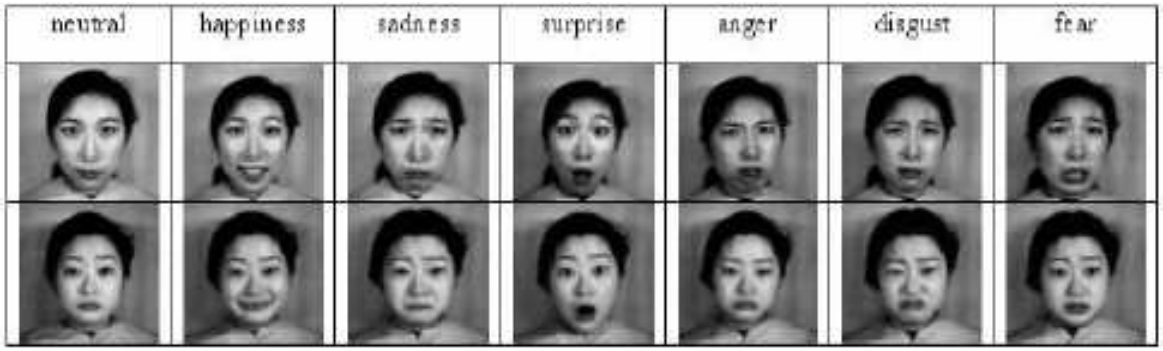
Overview of our approach

- Face Alive Icons (FAI) are small size realistic facial images
- Facial Icon Profile (FIP) is compact in size;
- Expression code (ECode) is just one byte.
- Expression synthesis is a light load operation.

Step	Host	Input	Output
Step1: Expression Decomposition	Servers	Photograph	Facial Icon Profile
Step2: Icon Synthesis	Terminal Devices	Facial Icon Profile	Face Alive Icons



Expression Decomposition



- Observations from Japanese Female Facial Expression (JAFFE) database(**216 images, 10 persons**):
  - **Inactive** and **active parts**
  - Parts share the **same** appearances across expressions
- Decomposition rule FAI:=FF+EF
  - FF: Facial Features, the inactive parts of face in expressions
  - EF: Expressional Features, the active parts of face in expressions
- FAI:=FF+<left eye>+<right eye>+<mouth>

Discrete Model of EF

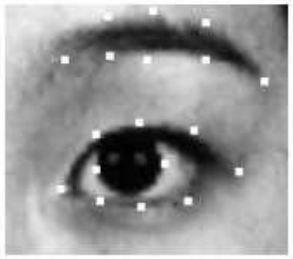


Figure : The Eye Area with 18 Landmark Points

Built through Principle Component Analysis (PCA) –**Distribution** is investigated and **standard states** are defined.

Assuming *n* sample data items in training set and *m* variables for landmark points of each item. The i*th* data item *Xi*

$$X_i = (x_{i,1}, x_{i,2}, \cdots, x_{i,m})$$

Where $x_{i,k}$ could be either the coordinates or grayscales of the landmark points

$$X_i = \overline{X} + P \cdot b$$

In which,

- *X* is the average of the training samples
- *P* is the matrix of unit eigenvectors of the covariance of deviation.
- *b* is a vector of eigenvector weights referred as Model Parameters.

Discrete Model of EF

- Categorized by expressions, the vector *b* of features for e_i can be represented by the averages

$$S = \{\overline{b}_{e1}, \overline{b}_{e2}, \cdots, \overline{b}_{ep}\}$$

- Assuming uniform distribution,

$$\text{delta} = \max |b_i - b_j| / (n-1)$$

- Keep merging the closest items b_p and b_q if $|b_p - b_q| <$ delta
- A unique semantic name is given for each of the standard states according their appearances.

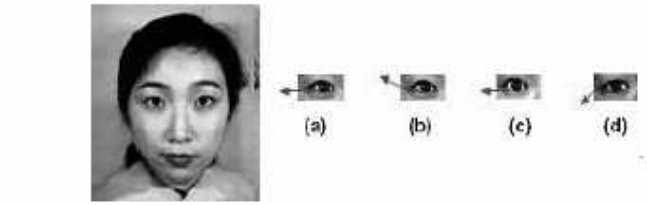


Figure States of the Right Eye: (a) b_1: normal (b) b_2: up (c) b_3: wide-open (d) b_4: down

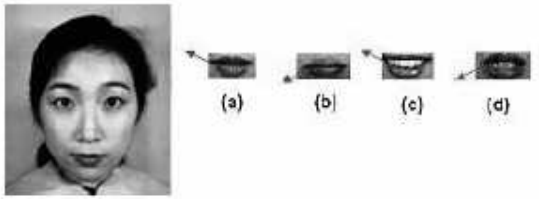


Figure The States of the mouth: (a) b_1: normal (b) b_2: down-close (c) b_3: up-open (d) b_4: down-open

FAI Synthesis

- Synthesis rules are determined by statistical analysis on the distance from the standard states.
- Face Icon Profile (FIP) includes only one facial image and several standard states of expressional features

		NEU	HAP	SAD	SUR	ANG	DIS	FEA
left eye	normal	0.27	1.94	1.65	2.42	3.09	3.32	2.34
	up	2.47	0.45	3.24	2.06	2.21	2.90	2.97
	wideopen	3.00	2.89	2.73	0.75	1.05	2.47	1.30
	down	2.29	3.45	0.96	2.75	4.27	1.23	4.08
right eye	normal	0.33	1.87	1.34	2.36	3.28	3.39	2.35
	up	2.28	0.53	3.57	2.11	2.24	2.99	3.00
	wideopen	2.77	3.08	2.88	0.89	1.11	2.43	1.23
	down	2.06	3.25	0.83	3.07	3.77	1.21	3.92
mouth	normal	0.18	2.35	2.27	3.07	4.01	1.50	2.25
	down-close	2.87	3.45	0.43	2.23	2.25	3.56	1.06
	up-open	3.94	0.94	3.77	0.54	3.17	2.76	3.86
	down-open	3.25	1.95	2.06	2.09	1.23	2.47	3.02

The average distance D distance between training data and the standard states

	left eye	right eye	mouth
NEU	normal	normal	normal
HAP	up	up	up-open
SAD	down	down	down-close
SUR	wide-open	wide-open	up-open
ANG	wide-open	wide-open	down-open
DIS	down	down	normal
FEA	wide-open	wide-open	down-close

Figure The synthesis rules

Prototype System

- An experimental prototype has been implemented.
  - decomposition process: simulated on PC (Window XP)
  - synthesis process: simulated on Palm Zire (Palm OS)
- More than 30 meaningful expressions can be generated.
- Suppose
  - 20 buddies on you list
  - 30 expressions

  Our approach

  FIP size = 20*128K= 2.56 Mbytes

  Approaches which transfer Images in JPEG files:

  Total size = 20 *30* 30K = 18 Mbyes

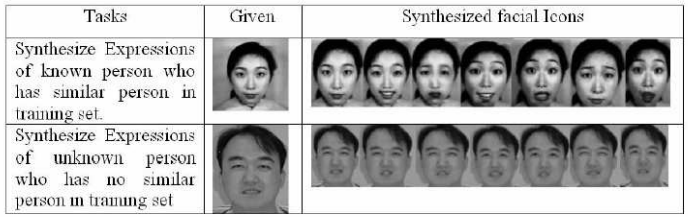


Figure Facial expression synthesis experiments. Facial expression from left to right: neutral, happiness, sadness, surprise, anger, disgust, and fear.

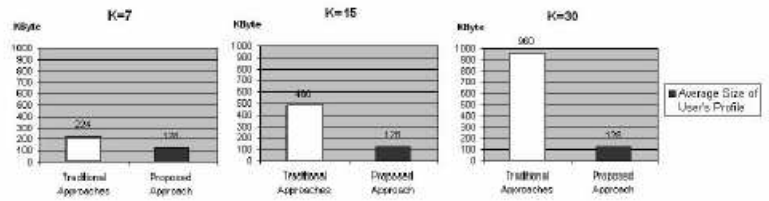


Figure Comparison of the average size of user expression profiles for K expressions per person

Contribution

- Propose an system to synthesize the realistic facial expressions from photographs for portable devices in the “*LLL*” environment.

- Reveal the inner relations among expressions by the semantic synthesis rules assessed through the statistical analysis on training data.

Acknowledgement

- This work is advised by Prof. Shi-Kuo Chang.

- We would like to thank Dr. Lyons who kindly provides the JAFFE database, ChiehChih Chang and Jui-Hsin Huang for his work on implementation of the experimental prototype system, and Dr. Carl Kuo for the valuable comments.

- The work is pending to the patents in Taiwan and US.